AIMS Biophysics

AIMS Biophysics, 12(3): 333−374.
DOI: 10.3934/biophy.2025018
Received: 14 May 2025
Revised: 17 July 2025
Accepted: 22 July 2025
Published: 01 August 2025

http://www.aimspress.com/journal/biophysics

Review

Protein biophysics: current limitations and prospects

Giovanni Colonna*

Unit of Medical Informatics—AOU-L. Vanvitelli, University of Campania, 80138 Naples, Italy

* **Correspondence:** Email: giovanni.colonna@unicampania.it.

Abstract: This review critically examines the current landscape of protein biophysics, highlighting a significant epistemological divide emerging from the rapid advancement of predictive algorithms compared to the slower pace of experimental validation. While artificial intelligence models and inferential statistical methods can generate millions of in silico protein interactions, experimental biophysics remains limited by physiological and technical challenges, leading to a scarcity of confirmed, functional data. This imbalance creates an environment where hypotheses and large-scale interactome networks proliferate based on unverified interactions. This fosters a false sense of causal understanding, potential misdirection in therapeutic development, and overinterpreting omics datasets. The challenge becomes even more urgent with complex, multifactorial diseases such as cancer and viral infections. Recognizing this subtle but crucial disparity is essential for researchers in protein biophysics to navigate the evolving landscape effectively and to focus validation efforts that support meaningful scientific and clinical progress. The review emphasizes the critical need for rigorous experimental validation alongside rapid computational advances. It highlights key challenges, including complexities in protein folding, dynamics, membrane interactions, and quantum effects. It discusses emerging technologies such as AI-based structure prediction (e.g., AlphaFold), cryo-electron microscopy, and integrated biophysical approaches. The importance of multidisciplinary collaboration and orthogonal validation methods is underscored to enhance the reliability of protein interaction data. While technological advances promise to deepen our understanding of protein functions and their roles in health and disease, the review advocates for cautious integration of predictive models with meticulous experimental verification, ensuring the development of accurate, biologically meaningful insights poised to advance medicine and biotechnology.

Keywords: Protein biophysics; protein–protein interactions; orthogonal validation; challenges in multidisciplinary collaborations; the where, how, and when of proteins

1. Introduction

Protein biophysics is a crucial field at the crossroads of biology, physics, and chemistry. It offers the quantitative tools to understand the complex mechanisms that control life. This discipline explains fundamental processes, from how individual polypeptide chains fold to the intricate interactions within cellular machinery, and has significant implications for advances in medicine and biotechnology [1].

Over the past decades, we have made significant progress in determining protein structures and understanding their functions, thanks to technological innovations and theoretical breakthroughs [1,2]. However, despite these achievements, major limitations still exist that hinder a full and dynamic understanding of protein behavior in the complex and constantly changing environment of a living cell [3]. This report seeks to analyze these current challenges in various areas of protein biophysics research and to examine promising opportunities and emerging technologies that are likely to overcome these obstacles, leading to a deeper understanding of the molecular foundations of life.

Understanding the deep molecular mechanisms that cause cancer, viral, and many chronic diseases is the current frontier of human knowledge in the fight against these conditions [4]. One of the few reliable tools for studying metabolism at the molecular level is interactomics of protein–protein interactions. Protein biophysics is essential in this context. However, indirect algorithmic methods are significantly faster and more efficient than experimental approaches. As time progresses, the number of interactions identified through computational methods increasingly outnumbers those obtained through experiments. The issue with indirect methods is that they often contain a high level of misinformation, as they are based on predicted hypotheses that require experimental validation. This situation has created a speculative science filled with misinformation and unverified data, which consequently hinders progress in fields like cancer and viral disease research. [5]. What is the point of developing new techniques if they do not assist with the reckless use of methods that explain deep molecular mechanisms through unvalidated predictions?

Predictive algorithms can generate millions of protein interactions in silico, and the speed difference arises because experimental biophysics relies on human intervention, which is significantly slower and more costly. This discrepancy creates an epistemological divide: while we have an abundance of hypotheses and vast interactome networks, there is a scarcity of confirmations and actual functional studies [6]. This asymmetry is leading to genuine scientific speculation resembling economic speculative bubbles. When articles, grants, and models are based on untested hypothetical interactions, they foster a false sense of causal knowledge, resulting in illusory therapeutic options and overinterpreting omics data. In the context of cancer and viral diseases, with highly multifactorial and dynamic conditions, this issue becomes even more critical. We should question why biophysical techniques are not "keeping pace" with predictive algorithms. The reasons are concrete and significant:

1. Precision experimentation: Validating an interaction requires isolating it, reconstructing it in vitro or in cells, and quantifying its affinity and dynamics.

2. Biological context: Interactions are often transient, conditional, or multiple, making them difficult to reproduce in the laboratory.

3. Instrumental limitations: Even emerging techniques such as cryo-EM, in-cell NMR, and HDX have specific applicability ranges and do not always capture the complexity of the proteome.

All stakeholders must recognize that technologies are tools to limit ambiguity, not shortcuts to finding ambiguity. These tools assist in revealing relevant and reproducible interactions, but do not verify every algorithmic output. Protein biophysics progresses slowly because it is meticulous. It

emphasizes molecular verification over speed, which is crucial for medical progress. If we do not limit the algorithmic creation of unvalidated interactomes, we risk developing a science that is more like science fiction. Progress is the accumulation of data and the extraction of truth from potential. Biophysics functions as one of the defenses working to preserve this vital boundary.

The "proteome'' is a fascinating and complex concept often oversimplified to around 20,000 gene products. However, the true complexity of the human proteome is much greater! It involves a dynamic interplay of alternative splicing, post-translational modifications (PTMs), diverse isoforms, and various states of molecular activity [7]. These factors vary significantly depending on the tissue, cellular environment, and time, generating a huge number of proteoforms [8]. Each proteoform must have a specific location, a method of function, and the right timing to pursue particular functions [7]. Excitingly, the estimated number of unique proteoforms exceeds what standard databases suggest, with some estimates indicating numbers greater than 10^6 up to 10^27. However, we should have an average of 100 proteoforms for each native protein [7]. Let us examine projects like HuRI (Human Reference Interactome) and the BioSNAP (Stanford Biomedical Network Dataset Collection) database. These initiatives gather interactions of biomedical interest by documenting experimentally validated physical interactions. We find that, on average, only about 10% of the collected interactions are validated. Which means that while a small number of interactions meet reliability requirements, they are insignificant when compared to the total number of existing proteoforms. All this illustrates not only the vastness of the proteome but also emphasizes the risks of reducing "the protein" to a simple term, which has led to some speculative bubbles in molecular science. It is essential to recognize the existence and risks of this bubble. If an interaction is not functionally or biophysically validated, it cannot be considered part of reliable molecular knowledge. While it may contribute to a hypothesis, a hypothesis alone is not a sufficient basis for therapy. Factors fueling this bubble include academic pressure to produce extensive data and impressive interactome networks, funding models prioritizing quantity over quality, and the misconception that omic correlations imply biological significance and causation.

2. State of the art in protein biophysics

Protein biophysics is one of the most dynamic and interdisciplinary fields in modern science [9]. Here is an overview of the key points of the current state of the art:

1. Integration of structure, dynamics, and function: Proteins are not static entities; flexibility and conformational dynamics are essential to their function. Next-generation methods (cryo-EM, NMR, HDX, SAXS/SANS) enable the study of proteins under near-physiological conditions.

2. Studying proteins in their native context: Techniques such as in-cell NMR and cryo-electron tomography facilitate in vivo analysis, revealing molecular interactions within the cell. The goal is to overcome the limitations of purified systems and to approach real biological conditions.

3. Emerging technologies and multimodality: New platforms combine multiple methods to reconstruct energetic landscapes and interaction networks. Integrating paramagnetic NMR, spin labeling, and advanced imaging allows investigation of transient interactions and functional states.

4. Intrinsically disordered proteins (IDPs). Thanks to IDPs, the traditional "structure-function" paradigm has been reformulated. These proteins perform structureless functions and are often involved in signaling, regulation, and neurodegenerative diseases.

5. Role of artificial intelligence and computational models: Algorithms like AlphaFold have

transformed structural prediction. Big data analysis in proteomics and biophysics aids in identifying new therapeutic targets and understanding complex protein networks.

In short, protein biophysics today is becoming increasingly interdisciplinary, quantitative, and contextual. It is evolving into a holistic-systemic view of the protein as a dynamic biological actor immersed in a highly variable environment [10]. The biomedical implications are significant. Protein biophysics is making an increasingly profound contribution to clinical research, transforming how we understand, diagnose, and treat diseases. Key areas where the impact is already evident include:

1. Understanding molecular diseases: Biophysical techniques enable us to study pathological protein mutations, such as those involved in cancer, Alzheimer's, Parkinson's, and others. Through structural and dynamic analysis, previously undiscovered pathogenetic mechanisms are identified.

2. Development of targeted therapies (drug design): New technologies facilitate precise modeling of drug–protein interactions. This rational approach helps optimize drug efficacy and reduce side effects.

3. Biomarkers and advanced diagnostics: Changes in protein conformation or molecular dynamics can serve as early biomarkers. Techniques such as MS/HDX and NMR detect altered proteins in blood or tissue.

4. Personalized therapies: Combining protein biophysics with omics and AI approaches enables precision medicine tailored to the individual patient's molecular profile. These methods help identify specific and dynamic therapeutic targets, even in multifactorial diseases.

5. Real-time monitoring: In-cell and real-time NMR techniques provide tools to observe treatment effects directly in cells under physiological conditions. They enable rapid assessment of therapy effectiveness.

Therefore, the future direction of protein biophysics in the clinical field increasingly points toward dynamic molecular medicine [11], shifting from mere observation to active interaction with the target protein.

2.1. Inherent limitations facing protein biophysics

Despite its impressive progress, protein biophysics faces several fundamental issues and hurdles. Here are some thoughts on the field's inherent limitations:

1. Complexity of biological systems: Proteins operate within intricate cellular environments, where many variables are beyond experimental control. Cells are complex systems, making precise predictions inherently difficult. Consequently, accurate reconstructions of molecular mechanisms are often incomplete or probabilistic [12].

2. High costs and extended timelines: Techniques like cryo-electron microscopy (cryo-EM), advanced nuclear magnetic resonance (NMR), and electron tomography require significant financial resources and specialized expertise. This complexity can result in lengthy data collection and analysis, slowing progress and complicating widespread clinical use.

3. Difficulties in data interpretation: Results from these techniques are complex and multidimensional [13]. Substantial computational effort is usually needed to interpret the data, which can introduce errors or biases.

4. Technological and solution constraints: Certain methods may not be suitable for studying highly dynamic, disordered, or low-concentration proteins. Consequently, artificial in vitro conditions might alter the proteins' natural behavior [14].

5. Risks of overinterpretation: Pursuing predictive models or "virtual" structures can lead to premature conclusions. Enthusiasm for tools like AlphaFold, without experimental validation, can create a false sense of understanding. It is important to remember that proteins are dynamic, utilizing their intrinsic flexibility for functions; AlphaFold models depict static snapshots.

6. Interdisciplinary barriers: Effective collaboration between physicists, biologists, chemists, and clinicians is vital in protein biophysics. However, cultural and linguistic differences across disciplines can obstruct collaborative progress.

Concluding, protein biophysics is a promising yet sensitive field that requires rigor, balance, and critical thinking. The goal is not merely to understand life at the molecular level but to do so systematically and ethically.

3. Advances in protein biophysics

3.1. Advances in understanding protein dynamics and function

Analyzing protein dynamics involves a range of sophisticated techniques. Single-molecule FRET (smFRET) allows researchers to examine conformational changes and interactions at the individual molecule level, offering insights into dynamic processes that are often hidden in ensemble measurements [15]. High-speed atomic force microscopy (HS-AFM) enables real-time visualization of protein dynamics, capturing conformational changes and interactions as they happen [16,17]. NMR spectroscopy is a powerful tool for characterizing protein dynamics across different timescales, providing detailed information about backbone and side-chain motions [18]. Computational methods and molecular dynamics simulations facilitate the modeling and analysis of protein motions and conformational changes at atomic resolution, offering a theoretical framework for interpreting experimental data [19,20]. This combination of experimental and computational approaches gives a comprehensive view of the complex dynamics that regulate protein function.

There is a growing emphasis in the field on understanding the complex relationships between protein structure, dynamics, and function, moving beyond the traditional focus on static structures. New computational methods are quickly and cost-effectively predicting protein dynamics, offering valuable tools for pharmacology and drug discovery by showing how proteins change shape during cellular processes [21,22]. Integrative approaches that combine evolutionary analysis with studies of protein dynamics [23] provide deeper insights into how mutations impact protein function, as seen in research on viral variants [24]. These advancements highlight the increasing recognition that protein flexibility and movement are essential to their function and are vital for developing a comprehensive understanding of biological systems.

3.2. Cutting-edge techniques for functional characterization

Characterizing protein function involves various innovative biophysical techniques. Native mass spectrometry (nMS) enables the study of protein complexes and their interactions under near-native conditions, providing information about stoichiometry and binding partners [25]. Surface plasmon resonance (SPR) is a label-free technique used to analyze biomolecular interactions in real-time, offering kinetic and affinity data [26]. Isothermal titration calorimetry (ITC) measures the thermodynamic parameters of biomolecular interactions, giving insight into the forces driving binding [27].

Fluorescence spectroscopy remains a versatile method for examining protein conformation, dynamics, and interactions, often used to explore local environments and conformational changes [28]. These techniques, together with others like enzyme activity and cell-based assays, form a comprehensive toolkit for understanding the diverse functions of proteins.

3.3. Understanding cell migration mechanisms at the atomic level

Recent research has made a breakthrough in understanding how cells move by revealing the atomic-level assembly of filopodia, finger-like protrusions crucial for cell movement and linked to cancer spread [29,30]. Using advanced cryo-EM and sophisticated computational image analysis, scientists have shown the flexible and adaptive way that the protein fascin bundles actin filaments to create these structures [30,31]. This discovery solves a long-standing puzzle in the field and has important implications for developing more effective cancer treatments by targeting fascin's role in cell migration. Being able to visualize such a complex protein assembly at the atomic level [32] marks a significant technological advance and opens the door for future studies of complex cellular components.

3.4. Progress in single-molecule protein analysis

Single-molecule techniques, including Förster resonance energy transfer (FRET), optical tweezers, and atomic force microscopy (AFM), play a vital role in protein biophysics by enabling the study of individual protein molecules. These methods reveal the heterogeneity and dynamic behaviors often hidden in ensemble averaging experiments [33−35]. Advances in fluorophore technology and customized labeling strategies improve the sensitivity and resolution of single-molecule fluorescence microscopy, allowing for more accurate measurements of biomolecular interactions and dynamics [36]. Single-molecule force spectroscopy techniques reveal the mechanical properties of proteins and their conformational changes under force, clarifying their stability and function [37]. These single-molecule approaches are essential for moving beyond population averages and uncovering the inherent diversity and dynamic nature of protein behavior at the most basic level. Ongoing developments in these tools and methods will broaden the range of questions that can be addressed, from detailed conformational dynamics to interactions within complex biological systems.

3.5. Computational methods and AI-driven breakthroughs

Computational methods, those leveraging the power of machine learning and artificial intelligence (AI), are revolutionizing various aspects of protein biophysics research [38]. AI algorithms, such as AlphaFold, have achieved remarkable accuracy in predicting protein structures from their amino acid sequences, addressing a long-standing challenge in the field [39]. Beyond structure prediction, AI is also applied to predict protein–protein interactions, design novel proteins with desired functions, and analyze the vast datasets generated from biophysical experiments [40−42]. Machine learning techniques are also being used to interpolate and extrapolate RNA fitness landscapes based on limited experimental data, providing a more comprehensive view of sequence-function relationships [43]. AI-powered methods are being developed to predict not only static structures but also protein dynamics and the ensemble of conformations that a protein can adopt, addressing a key

limitation of earlier prediction methods [44]. Applying AI is proving to be a transformative force in protein biophysics, enabling researchers to tackle intractable problems and opening up new avenues for discovery and protein engineering. AI is also being employed to predict protein interactions and to design new proteins with specific functions, opening new avenues for therapeutic development and other applications [46−48].

4. Advancements in structural biology techniques: future prospects and emerging technologies in protein biophysics

4.1. Advanced techniques for structure determination

Structural biology continually advances, with key techniques that address many current limitations [49−51]. Today's state-of-the-art protein structure determination depends on a powerful combination of advanced experimental methods. Cryo-electron microscopy (Cryo-EM) consistently improves its resolution, automation, and sample preparation techniques. This progress broadens its ability to study various proteins and complexes, including those in different dynamic states and challenging targets like membrane proteins and spliceosomes [52]. Nuclear magnetic resonance (NMR) spectroscopy remains essential, offering unique insights into protein structures and dynamics in solution, especially for smaller proteins [50]. This data complements information from other structural methods. Time-resolved crystallography, which employs intense, ultrashort pulses generated by X-ray free-electron lasers (XFELs) and advanced synchrotrons, provides exceptional temporal resolution [53]. This technique enables researchers to capture dynamic processes, such as enzyme catalysis and conformational changes at the molecular level [54].

The growth of integrative structural biology involves combining data from multiple experimental techniques, such as Cryo-EM, X-ray crystallography, NMR, and mass spectrometry, with computational modeling [55]. This approach promotes a more comprehensive understanding of complex biological systems in their cellular environments. These improvements allow researchers to study intricate and dynamic systems more precisely and under physiologically relevant conditions. Integrating experimental methods with computational modeling provides a powerful way to determine complex protein structures. Highly accurate AI tools like AlphaFold represent a breakthrough in protein science, offering detailed structural information and transforming how biologists analyze proteins [55]. These advancements also facilitate the prediction of protein interactions and the design of new proteins, opening new possibilities for therapeutic development and other applications.

4.2. Emerging biophysical methodologies

Several emerging biophysical methodologies will enhance the field. Mass photometry, a new technique, allows for the characterization of biomolecules, including proteins and mRNA, at the intact level by measuring their mass through light scattering [56]. This method offers insights into biomolecule size, stoichiometry, and interactions, even for large and heterogeneous samples [57]. There is also a growing focus on automation and miniaturization in biophysical techniques, with high-throughput methods and microfluidics being explored to improve the efficiency and capacity of sample analysis [58−60]. Advances in mass spectrometry, such as charge detection mass spectrometry, enable the characterization of large and complex biomolecules, including next-generation bio-therapeutics [61].

New tools and approaches are being developed for high-throughput and comprehensive analysis of protein stability, a crucial parameter for both fundamental research and pharmaceutical development [62]. These emerging methodologies expand the biophysical toolkit, allowing researchers to study challenging systems and properties more efficiently and in greater detail. However, other technologies are still emerging and providing support for advanced protein structural and functional biophysics. Let us briefly examine them.

Electron cryo-tomography (ECT or Cryo-ET) is an imaging technique that enables the 3D reconstruction of biological samples, such as macromolecular complexes, organelles, or even entire small cells, in a near-native, vitrified state [63]. ECT provides unprecedented insights into the cell's "molecular sociology", allowing visualization of biomolecules within their native cellular environment without extraction or crystallization. It determines the 3D organization of macromolecular complexes, organelles, and their interactions within cells, capturing different conformational states of dynamic biological systems. This technique can visualize complexes within their crowded cellular environment at a molecular resolution of approximately 1–4 nm.

SAXS/SANS (small-angle x-ray scattering/small-angle neutron scattering) are complementary small-angle scattering techniques used to analyze the overall size, shape, and quaternary structure of biological macromolecules (such as proteins, nucleic acids, and complexes) in solution, as well as their interactions and dynamics [64,65]. They examine structures at a mesoscopic scale, typically 1–100 nm. SAXS commonly determines the overall shapes of proteins and nucleic acids, oligomerization states, and conformational changes. SANS is especially suitable for studying large assemblies, membrane proteins (by matching out lipids), protein–ligand interactions, and hydration layers around biomolecules. Both techniques study molecules in solution under near-physiological conditions and can characterize flexible and heterogeneous systems [64,65]. SANS provides unique contrast variation capabilities. However, they offer low-resolution structural information, lacking atomic detail. They may face challenges such as aggregation issues and require specialized facilities, like synchrotrons for SAXS and neutron sources for SANS.

In-cell NMR is an NMR spectroscopy technique that studies the structure, dynamics, and interactions of biomolecules inside living cells [66]. This method overcomes the limitations of traditional in vitro NMR, where scientists examine molecules in isolated, highly purified solutions, often far from their physiological environment. Using this technique, we can gain insights into native environment studies: post-translational modifications, drug delivery and target engagement, protein–protein interactions (identifying and characterizing transient or weak interactions within the cellular milieu), and protein folding or misfolding [67]. It provides atomic-level information in a near-native setting and allows direct observation of biomolecular processes within living cells [68]. However, cellular crowding causes signal broadening and sensitivity issues, and the technique is limited to relatively small-to-medium-sized proteins or highly soluble domains.

Real-time NMR (RT-NMR) uses NMR spectroscopy to observe biochemical or biophysical processes as they happen. This allows for direct monitoring of reaction kinetics, conformational changes, or interaction dynamics [69]. NMR spectroscopy supports the study of enzyme kinetics, tracking enzymatic reactions, identifying transient intermediates, and examining protein folding and unfolding. It also helps analyze the kinetics of folding or unfolding pathways, ligand binding, and metabolic activities. This technique provides high-resolution, real-time insights into dynamic processes and captures often difficult-to-detect transient intermediates [69]. However, it is best suited for relatively fast processes that occur within the NMR acquisition time, while low-concentration

species or rapid events might pose sensitivity challenges.

Paramagnetic NMR/spin labeling for studying dynamics and transient interactions is a technique that combines NMR spectroscopy with the introduction of paramagnetic centers (unpaired electrons) into biomolecules [70]. The presence of these paramagnetic centers significantly alters the NMR signals of nearby nuclei, providing unique long-range structural and dynamic information that is difficult to obtain with conventional NMR [71]. It enables the study of long-range distance restraints (overcoming the distance limitations of traditional NMR) for determining structures of larger proteins and complexes, examining dynamics, mapping interaction interfaces (identifying binding sites and transient interactions), characterizing disordered regions, and exploring the structure and dynamics of membrane proteins, where crystallization or high-resolution NMR can be challenging [72]. This technique is sensitive to weak and transient interactions and can reveal low-population conformational states. Some limitations are noteworthy: it requires site-specific introduction of the spin label, often through cysteine mutations, and data interpretation can be complex, especially for highly dynamic systems [73].

HDX (hydrogen-deuterium exchange) detected by MS and NMR is a biophysical technique that measures the exchange rate of backbone amide hydrogens with deuterons from the solvent [74]. The solvent accessibility and hydrogen bonding of these hydrogens influence this rate, providing insights into protein conformation, dynamics, and interactions. Either mass spectrometry (MS) [75] or nuclear magnetic resonance (NMR) [76] can detect HDX. Researchers use HDX-MS to study conformational changes, responses to ligand binding, protein–protein interactions, protein folding, stability, and epitope mapping (identifying regions of a protein that interact with antibodies). They also use it to investigate protein dynamics by examining the intrinsic flexibility and solvent accessibility of various protein regions [77]. Researchers employ HDX-NMR to analyze residue-specific dynamics, hydrogen bond stability (assessing the stability of individual hydrogen bonds in protein secondary structures), folding intermediates, and allosteric effects. HDX-MS can study large proteins and complexes but requires relatively small sample amounts. At the same time, HDX-NMR offers atomic-resolution insights into dynamics and hydrogen bonding, complementing HDX-MS by providing more detailed, site-specific information. Both techniques require careful control of experimental parameters. Data analysis can be computationally intensive for complex datasets, and detailed experiments may be time-consuming [75,76]. Combining these biophysical techniques provides powerful, complementary methods to study the complex world of biological macromolecules, transitioning from static structures to dynamic processes in more native environments.

4.3. Advances in membrane protein biophysics

Membrane proteins, which make up a significant portion of the proteome and are crucial for various cellular processes such as transport, signaling, and maintaining cellular integrity, present unique challenges to biophysical studies [78,79]. Their inherent hydrophobicity, required for embedding within the lipid bilayer, makes them difficult to express, purify, and stabilize outside their native environment [78]. Getting sufficient quantities of pure and functional membrane proteins for biophysical analysis remains a major hurdle [80]. Traditional techniques like X-ray crystallography often struggle with membrane proteins because of the difficulty in forming well-diffracting crystals [80]. While cryo-electron microscopy (cryo-EM) has emerged as a powerful tool for studying these proteins [81], it can face limitations with smaller membrane proteins [82]. The size of the protein-

membrane complex often restricts NMR spectroscopy, another valuable technique. Mimicking the native lipid environment in vitro, which is crucial for maintaining the structure and function of membrane proteins, remains a significant challenge [83,84]. The dynamic nature and potential heterogeneity of membrane proteins also complicate their analysis [85,86].

Advancements across various biophysical techniques [80,86] revolutionized the study of membrane proteins. Cryo-EM has become powerful for determining high-resolution membrane proteins and complex structures, often in near-native states, overcoming limitations of X-ray crystallography. Detergent-free methods for extracting and stabilizing membrane proteins, such as using nanodiscs with a wide variety of lipids and apolipoproteins, are improving the quality and functional relevance of samples for biophysical studies [86]. Integrative approaches combining structural biology, computational modeling, and functional assays provide deeper insights into the dynamics and interactions of membrane proteins within lipid bilayers. AI-based protein structure prediction tools, including those designed for membrane proteins, are accelerating the discovery and characterization of these challenging molecules [87]. Emerging techniques like mass photometry and advanced fluorescence microscopy provide a more comprehensive understanding of membrane protein behavior (other details in § 4.2).

4.4. Exploring quantum mechanical effects in proteins

Applying quantum mechanics to understand protein function is an emerging and exciting frontier in biophysics [88,89]. Researchers are exploring phenomena such as quantum tunnelling in enzyme catalysis, quantum coherence in energy transfer processes within proteins, and the potential role of quantum entanglement in biological systems [90,91]. Advances in experimental techniques, including ultrafast spectroscopy and single-molecule manipulation, are enabling the investigation of these subtle quantum effects [92,93]. Computational methods, like hybrid quantum-classical approaches, are being developed to model and simulate quantum phenomena in proteins, providing insights into their functional significance [94]. The field of quantum biology seeks to uncover how these complex quantum mechanical effects might contribute to the remarkable efficiency and specificity observed in biological processes involving proteins [95]. Quantum biology stands as one of the most interdisciplinary fields ever created. Protein quantum biophysics research will use quantum computing for protein structure prediction, leading to innovative projects [96,97]. Combining AI and biophysical modeling will develop novel approaches to improve protein design for therapeutics such as antibodies and vaccines [86,87]. Although predicting single amino acid changes remains challenging, ongoing research continues to improve methods applicable to vaccine development [96]. Future studies in this area aim to reveal new mechanisms behind protein function and inspire advancements in fields like medicine and materials science.

5. Current limitations in protein biophysics research

5.1. Challenges in deciphering the protein folding problem

The protein folding problem, which involves understanding the thermodynamic principles that control a protein's native structure, explaining the kinetic pathways it takes to reach this state, and predicting its three-dimensional structure from its amino acid sequence, remains a key challenge in

biophysics [98,99]. However, significant difficulties still exist with larger, multi-domain proteins or those embedded in cell membranes [98−100], even though researchers have made notable progress in predicting the structures of smaller proteins.

The cellular environment, characterized by a high concentration of macromolecules, significantly influences the folding process, which is often difficult to replicate and study in vitro [101,102]. The many conformations, even for a medium-sized protein as described by Levinthal's paradox [103], highlight the complexity of folding. This shows that proteins must follow complex pathways to reach their native state within biologically relevant timeframes. We still lack a full understanding of these pathways and the factors that drive them [101,102]. As a result, achieving high accuracy, especially for complex systems, remains challenging [104] even though computational methods have made substantial progress in predicting protein structures [105]. This advancement will persist until training data encompasses a broader range of proteins and AI's role expands from prediction to the actual study of protein folding.

5.2. Limitations of current biophysical techniques for protein analysis

Investigating proteins depends on various biophysical techniques, each with strengths and limitations [106]. Cryo-electron microscopy (cryo-EM) has become a groundbreaking tool for determining the structures of large biomolecular complexes, often at near-atomic resolution [107]. However, it can sometimes offer lower resolution compared to X-ray crystallography, and issues related to sample preparation—such as protein adsorption and denaturation at the air–water interface—can still affect the quality of results [108−110]. X-ray crystallography, a core technique in structural biology, requires proteins to form well-ordered crystals, a condition that not all proteins meet, especially those that are fluctuating or membrane-associated [111]. Smaller and medium-sized proteins are suitable for nuclear magnetic resonance (NMR) spectroscopy, a non-invasive method that provides valuable information about protein dynamics in solution [112,113]. Studying membrane proteins and receptors, key drug targets, poses significant challenges for many biophysical measurements because of their hydrophobic nature and the need to maintain a lipid environment mimicking biological membranes [114,115]. Fully characterizing heterogeneous complexes and bio-therapeutics remains difficult, as current analytical methods often cannot provide a complete picture [116]. The limitations inherent to each technique highlight the importance of integrated approaches to achieve a more comprehensive understanding of protein behavior. Reproducing the native cellular environment in vitro, especially for techniques sensitive to buffer conditions and molecular crowding [117], remains a persistent challenge. In Table 1, we compare some fundamental methods and their key strengths and limitations.

Table 1. Strengths and limitations of current biophysical techniques.

Technique	Key strengths	Key limitations
Cryo-EM	Large, dynamic, membrane-associated molecules can be resolved at near-atomic resolution.	Can have lower resolution than X-ray crystallography, specialized equipment, and sample preparation required.
X-ray crystallography	High resolution, well-developed, and widely adopted.	Not all proteins crystallize well; not suited for large, dynamic, or membrane-associated molecules.
NMR spectroscopy	Non-invasive, provides dynamic information, well-suited for small/medium proteins.	Limited to smaller-to-medium-sized molecules.
Mass spectrometry	The method features high sensitivity, high specificity, high throughput, and provides both qualitative and quantitative data.	There is no single method for identifying all proteins; hydrophobic and basic peptides pose challenges.
Single-molecule FRET	Studies individual molecules, reveals heterogeneity and dynamics.	Can be complex to establish and interpret; limited distance range.

5.3. Difficulties in studying protein dynamics

A significant frontier in protein biophysics is the comprehensive characterization of protein structural dynamics, encompassing the wide range of motions from rapid local fluctuations to slower global conformational changes [106,118]. Many current computational tools excel at predicting static protein structures but often cannot capture the inherent dynamic nature of these molecules, which is essential for their biological function [119]. Understanding how these observed dynamics relate to specific biological functions and mapping the underlying energy landscapes that govern these movements are complex endeavors. Modeling the unfolded state of a protein and characterizing the ensemble of conformations it adopts along the folding pathway also remain open questions in the field [120]. The effects of macromolecular crowding, phase transitions, and post-translational modifications on protein dynamics, as part of the intricate intracellular environment, remain unexplained [121,122]. The field is transforming "dynamic structural biology", recognizing the critical role of protein motion [123]. However, the tools and methodologies required to capture and interpret this dynamism [124], especially within the complex cellular context, are still under development. Deciphering the energy landscape that dictates these dynamics is also a significant challenge in accounting for entropic contributions and the complexities introduced by the cellular milieu.

5.4. Challenges in predicting protein interactions and assemblies

Predicting how proteins interact and form larger assemblies is crucial for understanding cellular

processes, yet it remains a significant challenge even with recent advancements in AI-driven structure prediction [125]. While tools like AlphaFold have shown remarkable success in predicting individual protein structures, modeling the interfaces between interacting proteins and predicting the structures of protein complexes—especially those involving substantial conformational changes upon binding—is more difficult [126,127] because it requires knowledge of structural dynamics. Current AI-based methods often struggle with predicting multi-chain assemblies, only providing a snapshot of the structure. They lack information about ligands, ions, and cofactors that are usually essential for function. Capturing the dynamic nature of protein–protein interactions, including conformational changes during binding, presents a significant challenge for computational approaches [128]. The inaccuracy of energy functions and limitations on the time and length scales sampled often hinders the reliability of these predictions [129,130]. The accurate prediction of dynamic interactions and assemblies that underpin biological functions remains a major obstacle, despite significant progress in static structure prediction [131,132].

5.5. Protein interactions in the cell

The protein–protein interaction within the cell is an essential but less-perceived point [133,134]. Understanding who interacts with whom at different spatial and temporal scales is crucial for medicine. The proteins encoded execute a genetic program, but the proteins produced by a cell at a time and in each location make up its "proteome", whose components differ from native proteins because they are chemically modified. The proteome is much more dynamic than the genome: it changes during development and in response to external stimuli. Proteins form large interaction networks that regulate and support each other. Protein interaction networks (PPINs) show physical interactions between different proteins, creating a complex network that graphs can analyze [135,136]. Critical conditions such as cellular stress or disease can affect them, revealing how proteins respond and adapt to such situations [137]. Many studies that build models of protein–protein interactions rely on techniques such as co-immunoprecipitation, yeast two-hybrid, or computational approaches. These instruments often cannot accurately reflect the biological setting despite their potency. Lacking direct experimental validation, interpretations of these interactions risk being speculative, especially within complex systems such as human metabolism [138]. This issue raises a fundamental question in the biophysics of complex biological systems: how much can we trust theoretical models, and how can we balance their use with concrete experimental data? Lack of validation can lead to biases in the proposed metabolic pathways and affect understanding of the biological significance of signaling pathways. From a critical perspective, this also limits the development of targeted therapies because drugs designed for unvalidated interactions risk being less effective or having unpredictable effects.

Listing proteins is not enough to understand the cellular mechanism; it is also necessary to outline all the interactions between them. Proteins work in groups and all together to perform a biological function. Only the reliability and certainty of their interactions allow accurate prediction of the causal molecular mechanisms of metabolism and pathological states, for example, through interactomics. Therefore, researchers should study the interactions between individual proteins using genetic, biochemical, and biophysical techniques. In this way, we get certain and reliable information. However, the speed with which computational methods predict new proteins has created the need for high-throughput interaction detection methods [139,140]. In response to these identified needs, researchers have developed comprehensive digital methodologies that have recently generated substantial volumes

of data on interactions. These advancements are instrumental in enhancing our understanding of these complex dynamics. These interactions are indirect and thus lack the certainty and reliability that experimental methods offer. They can often be reliable hypotheses, but experiments must validate them. To date, we still have little knowledge about where, when, with whom, and how a protein functions in the complex system of metabolic departments and cellular compartments, not forgetting existing membraneless organelles. Therefore, a preliminary critical assessment of their accuracy, biases, overlaps, and complementarities is essential for using these data.

5.6. Critical issue in protein interactions in the cell

Today's molecular biology and omics research suffers from a critical flaw: insufficient use of protein biophysics to validate hypothesized interactions. The dynamic complexity of the proteome, much more elusive than the static genome, makes the entire field of protein interaction networks (PPINs) a vast and still little-explored territory. The critical issues in today's research are:

1. Proteome and native proteins: A cell's proteome varies depending on the context because decoded proteins can differ from the "native" ones. This reflects a vast gap in our understanding: quantifying and characterizing proteins under specific physiological conditions is still challenging.

2. PPIN and experimental validation: While a powerful approach, protein interaction networks suffer from limited experimental validation. Direct experiments have validated around 10% of the interactions identified in the human proteome [141]. This introduces a lot of "noise" in our theoretical models and risks leading us down potentially incorrect interpretative paths.

3. "Wet" vs. computational laboratories: Computational tools and AI possess immense strength, yet they rely on a sturdy foundation of experimental knowledge, particularly in protein biophysics methods, to be trustworthy [142]. Without it, we risk overestimating the ability of models to predict and explain real biological phenomena. The AI infrastructure layer includes data, storage, computing power, ML algorithms, and the AI framework. However, they still lack critical decisions based on the fundamental principles that have guided scientific research for centuries. Based on experience and verifiable facts, scientific knowledge uses observation and experimentation to collect data, interpreting phenomena by mathematics to arrive at demonstrable conclusions [143].

4. Balanced management between AI and experimental research: We cannot effectively manage the integration of computational and experimental science consciously [144]. Unlike other fields, automation and AI in biomedicine involve both physical and data processing. This process will take time, but it also requires a change of mindset and significant investments in training and research methods. We need to return to the scientific rigor of experimentation from the last century. However, at the core of everything is the same question Alan Turing asked years ago: Can machines think? [145]. A fascinating issue whose answer is beyond the scope of this review.

5.7. The significance of high-quality protein–protein interactions in cancer

For the sake of clarity, I would like to give an example. A PubMed search for "HCC" OR "hepatocarcinoma" OR "hepatocellular carcinoma" yielded 400,277 articles since 1950. However, despite 75 years of research, not all HCC researchers agree that we are on the right track to an effective cure. Let us consider the main molecular signaling pathways that characterize hepatocellular carcinoma (HCC) and that contribute to its progression and aggressiveness. They include the TGF-B

pathway, PD-L1, HDAC2-AS2, Wnt/β-catenin, MAPK/ERK, PI3K/AKT/mTOR, the Notch pathway, and others that we can exemplify in the following set of 27 proteins: VEGF, IFNG, Ras, Raf, MAPK, ERK, PI3K, Akt, mTOR, CTNNB1, WNT, JAK, STAT, Hippo, YAP, TAZ, SHH, PTCH1, SMO, NOTCH1, NOTCH2, NOTCH3, NOTCH4, Dll4, TGFB, PDL1, HDAC2. These proteins are involved in HCC and its progression [146−149]. But they are also all proteins we can input into STRING, one of the most potent online computational analysis platforms [150,151]. The 27 proteins, if used as functional seeds, can extract from the human proteome all the counterparts with which they have functional relationships.

Before using the analytical approach (enrichment) to reveal the most significant and functionally relevant portions of the network of molecular interactions in this biological context, it is instructive to understand which functional relationships exist between these proteins. Since proteins operate in the same pathology, we expect they should show strong functional relationships with each other. Here, we are not interested in the characteristics of the interactome itself, but in an accessory evaluation performed by STRING. We will consider only the quality of interactions of these proteins, ignoring the properties of their edges and nodes and their architecture. This way, we can better glimpse some organizational principles underlying these systems. In Figure 1, I report two graphs of the 27 proteins calculated with two confidence scores, 0.400 and 0.900. The first, of average value, being less significant, does not filter most of the heterogeneous information present in the literature on these proteins. The second collects only the most critical interactions (with the highest confidence). In this way, we can distinguish the specific relationships that exist between them.

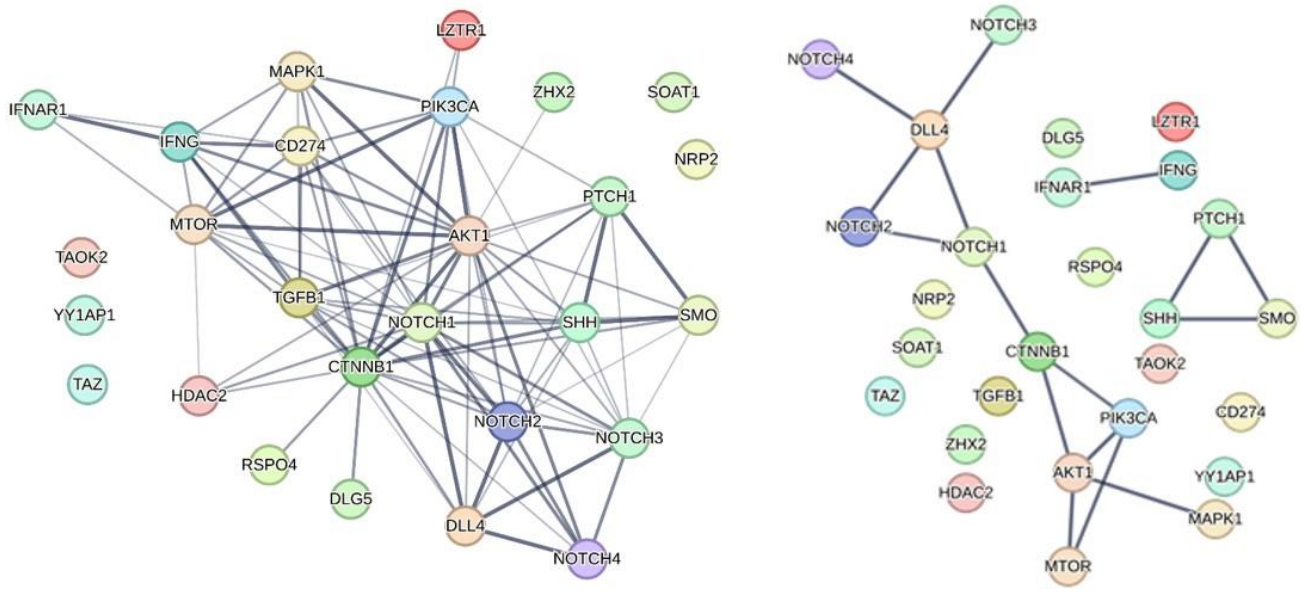

Figure 1. Confidence score graphs of 27 proteins contributing to HCC progression and aggressiveness.

Left graph: relationships between the 27 proteins with a confidence score of 0.400. The graph shows a compact central part and 5 unconnected proteins. Thin edges identify weak and poorly defined interactions. The key parameters are as follows: number of nodes, 27; number of edges, 92; average node degree, 6.81; avg. local clustering coefficient, 0.666; expected number of edges, 34; PPI

enrichment p-value, 1.11e-16.

Right graph: relationships between the 27 proteins with a confidence score of 0.900. The graph shows three unconnected subgraphs and 12 disconnected proteins. Topological parameters are as follows: number of nodes, 27; number of edges, 16; average node degree, 1.19; avg. local clustering coefficient, 0.438; expected number of edges, 3; PPI enrichment p-value, 9.18e-07. There is a difference in the relationships between proteins when we require greater significance.

STRING allows the export of a TSV file that reports each protein–protein interaction for each source channel (seven channels) in terms of a confidence score. Through this analysis, it is possible to evaluate both the composite origin of each interaction and the contribution of each channel to the definition of the total interaction score. One of these channels contains data related to the experimental determinations. The short-TSV-interactions file shows the individual interactions (one by one). The rigor of the criteria used to identify the interactions varies between the source channels; therefore, they differ in content and quality. From this file, we can evaluate the overall number of interactions validated by the experiments and their reliability. To achieve this, we need to isolate the most statistically significant experimental interactions by selecting them using the confidence score of 0.900 (maximum confidence) and 0.700 (high confidence) as filters. We can classify these interactions as certain and reliable. Table 2 shows that the experiments validate only four interactions as reliable and certain. What is surprising is the high score (0.900) in the two channels of Text Mining and annotated database, even in the absence of experimental data. The question to ask is: What is the origin of this data without validation? As an example, I suggest checking the interaction IFNAR1-IFNG.

Table 2. Distribution of interactions between crucial proteins for HCC.

#node1	node2	neighborhood_on_chromosome	gene_fusion	phylogenetic_cooccurrence	homology	coexpression	experimental	database_annotated	automated_textmining	combined_score
AKT1	MAPK1	0	0	0	0.593	0.118	0.110	0.900	0.872	0.988
AKT1	PIK3CA	0	0	0	0	0.060	0.738	0.900	0.951	0.998
AKT1	**MTOR**	**0**	**0**	**0**	**0**	**0.060**	**0.925**	**0.900**	**0.954**	**0.999**
AKT1	CTNNB1	0	0	0	0	0.074	0.539	0.400	0.930	0.979
CTNNB1	PIK3CA	0	0	0	0	0.049	0.130	0	0.890	0.901
CTNNB1	NOTCH1	0	0	0	0	0.083	0.328	0.500	0.931	0.976
DLL4	NOTCH4	0	0	0	0.645	0.181	0.087	0.900	0.988	0.998
DLL4	NOTCH3	0	0	0	0.654	0.050	0	0.900	0.986	0.998
DLL4	**NOTCH1**	**0**	**0**	**0**	**0.665**	**0.062**	**0.824**	**0.900**	**0.987**	**0.999**
DLL4	NOTCH2	0	0	0	0.667	0	0	0.900	0.990	0.999
IFNAR1	IFNG	0	0	0	0	0	0	0.900	0.898	0.989
MTOR	PIK3CA	0	0	0	0.548	0.060	0.166	0.900	0.592	0.963
NOTCH1	NOTCH2	0	0	0.065	0.950	0.145	0.180	0.900	0.122	0.982
PTCH1	**SMO**	**0**	**0**	**0**	**0**	**0.043**	**0.736**	**0.900**	**0.999**	**0.999**
PTCH1	**SHH**	**0**	**0**	**0**	**0**	**0.060**	**0.993**	**0.900**	**0.999**	**0.999**
SHH	SMO	0	0	0	0	0.045	0.162	0.400	0.999	0.999

Enriching with 500 first-order proteins (only direct interactions) and no specific filtering, we can calculate an interactome (not shown), which is based on over 10,000 scientific articles that the STRING AI has extracted from PubMed (all downloadable). We can notice a significant p-value (< 1.0e-16) and a compact architecture of 527 nodes that produce 31,869 interactions with 9085 functionally enriched terms in 15 functional categories. The network architecture reflects cellular localization, biological process, and molecular function, enabling the functional characterization of the proteins. Based on current scientific literature, this graph represents the entire molecular knowledge of HCC that induces disease progression in humans. But as before, we are not interested in the characteristics of the interactome itself. From the downloadable file about interactions, we can extract 285 interactions of the highest confidence and 230 interactions of high confidence. These interactions represent 0.89% and 0.72% of the total, thus showing that only 1.61% of the interactions

have experimental validation, which means nothing. This explains why HCC remains a deadly cancer even today. Verification of interactions should be a mandatory prerequisite for researchers before starting any scientific study of protein interactions in human metabolic pathways. These considerations related to HCC are not novel [5]. However, it is important to note that many studies discuss the molecular causes of human disease, analyzing many interactions, but without considering how many of these are certain, as confirmed by experiments [152]. Therefore, the emphasis on real experimental data is essential to provide the clarity and precision necessary to improve scientific rigor in this field.

5.8. High false positive and false negative rates in high-throughput studies

5.8.1. Yeast two-hybrid (Y2H) screens and affinity purification-mass spectrometry (AP-MS)

These are foundational high-throughput methods for discovering PPIs. However, they have significant rates of false positives (interactions detected that do not occur physiologically) and false negatives (true interactions that are missed) [153,154].

1. Case study implication: Later comparisons reveal that many early interactome maps generated by these methods had limited overlap with maps generated by different methods or maps whose individual interactions underwent rigorous validation. This led to a consensus that *no single high-throughput method suffices for defining a reliable interactome*, and orthogonal validation methods are critical [155].

2. Example: Studies comparing early large-scale Y2H datasets for *Saccharomyces cerevisiae* (yeast) revealed relatively poor overlap, leading to debates about the "true" interactome size and the reliability of individual interactions. Researchers often found that only a fraction of interactions were corroborated by multiple methods [156,157].

5.8.2. Context-dependency and transient interactions

1. Physiological vs. experimental conditions: Proteins interact differently depending on the cellular context (e.g., cell type, developmental stage, presence of other molecules, post-translational modifications, cellular localization). An interaction observed in vitro might not be physiologically relevant in vivo [158].

2. Case study implication: Many PPIs are transient, forming and dissociating rapidly to regulate cellular processes (e.g., signaling pathways). These are notoriously difficult to capture by many experimental methods, which often favor stable complexes. If these are missed or misinterpreted, understanding dynamic cellular processes is flawed [159,160].

3. Example: The interactions involved in signal transduction pathways (e.g., kinases and their substrates) are often transient. Capturing these precisely requires specialized techniques to monitor dynamic binding events, and relying solely on methods that capture stable complexes can lead to an incomplete picture [161,162].

5.8.3. Indirect vs. direct interactions

1. Mediated interactions: Some experimental methods (especially co-immunoprecipitation followed by mass spectrometry) detect proteins that are part of the same complex, but not necessarily

in direct physical contact [163]. Without further validation, by using methods like yeast two-hybrid, FRET, or structural biology, researchers might mistakenly interpret an interaction as a direct binary link.

2. Case study implication: Assuming all co-purified proteins directly interact can lead to incorrect network topologies and functional assignments [164].

3. Example: A protein might pull down a large complex, and while all components are part of the complex, only a few directly interact with the "bait" protein. Without follow-up binary interaction assays, one might falsely infer direct interactions between all components [165,166].

5.8.4. Intrinsic disorder and promiscuity

1. Intrinsically disordered regions (IDRs): Many proteins contain IDRs that are highly flexible and can adopt different conformations to interact with multiple partners. This "fuzziness" makes their interactions complex and sometimes promiscuous [167].

2. Case study implication: Predicting and experimentally validating interactions involving IDRs is particularly challenging. Traditional structural methods might not capture the full range of binding modes, and computational predictions (even with advanced tools like AlphaFold-Multimer, though it is improving) can struggle with the dynamic nature of these interactions. Incorrect assumptions about rigid body interactions can lead to errors [168].

3. Example: The p53 tumor suppressor protein has extensive IDRs that allow it to interact with a vast array of proteins and nucleic acids, mediating its diverse cellular roles. Characterizing the specific binding modes and partners of p53's IDRs for each function requires careful validation to avoid misinterpreting its complex regulatory network [169].

5.8.5. Technical limitations and artifacts

1. Method-specific biases: Each PPI detection experimental method has its limitations and biases. For instance, Y2H can suffer from protein expression, localization, and auto-activation issues. AP-MS biases results toward abundant proteins or those forming stable complexes.

2. Case study implication: Over-reliance on a single method, or a lack of understanding of its limitations, can lead to the propagation of artifacts into databases and subsequent downstream analyses [170,171].

3. Example: Early Y2H screens occasionally reported interactions that were later found to be due to self-activating bait proteins or nonspecific binding, rather than true biological interactions [156,172].

5.8.6. The drug BIA 10-2474 case

Unfortunately, some clinical trials have resulted in serious errors, such as the case with the drug BIA 10-2474. This drug, which was an FAAH enzyme inhibitor, was being developed for chronic pain and neurological disorders. Initial predictions using in silico models suggested selective and safe interactions; however, both biophysics and functional interactomics were either not utilized or underestimated, and potential off-target proteins were not taken into account. The consequences were severe. During the clinical trials, one patient tragically died, while others suffered significant neurological damage. Proper biophysical analysis could have identified off-target interactions with

critical proteins [173,174].

5.8.7. SARS-CoV-2 and the exploded interactomes

An emblematic case is what occurred following the spread of the SARS-CoV-2 viral infection. At the start of the COVID-19 pandemic, numerous studies predicted hundreds of virus–cell interactions [175]. Biophysical techniques (NMR, SPR, MS) confirmed only a very small portion. The results caused confusion about therapeutic targets and led to many drug repurposing failures. Rushing too quickly resulted in unstable and sometimes inconsistent outcomes [176,177].

6. The need for validation

A key point in many current fields is the absolute need for experimental validation of protein–protein interactions (PPIs) [178,179]. Without it, computational predictions or indirect associations can cause misleading functional interpretations. The field abounds with examples showing these complexities and the challenges of accurately mapping PPIs. Various areas and types of "errors" or complexities exist, which many case studies can show [180]. While experimentation is essential, computational approaches like machine learning models assist in predicting protein binding affinities [181,182]. However, these predictions require experimental validation. This validation process makes it easier to visualize protein structures, improving our understanding of their functions [183]. Despite the insights provided by computational models, it is essential to experimentally verify protein binding affinities to ensure their accuracy and reliability, as Wang notes. This validation step is fundamental to molecular biology [184]. Experimental validation is crucial in biomedical research for multiple reasons. While machine learning models can generate promising predictions, it is essential to verify that these predictions hold true in real-world settings. Conducting experiments helps identify and correct errors in data or models, ensuring that conclusions are based on dependable evidence. Therefore, experimental validation is vital for translating scientific discoveries into clinical applications, such as new treatments or diagnostics, by confirming their effectiveness and safety for humans. Moreover, biomedical research often must comply with strict regulatory standards; obtaining approval from agencies like the FDA requires comprehensive validation.

Furthermore, the validation process allows researchers to explore new ideas and approaches, promoting innovation in biomedicine. Validating results builds trust in biomedical research, fostering confidence among researchers, the public, and funders. It is essential to ensure that scientific discoveries are not only theoretically sound but also practical and safe in a clinical setting. In particular, experimental data on protein–protein interactions are critical for research because they provide some of the few reliable tools to thoroughly investigate the molecular mechanisms of metabolism at the microscopic level [185,186]. However, limitations exist in computational methods for predicting virus–host interactions, which must also be addressed. These include:

1. Data quality

Bias and incompleteness: Datasets often contain biases or lack information, which can lead to inaccurate predictions.

Data variability: Variations in viral strains and variants can influence predicted interactions.

2. Inadequate models

Model simplicity: Many computational models fail to fully capture the complexity of biological

interactions.

Incorrect assumptions: Assumptions within models may not accurately reflect biological reality, resulting in misleading conclusions.

3. Cell dynamics

Dynamic interactions: Interactions between viruses and cells are fluid and can change over time, making accurate predictions more difficult.

Microenvironment: Factors from other cells or the surrounding environment can influence interactions but are often overlooked in models.

4. Computational limitations

Processing capacity: Complex simulations demand significant computational resources, which can limit the scale and precision of analyses.

Computational time: Running complex models can take a considerable amount of time, delaying the production of results.

5. Experimental validation

Verification difficulties: Validating predicted results through in vitro or in vivo experiments is often challenging, which can reduce confidence in these predictions.

Discrepancies between models and experiments: Differences between computational and experimental results can undermine trust in computational methods.

In conclusion, despite these challenges, computational approaches remain a valuable tool for studying virus–host interactions. However, they should be complemented with experimental methods to improve prediction accuracy.

6.1. The solution and the importance of validation

Understanding these complexities has driven the scientific community to emphasize:

1. Orthogonal validation: Using multiple, independent experimental methods to confirm a predicted or initially detected interaction [187,188].

2. Quantitative assays: Moving beyond binary "yes/no" interactions to measure binding affinities (Kd), kinetics (on/off rates), and stoichiometry [189].

3. Contextualization: Performing experiments in physiologically relevant conditions (e.g., in cells, in specific subcellular compartments, considering post-translational modifications) [190].

4. Structural characterization: Determining the atomic-level structures of protein complexes to provide definitive proof of direct interaction and insights into the binding interface [191].

Although the scientific process corrects these errors and retractions are infrequent, the broader story of the field's development in PPI research implicitly shows the distortion of published functional results because of unvalidated PPIs. The ongoing refinement of high-throughput methods, the push for more rigorous validation, and the creation of advanced structural and biophysical techniques are all responses to the inherent complexities and the potential for misinterpretation in protein interaction data.

7. The gap between in vitro and in vivo studies

The interior of a living cell is a crowded and dynamic environment, filled with a vast array of macromolecules and undergoing fluctuations [133]. This complex environment can affect protein behavior in ways that are not always reflected in the dilute solutions used for in vitro biophysical

studies [83,133]. Understanding how proteins function within this native cellular context, including the impact of macromolecular crowding and confinement on their structure, dynamics, and interactions, is essential for a complete picture of their roles in biological systems [83]. Bridging the gap between the well-controlled conditions of in vitro experiments and the inherent complexity of in vivo systems requires innovative approaches and techniques that can probe protein behavior within the cellular milieu. In vitro studies provide invaluable fundamental insights into protein properties, while in vivo, the intricate cellular context often limits their direct relevance to the actual behavior of these molecules within the cell. Developing methods to study proteins in their native environment is crucial for a more accurate understanding of their function in living organisms.

8. The critical example of a tiny membrane protein, ORF7b of SARS-CoV-2

We have examples where the distortion of a biomolecule's biological role produces a chain of incorrect conclusions in the scientific literature. This can happen for several reasons, including, for example, the use of information without direct experimental controls or the misleading interpretation of data. Unfortunately, these distortions can persist for a considerable time, affecting the understanding of that area and delaying scientific progress. The example I would like to give falls in the biophysics of membrane proteins and is creeping and subtle but quite common, even if not perceived by researchers. Excessive reliance on initial, unvalidated studies leads to a lack of control in defining a scientific project.

ORF7b is a small accessory protein of SARS-CoV-2 [192]. In analogy to its SARS-CoV homologous counterpart, researchers describe this protein as "highly hydrophobic" and classify it as a single-spanning transmembrane protein located in the Golgi, as a recent textbook states. However, experiments showed that the SARS-CoV homologous protein required the transmembrane domain for Golgi complex retention, unlike the SARS-CoV-2 protein [62]. Therefore, despite recent findings, confusion remains between the two proteins, whose structures are still unknown. The physico-chemical properties of SARS-CoV-2 ORF7b reveal a 43-residue protein with a hydrophobic helical segment spanning residues 9–29 and very mobile, negatively charged N- and C-terminal tails [193]. These characteristics give the protein a strong net negative charge of -4 at pH 7, with a diffuse negative charge throughout the structure [193]. Thus, it behaves in solution like a weak negative polyelectrolyte or, more precisely, a polyanion. These undeniable physico-chemical properties suggest a peripheral protein [194], but do not exclude the possibility of it being a transmembrane protein or even a protein targeting lipid droplets [193]. All these environments have very different membrane characteristics, and we cannot exclude the protein being in the cytoplasm because its physico-chemical properties allow that potential. However, other data make us think differently. BioGRID, one of the most reputable computational analysis platforms [195], shows that through in vivo experiments, the protein has physically interacted with over 1,700 human proteins. One or more biophysical and non-biophysical experimental methods validate each interaction. Considering only the most significant interactions, according to BioGRID, we can select about 75 interactions spread throughout the cell, including the nucleus and mitochondria [192]. The interactomic analysis shows that starting from these interactions, ORF7b gives rise to 5057 functional terms across 15 categories. These data suggest that the protein must have one or more mechanisms to interact with these proteins in the various cellular compartments where they are located.

Evidence suggests that ORF7b, instead of being confined to the Golgi apparatus, operates within

a dynamic cellular environment, interacting with many proteins in various locations. We have a small biological object with a broad range of cellular actions, which is challenging to characterize. Experimental and computational data still do not confirm whether the protein is membrane-bound or has free mobility in the cytosol and between cellular compartments to carry out its wide-ranging functions. Neither possibility is definitively ruled out. This remains elusive today. However, it is crucial to recognize that only by integrating biological and biophysical knowledge through computational, experimental, and biological methods, considering the evolutionary aspects of the sequence and functions, we will find the solution for this small elusive protein. AI can facilitate this integration.

9. Limitations in understanding quantum mechanical effects of proteins

While proteins are quantum mechanical entities, understanding and characterizing the role of quantum effects in their biological functions presents significant limitations [196,197]. Explicit quantum mechanical calculations are expensive and are workable only for small systems, falling short of the size of most proteins [196,197]. While semi-classical models are often used to study protein structure and dynamics, they may not capture subtle but important quantum phenomena [198]. Observing and manipulating quantum effects in a living cell's "warm, wet", and noisy environment is also challenging, as these effects are often fragile and short-lived [93]. Distinguishing genuine quantum effects from classical behavior in complex biological systems requires sophisticated experimental techniques and theoretical frameworks, which are still under development [93,95]. The biological relevance and functional implications of many observed quantum phenomena in proteins are still being investigated [199−201].

10. Collaboration

The study of protein biophysics is a constantly developing field where multidisciplinary collaboration is key to success. This includes expertise in molecular biology, biochemistry, bioinformatics, statistics, engineering, and medicine, and its future depends on combining knowledge and techniques from multiple disciplines [202,203]. Merging protein biophysics with molecular evolution helps us better understand how proteins have developed to perform their functions and how mutations can change their properties [204]. Connecting biophysics with systems biology approaches offers a more comprehensive view of how proteins work within the complex networks of cellular processes [205]. By applying biophysical principles to analyze evolutionary data and using evolutionary information to predict biophysical behaviors, researchers can create more accurate models for biological understanding and prediction [206,207]. This interdisciplinary approach is essential for addressing the most complex biological questions and for turning fundamental biophysical insights into practical applications. Here are some examples of successful collaborations and the challenges they face.

10.1. Challenges in multidisciplinary collaborations

Despite the clear advantages of multidisciplinary collaborations, they present many diverse challenges:

10.1.1. Linguistic and conceptual barriers

Researchers from different disciplines often use specific terminology and concepts that may be difficult for others to understand. For instance, a "model" for a biologist might refer to a cell culture, while for a bioinformatician, it could mean a computational algorithm. Overcoming these barriers requires a firm commitment to communication and mutual training, often involving joint workshops or the assignment of "bridge scientists"—individuals with a hybrid background who can facilitate the exchange of ideas.

10.1.2. Different methodologies and data expectations

Each discipline has methodologies, data quality standards, and expectations regarding reproducibility. For example, a biological experiment may exhibit greater intrinsic variability than a computational calculation, leading to potential misunderstandings. Addressing this requires clearly defining protocols, establishing quality standards, and setting success metrics for all data generated at the project's outset. Creating integrated workflows and sharing standardized technology platforms can facilitate collaboration.

10.1.3. Complexity in project and intellectual property management

Coordinating teams with diverse backgrounds can be challenging. Managing intellectual property (such as patents and publications) becomes complex when different institutions or commercial entities contribute to the project. Strong and well-defined project leadership is crucial to establish clear agreements regarding intellectual property and publishing from the beginning, often with the support of legal or technology transfer departments.

10.1.4. Cultural and academic differences

Research cultures can vary significantly. For example, bioinformaticians may prefer open-source approaches and rapid publication, while experimental biologists might have longer turnaround times for experiments and prioritize journals with high impact factors. Therefore, fostering an environment of mutual respect and understanding of differing priorities is vital. Jointly celebrating successes and recognizing the contributions of all team members can strengthen a sense of belonging and motivation.

In summary, multidisciplinary collaborations are essential for addressing complex challenges. They require careful planning, effective communication, and flexible management to overcome the inevitable challenges that arise.

10.2. Examples of successful collaborations in interactomics

10.2.1. Human Proteome Project (HPP) and l'International Protein Interactome Database (iPID)

This is a large international project involving thousands of scientists from international institutions, including molecular biologists, analytical chemists (mass spectrometry), bioinformaticians, geneticists, and statisticians. iPID (which integrates data from databases such as

BioGRID, IntAct, etc.) results from collaborative efforts to standardize and make accessible data on human PPIs. The goal is to map the entire human proteome and its interactions to better understand biology and disease. It is a success story because this international collaboration has identified and validated many new PPIs. The Characterization of the Human Proteome Project (Chromosome-Centric HPP, C-HPP) has led to the discovery of "missing proteins" and their functional characterization, often through integrating proteomic data with genetic and transcriptional data. This collaboration directly leads to the publication of comprehensive datasets and the development of open bioinformatics tools. These projects required data standardization across many laboratories. Building strong computational infrastructures for integrating diverse datasets was a major challenge, which was resolved through shared protocols and worldwide platforms.

10.2.2. Disease-specific research consortia (e.g., neurodegenerative diseases)

Many research consortia focused on diseases such as Alzheimer's and Parkinson's and involve a diverse group of experts. These include neurologists, pathologists, neuroscientists working with animal and cellular models, and biochemists studying dysfunctional proteins such as Tau or alpha-synuclein. Imaging specialists (working with PET/MRI technologies) and bioinformaticians (analyzing altered protein–protein interaction networks) play a crucial role in these efforts. Initiatives such as the Alzheimer's Disease Neuroimaging Initiative (ADNI) and the Parkinson's Progression Markers Initiative (PPMI) are prominent examples. While these projects do not focus only on interactomics, they jointly collect and analyze multimodal data, including genetics, cerebrospinal fluid proteomics, and imaging. Identifying interactions between specific proteins, such as Tau, and their interactions with other cytoskeletal proteins often highlights potential therapeutic targets. This identification typically results from complex bioinformatics analyses applied to data generated through various experimental methods. Integrating heterogeneous clinical data, such as imaging, fluid biomarkers, and genetic information, with "pure" molecular data (protein–protein interactions) represents a significant analytical challenge. Bioinformaticians and statisticians addressed this complexity using systems biology approaches and machine learning algorithms.

10.2.3. Development of PPI-based therapies in oncology

This section highlights a typical collaboration among various specialists in oncology. Clinical oncologists, cell biologists, biochemists, medicinal chemists, and pharmacists collaborate to develop therapies targeting protein–protein interactions (PPIs). The primary goal of these international teams is to develop PPI inhibitors for cancer treatment successfully.

For instance, researchers focus on drugs that target the interaction between MDM2 and p53, a crucial protein–protein interaction for tumor suppression. The team that pioneered the development of MDM2 inhibitors, such as Nutlin-3 and Idasanutlin, included computational chemistry specialists for drug design, biologists for both in vitro and in vivo validation, and clinical teams for the experimental testing phases.

One of the key challenges in this field is identifying druggable "hotspots" on protein surfaces and designing molecules that can selectively modulate these interactions. However, advances in molecular modeling techniques and high-throughput screening—often assisted by robotics and automation—have successfully addressed these challenges and achieved critical therapeutic goals.

10.3. A practical example of how to design a collaboration: Search for "Pathogenic" proteoforms and "neutral" proteoforms

Distinguishing between "pathogenic" and "neutral" proteoforms is the key step in translating proteomics into precision medicine. Functional proteomics, as it is evolving, is one of the most effective tools for this. We can consider this type of scientific planning as a practical scheme to be implemented for a multidisciplinary collaboration. In short, this could be the strategy.

10.3.1. Definition and mapping of proteoforms

It starts by recognizing that a single gene sequence can produce hundreds of proteoforms (alternative splicing, PTMs, proteolytic cleavage, etc.). Functional proteomics uses techniques such as top-down mass spectrometry, middle-down MS, and multi-omics combinations to precisely identify these variants in the cellular context; in many cases, work is carried out under physiological conditions, not in vitro.

10.3.2. Association with cellular or tissue phenotypes

Pathological conditions (tumor, inflammation, infection) are analyzed, and the observed proteoforms are correlated with specific functional outputs (proliferation, apoptosis, invasiveness, etc.); techniques such as single-cell proteomics and laser microdissection allow us to study specific microenvironments, such as cirrhotic liver or tumor nodules.

10.3.3. Validation through biophysical and biochemical approaches

Suspected proteoforms are isolated and tested with SPR (surface plasmon resonance) for affinity and kinetics, HDX-MS and cross-linking MS for structural interactions, and paramagnetic NMR and cryo-EM for conformational states. This confirms whether the proteoform has pathological interactions and aberrant functions or is merely a functionally neutral variation.

10.3.4. Integration with computational models

Interactomic networks are filtered and refined using experimental data; AI and machine learning algorithms (with functional proteomic input) allow pathogenic proteoforms to be prioritized for clinical studies. Systems such as DeepPTM, Prosit, and MSFragger help "reconstruct" the proteoform in its actual cellular physiology.

10.3.5. Clinical applications

Identification of specific biomarkers for hepatocellular carcinoma (HCC), viral diseases with tissue tropism, and neurodegenerative diseases with selective protein accumulation; definition of therapeutic targets based on pathogenic proteoforms that are actually active, not simply present; and the design of drugs that stabilize, inhibit, or correct the abnormal proteoform.

11. Conclusions

In today's scientific landscape, researchers often find themselves limited to narrow areas of expertise, which can impair their critical thinking skills when they need to explore topics outside their specific fields. As a result, gaining a deeper understanding of the complex molecular mechanisms underlying many diseases represents a crucial frontier that requires extensive multidisciplinary research.

A widely recognized yet often underestimated issue is that only about 10% of all known protein interactions have been experimentally validated through methods in protein biophysics or biochemistry. Researchers are increasingly relying on indirect algorithmic methods, which have proven to be much more efficient than experimental approaches. However, this reliance has led to a growing number of unverified interactions over time. These factors have created a speculative scientific environment filled with misinformation, which hinders scientific progress and contributes to a phenomenon known as knowledge pollution. It is essential to understand why biophysical techniques have lagged behind predictive algorithms and to examine the concrete and significant reasons for this disparity. The reasons are concrete and significant:

1. Precision experimentation: Validating an interaction requires isolating it, reconstructing it in vitro or in cells, and quantifying its affinity and dynamics.
2. Biological context: Interactions are often transient, conditional, or multiple, making them difficult to reproduce in the laboratory.
3. Instrumental limitations: Even emerging techniques such as cryo-EM, in-cell NMR, and HDX have specific applicability ranges and do not always capture the complexity of the proteome.

All stakeholders must recognize that technologies are not shortcuts to the truth but tools designed to limit ambiguity. These tools help reveal relevant and reproducible interactions, not to confirm every algorithmic output. Protein biophysics is slow because it is meticulous. It prioritizes molecular verification over speed, which is essential for medical advancement. If we do not curb the algorithmic generation of unvalidated interactomes, we risk creating a science that resembles science fiction. Progress is not merely the accumulation of data but the distillation of truth from potential. Biophysics serves as one of the bulwarks striving to maintain this critical boundary.

11.1. The where, how, and when of proteins

The future of the field looks promising, with emerging technologies and ongoing advances that aim to overcome many challenges. All of this could transform our understanding of fundamental biological systems and lead to major innovations in medicine and biotechnology. However, current protein biophysics research still faces significant challenges in fully understanding the complexity of protein folding, the dynamic behavior of proteins in the cellular environment, and the intricate interactions that control biological processes. Table 3 outlines future solutions to some existing limitations in key areas of protein biophysics knowledge. While individual biophysical techniques are powerful, each has inherent limitations that make integrative approaches necessary. Increasing recognition of interdisciplinary collaborations, along with ongoing improvements in experimental and computational tools, will be essential for achieving a complete understanding of protein biophysics.

Table 3. Emerging technologies in protein biophysics.

Limitation category	Addressed by future prospects/emerging technologies
Challenges in protein folding	AI-based structure prediction (AlphaFold, RoseTTAFold), advanced computational methods for simulating folding pathways, integrative structural biology to study folding intermediates.
Limitations of biophysical techniques	Advancements include higher-resolution cryo-EM, automated cryo-EM, time-resolved crystallography, integrative structural biology combining multiple techniques, mass photometry for intact biomolecule analysis, automation, and miniaturization for high-throughput studies.
Difficulties in studying protein dynamics	AI-powered methods for predicting protein dynamics and multiple conformations, single-molecule techniques (FRET, HS-AFM), advanced NMR techniques, and enhanced computational power for molecular dynamics simulations.
Challenges in predicting protein interactions	AI-based methods for predicting protein–protein interactions (e.g., AlphaFold-Multimer), integrative structural biology to combine computational predictions with experimental data, advancements in computational modeling of binding dynamics.
The gap between in vitro and in vivo studies	In-cell biophysical techniques are developing, along with advancements in imaging technologies for studying proteins in their native environments and computational modeling that incorporates cellular crowding and other in vivo factors.
Challenges in studying membrane proteins	Advancements include cryo-EM, detergent-free membrane protein stabilization (nanodiscs and SMALPs), AI-based structure prediction, integrative structural biology, and emerging biophysical techniques.
Limitations in understanding quantum mechanical effects	Advanced spectroscopic techniques, computational methods for modeling quantum phenomena, interdisciplinary research in quantum biology, and exploration of the functional relevance of quantum effects in biological processes are all advancing.

Proteins are biomolecules that dynamically perform the functional activities of cells. They do not act as native encoded proteins [208−210] but are almost always modified to go to a specific place in the cell where they will perform their function in a particular way. Much of the information on these space-temporal aspects is fragmentary and poorly understood even today [211−214]. We can summarize all this in one sentence: We do not know the where, when, and how of individual proteins.

The "where" refers to the subcellular localization of proteins, which is the specific compartment of the cell where a protein is found and performs its function. This could be the nucleus, cytoplasm, cell membrane, or organelles such as the endoplasmic reticulum, lipid droplets, or even mitochondria. The "when" shows when a protein is active or expressed. Cells synthesize proteins in response to cellular signals; post-translational modifications (e.g., phosphorylation or ubiquitination) activate or deactivate them; and cells degrade proteins when unnecessary. The "how" concerns the mechanism by which a protein carries out its function. This includes its three-dimensional structure, interactions with

other molecules, chemical modifications that regulate its activity, and how it contributes to cellular processes.

These three aspects are still the subject of intense research, since understanding protein dynamics is fundamental to the future of biology and medicine. Studying the "where", "when", and "how" of proteins requires a multidisciplinary approach and the use of those advanced techniques that we have already discussed.

However, these techniques highlight activities involving a protein but do not explain its function directly. The function of a protein is determined not only by the interactions it engages in but also by the biological context in which it operates [215]. This requires other methodological approaches typical of different fields, such as loss and gain of function studies (e.g., CRISPR-Cas9), in vivo functional analysis with animal or cellular models, genetic screening, RNA interference, metabolic analyses, and pathway mapping. In summary, the functional role of a protein arises from the close integration of structural, dynamic, and functional data. Modern biology needs to move away from a deterministic approach and adopt a more systemic and indeterministic perspective, where functions emerge from complex molecular networks and dynamic interactions. Causality is a key concept: cellular functions do not originate from single proteins but from the simultaneous interaction of multiple components operating within a specific space and time [216]. This is the core of systems biology, which encourages approaches like computational models of metabolic networks, to simulate the emergence of cellular functions; single-cell proteomics, to analyze the dynamic variability of proteins in individual cells; and multi-omics analysis, which integrates genomics, transcriptomics, proteomics, and metabolomics to understand emergent causality.

However, the approaches mentioned remain valid only if we validate each piece of information or data. This is always a critical point. Experimental validation of protein interactions remains limited compared to the system's complexity. Less than 10% of interactions have been experimentally confirmed. This means that many analyses rely on hypothetical data, risking misleading results. Molecular indeterminism presents a significant challenge because it suggests that the relationships between proteins are not simply linear or deterministic but arise from a dynamic and probabilistic system [216]. Quantum physics influences the behavior of biomolecules, making it even harder to establish direct cause-and-effect links. The only ways to address this complexity are increasing experimental validation of protein interactions with techniques such as high-resolution mass spectrometry and advanced microscopy, integrating multi-omics approaches to gain a more comprehensive view of molecular networks, and using computational models built on experimental data.

Addressing the challenges in protein research requires a broader focus beyond specific topics, as a much larger problem is at hand. Current protein databases are outdated and contain highly heterogeneous data, which requires thorough cleaning to ensure quality and reliability. Without this effort, there is a risk of misinterpreting results and drawing incorrect conclusions. Advances in structural biology, bioinformatics, and imaging techniques offer deeper insights into biomolecular functions. Artificial intelligence transforms protein structure prediction, enabling us to create more accurate models by integrating structural, functional, and dynamic data. Combining computational biology, big data, and AI could almost fully explain biomolecules, thus paving the way for advancements in personalized medicine and biotechnology. However, the chaotic nature of information in big data systems can cause confusion if not properly managed. A key challenge is data cleaning, a complex task that requires experimental validation and a return to traditional laboratory

biology. Combining computational models with direct experiments is essential to overcoming this hurdle. Although AI reveals patterns in the data, we must experimentally confirm these findings to avoid purely theoretical constructs without a biological basis. Techniques such as protein chemistry, mass spectrometry, and advanced imaging are vital for accurately understanding how proteins are modified to function in specific contexts. Ultimately, the real challenge is knowing a protein's structure and understanding its dynamic function under actual conditions. Though this task is monumental, a balanced approach between computational biology and laboratory work can help us achieve more precise answers.

Creating new Big Data repositories that are well-organized and agreed upon by the scientific community could significantly speed up the understanding of proteins. Many existing databases suffer from inconsistencies, redundant data, or unverified information. Without a strict classification system, building studies on unstable foundations is risky.

For an effective repository, protein classification should be based on well-defined characteristics, including post-translational modifications, molecular interactions, and cellular conditions. Experimental verification from laboratory work, rather than unverified theoretical data, should support all data points. This creates a well-structured, accessible, and regularly updated scientific archive, preventing obsolescence. Such a system could significantly reduce the time and uncertainty involved in studying proteins in their natural context. A well-trained AI could analyze data in real time, identify recurring patterns, and even predict protein modifications based on cellular conditions. AI could assist by filtering out duplicate or unverified information, categorizing each protein according to agreed-upon parameters, suggesting possible post-translational modifications and functions based on existing data, and connecting laboratory results with computational models for quicker validation. The future depends on integrating disciplines that may seem distant, as everything hinges on being understood through a common language, much like the ancient Greeks' "koinè diàlektos" (κοινή διάλεκτος).

Use of generative-AI tools declaration

The author declares he have not used Artificial Intelligence (AI) tools in the creation of this article.

Conflict of interest

The author declares no conflicts of interest.

References

1. Hille B (2022) A Life of Biophysics. *Annu Rev Biophys* 51: 1−17. https://doi.org/10.1146/annurev-biophys-120121-074034
2. Dill KA, Ozkan SB, Weikl TR (2007) The protein folding problem: when will it be solved? *Curr Opin Struc Biol* 17: 342–346. https://doi.org/10.1016/j.sbi.2007.06.001
3. Murugan A, Husain K, Rust MJ, et al. (2021) Roadmap on biology in time varying environments. *Phys Biol* 18: 041502. https://doi.org/10.1088/1478-3975/abde8d

4. Bustin SA, Jellinger KA (2023) Advances in molecular medicine: unravelling disease complexity and pioneering precision healthcare. *Int J Mol Sci* 24: 14168. https://doi.org/10.3390/ijms241814168
5. Sharma A, Colonna G (2021) System-Wide Pollution of Biomedical Data: Consequence of the Search for Hub Genes of Hepatocellular Carcinoma Without Spatiotemporal Consideration. *Mol Diagn Ther* 25: 9−27. https://doi.org/10.1007/s40291-020-00505-3
6. Chautard E, Thierry-Mieg N, Ricard-Blum S (2009) Interaction networks: from protein functions to drug discovery. a review. *Pathol Biol* 57: 324−333. https://doi.org/10.1016/j.patbio.2008.10.004
7. Aebersold R, Agar JN, Amster IJ, et al. (2018) How many human proteoforms are there? *Nat Chem Biol* 14: 206–214. https://doi.org/10.1038/nchembio.2576
8. Naryzhny S (2024) Puzzle of proteoform variety—Where is a key? *Proteomes* 12: 15. https://doi.org/10.3390/proteomes12020015
9. Ashrafuzzaman M (2023) *Introduction to Modern Biophysics*. CRC Press https://doi.org/10.1201/9781003287780
10. Lukyanenko R, Storey VC, Pastor O (2022) Conceptual modelling for life sciences based on systemist foundations. BMC bioinformatics 23: 574. https://doi.org/10.1186/s12859-023-05287-z
11. Wang X, Fan J (2021) Spatiotemporal molecular medicine: a new era of clinical and translational medicine. Clinical and Translational Medicine 11: e294. https://doi.org/10.1002/ctm2.294
12. Nakajima T (2013) Probability in biology: overview of a comprehensive theory of probability in living systems. *Prog Biophys Mol Biol* 113: 67−79. https://doi.org/10.1016/j.pbiomolbio.2013.03.007
13. Saeed N, Nam H, Haq M (2018) A survey on multidimensional scaling. *ACM Comput Surv* 51: 1−25. https://doi.org/10.1145/3178155
14. Buxbaum JN (2003) Diseases of protein conformation: what do in vitro experiments tell us about in vivo diseases? Trends in biochemical sciences 28: 585−592. https://doi.org/10.1016/j.tibs.2003.09.009
15. Mazal H Haran G (2019) Single-molecule FRET methods to study the dynamics of proteins at work. *Curr Opin Biomed Eng* 12: 8−17. https://doi.org/10.1016/j.cobme.2019.08.007
16. Uchihashi T, Scheuring S (2018) Applications of high-speed atomic force microscopy to real-time visualization of dynamic biomolecular processes. *BBA-Gen Subjects* 1862: 229−240. https://doi.org/10.1016/j.bbagen.2017.07.010
17. Ando T, Fukuda S, Ngo KX, et al. (2024). High-speed atomic force microscopy for filming protein molecules in dynamic action. *Annu Rev Biophys* 53: 19−39. https://doi.org/10.1146/annurev-biophys-030722-113353
18. Karunanithy G, Shukla, VK, Hansen DF (2021) Methodological advancements for characterising protein side chains by NMR spectroscopy. *Curr Opin Struct Biol* 70: 61−69. https://doi.org/10.1016/j.sbi.2021.04.002
19. Huggins DJ, Biggin PC, Dämgen MA, et al. (2019). Biomolecular simulations: from dynamics and mechanisms to computational assays of biological activity. *Wires: Comput Mol Sci* 9: e1393. https://doi.org/10.1002/wcms.1393
20. Nelson LK (2020) Computational grounded theory: a methodological framework. *Sociol Method Res* 49: 3−42. https://doi.org/10.1177/0049124117729703

21. Monteiro da Silva G, Cui JY, Dalgarno DC, et al. (2024) High-throughput prediction of protein conformational distributions with subsampled AlphaFold2. *Nat Commun* 15: 2464. https://doi.org/10.1038/s41467-024-46715-9
22. Sadybekov AV, Katritch V (2023) Computational approaches streamlining drug discovery. *Nature* 616: 673–685. https://doi.org/10.1038/s41586-023-05905-z
23. Wang M, Ma A, Wang H, et al. (2024) Atomic molecular dynamics simulation advances of de novo-designed proteins. *Q Rev Biophys* 57: e14. https://doi.org/10.1017/S0033583524000131
24. Ose NJ, Campitelli P, Modi T, et al. (2024) Some mechanistic underpinnings of molecular adaptations of SARS-COV-2 spike protein by integrating candidate adaptive polymorphisms with protein dynamics. *eLife* 12: RP92063. https://doi.org/10.7554/eLife.92063.3
25. Fabbrizi E, Fiorentino F, Casano F, et al. (2025) Native mass spectrometry for proximity-inducing compounds: a new opportunity for studying chemical-induced protein modulation. Expert Opin Drug Dis 20: 643−657. https://doi.org/10.1080/17460441.2025.2486146
26. Helmerhorst E, Chandler DJ, Nussio M, et al. (2012) Real-time and label-free bio-sensing of molecular interactions by surface plasmon resonance: a laboratory medicine perspective. *Clin Biochem Rev* 33: 161−173.
27. Cavalcanti IDL, Junior FHX, Magalhães NSS, et al. (2023) Isothermal titration calorimetry (ITC) as a promising tool in pharmaceutical nanotechnology. *Int J Pharm* 641: 123063. https://doi.org/10.1016/j.ijpharm.2023.123063
28. dos Santos Rodrigues FH, Delgado GG, da Costa TS, et al. (2023) Applications of fluorescence spectroscopy in protein conformational changes and intermolecular contacts. *BBA Adv* 3: 100091. https://doi.org/10.1016/j.bbadva.2023.100091
29. AlQuraishi M, Sorger PK (2021) Differentiable biology: using deep learning for biophysics-based and data-driven modeling of molecular mechanisms. *Nat Methods* 18: 1169−1180. https://doi.org/10.1038/s41592-021-01283-4
30. Gong R, Reynolds MJ, Carney KR, et al. (2025) Fascin structural plasticity mediates flexible actin bundle construction. *Nat Struct Mol Biol* 32: 940−952. https://doi.org/10.1038/s41594-024-01477-2
31. Piper SJ, Johnson RM, Wootten D, et al. (2022) Membranes under the magnetic lens: a dive into the diverse world of membrane protein structures using Cryo-EM. *Chem Rev* 122: 13989−14017. https://doi.org/10.1021/acs.chemrev.1c00837
32. Duckhoe K, and Sahin O (2015) Imaging and three-dimensional reconstruction of chemical groups inside a protein complex using atomic force microscopy. *Nat Nanotechnol* 10.3: 264−269. https://doi.org/10.1038/nnano.2014.335
33. Uzorka A, (2025) Single-molecule biophysics: probing biological processes at the nanoscale. *Biophys Rev Lett* 20: 13−29. https://doi.org/10.1142/S1793048025300014
34. Sonar R, Mishra PP (2025) Exploring dynamics at the mmolecular sscale: advances in single-molecule biophysics, *Biochemical and Biophysical Methods in Molecular and Cellular Biology*, 163−184. Singapore: Springer Nature Singapore. https://doi.org/10.1007/978-981-96-2088-3_7
35. Miller H, Zhou Z, Shepherd J, et al. (2017) Single-molecule techniques in biophysics: a review of the progress in methods and applications. *Rep Prog Phys* 81: 024601. https://doi.org/10.48550/arXiv.1704.06837

36. Zhang L, Shen D, Yang J (2025) Advancing Single-Molecule Biophysics: Next-Generation Organic Fluorophores with Tailored Labeling Strategies. *Chem Biomed Imag* https://doi.org/10.1021/cbmi.5c00007
37. Czajkowsky DM, Sun J, Shen Y, et al. (2015) Single molecule compression reveals intra-protein forces drive cytotoxin pore formation. *elife* 4: e08421. https://doi.org/10.7554/eLife.08421
38. Kopac T (2025) Leveraging artificial intelligence and machine learning for characterizing protein corona, nanobiological interactions, and advancing drug discovery. *Bioengineering* 12: 312. https://doi.org/10.3390/bioengineering12030312
39. Jumper J, Evans R, Pritzel A, et al. (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* 596: 583–589. https://doi.org/10.1038/s41586-021-03819-2
40. Abramson J, Adler J, Dunger J, et al. (2024) Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* 630: 493–500. https://doi.org/10.1038/s41586-024-07487-w
41. Brotzakis ZF, Zhang S, Murtada, MH, et al. (2025) AlphaFold prediction of structural ensembles of disordered proteins. *Nat Commun* 16: 1632. https://doi.org/10.1038/s41467-025-56572-9
42. Ruffolo JA, Chu LS, Mahajan SP, et al. (2023) Fast, accurate antibody structure prediction from deep learning on massive set of natural antibodies. *Nat Commun* 14: 2389. https://doi.org/10.1038/s41467-023-38063-x
43. Gasparyan L, Mazo I, Simonyan V, et al. (2019) DNA Sequencing: current state and prospects of development. *Open J Biophys* 9: 169−197. https://doi.org/10.4236/ojbiphy.2019.93013
44. Monteiro da Silva G, Cui JY, Dalgarno DC, et al. (2024) High-throughput prediction of protein conformational distributions with subsampled AlphaFold2. *Nat Commun* 15: 2464. https://doi.org/10.1038/s41467-024-46715-9
45. Liu H, Zhuo C, Gao J, et al. (2024) AI-integrated network for RNA complex structure and dynamic prediction. *Biophys Rev* 5: 041304. https://doi.org/10.1063/5.0237319
46. Castaño-Díez D, Danev R, Baumeister W (2020) Progress in cryo-EM aided by new technological advances. *Curr Opin Struct Biol* 64: 37−48. https://doi.org/10.1016/j.sbi.2020.01.006
47. Lee M (2023) Recent advances in deep learning for protein-protein interaction analysis: a comprehensive review. *Molecules* 28: 5169. https://doi.org/10.3390/molecules28135169
48. Schauperl M, Denny RA (2022) AI-based protein structure prediction in drug discovery: impacts and challenges. *J Chem Inf Model* 62: 3142−3156. https://doi.org/10.1021/acs.jcim.2c00026
49. Nussinov R, Tsai CJ, Shehu A, et al. (2019) Computational structural biology: successes, future directions, and challenges. *Molecules* 24: 637. https://doi.org/10.3390/molecules24030637
50. Dobson CM (2019) Biophysical techniques in structural biology. *Annu Rev Biochem* 88: 25−33. https://doi.org/10.1146/annurev-biochem-013118-111947
51. Miller H, Zhou Z, Shepherd J, et al. (2017) Single-molecule techniques in biophysics: a review of the progress in methods and applications. *Rep Prog Phys* 81: 024601. https://iopscience.iop.org/article/10.1088/1361-6633/aa8a02
52. Zhang Z, Kumar V, Dybkov O, et al. (2024) Cryo-EM analyses of dimerized spliceosomes provide new insights into the functions of B complex proteins. *EMBO Journal* 43: 1065−1088. https://doi.org/10.1038/s44318-024-00052-1
53. Yeukuang H, and Margaritondo G (2021) Synchrotron radiation and X-ray free-electron lasers (X-FELs) explained to all users, active and potential. *Synchrotron Radiat* 28: 1014−1029. https://doi.org/10.1107/S1600577521003325

54. Sperger T, Sanhueza IA, and Schoenebeck F (2016) Computation and experiment: a powerful combination to understand and predict reactivities. *Acc Chem Res* 49: 1311−1319. https://doi.org/10.1021/acs.accounts.6b00068
55. Chang L, Mondal A, Singh B, et al. (2024) Revolutionizing peptide-based drug discovery: advances in the post-AlphaFold era. Wires: *Comput Mol Sci* 14: e1693. https://doi.org/10.1002/wcms.1693
56. Sonn-Segev A, Belacic K, Bodrug T, et al. (2020) Quantifying the heterogeneity of macromolecular machines by mass photometry. *Nat Commun* 11: 1772. https://doi.org/10.1038/s41467-020-15642-w
57. Rathore AS, Sarin D (2024) What should next-generation analytical platforms for biopharmaceutical production look like? *Trends Biotechnol* 42: 282−292. https://doi.org/10.1016/j.tibtech.2023.08.008
58. Jeon H, Han J (2025) Microfluidics with machine learning for biophysical characterization of cells. *Annu Rev Anal Chem* 18: 447−472. https://doi.org/10.1146/annurev-anchem-061622-025021
59. Silva TC, Eppink M, Ottens M (2022) Automation and miniaturization: enabling tools for fast, high-throughput process development in integrated continuous biomanufacturing. *J Chem Technol Biotechnol* 97: 2365−2375. https://doi.org/10.1002/jctb.6792
60. Krishnan M, Darling S (2020) Meeting analytical challenges in a brave new world of next generation biotherapeutics. *Drug Discov* 21: 31−39.
61. Capelle MA, Gurny R, Arvinte T (2007) High throughput screening of protein formulation stability: practical considerations. *Eur J Pharm Biopharma* 65: 131−148. https://doi.org/10.1016/j.ejpb.2006.09.009
62. Kwan TO, Kolek SA, Danson, AE, et al. (2022) Measuring protein aggregation and stability using high-throughput biophysical approaches. *Front Mol Biosci* 9: 890862. https://doi.org/10.3389/fmolb.2022.890862
63. Lindsey YN, Villa E (2023) Bringing structure to cell biology with cryo-electron tomography. *Annu Rev Biophys* 52: 573−595. https://doi.org/10.1146/annurev-biophys-111622-091327
64. Petoukhov MV, Svergun DI (2012) Small angle X-ray scattering/small angle neutron scattering as methods complementary to NMR, *NMR of Biomolecules: Towards Mechanistic Systems Biology*, 562−574. https://doi.org/10.1002/9783527644506.ch35
65. Ashkar R, Bilheux H, Bordallo H, et al. (2018) Neutron scattering in the biological sciences: progress and prospects. *Acta Cryst Struct Biol* 74: 1129−1168. https://doi.org/10.1107/S2059798318017503
66. Barbieri L, Luchinat E, Banci L (2016) Characterization of proteins by in-cell NMR spectroscopy in cultured mammalian cells. *Nat Protoc* 11: 1101−1111. https://doi.org/10.1038/nprot.2016.061
67. Serber Z, Corsini L, Durst F, et al. (2005) In-cell NMR spectroscopy. *Method Enzymol* 394: 17−41. https://doi.org/10.1016/S0076-6879(05)94002-0
68. Smith MJ, Marshall CB, Theillet FX, et al. (2015) Real-time NMR monitoring of biological activities in complex physiological environments. *Curr Opin Struct Biol* 32: 39−47. https://doi.org/10.1016/j.sbi.2015.02.003
69. Breindel L, Burz DS, Shekhtman A (2020) Active metabolism unmasks functional protein–protein interactions in real time in-cell NMR. *Commun Biol* 3: 249. https://doi.org/10.1038/s42003-020-0976-3

70. Xiang X, Hansen AL, Bruschweiler-Li L, et al. (2025) Detection of intramolecular protein dynamics on nanosecond-to-microsecond timescales by nanoparticle-assisted NMR spin relaxation (NASR). *Nat Protoc* 434. https://doi.org/10.1038/s41596-025-01177-1
71. Novotny J, Komorovsky S, Marek R (2024) Paramagnetic effects in NMR spectroscopy of transition-metal complexes: principles and chemical concepts. *Acc Chem Res* 57: 1467−1477. https://doi.org/10.1021/acs.accounts.3c00786
72. Schreier S, Polnaszek CF, Smith IC (1978). Spin labels in membranes problems in practice. *BBA-Rev Biomemb* 515: 395−436. https://doi.org/10.1016/0304-4157(78)90011-4
73. Bren KL (2015) NMR analysis of spin densities, *Spin States in Biochemistry and Inorganic Chemistry: Influence on Structure and Reactivity,* Wiley ch16: 409−434. https://doi.org/10.1002/9781118898277.ch16
74. Masson GR, Jenkins ML, Burke JE (2017) An overview of hydrogen deuterium exchange mass spectrometry (HDX-MS) in drug discovery. *Expert Opin Drug Dis* 12: 981−994. https://doi.org/10.1080/17460441.2017.1363734
75. Damont A, Legrand A, Cao C, et al. (2023) Hydrogen/deuterium exchange mass spectrometry in the world of small molecules. *Mass Spectrom Rev* 42: 1300−1331. https://doi.org/10.1002/mas.21765
76. Lopez del Amo J, Uwe F, Bernd R (2010) Quantification of protein backbone hydrogen-deuterium exchange rates by solid state NMR spectroscopy. *J Biomol NMR* 48: 203−212. https://doi.org/10.1007/s10858-010-9450-8
77. Ziegler SJ, Mallinson SJB, St. John PC, et al. (2021) Advances in integrative structural biology: Towards understanding protein complexes in their cellular context. *Comput Struct Biotechnol J* 19: 214−225. https://doi.org/10.1016/j.csbj.2020.11.052
78. White SH (2023) Fifty years of biophysics at the membrane frontier. *Annu Rev Biophys* 52: 21−67. https://doi.org/10.1146/annurev-biophys-051622-112341
79. Zhang L, Shen D, Yang J (2025) Advancing single-molecule biophysics: next-generation organic fluorophores with tailored labeling strategies. *Chem Biomed Imag* https://doi.org/10.1021/cbmi.5c00007
80. D'Imprima E, Kühlbrandt W (2021) Current limitations to high-resolution structure determination by single-particle cryoEM. *Quart Rev Biophys* 54: 1–15. https://doi.org/10.1017/S003358352100002
81. Piper SJ, Johnson RM, Wootten D, et al. (2022) Membranes under the magnetic lens: a dive into the diverse world of membrane protein structures using Cryo-EM. *Chem Rev* 122: 13989−14017. https://doi.org/10.1021/acs.chemrev.1c00837
82. Yael P, Schreiber G (2013) Formation of protein complexes in crowded environments–from in vitro to in vivo. *FEBS Lett* 587: 1046−1052. https://doi.org/10.1016/j.febslet.2013.01.007
83. Vikulina AS, Skirtach AG, Volodkin D (2019) Hybrids of polymer multilayers, lipids, and nanoparticles: mimicking the cellular microenvironment. *Langmuir* 35 8565−8573. https://doi.org/10.1021/acs.langmuir.8b04328
84. Moffat K (2025) Dynamics and kinetics in structural biology: the example of DNA photolyase. *Quart Rev Biophys* 58: 1–15. https://doi.org/10.1017/S0033583524000222
85. Dufourc EJ (2021) Bicelles and nanodiscs for biophysical chemistry. *BBA-Biomembranes* 1863: 183478. https://doi.org/10.1016/j.bbamem.2020.183478

86. Li H, Sun X, Cui W, et al. (2024) Computational drug development for membrane protein targets. *Nat Biotechnol* 42: 229−242. https://doi.org/10.1038/s41587-023-01987-2
87. Merz Jr KM (2014) Using quantum mechanical approaches to study biological systems. *Acc Chem Res* 47: 2804–2811. https://doi.org/10.1021/ar5001023
88. Zoler E, Meyer T, Bellón JS, et al. (2024) Promiscuous Janus kinase binding to cytokine receptors modulates signaling efficiencies and contributes to cytokine pleiotropy. *Sci Signal* 17: eadl1892. https://doi.org/10.1126/scisignal.adl1892
89. Mucs D, Bryce RA (2013) The application of quantum mechanics in structure-based drug design. *Expert Opin Drug Dis* 8: 263−276. https://doi.org/10.1517/17460441.2013.752812
90. Sousa SF, Ribeiro, AJ, Neves RP, et al. (2017) Application of quantum mechanics/molecular mechanics methods in the study of enzymatic reaction mechanisms. *Wires: Comput Mol Sci* 7: e1281. https://doi.org/10.1002/wcms.1281
91. Brinks D, Hildner R, van Dijk EMKP, et al. (2014) Ultrafast dynamics of single molecules. *Chem Soc Rev* 43: 2476−2491. https://doi.org/10.1039/C3CS60269A
92. Sharples F, Bremer A, Moss SM, et al. (2021) Quantum science concepts in enhancing sensing and imaging technologies. In Applications for Biology: Proceedings of a Workshop. Technologies: Applications for Biology: Proceedings of a Workshop. Washington, DC: The National Academies Press. https://doi.org/10.17226/26139
93. Shubin Q (2025) Comparison of classical and hybrid quantum models in secondary structure protein prediction. *Int J High School Res* 7: 87–96. https://doi.org/10.36838/v7i3.14
94. Melkikh AV, Khrennikov A (2015) Nontrivial quantum and quantum-like effects in biosystems: Unsolved questions and paradoxes. *Prog Biophys Mol Biol* 119: 137−161. https://doi.org/10.1016/j.pbiomolbio.2015.07.001
95. Aiello CD (2023) Quantum physics proposes a new way to study biology–and the results could revolutionize our understanding of how life works.
96. Ertelt M, Moretti R, Meiler J, et al. (2025) Self-supervised machine learning methods for protein design improve sampling but not the identification of high-fitness variants. *Sci Adv* 11: eadr7338. https://doi.org/10.1126/sciadv.adr7338
97. Dill KA, Ozkan SB, Weikl TR, et al. (2007) The protein folding problem: when will it be solved? *Curr Opin Struct Biol* 7: 342–346. https://doi.org/10.1016/j.sbi.2007.06.001
98. Murugan A, Husain K, Rust MJ, et al. (2021) Roadmap on biology in time varying environments. *Phys Biol* 18: 041502. https://doi.org/10.1088/1478-3975/abde8d
99. Dill KA, Ozkan SB, Shell MS, et al. (2008) The protein folding problem. *Annu Rev Biophys* 37: 289−316. https://doi.org/10.1146/annurev.biophys.37.092707.153558
100. Hansen J, Jain AR, Nenov P, et al. (2024) From transcriptomics to digital twins of organ function. *Front Cell Develop Biol* 12: 1240384. https://doi.10.3389/fcell.2024.1240384
101. Balchin D, Hayer-Hartl M, Hartl FU (2016) In vivo aspects of protein folding and quality control. *Science* 353: aac4354. https://doi.org/10.1126/science.aac4354
102. Zwanzig R, Szabo A, Bagchi B (1992) Levinthal's paradox. *P Natl A Sci* 89: 20−22. https://doi.org/10.1016/S1359-0278(97)00067-9
103. Wang L, Wen Z, Liu SW, et al. (2024) Overview of AlphaFold2 and breakthroughs in overcoming its limitations. *Comput Biol Med* 108620. https://doi.org/10.1016/j.compbiomed.2024.108620
104. David A, Islam S, Tankhilevich E, et al. (2022) The AlphaFold database of protein structures: a biologist's guide. *J Mol Biol* 434: 167336. https://doi.org/10.1016/j.jmb.2021.167336

105. Grandori R (2023) Protein structure and dynamics in the era of integrative structural biology. *Front Biophys* 1: 1219843. https://doi.org/10.3389/frbis.2023.1219843
106. Castaño-Díez D, Danev R, Baumeister W (2020) Progress in cryo-EM aided by new technological advances. *Curr Opin Struct Biol* 64: 37−48. https://doi.org/10.1016/j.sbi.2020.01.006
107. Chua EY, Mendez JH, Rapp M, et al. (2022) Better, faster, cheaper: recent advances in cryo–electron microscopy. *Annu Rev Biochem* 91: 1−32. https://doi:10.1146/annurev-biochem-032620-110705
108. D'Imprima E, Kühlbrandt W (2021) Current limitations to high-resolution structure determination by single-particle cryoEM. *Quart Rev Biophys* 54: 1–15. https://doi.org/10.1017/S003358352100002
109. Danev R, Yanagisawa H, and Kikkawa M (2019) Cryo-electron microscopy mmethodology: current aspects and future directions. *Trends Biochem Sci* 44: 837–848. https://doi.org/10.1016/j.tibs.2019.04.008
110. Wang HW, Wang JW (2017) How cryo-electron microscopy and X-ray crystallography complement each other. *Protein Sci* 26: 32–39. https://doi.org/10.1002/pro.3022
111. Marion D (2013) An introduction to biological NMR spectroscopy. *Mol Cell Proteomics* 12: 3006–3025. https://doi.org/10.1074/mcp.O113.030239
112. Sugiki T, Kobayashi N, Fujiwara T (2017) Modern technologies of solution nuclear magnetic resonance spectroscopy for three-dimensional structure determination of proteins open avenues for life scientists. *Comput Struct Biotechnol J* 15: 328–339 https://doi.org/10.1016/j.csbj.2017.04.001
113. Caffrey M (2009) Crystallizing membrane proteins for structure determination: use of lipidic mesophases. *Annu Rev Biophys* 38: 29–51. https://doi.org/10.1146/annurev.biophys.050708.133655
114. Vinothkumar KR, Henderson R (2010) Structures of membrane proteins. *Q Rev Biophys* 43: 65–158. https://doi.org/10.1017/S0033583510000041
115. Pörtner R (2023) *Biopharmaceutical Manufacturing*, Springer, https://doi.org/10.1007/978-3-031-45669-5
116. Zhou HX (2013) Influence of crowded cellular environments on protein folding, binding, and oligomerization: biological consequences and potentials of atomistic modeling. *FEBS Lett* 587: 1053−1061. https://doi.org/10.1016/j.febslet.2013.01.064
117. Singh DB, Tripathi T (2020) *Frontiers in Protein Structure, Function, and Dynamics*. Berlin: Springer. https://doi.org/10.1007/978-981-15-5530-5
118. Son A, Kim W, Park J, et al. (2024) Utilizing molecular dynamics simulations, machine learning, Cryo-EM, and NMR spectroscopy to predict and validate protein dynamics. *Int J Mol Sci* 25: 9725. https://doi.org/10.3390/ijms25179725
119. Voelz VA, Singh VR, Wedemeyer WJ, et al. (2010) Unfolded-state dynamics and structure of protein L characterized by simulation and experiment. *J Am Chem Soc* 132: 4702−4709. https://doi.org/10.1021/ja908369h
120. Kuznetsova IM, Turoverov KK, Uversky VN (2014) What macromolecular crowding can do to a protein. *Int J Mol Sci* 15: 23090−23140. https://doi.org/10.3390/ijms151223090
121. Xin F, Radivojac P (2012) Post-translational modifications induce significant yet not extreme changes to protein structure. *Bioinformatics* 28: 2905−2913. https://doi.org/10.1093/bioinformatics/bts541

122. Van Den Bedem H, Fraser JS (2015) Integrative, dynamic structural biology at atomic resolution—it's about time. *Nat Methods* 12: 307−318. https://doi.org/10.1038/nmeth.3324
123. Lerner E, Cordes T, Ingargiola A, et al. (2018) Toward dynamic structural biology: Two decades of single-molecule Förster resonance energy transfer. *Science* 359: eaan1133. https://doi.org/10.1126/science.aan1133
124. Varadi M, Tsenkov M, Velankar S (2025) Challenges in bridging the gap between protein structure prediction and functional interpretation. *Proteins* 93:400−410. https://doi.org/10.1002/prot.26614
125. Harmalkar A, Lyskov S, Gray JJ. (2024) Reliable protein-protein docking with AlphaFold, Rosetta, and replica-exchange. *eLife* 13: RP94029. https://doi.org/10.7554/eLife.94029.2
126. Wang L, Wen Z, Liu S, et al. (2024) Overview of AlphaFold2 and breakthroughs in overcoming its limitations. *Comput Biol Med* 176: 108620. https://doi.org/10.1016/j.compbiomed.2024.108620
127. Yuan R, Zhang J, Zhou J, et al. (2025) Recent progress and future challenges in structure-based protein-protein interaction prediction. *Mol Ther* 33: 2252−2268. https://doi.org/10.1016/j.ymthe.2025.04.003
128. Li J, Liu Y (2023) Data quality, data sampling and data fitting: a tutorial guide for constructing full-dimensional accurate potential energy surfaces (PESs) of molecules and reactions, *Machine Learning in Molecular Sciences,* 36: 161−201. Springer, Cham. https://doi.org/10.1007/978-3-031-37196-7_6
129. Krokidis MG, Dimitrakopoulos GN, Vrahatis AG, et al. (2024) Challenges and limitations in computational prediction of protein misfolding in neurodegenerative diseases. *Front Comput Neurosci* 17: 1323182. https://doi.org/10.3389/fncom.2023.1323182
130. Wei G, Xi W, Nussinov R., et al. (2016) Protein ensembles: how does nature harness thermodynamic fluctuations for life? The diverse functional roles of conformational ensembles in the cell. *Chem Rev* 116: 6516−6551. https://doi.org/10.1021/acs.chemrev.5b00562
131. Beck M, Covino R, Hänelt I, et al. (2024) Understanding the cell: future views of structural biology. *Cell* 187: 545−562. https://doi.org/10.1016/j.cell.2023.12.017
132. Hribar-Lee B, Lukšicˇ M (2024) Biophysical principles emerging from experiments on protein–protein association and aggregation. *Annu Rev Biophys* 53: 1–18. https://doi.org/10.1146/annurev-biophys-030722-111729
133. De Las Rivas J, Fontanillo C (2012) Protein–protein interaction networks: unraveling the wiring of molecular machines within the cell. *Brief Funct Genomics* 11: 489−496. https://doi.org/10.1093/bfgp/els036
134. Barabási AL, Bonabeau E (2003) Scale-free networks. *Sci Am* 288: 60−69. https://www.jstor.org/stable/26060284
135. Loscalzo J, Barabasi AL (2011) Systems biology and the future of medicine. Wires: *Syst Biol Med* 3: 619−627. https://doi.org/10.1002/wsbm.144
136. Goh KI, Cusick ME, Valle D, et al. (2007) The human disease networks. *P Natl A Sci* 104: 8685−8690. https://doi.org/10.1073/pnas.0701361104
137. Braun P, Gingras AC (2012) History of protein–protein interactions: from egg-white to complex networks. *Proteomics* 12: 1478−1498. https://doi.org/10.1002/pmic.201100563
138. Bajorath J (2002) Integration of virtual and high-throughput screening. *Nat Rev Drug Discov* 1: 882−894. https://doi.org/10.1038/nrd941

139. Jessulat M, Pitre S, Gui Y, et al. (2011) Recent advances in protein–protein interaction prediction: experimental and computational methods. *Expert Opin Drug Dis* 6: 921−935. https://doi.org/10.1517/17460441.2011.603722
140. Huttlin EL, Ting L, Bruckner RJ, et al. (2015) The BioPlex network: a systematic exploration of the human interactome. *Cell* 162: 425−440. https://doi.org/10.1016/j.cell.2015.06.043
141. Cizauskas C, DeBenedictis, Kelly P (2025) How the past is shaping the future of life science: The influence of automation and AI on biology. *New Biotechnol* 88: 1−11. https://doi.org/10.1016/j.nbt.2025.03.004.
142. Colonna G (2025) How digitalization and algorithms are changing the knowledge in biomedicine, preprints. https://www.preprints.org/manuscript/202507.1086/v1
143. Gendler TS (1998) Galileo and the indispensability of scientific thought experiment. *Brit J Philos Sci* 49: 397−424. https://doi.org/10.1093/bjps/49.3.397
144. Turing AM (2021) Computing machinery and intelligence (1950). *Mind* 59: 433–460. https://doi.org/10.1093/oso/9780198250791.003.0017
145. Wu ST, Zhu L, Feng XL, et al. (2025) Strategies for discovering novel hepatocellular carcinoma biomarkers. *World J Hepatol* 17: 101201. http://dx.doi.org/10.4254/wjh.v17.i2.101201
146. Gao Q, Zhu H, Dong L, et al. (2019) Integrated proteogenomic characterization of HBV-related hepatocellular carcinoma. *Cell* 179: 561−577. https://doi.org/10.1016/j.cell.2019.08.052
147. Costantini S, Capone F, Guerriero E, et al. (2010) Serum cytokine levels as putative prognostic markers in the progression of chronic HCV hepatitis to cirrhosis. *Eur Cytokine Netw* 21: 251−256. https://doi.org/10.1155/2015/707254
148. Capone F, Guerriero E, Colonna G, et al. (2014) Cytokinome profile evaluation in patients with hepatitis C virus infection. *World J Gastroentero* 20: 9261−9269. https://doi.org/10.3748/wjg.v20.i28.9261
149. Szklarczyk D, Gable AL, Nastou KC, et al. (2020) The STRING database in 2021: customizable protein–protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res* 49: D605–D612. https://doi.org/10.1093/nar/gkaa1074
150. Szklarczyk D, Kirsch R, Koutrouli M, et al. (2022) The STRING database in 2023: protein–protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res* 51: D638–D646. https://doi.org/10.1093/nar/gkac1000
151. Colonna G (2025) Unravelling the barriers: current limitations in cancer biology research and how to overcome them, preprints. https://doi.org/10.20944/preprints202505.1837.v1
152. Deane CM, Salwiński Ł, Xenarios I, et al. (2002) Protein interactions: two methods for assessment of the reliability of high throughput observations. *Mol Cell Proteom* 11: 349−356. https://doi.org/10.1074/mcp.M100037-MCP200
153. Braun P (2012) Interactome mapping for analysis of complex phenotypes: Insights from benchmarking binary interaction assays. *Proteomics* 12 1499−1518. https://doi.org/10.1002/pmic.201100598
154. Westermarck J, Ivaska J, Corthals GL (2013) Identification of Protein Interactions Involved in Cellular Signaling. *Mol Cell Proteom* 12: 1752. https://doi.org/10.1074/mcp.R113.027771
155. Brückner A, Polge C, Lentze N, et al. (2009) Yeast two-hybrid, a powerful tool for systems biology. *Int J Mol Sci* 10: 2763−2788. https://doi.org/10.3390/ijms10062763
156. Ito T, Chiba T, Ozawa R, et al. (2001) A comprehensive two-hybrid analysis to explore the yeast protein interactome. *P Natl A Sci* 98: 4569−4574. https://doi.org/10.1073/pnas.061034498

157. Ozbabacan SEA, Engin HB, Gursoy A, et al. (2011) Transient protein–protein interactions. *Protein Eng Des Sel* 24: 635−648. https://doi.org/10.1093/protein/gzr025
158. Morris OM, Torpey JH, Isaacson RL (2021) Intrinsically disordered proteins: Modes of binding with emphasis on disordered domains. *Open Biol* 11: 210222. https://doi.org/10.1098/rsob.210222
159. Bakker MJ, Svensson O, Sørensen HV, et al. (2024) Exploring the functional landscape of the p53 regulatory domain: the stabilizing role of post-translational modifications. *J Chem Theory Comput* 20: 5842–5853. https://doi.org/10.1021/acs.jctc.4c00570
160. Hlavacek WS, Faeder JR, Blinov ML, et al. (2003) The complexity of complexes in signal transduction. *Biotechnol Bioeng* 84: 783−794. https://doi.org/10.1002/bit.10842
161. Johnson S, Hunter T (2005) Kinomics: methods for deciphering the kinome. *Nat Methods* 2: 17–25. https://doi.org/10.1038/nmeth731
162. Poluri KM, Gulati K, Sarkar S (2021) Experimental methods for determination of protein–protein interactions, *Protein-Protein Interactions*, Springer, Singapore. https://doi.org/10.1007/978-981-16-1594-8_5
163. Snider J, Kotlyar M, Saraon P, et al. (2015) Fundamentals of protein interaction network mapping. *Mol Syst Biol* 11: 848. https://doi.org/10.15252/msb.20156351
164. Louche A, Salcedo SP, Bigot S (2017) Protein–protein interactions: pull-down assays, *Bacterial Protein Secretion Systems: Methods and Protocols*, New York: Springer, 247−255 https://doi.org/10.1007/978-1-4939-7033-9_20
165. Brymora A, Valova VA, Robinson PJ (2004) Protein-protein interactions identified by pull-down experiments and mass spectrometry. *Curr Protoc Cell Biol* 22: 17−25. https://doi.org/10.1002/0471143030.cb1705s22
166. Fuxreiter M (2018) Fuzziness in protein interactions—a historical perspective. *J Mol Biol* 430: 2278−2287. https://doi.org/10.1016/j.jmb.2018.02.015
167. Gentili PL (2018) The fuzziness of the molecular world and its perspectives. *Molecules* 23: 2074. https://doi.org/10.3390/molecules23082074
168. Uversky VN (2016) p53 proteoforms and intrinsic disorder: an illustration of the protein structure–function continuum concept. *Int J Mol Sci* 17: 1874. https://doi.org/10.3390/ijms17111874
169. Inae E, Liu G, Jiang M (2023) Motif-aware attribute masking for molecular graph pre-training, preprint. https://doi.org/10.48550/arXiv.2309.04589
170. Vigo M, Brown J, Conway V (2013) Benchmarking web accessibility evaluation tools: measuring the harm of sole reliance on automated tests. Proceedings of the 10th international cross-disciplinary conference on web accessibility, 1−10. https://doi.org/10.1145/2461121.2461124
171. Rajagopala SV, Uetz P (2011) Analysis of protein–protein interactions using high-throughput yeast two-hybrid screens. *Methods Mol Biol* 781: 1−29 https://doi.org/10.1007/978-1-61779-276-2_1
172. Moore N (2016) Lessons from the fatal French study BIA-10-2474. *Bmj* 353. https://doi.org/10.1136/bmj.i2727
173. Kaur R, Sidhu P, Singh S (2016) What failed BIA 10–2474 Phase I clinical trial? Global speculations and recommendations for future Phase I trials. *J Pharmacol Pharmaco* 7: 120−126. https://doi.org/10.4103/0976-500X.189661

174. Burak AT, Turkoglu I (2022) Prediction of viral-host interactions of COVID-19 by computational methods. *Chemometr Intell Lab Syst* 228: 104622. https://doi.org/10.1016/j.chemolab.2022.104622
175. Iuchi H, Kawasaki J, Kubo K, et al. (2023) Bioinformatics approaches for unveiling virus-host interactions. *Comput Struct Biotechnol J* 21: 1774−1784. https://doi.org/10.1016/j.csbj.2023.02.044
176. Domenech de Cellès M, Goult E, Casalegno JS, et al. (2022) The pitfalls of inferring virus-virus interactions from co-detection prevalence data: application to influenza and SARS-CoV-2. *Proc Royal Soc Biol Sci* 289: 20212358. https://doi.org/10.1098/rspb.2021.2358
177. Alvarez-Ponce D (2017) Recording negative results of protein–protein interaction assays: an easy way to deal with the biases and errors of interactomic data sets, Briefings in Bioinformatics 18: 1017–1020. https://doi.org/10.1093/bib/bbw075
178. Geva G, Sharan R (2010) Identification of protein complexes from co-immunoprecipitation data. *Bioinformatics* 27: 111−117. https://doi.org/10.1093/bioinformatics/btq652
179. Zhou M, Li Q, Wang PR (2016) Current experimental methods for characterizing protein–protein interactions. *Chemmedchem* 11: 738. https://doi.org/10.1002/cmdc.201500495
180. Siebenmorgen T, Zacharias M (2020) Computational prediction of protein–protein binding affinities. *Wires: Computat Mol Sci* 10: e1448. https://doi.org/10.1002/wcms.1448
181. Chaves EJ, Sartori J, Santos WM, et al. (2025) Estimating absolute protein–protein binding free energies by a super learner Model. *J Chem Inf Model* 65: 2602–2609. https://doi.org/10.1021/acs.jcim.4c01641
182. Pushkaran AC, Arabi AA (2024) From understanding diseases to drug design: can artificial intelligence bridge the gap? *Artif Intell Rev* 57: 86. https://doi.org/10.1007/s10462-024-10714-5
183. Wang H (2024) Prediction of protein-ligand binding affinity via deep learning models. *Brief Bioinform* 25: bbae081. https://doi.org/10.1093/bib/bbae081
184. Wu X, Xu M, Geng M, et al. (2023) Targeting protein modifications in metabolic diseases: molecular mechanisms and targeted therapies. *Sig Transduct Target Ther* 8: 220. https://doi.org/10.1038/s41392-023-01439-y
185. Hachiya R, Tanaka M, Itoh, M, et al. (2022) Molecular mechanism of crosstalk between immune and metabolic systems in metabolic syndrome. *Inflamm Regen* 42: 13. https://doi.org/10.1186/s41232-022-00198-7
186. Malinverni D, Babu MM (2023) Data-driven design of orthogonal protein-protein interactions. Science signaling 16: eabm4484. https://www.science.org/doi/full/10.1126/scisignal.abm4484
187. Hardy LW, and Norton PP (2004) The multiple orthogonal tools approach to define molecular causation in the validation of druggable targets. *Drug Discov Today* 9: 117−126. https://doi.org/10.1016/S1359-6446(03)02969-6
188. Day ES, Capili AD, Borysenko CW, et al. (2013) Determining the affinity and stoichiometry of interactions between unmodified proteins in solution using Biacore. *Anal Biochem* 440: 96−107. https://doi.org/10.1016/j.ab.2013.05.012
189. Navarro-Lérida I, Sánchez-Álvarez M, Del Pozo MA (2021) Post-translational modification and subcellular compartmentalization: emerging concepts on the regulation and physiopathological relevance of RhoGTPases. *Cells* 10: 1990. https://doi.org/10.3390/cells10081990
190. Pan AC, Jacobson D, Yatsenko K, et al. (2019) Atomic-level characterization of protein–protein association. *P Natl Acad Sci* 116: 4244−4249. https://doi.org/10.1073/pnas.1815431116

191. Mansueto G, Fusco G, Colonna G (2024) A tiny viral protein, SARS-CoV-2-ORF7b: functional molecular mechanisms. *Biomolecules* 14: 541. https://doi.org/10.3390/biom14050541
192. Colonna G (2025) A tiny viral protein, SARS-CoV-2-ORF7b: structural features, preprints. https://doi.org/10.20944/preprints202304.0522.v2
193. Whited AM, Johs A (2015) The interactions of peripheral membrane proteins with biological membranes. *Chem Phys Lipids* 192: 51−59. https://doi.org/10.1016/j.chemphyslip.2015.07.015
194. Oughtred R, Stark C, Breitkreutz BJ, et al. (2019) The BioGRID interaction database: 2019 update. *Nucleic Acids Res* 47: D529−D541. https://doi.org/10.1093/nar/gky1079
195. Lambert N, Chen YN, Cheng YC, et al. (2013) Quantum biology. *Nat Phys* 9: 10−18. https://doi.org/10.1038/nphys2474
196. Kim Y, Bertagna F, D'souza EM, et al. (2021) Quantum biology: an update and perspective. *Quantum Rep* 3: 80−126. https://doi.org/10.3390/quantum3010006
197. Stöhr M, Tkatchenko A (2019) Quantum mechanics of proteins in explicit water: the role of plasmon-like solute-solvent interactions. *Sci Adv* 5: eaax0024. https://www.science.org/doi/full/10.1126/sciadv.aax0024
198. Layton I, Oppenheim J, Weller-Davies Z (2024) A healthier semi-classical dynamics. *Quantum* 8: 1565. https://doi.org/10.22331/q-2024-12-16-1565
199. Basieva I, Khrennikov A, Ozawa M (2021) Quantum-like modeling in biology with open quantum systems and instruments. *Biosystems* 201: 104−328. https://doi.org/10.1016/j.biosystems.2020.104328
200. Goh BH, Tong ES, Pusparajah P (2020) Quantum biology: Does quantum physics hold the key to revolutionizing medicine? *Prog Drug Discov Biomed Sci* 3: a0000130. https://doi.org/10.36877/pddbs.a0000130
201. Alsteens D (2023) Grand challenges in biophysics. *Front Biophys* 1: 1215594. https://doi.org/10.3389/frbis.2023.1215594
202. Hatters DM (2023) Grand challenges in biomolecular condensates: structure, function, and formation. *Front Biophys* 1: 1208763. https://doi.org/10.3389/frbis.2023.1208763
203. Sikosek T, Chan HS (2014) Biophysics of protein evolution and evolutionary protein biophysics. *J R Soc Interface* 11: 100. https://doi.org/10.1098/rsif.2014.0419
204. Mateus A, Savitski MM, Piazza I (2021) The rise of proteome-wide biophysics. *Mol Syst Biol* 17: e10442. https://doi.org/10.15252/msb.202110442
205. Echave J, Wilke CO (2017) Biophysical models of protein evolution: understanding the patterns of evolutionary sequence divergence. *Annu Rev Biophys* 46: 85−103. https://doi.org/10.1146/annurev-biophys-070816-033819
206. Zur H, Tuller T (2016) Predictive biophysical modeling and understanding of the dynamics of mRNA translation and its evolution. *Nucleic Acids Res* 44: 9031−9049. https://doi.org/10.1093/nar/gkw764
207. Skolnick J, Gao M, Zhou H (2016) How special is the biochemical function of native proteins? *F1000Research* 5: F1000-Faculty. https://doi.org/10.12688/f1000research.7374.1
208. Yadav R, Gupta DP, Singh C (2025) Lock and key: quest to find the most compatible membrane mimetic for studying membrane proteins in native environment. BBA-Biomembranes 1867: 184414. https://doi.org/10.1016/j.bbamem.2025.184414

209. Staniak M, Huang T, Figueroa-Navedo AM, et al. (2025). Relative quantification of proteins and post-translational modifications in proteomic experiments with shared peptides: a weight-based approach. *Bioinformatics* 41: btaf046. https://doi.org/10.1093/bioinformatics/btaf046
210. Fumagalli C, Barberis M (2021) Breast cancer heterogeneity. *Diagnostics* 11: 1555. https://doi.org/10.3390/diagnostics11091555
211. Ribeiro AL, Dallagiovanna B (2025) The role of long non-coding RNAs in human endoderm differentiation. *Non-Coding RNA* 11: 29. https://doi.org/10.3390/ncrna11020029
212. Costantini S, Sharma A, Raucci R, et al. (2013) Genealogy of an ancient protein family: the Sirtuins, a family of disordered members. *BMC Evol Biol* 13: 60. https://doi.org/10.1186/1471-2148-13-60
213. Hu KH, Kuhn NF, Courau T, et al. (2023) Transcriptional space-time mapping identifies concerted immune and stromal cell patterns and gene programs in wound healing and cancer. *Cell Stem Cell* 30: 885−903.e10. https://doi.org/10.1016/j.stem.2023.05.001
214. Droin C, Kholtei JE, Bahar Halpern K, et al. (2021) Space-time logic of liver gene expression at sub-lobular scale. *Nat Metab* 3: 43−58. https://doi.org/10.1038/s42255-020-00323-1
215. Ross LN (2021) Causal concepts in biology: How pathways differ from mechanisms and why it matters. *Brit J Philos Sci* 72: 131−158. https://doi.org/10.1093/bjps/axy078
216. Bishop RC, Atamanspacher H, Bishop R (2002) Deterministic and indeterministic descriptions, Between Chance and Choice: Interdisciplinary Perspectives on Determinism.

AIMS Press