



HAL
open science

AlphaFold2 Update and Perspectives

Sébastien Tourlet, Ragousandirane Radjasandirane, Julien Diharce, Alexandre G de Brevern

► **To cite this version:**

Sébastien Tourlet, Ragousandirane Radjasandirane, Julien Diharce, Alexandre G de Brevern. AlphaFold2 Update and Perspectives. *BioMedInformatics*, 2023, 3, pp.378 - 390. <10.3390/biomedinformatics3020025>. <inserm-04258522>

HAL Id: inserm-04258522

<https://inserm.hal.science/inserm-04258522v1>

Submitted on 25 Oct 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Copyright - All rights reserved



Perspective

AlphaFold2 Update and Perspectives

Sébastien Tourlet ¹, Ragousandirane Radjasandirane ², Julien Diharce ² and Alexandre G. de Brevern ^{2,*} ¹ Capgemini Invent, 92130 Issy-Les-Moulineaux, France² Department of Biological Research on the Red Blood Cells, Université Paris Cité and Université des Antilles and Université de la Réunion, INSERM, BIGR, DSIMB Bioinformatics Team, F-75014 Paris, France

* Correspondence: alexandre.debrevern@univ-paris-diderot.fr; Tel.: +33-1-44493000

Abstract: Access to the three-dimensional (3D) structural information of macromolecules is of major interest in both fundamental and applied research. Obtaining this experimental data can be complex, time consuming, and costly. Therefore, in silico computational approaches are an alternative of interest, and sometimes present a unique option. In this context, the Protein Structure Prediction method AlphaFold2 represented a revolutionary advance in structural bioinformatics. Named method of the year in 2021, and widely distributed by DeepMind and EBI, it was thought at this time that protein-folding issues had been resolved. However, the reality is slightly more complex. Due to a lack of input experimental data, related to crystallographic challenges, some targets have remained highly challenging or not feasible. This perspective exercise, dedicated to a non-expert audience, discusses and correctly places AlphaFold2 methodology in its context and, above all, highlights its use, limitations, and opportunities. After a review of the interest in the 3D structure and of the previous methods used in the field, AF2 is brought into its historical context. Its spatial interests are detailed before presenting precise quantifications showing some limitations of this approach and finishing with the perspectives in the field.

Keywords: molecular modelling; protein sequences; protein structures; comparative modelling; threading; de novo; meta-servers; deep learning; CASP



Citation: Tourlet, S.; Radjasandirane, R.; Diharce, J.; de Brevern, A.G. AlphaFold2 Update and Perspectives. *BioMedInformatics* **2023**, *3*, 378–390. <https://doi.org/10.3390/biomedinformatics3020025>

Academic Editors: Pan Zheng and Bin Wang

Received: 15 March 2023

Revised: 18 April 2023

Accepted: 28 April 2023

Published: 9 May 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Foreword

The idea for this short perspective comes from multiple discussions about the real impact of AlphaFold2 (AF2) with fellow specialists, biologists, and students. We provide a simple but comprehensive overview including the expertise of researchers who deal with AF2 on a regular basis, for non-specialists such as medical doctors. AF2 is has various users. It is a method that has been discussed in an unparalleled way in recognized scientific journals (method of the year for *Nature Methods* [1], with a \$3 million award for its designers [2]) and has impacted non-specialists (e.g., the Times best inventions 2022 [3]). Statements asserting that ‘It will change everything’ [4] or ‘DeepMind AI cracks 50-year-old problem of protein folding’ [5] bring questions, especially when the reality and impact of the results differ from one research lab to another.

This strategic perspective exercise is articulated in four parts. First, we outline for the record the issues of interest in protein structure and the history of the field of three-dimensional (3D) structural model prediction. Second, we discuss more specifically the deep learning approaches in Structural Bioinformatics. Third, we present our ideas on the contributions and limitations of AF2. Finally, we conclude with perspectives for the evolution of the field.

2. Introduction

2.1. Proteins and 3D Structures

Proteins are composed of a succession of amino acids, essential biological molecules that are the building blocks of macromolecules. With 20 different types, these amino acids

trigger the folding of the three-dimensional structure. The latter support the biological functions of these entities. When the protein is small in size, it is called a peptide. In this case, its structure is often highly flexible and adaptable. In other cases, proteins can be of very large size (average size of human protein is about 800 residues). By knowing the protein structure, one can better understand what happens at the atomic level, e.g., interaction between a parasite protein and a human cell or in the case of a drug, with its target [6].

Figure 1 shows an example of an FDA-approved drug (namely Jakavi) associated with the Janus Kinase 2 protein [7,8]. This structural information allows us to precisely understand which atoms are interacting and, thus, which are essential to consider for drug design purposes.

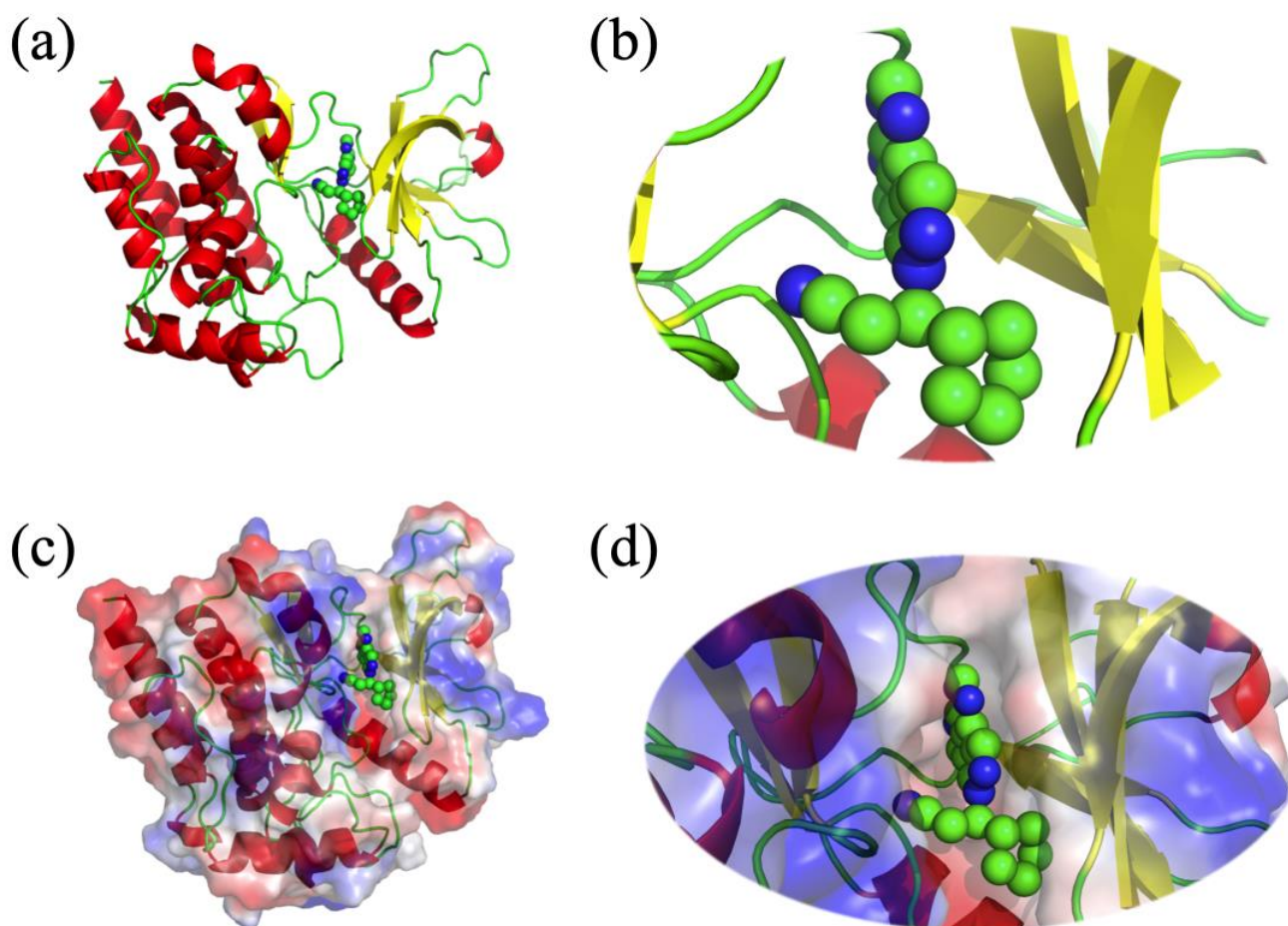


Figure 1. Jakavi™ interacting with JAK2. Taken from the Protein DataBank, the PDB id 6VGL is the Janus Kinase 2 (JAK2) JH1 domain in complex with Ruxolitinib (commercial name Jakavi) [7]. (a) Cartoon representation of JAK2 JH1 while Ruxolitinib is in sphere representation, with (b) a zoom on the interaction zone, (c) adding electrostatics and (d) zoom (red: negative charge; blue: positive charge). JAK2 variant is a driver of myeloproliferative neoplasms (MPNs); this drug has been approved by FDA against MPNs. Visualisation is done with PyMOL software [9].

Protein 3D structures have been deposited for more than 50 years in the Protein Data Bank (<https://www.rcsb.org>, accessed on 15 March 2023) [10–12]. Two other key databases are (i) the UniProt database (<https://www.uniprot.org>, accessed on 15 March 2023) [13] and (ii) the EMBL's European Bioinformatics Institute website (EBI, <https://www.ebi.ac.uk>, accessed on 15 March 2023) [14]. UniProt is particularly well known because of the large amount of information. It includes almost 555,000 experimentally manually curated pro-

teins. They are the reference proteins for biology, and since 2021, Uniprot has listed also the best AlphaFold model generated for each entry in the “Structure” section. EBI has a large number of tools and databases, mainly concerning protein sequences. It has been central to the diffusion of the AlphaFold2 models.

2.2. Protein Structure Prediction

High throughput sequencing approaches allow for obtaining particularly large masses of sequences. However, obtaining experimental structures remains highly complex (expensive and sometimes not possible, for example, transmembrane proteins). In consequence, the number of available structures remains highly limited compared to the tremendous number of currently available sequences.

For more than 40 years, techniques have been developed to predict 3D structural models of proteins from their sequences alone. The first approaches were rather simple, namely homology/comparative modelling; the most well-known and successful approach in this field is the Modeller from Sali’s lab [15]. It has been maintained and improved since 1993 (<https://salilab.org/modeller/>, accessed on 15 March 2023) [16,17]. Its main principles are, considering a sequence of interest without known structures, (i) to find a sequence sharing a good sequence identity or similarity, i.e., that can be aligned; (ii) to use this sequence alignment to build a local structural analogue (performing a copy/paste from the sequence with a 3D structure to the query one); and (iii) to optimize the global fold. The closer the sequences are, the easier the process is. Similarly, the fewer the gaps (insertions/deletions), the easier it is. Specialists often manually adapt the alignment and add extra constraints, such as multiple template processes or partial structures for specific portions of the query sequence [18]. The modeller is free for academics and has been incorporated in commercial solutions. An effective comparable approach, named SwissModel, is also freely available (<https://swissmodel.expasy.org>, accessed on 15 March 2023) [19,20].

This approach is based on the hypothesis of accumulation of neutral mutations during evolution, leading to the famous sentence: ‘sequence is less conserved than the structure’, as the structure must conserve its properties to provide the protein biological functions. However, as mutations accumulate, the alignment becomes impossible, even if the fold is preserved.

Following this idea, the next generation of Protein Structural Prediction (PSP) was developed and named the ‘threading approach’. Using more complex compatibility measures, threading approaches are able to find structural compatibility. PSP calculates the energetic compatibility of the fold when no alignments are possible using classical approaches. An iconic example of this method is Phyre2 (<http://www.sbg.bio.ic.ac.uk/phyre2/>, accessed on 15 March 2023) [21]. There are also some ab initio approaches based only on physical parameters, i.e., trying to mimic the protein folding process and parameters; however, these are mainly limited to small proteins [22].

A first generation of web servers providing consensus approaches of homology/comparative modelling, threading and ab initio, have been proposed, e.g., ‘Frankenstein’s monster’, which is aptly named [23]. However, it was the second generation of web servers with de novo approaches that resulted in significant advances [24]. The de novo approaches used a large series of evolution, energetic, and optimisation processes. They necessitated large clusters and CPU powers. Two main groups were leaders in this field: I-Tasser from Zhang’s lab (<https://zhanggroup.org/I-TASSER/>, accessed on 15 March 2023) [25–27] and the Rosetta/Robetta suite from Baker’s lab (<https://rosetta.bakerlab.org>, accessed on 15 March 2023) [28–30]. These approaches rely on complex computational searches to find distant associations.

2.3. Recent Protein Structure Prediction Methods

Since the beginning of the 2010s, artificial neural networks (ANNs)/machine learning approaches have undergone a phenomenal evolution, with the emergence of deep learning (DL). ANNs were known to have a succession of a small number of layers with thousands of parameters to be optimized. DL, thanks to new Graphical Process Units (GPUs) and innovative developments, allowed the design of highly complex architectures with dozens of layers and tens of millions of parameters to be optimized. DL was first known through Go and chess play of human versus machine as well as through automatic recognition of faces as proposed by Facebook (now called Meta) in 2011.

In 2018, AlphaFold1, developed by DeepMind (a subsidiary of Alphabet Inc., Mountain View, CA, USA), was the first complex DL applied to PSP. It was perfectly comparable to most of the current approaches [31,32], such as I-Tasser [27]. Two years later, the second version showed an exceptional improvement [33–35]. It moved beyond the classical methods, which is why we will discuss it in the following sections. Since then, several other similar approaches have been made available, such as RoseTTAfold by Baker's group [36] and by Meta Platforms, Inc. [37]. However, they are considered less efficient.

3. Deep Learning in Structural Bioinformatics

Every two years since 1994, a competition known as the Critical Assessment of Structural Prediction (CASP, <https://predictioncenter.org>, accessed on 15 March 2023) has evaluated the methods of structural model prediction. For this reason, researchers who obtain new protein structures in this two-year period do not deposit them directly in the PDB. They provide their sequences to the committee, and competitors have access to them, in order to propose a structural model. Obviously, the latter must be close to the experimental structure, not yet revealed. They deposit their proposal on a dedicated website, which is evaluated by the committee [29,38].

In 2018, AlphaFold1 was the first Deep Learning method present at CASP [31,32]. It showed very positive results, but older approaches gave comparable structural models in most of the cases. Two years later, DeepMind completed a large amount of refinement work by modifying several new pieces of information in their model architecture [39]. Surprisingly, the information and general parameters used were very similar to those of de novo approaches, i.e., sequence evolution and optimization of local and global folding [40]. These Deep Learning approaches derive from historical neural network approaches, but at higher level (millions of parameters optimized, thanks to the power of GPUs) [41].

The results of AlphaFold2 are very significant; it outdoes the second-best approach (from David Baker's group) by a large margin, and above all, it proposes models that are sometimes of atomistic quality (equivalent to the experimental structure) and thus useful for numerous purposes [32,33,35,42]. For example, AlphaFold2 models may be used for protein engineering and prediction of entirely computational sequences. In addition, drug discovery could be an important field of application of this methodology. Because of their quality, models coming from this software could also be good starting points for molecular dynamics simulations or docking calculations. These few examples and the corresponding figures quickly spread on Twitter, then the scientific media, and finally the general media. AlphaFold2 was named method of the year by Science and Nature [1,43]. The original paper had 8965 citations in less than 2 years (Google Scholar, accessed 1 March 2023). The researchers behind the AlphaFold artificial-intelligence (AI) system won one of this year's US \$3 million Breakthrough prizes—the most lucrative award in science [2]. Nowadays, AlphaFold2 also proposes the prediction of protein–protein multimeric complexes, called AlphaFold-multimers, with interesting results [44].

Another interesting point that contributes to its success is the possibility to locally install the program on a computer. Though the device must have a good GPU and memory, this is attainable even for a small academic laboratory. Moreover, AlphaFold2 is faster than other methods. Indeed, it can propose a model within days, whereas some methods take weeks to produce a prediction [34]. In addition, researchers have made their AlphaFold2

personal installation available for free to the scientific community (CollabFold based on the Google forms, https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/beta/AlphaFold2_advanced.ipynb, accessed on 15 March 2023 [45]), with great success. Similarly, the European Bioinformatics Institute (EBI) has proposed dedicated databases for all human proteins, as for other organisms. Those models were generated by the AlphaFold2 approach; eventually, the whole UniProt database encompassed, in the structure section, the AlphaFold results of about 200 million structural models (<https://www.alphafold.ebi.ac.uk>, accessed on 15 March 2023) [46,47].

More recently, competing approaches have been made available, such as RoseTTAfold [36,48] and ESMFold by Meta [37]. This latter also made available more than 600 million structural models coming from various meta-genomes [37]. These approaches are very similar to AF2 but, on average, perform less well.

4. Current Identified AlphaFold2 Limits

4.1. Introduction

The emergence of the AF2 algorithm was indeed game-changing in PSP, but there remain significant issues, despite the strong statements proposed by the non-specialist media [49]. Indeed, one question that may highlight some flaws of AF2 is as follows: does AF2 propose quality models for every potential protein? Here, we aim to answer this question by considering different types of proteins and their complexities. In addition, we look at the different precise quantifications that will help us to have a better and less biased view of AF2.

4.2. Law and Order

An essential first point to remember is that 15 to 20% of the residues are not associated with an ordered region, experimentally defined, but are disordered (highly flexible and therefore not experimentally characterized) [50].

AF2 provides a confidence index (pLDDT), informing about the quality of the folding at a level of the residue. Numerous works have shown a correlation between low pLDDT and disorder [51–53]. Nonetheless, studies have shown that low pLDDT is not only associated with disorder but also with poor-quality prediction [54,55].

It remains important to remember that AlphaFold2 learns the experimental data and, therefore, potentially the errors that are associated. Three classes of proteins exist: (i) globular (98% of the PDB), (ii) transmembrane (2%), and (iii) fibber (almost nothing). Only the globular proteins are strongly represented. For transmembrane proteins, this low number (and also with much redundancy) is related to the lipid bilayer. By consequence, it is an unbalanced learning.

Finally, the number of targets to be reached, i.e., a very low number, always limits the CASP competition. It is also difficult to simply and directly generalize the results on all protein datasets.

4.3. Few Pertinent Quantifications

Four primary studies provide a better understanding of the general quality of AlphaFold2 models:

- (i) The first work, coming directly from the EBI with the help of DeepMind, made available structural models of 220 million sequences (including proteomes of interest, UniProt, and other sequences) [46,47]. With the analysis of the pLDDT, they clearly show that a third of the human proteome is of atomistic quality; 58% corresponds to a correct fold, and there is no clear information on remaining 42%. This tendency is also found on several of the other available proteomes. Of course, this result may be partly due to disordered areas, but they are considered in smaller numbers. Clearly, a significant number of transmembrane protein structures are not correctly predicted [56]. It was recently confirmed with the analysis of nearly 700,000 domains

provided by AF2 that only 52% of models were appropriate for analysis with the new CATH-annotation tool [57].

- (ii) The second work came from a large academic consortium that has independently evaluated the advancements provided by this methodology compared to a recognized comparative modelling tool, namely SwissModel [19], with its own repository [58]. They selected 21 model species, corresponding to more than 365,000 proteins, i.e., twice the number of experimental structures and six times the number of unique proteins in PDB. They analysed the SwissModel repository for 11 model species and compared it with the AF2 database. On average, the predicted models of AF2 provide longer predictions (+44% of residues). Looking at high-quality regions (pLDDT > 90), an average of around 25% of the residues of the proteomes of the 11 model species are covered by AF2 with novel (not present in SwissModel repository) and confident predictions [59]. This very elegant and rigorous study also shows, similar to the previous analysis, that a large number of proteins are still not reachable. The surprise for non-specialists is that they are not only transmembrane proteins but also globular ones.
- (iii) The third study comprises the analysis of the local conformation of proteins and shows that, globally, the results are very good. However, in a surprising way, some local conformations observed in a recurrent way within all the proteins are, in a way, systematically associated with particularly low confidence scores [60]. These conformations are PolyProline II helices (important for protein-proline interaction) [61], γ -turns (present in many loops), and ω angles in cis conformation (often associated with Proline) [62]. This analysis also shows that there would be an under-representation of sheets and beta compared to what should be observed. In addition, those β -like forms are present in large numbers and would only ask to be able to form sheets.
- (iv) The fourth study focuses on the position of the side chains, a complex subject due to their large panel of motions. The quality of the predictions at this level is still largely perfectible [63].

As a partial conclusion of these four representative studies, we can assert the following:

- (a) An undeniable strength of AlphaFold 2 compared to its previous version is that it is made available in the form of a usable and stable GitHub, which does not require overly expensive and powerful computers. In addition, several academic groups provide their own AlphaFold system, called CollabFold, which can be used free of charge by the scientific community, but with a less rich database of protein structures compared to the real AlphaFold.
- (b) AlphaFold2 can quickly model more protein than previous approaches and, on average, with better quality. However, the attainable/usable protein number is lower than we would have expected from the assertion that “a 50-year-old problem was solved”.
- (c) One AF2 limitation that strongly affects the biomedical field is the poor quality of transmembrane protein models, whereas the confidence indices of transmembrane segments can be of good quality. The overall predicted topology is not compatible with its insertion within a membrane bilayer. Figure 2 shows the AF2 model proposed for Atypical Chemokine Receptor 1 (ACKR1), previously named Duffy Antigen for Chemokine. ACKR1 is a seven TM protein associated with malarial *Plasmodium vivax* infection [64,65]. This model is incomplete, but segments are present. However, it is not possible to insert it in a membrane bilayer, e.g., with CHARMM-GUI [66]. Indeed, its topology does not allow any recognition of this portion as transmembrane by the webserver.



Figure 2. AF2 model of ACKR1. Model taken from UniProt website [67].

- (d) Another limitation is that most proteins have ions and co-factors (such as FAD and NADPH); however, AlphaFold 2 has not been trained to take them into account. In addition, in a certain number of cases, it is impossible to add them or to dock them. Thus, an external tool, called AlphaFill, has been dedicated to place them and add post-translational modifications such as glycosylation (often essential for the protein functions) on the models [68].
- (e) As noticed, AlphaFold2 is also pertinent for underlying intrinsic disordered regions (IDRs). Nonetheless, sometimes AF2 provides protein models with regions that look like IDRs but are in reality not disordered. Figure 3 presents the E3 ubiquitin-protein ligase PPP1R11 (UniProt ID O60927) AF2 model underlined by Thornton and collaborators in [69]. This model is considered as a poor-quality model because this protein is a globular one, and the model presents what looks like a disordered protein.

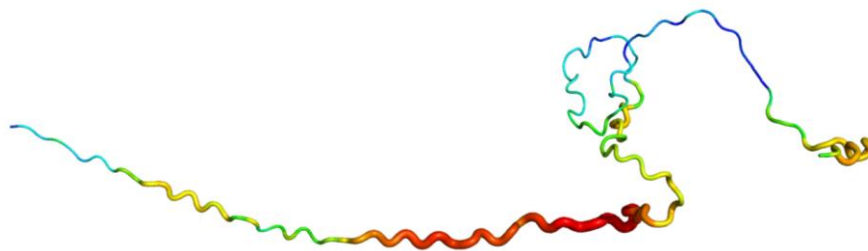


Figure 3. AF2 model of the E3 ubiquitin-protein ligase PPP1R11 (UniProt ID O60927). Model taken from UniProt website [67].

5. Perspectives

Future developments are not easy to predict, as all the structural information has been used to develop the AlphaFold models. For example, DeepMind did not participate in the last CASP competition, and results have not shown new cutting-edge data. Many academic teams have used AlphaFold code and adapted it.

In fact, current and future concerns are around the proper use of AlphaFold2 models. As in the previous approaches, it is always advisable to evaluate the quality of the model obtained. The confidence index (pLDDT) allows this, in part. Thus, a sequence having a structural model whose pLDDT exceeds 90 is particularly useful and undoubtedly of atomistic quality, but this is not always the case.

For instance, Figure 4 shows a recent example of the Scianna blood group. The AlphaFold2 model (see Figure 4a) provides a too-long helix, and some domains known

from the literature are not properly designed. On the other hand, a manual human-supervised model (see Figure 4b) encompasses each region in a better way. Obviously, a longer time was needed to design it (analysis of different domains, evolution, prediction of properties, multiple optimisations, etc.). It should be noted that all other automatic approaches have the same issues [70].

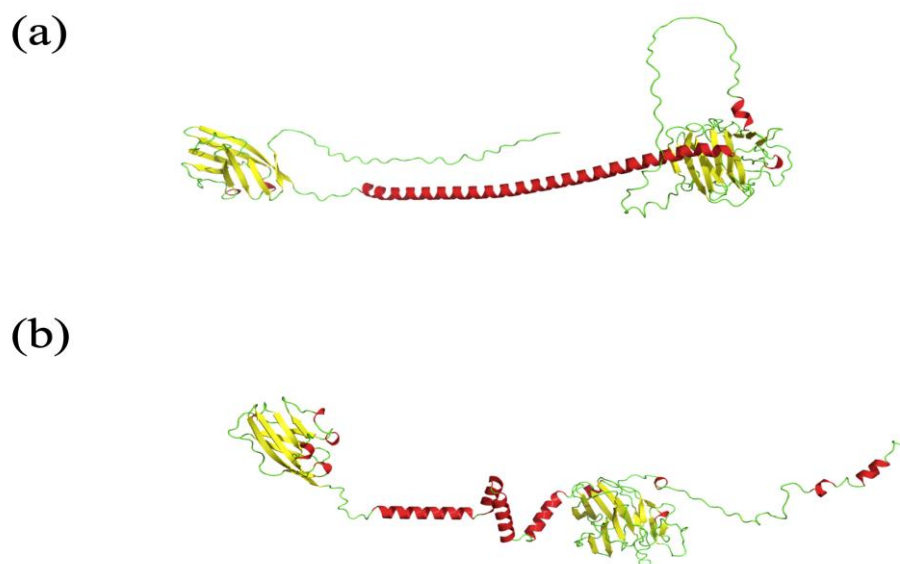


Figure 4. Human-supervised vs. AF2 model. (a) Model obtained from AF2; (b) human-supervised model made with Modeller [15] and taken from [70].

Another important point is that the structural model, even of excellent quality, is not final. For instance, single nucleotide polymorphisms (i.e., changes the amino acids) could be the driving force behind pathologies. A recent study showed that these point mutations placed on very good quality models do not provide information on the mechanism [71] but need extra treatments [72]. Similarly, AF2 models are used to establish links with physico-chemical parameters and experiments [73,74]. TAGPPI extracts multi-dimensional features by employing 1D convolution operation on protein sequences and graph learning methods on contact maps constructed from AlphaFold [75]. However, it is often indirect; e.g., PROST extracts several descriptors from the most promising sequence-based predictors, such as BoostDDG, SAAFEC-SEQ, and DDGun, but also from iFeature and AlphaFold2 descriptors [76]. Protein dynamics must be taken into account, as some effects can be of long range [77–79]. Some new DL developments are in progress [80,81]. It is essential to note that AF2 provides only a single topology, without giving any indication of conformational change, and that the notions of dynamics are to be taken into account to answer biological questions [82,83].

A question arises: Will AF2 have a negative impact on Structural Biology? Because of its existence, would it no longer be necessary to conduct X-ray crystallography or Nuclear Magnetic Resonance? The contribution of AF2 is twofold and positive in the short and medium term. First, AF2 helps experimenters to obtain their structures. AF2 models can be used directly in electron density maps and can solve complex cases [84–86]. Second, AF2 has brought back to the forefront the problem of proteins that are not yet predictable by AF2. More experimental data are needed to make advancements in the space of known folding. Hence, crystallographic strategies and high-accuracy models are being continually developed. The AlphaFold framework helps crystallographers to concentrate their work on the most unresolved and challenging structures, e.g., coiled—coils inducing modulations and challenges in crystallographic method development.

A final question is the possibility of improvement of AF2 and its competitors. A first point to take into account is the question of the topology of deep learning approaches.

New developments offer relative improvements and can allow small gains. During the last CASP, AF2 was not present. Self-based or similar approaches failed to show significant gains. The second essential point is the need for experimental data to increase our body of knowledge on difficult targets.

6. Conclusions

The interest in 3D structures includes the knowledge of the topology of proteins and the possibility to access the positions of the cofactors, ions, ligands, DNA, RNA, etc., that interact with them. In particular, 3D structures can facilitate the design of drugs (simple peptide, modified, or exclusively chemical types), bringing explanatory power in the functioning of the therapeutic molecule (which is currently required by the FDA).

AF2 has changed many things. However, as seen with the ACKR1 model of Figure 2 or the Scianna model of Figure 4, modelling research design has not changed greatly. Figure 5a shows the different steps of the classical protocol to provide correct 3D structural models, e.g., [70,87–89]. Figure 5b shows the actual protocol; AF2 has not replaced everything but is added as a new step in the design.

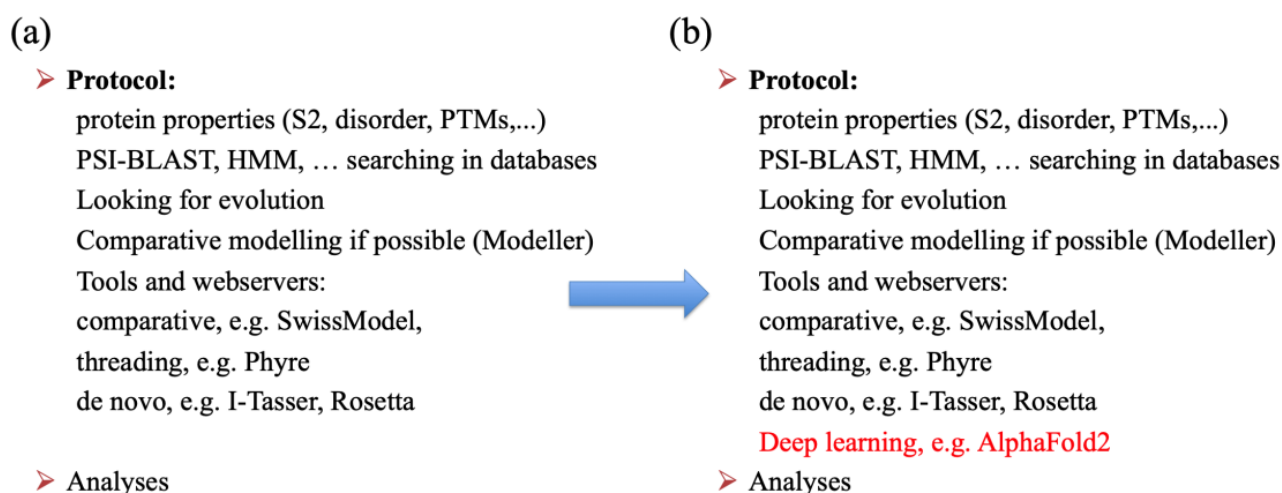


Figure 5. Example of PSP protocols. (a) A classical protocol and (b) a new protocol encompassing the use of AF2. At first, it is essential to analyse the protein sequence and its properties, then, looking at the information available in the different databases, the evolution is determined. Following this, the first building of structural models can begin, with (if possible) comparative modelling, threading, and de novo approaches. The most important point in the analyses is that it must be critical.

To conclude, it has been proven that experimentally determined X-ray structures (3.5 Å resolution or better) are more reliable than AlphaFold2-computed structure models. Experimental structure should be used preferentially whenever possible. Improvement of AF2 models, even the top ones, is possible in the future [90].

Author Contributions: Conceptualization, S.T. and A.G.d.B.; methodology, R.R., J.D. and A.G.d.B.; resources, S.T. and A.G.d.B.; writing—original draft preparation, A.G.d.B.; writing—review and editing, S.T., R.R., J.D. and A.G.d.B.; visualization, R.R. and A.G.d.B.; supervision, A.G.d.B.; project administration, A.G.d.B.; funding acquisition, A.G.d.B. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by grants from the Ministry of Research (France), Université Paris Cité (formerly University Paris Diderot, Sorbonne, Paris Cité, France and formerly Université de Paris), Université de la Réunion, National Institute for Blood Transfusion (INTS, France), National Institute for Health and Medical Research (INSERM, France), IdEx ANR-18-IDEX-0001 and labex GR-Ex. The labex GR-Ex, reference ANR-11-LABX-0051, is funded by the program “Investissements d’avenir” of the French National Research Agency, reference ANR-11-IDEX-0005-02.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: We would like to thank Catherine Etchebest and Jean-Christophe Gelly for fruitful discussions.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Service, R. Breakthrough of the Year—Protein Structures for All. *Science*, 16 December 2021. Available online: <https://www.science.org/content/article/breakthrough-2021> (accessed on 15 March 2023).
2. Knapp, A. 2023 Breakthrough Prizes Announced: Deepmind's Protein Folders Awarded \$3 Million. *Forbes*, 22 September 2022. Available online: <https://www.forbes.com/sites/alexknapp/2022/09/22/2023-breakthrough-prizes-announced-deepminds-protein-folders-awarded-3-million/> (accessed on 15 March 2023).
3. Perrigo, B. Mapping Life—DeepMind AlphaFold. *Time*, 10 November 2022. Available online: <https://time.com/collection/best-inventions-2022/6229912/deepmind-alphafold/> (accessed on 15 March 2023).
4. Callaway, E. 'It will change everything': DeepMind's AI makes gigantic leap in solving protein structures. *Nature* **2020**, *588*, 203–204. [CrossRef]
5. Sample, I. DeepMind AI Cracks 50-Year-Old Problem of Protein Folding. *Guardian* 2020. Available online: <https://www.theguardian.com/technology/2020/nov/2030/deepmind-ai-cracks-2050-year-old-problem-of-biology-research> (accessed on 15 March 2023).
6. Rabbani, G.; Baig, M.H.; Ahmad, K.; Choi, I. Protein-protein Interactions and their Role in Various Diseases and their Prediction Techniques. *Curr. Protein Pept. Sci.* **2018**, *19*, 948–957. [CrossRef]
7. Davis, R.R.; Li, B.; Yun, S.Y.; Chan, A.; Nareddy, P.; Gunawan, S.; Ayaz, M.; Lawrence, H.R.; Reuther, G.W.; Lawrence, N.J.; et al. Structural Insights into jak2 Inhibition by Ruxolitinib, Fedratinib, and Derivatives Thereof. *J. Med. Chem.* **2021**, *64*, 2228–2241. [CrossRef]
8. McLornan, D.P.; Pope, J.E.; Gotlib, J.; Harrison, C.N. Current and future status of jak inhibitors. *Lancet* **2021**, *398*, 803–816. [CrossRef]
9. Delano, W.L. The Pymol Molecular Graphics System on World Wide Web. 2013. Available online: <http://www.pymol.org> (accessed on 15 March 2023).
10. Berman, H.M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.N.; Weissig, H.; Shindyalov, I.N.; Bourne, P.E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242. [CrossRef]
11. Berman, H.M.; Battistuz, T.; Bhat, T.N.; Bluhm, W.F.; Bourne, P.E.; Burkhardt, K.; Feng, Z.; Gilliland, G.L.; Iype, L.; Jain, S.; et al. The Protein Data Bank. *Acta Cryst. D Biol. Cryst.* **2002**, *58*, 899–907. [CrossRef]
12. Bernstein, F.C.; Koetzle, T.F.; Williams, G.J.; Meyer, E.F., Jr.; Brice, M.D.; Rodgers, J.R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. The Protein Data Bank: A computer-based archival file for macromolecular structures. *J. Mol. Biol.* **1977**, *112*, 535–542. [CrossRef]
13. Apweiler, R.; Bairoch, A.; Wu, C.H.; Barker, W.C.; Boeckmann, B.; Ferro, S.; Yeh, L.S. UniProt: The Universal Protein Knowledgebase in 2023. *Nucleic Acids Res.* **2023**, *51*, D523–D531.
14. Thakur, M.; Bateman, A.; Brooksbank, C.; Freeberg, M.; Harrison, M.; Hartley, M.; Keane, T.; Kleywegt, G.; Leach, A.; Levchenko, M.; et al. EMBL's European Bioinformatics Institute (EMBL-EBI) in 2022. *Nucleic Acids Res.* **2023**, *51*, D9–D17. [CrossRef]
15. Sali, A.; Blundell, T.L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **1993**, *234*, 779–815. [CrossRef]
16. Pieper, U.; Eswar, N.; Stuart, A.C.; Ilyin, V.A.; Sali, A. Modbase, a database of annotated comparative protein structure models. *Nucleic Acids Res.* **2002**, *30*, 255–259. [CrossRef] [PubMed]
17. Melo, F.; Sali, A. Fold assessment for comparative protein structure modeling. *Protein Sci.* **2007**, *16*, 2412–2426. [CrossRef] [PubMed]
18. Webb, B.; Sali, A. Comparative Protein Structure Modeling Using modeller. *Curr. Protoc. Bioinform.* **2016**, *54*, 5.6.1–5.6.37. [CrossRef] [PubMed]
19. Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F.T.; de Beer, T.A.P.; Rempfer, C.; Bordoli, L.; et al. Swiss-model: Homology modelling of protein structures and complexes. *Nucleic Acids Res.* **2018**, *46*, W296–W303. [CrossRef] [PubMed]
20. Schwede, T.; Kopp, J.; Guex, N.; Peitsch, M.C. Swiss-model: An automated protein homology-modeling server. *Nucleic Acids Res.* **2003**, *31*, 3381–3385. [CrossRef] [PubMed]
21. Kelley, L.A.; Mezulis, S.; Yates, C.M.; Wass, M.N.; Sternberg, M.J. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **2015**, *10*, 845–858. [CrossRef]

22. Bystroff, C.; Shao, Y. Fully automated ab initio protein structure prediction using i-sites, hmmstr and rosetta. *Bioinformatics* **2002**, *18* (Suppl. S1), S54–S61. [[CrossRef](#)]
23. Kosinski, J.; Cymerman, I.A.; Feder, M.; Kurowski, M.A.; Sasin, J.M.; Bujnicki, J.M. A “Frankenstein’s monster” approach to comparative modeling: Merging the finest fragments of Fold-Recognition models and iterative model refinement aided by 3D structure evaluation. *Proteins* **2003**, *53* (Suppl. S6), 369–379. [[CrossRef](#)]
24. Baker, D.; Sali, A. Protein structure prediction and structural genomics. *Science* **2001**, *294*, 93–96. [[CrossRef](#)]
25. Roy, A.; Kucukural, A.; Zhang, Y. I-tasser: A unified platform for automated protein structure and function prediction. *Nat. Protoc.* **2010**, *5*, 725–738. [[CrossRef](#)]
26. Zhang, Y. Interplay of I-tasser and quark for template-based and ab initio protein structure prediction in casp10. *Proteins* **2014**, *82* (Suppl. S2), 175–187. [[CrossRef](#)] [[PubMed](#)]
27. Yang, J.; Yan, R.; Roy, A.; Xu, D.; Poisson, J.; Zhang, Y. The i-tasser Suite: Protein structure and function prediction. *Nat. Methods* **2015**, *12*, 7–8. [[CrossRef](#)] [[PubMed](#)]
28. Bradley, P.; Malmström, L.; Qian, B.; Schonbrun, J.; Chivian, D.; Kim, D.E.; Meiler, J.; Misura, K.M.; Baker, D. Free modeling with Rosetta in casp6. *Proteins* **2005**, *61* (Suppl. S7), 128–134. [[CrossRef](#)] [[PubMed](#)]
29. Kinch, L.N.; Li, W.; Monastyrskyy, B.; Kryshchak, A.; Grishin, N.V. Evaluation of free modeling targets in casp11 and roll. *Proteins* **2016**, *84* (Suppl. S1), 51–66. [[CrossRef](#)]
30. Leman, J.K.; Weitzner, B.D.; Lewis, S.M.; Adolf-Bryfogle, J.; Alam, N.; Alford, R.F.; Aprahamian, M.; Baker, D.; Barlow, K.A.; Barth, P.; et al. Macromolecular modeling and design in Rosetta: Recent methods and frameworks. *Nat. Methods* **2020**, *17*, 665–680. [[CrossRef](#)]
31. AlQuraishi, M. AlphaFold at casp13. *Bioinformatics* **2019**, *35*, 4862–4865. [[CrossRef](#)]
32. Senior, A.W.; Evans, R.; Jumper, J.; Kirkpatrick, J.; Sifre, L.; Green, T.; Qin, C.; Židek, A.; Nelson, A.W.R.; Bridgland, A.; et al. Protein structure prediction using multiple deep neural networks in the 13th Critical Assessment of Protein Structure Prediction (casp13). *Proteins* **2019**, *87*, 1141–1148. [[CrossRef](#)]
33. Senior, A.W.; Evans, R.; Jumper, J.; Kirkpatrick, J.; Sifre, L.; Green, T.; Qin, C.; Židek, A.; Nelson, A.W.R.; Bridgland, A.; et al. Improved protein structure prediction using potentials from deep learning. *Nature* **2020**, *577*, 706–710. [[CrossRef](#)]
34. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Židek, A.; Potapenko, A.; et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596*, 583–589. [[CrossRef](#)]
35. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Židek, A.; Potapenko, A.; et al. Applying and improving AlphaFold at casp14. *Proteins* **2021**, *89*, 1711–1721. [[CrossRef](#)]
36. Baek, M.; DiMaio, F.; Anishchenko, I.; Dauparas, J.; Ovchinnikov, S.; Lee, G.R.; Wang, J.; Cong, Q.; Kinch, L.N.; Schaeffer, R.D.; et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **2021**, *373*, 871–876. [[CrossRef](#)] [[PubMed](#)]
37. Lin, Z.; Akin, H.; Rao, R.; Hie, B.; Zhu, Z.; Lu, W.; Smetanin, N.; Verkuil, R.; Kabeli, O.; Shmueli, Y.; et al. Evolutionary-scale prediction of atomic level protein structure with a language model. *bioRxiv* **2022**. [[CrossRef](#)] [[PubMed](#)]
38. Mosimann, S.; Meleshko, R.; James, M.N. A critical assessment of comparative molecular modeling of tertiary structures of proteins. *Proteins* **1995**, *23*, 301–317. [[CrossRef](#)] [[PubMed](#)]
39. Jumper, J.; Hassabis, D. Protein structure predictions to atomic accuracy with AlphaFold. *Nat. Methods* **2022**, *19*, 11–12. [[CrossRef](#)] [[PubMed](#)]
40. Marcu, Ș.B.; Tăbircă, S.; Tangney, M. An Overview of AlphaFold’s Breakthrough. *Front. Artif. Intell.* **2022**, *5*, 875587. [[CrossRef](#)] [[PubMed](#)]
41. Fersht, A.R. AlphaFold—A Personal Perspective on the Impact of Machine Learning. *J. Mol. Biol.* **2021**, *433*, 167088. [[CrossRef](#)]
42. Skolnick, J.; Gao, M.; Zhou, H.; Singh, S. AlphaFold 2: Why It Works and Its Implications for Understanding the Relationships of Protein Sequence, Structure, and Function. *J. Chem. Inf. Model.* **2021**, *61*, 4827–4831. [[CrossRef](#)]
43. Method of the Year 2021: Protein structure prediction. *Nat. Methods* **2022**, *19*, 1. [[CrossRef](#)]
44. Evans, R.; O’Neill, M.; Pritzel, A.; Antropova, N.; Senior, A.; Green, T.; Židek, A.; Bates, R.; Blackwell, S.; Yim, J.; et al. Protein complex prediction with AlphaFold-Multimer. *bioRxiv* **2021**. [[CrossRef](#)]
45. Mirdita, M.; Schütze, K.; Moriwaki, Y.; Heo, L.; Ovchinnikov, S.; Steinegger, M. ColabFold—Making protein folding accessible to all. *Nat. Methods* **2022**, *19*, 679–682. [[CrossRef](#)]
46. Tunyasuvunakool, K.; Adler, J.; Wu, Z.; Green, T.; Zielinski, M.; Židek, A.; Bridgland, A.; Cowie, A.; Meyer, C.; Laydon, A.; et al. Highly accurate protein structure prediction for the human proteome. *Nature* **2021**, *596*, 590–596. [[CrossRef](#)] [[PubMed](#)]
47. Varadi, M.; Anyango, S.; Deshpande, M.; Nair, S.; Natassia, C.; Yordanova, G.; Yuan, D.; Stroe, O.; Wood, G.; Laydon, A.; et al. AlphaFold Protein Structure Database: Massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* **2022**, *50*, D439–D444. [[CrossRef](#)]
48. Du, Z.; Su, H.; Wang, W.; Ye, L.; Wei, H.; Peng, Z.; Anishchenko, I.; Baker, D.; Yang, J. The trRosetta server for fast and accurate protein structure prediction. *Nat. Protoc.* **2021**, *16*, 5634–5651. [[CrossRef](#)]
49. Al-Janabi, A. Has DeepMind’s AlphaFold solved the protein folding problem? *BioTechniques* **2022**, *72*, 73–76. [[CrossRef](#)] [[PubMed](#)]
50. DeForte, S.; Uversky, V.N. Order, Disorder, and Everything in Between. *Molecules* **2016**, *21*, 1090. [[CrossRef](#)] [[PubMed](#)]
51. He, J.; Turzo, S.B.A.; Seffernick, J.T.; Kim, S.S.; Lindert, S. Prediction of Intrinsic Disorder Using Rosetta ResidueDisorder and AlphaFold2. *J. Phys. Chem. B* **2022**, *126*, 8439–8446. [[CrossRef](#)] [[PubMed](#)]

52. Ma, P.; Li, D.W.; Brüschweiler, R. Predicting protein flexibility with AlphaFold. *Proteins* **2023**, *91*, 847–855. [[CrossRef](#)]
53. Piovesan, D.; Monzon, A.M.; Tosatto, S.C.E. Intrinsic protein disorder and conditional folding in AlphaFoldDB. *Protein Sci.* **2022**, *31*, e4466. [[CrossRef](#)]
54. Bruley, A.; Bitard-Feildel, T.; Callebaut, I.; Duprat, E. A sequence-based foldability score combined with AlphaFold2 predictions to disentangle the protein order/disorder continuum. *Proteins* **2023**, *91*, 466–484. [[CrossRef](#)]
55. Bruley, A.; Mornon, J.P.; Duprat, E.; Callebaut, I. Digging into the 3D Structure Predictions of AlphaFold2 with Low Confidence: Disorder and Beyond. *Biomolecules* **2022**, *12*, 1467. [[CrossRef](#)]
56. Azzaz, F.; Yahi, N.; Chahinian, H.; Fantini, J. The Epigenetic Dimension of Protein Structure Is an Intrinsic Weakness of the AlphaFold Program. *Biomolecules* **2022**, *12*, 1527. [[CrossRef](#)] [[PubMed](#)]
57. Bordin, N.; Sillitoe, I.; Nallapareddy, V.; Rauer, C.; Lam, S.D.; Waman, V.P.; Sen, N.; Heinzinger, M.; Littmann, M.; Kim, S.; et al. AlphaFold2 reveals commonalities and novelties in protein structure space for 21 model organisms. *Commun. Biol.* **2023**, *6*, 160. [[CrossRef](#)] [[PubMed](#)]
58. Bienert, S.; Waterhouse, A.; de Beer, T.A.; Tauriello, G.; Studer, G.; Bordoli, L.; Schwede, T. The swiss-model Repository-new features and functionality. *Nucleic Acids Res.* **2017**, *45*, D313–D319. [[CrossRef](#)] [[PubMed](#)]
59. Akdel, M.; Pires, D.E.V.; Pardo, E.P.; Jänes, J.; Zalevsky, A.O.; Mészáros, B.; Bryant, P.; Good, L.L.; Laskowski, R.A.; Pozzati, G.; et al. A structural biology community assessment of AlphaFold2 applications. *Nat. Struct. Mol. Biol.* **2022**, *29*, 1056–1067. [[CrossRef](#)] [[PubMed](#)]
60. de Brevern, A.G. An agnostic analysis of the human AlphaFold2 proteome using local protein conformations. *Biochimie* **2022**, *207*, 11–19. [[CrossRef](#)]
61. Narwani, T.J.; Santuz, H.; Shinada, N.; Melarkode Vattekatte, A.; Ghouzam, Y.; Srinivasan, N.; Gelly, J.C.; de Brevern, A.G. Recent advances on polyproline II. *Amino Acids* **2017**, *49*, 705–713. [[CrossRef](#)]
62. Craveur, P.; Joseph, A.P.; Poulain, P.; de Brevern, A.G.; Rebehmed, J. Cis-trans isomerization of omega dihedrals in proteins. *Amino Acids* **2013**, *45*, 279–289. [[CrossRef](#)]
63. Shiono, D.; Yoshidome, T. AlphaFold-predicted Protein Structure vs Experimentally Obtained Protein Structure: An Emphasis on the Side Chains. *J. Phys. Soc. Jpn.* **2022**, *91*, 064804. [[CrossRef](#)]
64. Horuk, R. The Duffy Antigen Receptor for Chemokines darc/ackr1. *Front. Immunol.* **2015**, *6*, 279. [[CrossRef](#)]
65. de Brevern, A.G.; Wong, H.; Tournamille, C.; Colin, Y.; Le Van Kim, C.; Etchebest, C. A structural model of a seven-transmembrane helix receptor: The Duffy antigen/receptor for chemokine (darc). *Biochim. Biophys. Acta* **2005**, *1724*, 288–306. [[CrossRef](#)]
66. Jo, S.; Cheng, X.; Lee, J.; Kim, S.; Park, S.J.; Patel, D.S.; Beaven, A.H.; Lee, K.I.; Rui, H.; Park, S.; et al. Charmm-gui 10 years for biomolecular modeling and simulation. *J. Comput. Chem.* **2017**, *38*, 1114–1124. [[CrossRef](#)] [[PubMed](#)]
67. Apweiler, R.; Bairoch, A.; Wu, C.H.; Barker, W.C.; Boeckmann, B.; Ferro, S.; Yeh, L.S. UniProt: The universal protein knowledgebase in 2021. *Nucleic Acids Res.* **2021**, *49*, D480–D489.
68. Hekkelman, M.L.; de Vries, I.; Joosten, R.P.; Perrakis, A. AlphaFill: Enriching AlphaFold models with ligands and cofactors. *Nat. Methods* **2023**, *20*, 205–213. [[CrossRef](#)] [[PubMed](#)]
69. Thornton, J.M.; Laskowski, R.A.; Borkakoti, N. AlphaFold heralds a data-driven revolution in biology and medicine. *Nat. Med.* **2021**, *27*, 1666–1669. [[CrossRef](#)] [[PubMed](#)]
70. Floch, A.; Lomas-Francis, C.; Vege, S.; Burgos, A.; Hoffman, R.; Cusick, R.; de Brevern, A.G.; Westhoff, C.M. Two new Scianna variants causing loss of high prevalence antigens: Ermap model and 3D analysis of the antigens. *Transfusion* **2023**, *63*, 230–238. [[CrossRef](#)] [[PubMed](#)]
71. Diwan, G.D.; Gonzalez-Sanchez, J.C.; Apic, G.; Russell, R.B. Next Generation Protein Structure Predictions and Genetic Variant Interpretation. *J. Mol. Biol.* **2021**, *433*, 167180. [[CrossRef](#)] [[PubMed](#)]
72. McBride, J.M.; Plev, K.; Reinhartz, V.; Grzybowski, B.A.; Tlusty, T. AlphaFold2 can predict single-mutation effects on structure and phenotype. *arXiv* **2022**, arXiv:2022.04.14.488301.
73. Terwilliger, T.C.; Liebschner, D.; Croll, T.I.; Williams, C.J.; McCoy, A.J.; Poon, B.K.; Afonine, P.V.; Oeffner, R.D.; Richardson, J.S.; Read, R.J.; et al. AlphaFold predictions: Great hypotheses but no match for experiment. *bioRxiv* **2022**. [[CrossRef](#)]
74. Pak, M.A.; Markhieva, K.A.; Novikova, M.S.; Petrov, D.S.; Vorobyev, I.S.; Maksimova, E.S.; Kondrashov, F.A.; Ivankov, D.N. Using AlphaFold to predict the impact of single mutations on protein stability and function. *bioRxiv* **2021**. [[CrossRef](#)]
75. Song, B.; Luo, X.; Luo, X.; Liu, Y.; Niu, Z.; Zeng, X. Learning spatial structures of proteins improves protein-protein interaction prediction. *Brief. Bioinform.* **2022**, *23*, bbab558. [[CrossRef](#)]
76. Iqbal, S.; Ge, F.; Li, F.; Akutsu, T.; Zheng, Y.; Gasser, R.B.; Yu, D.J.; Webb, G.I.; Song, J. Prost: AlphaFold2-aware Sequence-Based Predictor to Estimate Protein Stability Changes upon Missense Mutations. *J. Chem. Inf. Model.* **2022**, *62*, 4270–4282. [[CrossRef](#)] [[PubMed](#)]
77. Swapna, L.S.; Mahajan, S.; de Brevern, A.G.; Srinivasan, N. Comparison of tertiary structures of proteins in protein-protein complexes with unbound forms suggests prevalence of allostery in signalling proteins. *BMC Struct. Biol.* **2012**, *12*, 6. [[CrossRef](#)] [[PubMed](#)]
78. Craveur, P.; Joseph, A.P.; Esque, J.; Narwani, T.J.; Noël, F.; Shinada, N.; Goguet, M.; Leonard, S.; Poulain, P.; Bertrand, O.; et al. Protein flexibility in the light of structural alphabets. *Front. Mol. Biosci.* **2015**, *2*, 20. [[CrossRef](#)] [[PubMed](#)]
79. Narwani, T.J.; Craveur, P.; Shinada, N.K.; Floch, A.; Santuz, H.; Vattekatte, A.M.; Srinivasan, N.; Rebehmed, J.; Gelly, J.C.; Etchebest, C.; et al. Discrete analyses of protein dynamics. *J. Biomol. Struct. Dyn.* **2020**, *38*, 2988–3002. [[CrossRef](#)]

80. Degiacomi, M.T. Coupling Molecular Dynamics and Deep Learning to Mine Protein Conformational Space. *Structure* **2019**, *27*, 1034–1040.e3. [[CrossRef](#)] [[PubMed](#)]
81. Sun, Y.; Jiao, Y.; Shi, C.; Zhang, Y. Deep learning-based molecular dynamics simulation for structure-based drug design against SARS-CoV-2. *Comput. Struct. Biotechnol. J.* **2022**, *20*, 5014–5027. [[CrossRef](#)]
82. Nussinov, R.; Zhang, M.; Liu, Y.; Jang, H. AlphaFold, Artificial Intelligence (AI), and Allostery. *J. Phys. Chem. B* **2022**, *126*, 6372–6383. [[CrossRef](#)]
83. Chakravarty, D.; Porter, L.L. AlphaFold2 fails to predict protein fold switching. *Protein Sci.* **2022**, *31*, e4353. [[CrossRef](#)]
84. Barbarin-Bocahu, I.; Graille, M. The X-ray crystallography phase problem solved thanks to AlphaFold and RoseTTAFold models: A case-study report. *Acta Crystallogr. Sect. D Struct. Biol.* **2022**, *78*, 517–531. [[CrossRef](#)]
85. Bond, P.S.; Cowtan, K.D. ModelCraft: An advanced automated model-building pipeline using Buccaneer. *Acta Crystallogr. Sect. D Struct. Biol.* **2022**, *78*, 1090–1098. [[CrossRef](#)]
86. Simpkin, A.J.; Thomas, J.M.H.; Keegan, R.M.; Rigden, D.J. MrParse: Finding homologues in the PDB and the EBI AlphaFold database for molecular replacement and more. *Acta Crystallogr. Sect. D Struct. Biol.* **2022**, *78*, 553–559. [[CrossRef](#)] [[PubMed](#)]
87. de Brevern, A.G.; Floch, A.; Barrault, A.; Martret, J.; Bodivit, G.; Djoudi, R.; Pirenne, F.; Tournamille, C. Alloimmunization risk associated with amino acid 223 substitution in the RhD protein: Analysis in the light of molecular modeling. *Transfusion* **2018**, *58*, 2683–2692. [[CrossRef](#)] [[PubMed](#)]
88. Floch, A.; Pirenne, F.; Barrault, A.; Chami, B.; Toly-Ndour, C.; Tournamille, C.; de Brevern, A.G. Insights into anti-D formation in carriers of Rhd variants through studies of 3D intraprotein interactions. *Transfusion* **2021**, *61*, 1286–1301. [[CrossRef](#)] [[PubMed](#)]
89. Floch, A.; Lomas-Francis, C.; Vege, S.; Brennan, S.; Shakarian, G.; de Brevern, A.G.; Westhoff, C.M. A novel high-prevalence antigen in the Lutheran system, luga (lu24), and an updated, full-length 3D bcam model. *Transfusion* **2023**, *63*, 798–807. [[CrossRef](#)]
90. Shao, C.; Bittrich, S.; Wang, S.; Burley, S.K. Assessing pdb macromolecular crystal structure confidence at the individual amino acid residue level. *Structure* **2022**, *30*, 1385–1394.e3. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.