

Protein Crystallography: Achievements and Challenges

Vladimir Timofeev ^{1,2,*} and Valeriya Samygina ^{1,2} 

¹ A.V. Shubnikov Institute of Crystallography, Federal Scientific Research Centre “Crystallography and Photonics”, Russian Academy of Sciences, 59, Leninskii Prospect, 119333 Moscow, Russia

² National Research Centre “Kurchatov Institute”, 1, Akademika Kurchatova pl., 123182 Moscow, Russia

* Correspondence: tostars@mail.ru

Abstract: Proteins are the most important biological macromolecules, and are involved in almost all aspects of life. Therefore, the study of the structure of proteins is of great practical and fundamental importance. On the one hand, knowledge of the spatial structure is necessary to study the basic principles of protein functioning; for example, the mechanisms of enzymatic reactions. On the other hand, knowledge of the spatial structure of proteins is used, for example, in biotechnology, for the design of enzymes with desired properties, as well as in drug design. Today, the main method for determining the spatial structure of a protein is X-ray structural analysis of protein crystals. The main difficulty in applying this method is in obtaining a perfect protein-crystal. This review is devoted to the successes and challenges of modern protein crystallography.

Keywords: protein crystallography; protein structure; nucleation; X-ray analysis

1. Introduction

Proteins are biopolymers consisting of amino acids linked by peptide bonds. A peptide bond is a type of amide bond that occurs during the formation of proteins and peptides as a result of the interaction of the α -amino group ($-\text{NH}_2$) of one amino acid with the α -carboxyl group ($-\text{COOH}$) of another amino acid. There are four levels of structural organization of proteins: primary, secondary, tertiary and quaternary structures [1]. Primary structure refers to the sequence of amino acids. Secondary structure is the local ordering of amino acids under the action of hydrogen bonds. Tertiary structure is the spatial structure of the polypeptide chain. Quaternary structure is the mutual arrangement of several polypeptide chains, relative to each other. Proteins are the most essential and ubiquitous components of any living organism. The role of proteins is very diverse. The main functions of protein are: catalysis, structure and motions, energy provision and the regulation of processes and transport. A number of proteins have a catalytic function. Over 5000 enzymes have been described to date [2]. Structural proteins perform a supporting function, connecting the tissues of the body to each other, acting as a framework for them [3]. Contractile proteins are proteins that provide the cell with motor function. Examples of such proteins are actin and myosin. These proteins are part of the muscles, providing the latter with the ability to contract [4,5]. A number of proteins have a signaling function, that is, the capability to transmit various signals between the cells. For example, cytokines regulate cell functions [6]. Transport proteins carry various compounds. An example of such a protein is hemoglobin, whose function is to transport oxygen [7]. It should be noted that the functions of proteins are not limited to the above. The significance of the study of protein structures also follows from this diversity. First of all, these are fundamental studies of the mechanisms of the functioning of protein molecules, which means an understanding of the principles of various physiological processes in the organisms of living beings. Of practical importance is the study of protein structures for medical applications. For example, knowledge of the structure of a number of proteins of pathogenic viruses allows us to elucidate complicated virus replication and evasion mechanisms, and create more effective and safe vaccines based



Citation: Timofeev, V.; Samygina, V. Protein Crystallography: Achievements and Challenges. *Crystals* **2023**, *13*, 71. <https://doi.org/10.3390/cryst13010071>

Academic Editor: Abel Moreno

Received: 27 November 2022

Revised: 28 December 2022

Accepted: 29 December 2022

Published: 1 January 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/>).

on peptides [8,9]. As a separate topic, we should look at the so-called drug-design method, or directed drug design. Currently, due to the rapid growth of computing power available to researchers, the design of drugs using molecular-modeling methods is a promising and dynamically developing area. The main directions of molecular modeling in drug design are methods based on knowledge of the ligand structure and methods based on knowledge of the target structure [10–12]. Recent examples of the application of such an approach are the structural studies of SARS-CoV-2 main protease (Mpro). The apoform of Mpro was solved at the beginning of 2020 [13]. The appearance of this structure in PDB gave rise to a series of research studies on the search for new Mpro inhibitors. More than 20 structures of Mpro complexes with different inhibitors have been deposited in the PDB, to date. In addition, it should be noted that one of the urgent problems of modern enzymology is the prediction of the nature of mutations necessary for a directed change in the specificity of an enzyme. Currently, substrate specificity is explained by the Koshland theory, which states that the topology of the active site corresponds to the topology of the substrate, according to the “key-lock” principle. Accordingly, substrate specificity can be influenced by changing the active site using site-directed mutagenesis. Knowledge of the spatial structure of the enzyme makes it possible to rationally construct mutant forms of the protein with the required substrate specificity. Thus, a number of mutants of enzymes with changed specificity were obtained, which have industrial or medical significance [14–17].

Currently, the main method for studying the spatial structure of a protein is X-ray diffraction analysis. Approximately 90% of the structures deposited in the Worldwide Protein Data Bank were obtained using X-ray diffraction analysis [18]. Crystals of suitable diffraction quality are required for the realization of this method. There are, however, other methods for studying the spatial structure of a protein. One of them is the method of nuclear magnetic resonance. This method does not require a crystal, but is applicable only to biomolecules with molecular weight < 70 kDa [19]. The resulting models are usually less accurate than the models obtained from X-ray diffraction analysis [20]. Small-angle X-ray scattering also does not require obtaining a protein crystal, although it only allows one to determine the shape of the surface of a macromolecule [21].

The first protein crystal was obtained in 1840 [22]. In 1851, a method was described for obtaining such crystals from erythrocytes [23]. This protein was later named hemoglobin [24]. The first diffraction pattern from a hemoglobin crystal was obtained in 1934 [25]. The first spatial structure of a protein was obtained for myoglobin in 1958 [26]. In October 1971, the Protein Data Bank (PDB) appeared. At first it had seven protein structures [27]. Every year, the number of three-dimensional protein structures deposited there grew rapidly. Currently, the number of structures deposited in the PDB, obtained by experimental methods, exceeds 195,000. It should be noted that at present the PDB also contains spatial structures of polynucleotides. In addition, from 2022, the PDB has contained the spatial structures of proteins obtained by computational methods. Currently, there are more than 1,000,000 such models in the PDB. In most cases, recombinant proteins are used in protein crystallography. At present, a typical experiment to determine the structure of a recombinant protein using X-ray diffraction analysis consists of several stages: obtaining a recombinant protein, its purification, crystallization, the X-ray-diffraction experiment, and solution and refinement of the protein structure. Since the late 1990s flash cooling of protein crystals has been widely used for X-ray data collection. It allows for the reduction of radiation damage, which is essential for achieving the suitable data-collection statistics. The use of flash cooling is absolutely necessary for single-crystal X-ray experiments in modern synchrotrons, where a reduced beam is used to prevent fast-diffraction degradation.

However, flash cooling can distort protein structure and mosaicity. The investigation of most dynamic processes is unrealized in the frozen state of crystals. However, domain motions can still be investigated, using techniques such as TLS (translation- libration- and screw-motion) analysis. Serial microcrystallography developments in recent years have helped to overcome these limitations. A method based on data collection from microcrystal suspensions at ambient temperature is used. Special methods of microcrystal delivery

are required. Three main sample-delivery methods are used: crystal-injection methods, fixed-target methods and hybrid delivery methods [28–31]. Such experiments are realized in the fourth generation synchrotrons or FELs [32,33]. The main achievements of this technique are the structures of membrane proteins and time-resolved experiments. In time-resolved experiments, lasers are used as chemical triggers, depending on the nature of the object [34–39]. The importance of anomalous dispersion in the development of protein crystallography should also be noted [25].

Crystallization is one of the longest and most unpredictable stages of crystallographic experiment. A lot of reviews are devoted to the problem of protein crystallization [40–45]. Protein crystals are grown from saline solutions, using precipitants to reduce solubility, i.e., concentrated salt solutions, amphiphilic organic compounds, soluble organic polymers, and mainly polyethylene glycols of various molecular weights. Crystals form when a certain level of supersaturation is reached. A protein globule, maintained by hydrogen bonds and hydrophobic interactions, retains its structure under a limited range of conditions. It is possible to obtain a crystal only from a native non-denatured molecule; therefore, during crystallization, the range of temperatures and pH values is limited by the intervals in which the protein molecule is sufficiently stable. The decrease in protein solubility at high ionic strength is explained by the competition of protein molecules for water: at a high salt-concentration, the water-protective layer around the protein molecule decreases, and the ability of the molecules to interact with each other increases. Organic solvents, reducing the value of the dielectric constant, enhance the interaction between charges on the surface of molecules and the attraction of molecules to each other. Soluble organic polymers cause both effects. Since the solubility of a protein also depends on the pH value and temperature of the solution, it is possible to decrease the solubility and induce crystal formation by varying these parameters [41]. The crystallization process includes two stages: nucleation and crystal growth. Nucleation is the initial stage of crystallization. Protein crystals, unlike crystals of small molecules, are nucleated at a very high level of supersaturation. The high level of supersaturation is apparently related to the high entropy barrier in the ordering of large molecules. Nucleation is preceded by an induction period, which is critical for the total duration of the process. As supersaturation increases, the probability of a simultaneous collision of two or more molecules, and the lifetime of the resulting cluster, increase; subsequent molecules can join the resulting cluster. When the cluster reaches a critical value, the attractive forces between molecules become stronger than the forces that cause dissociation. The proto-crystal, meeting with other solute molecules, attaches to them and becomes the center of nucleation. This is how homogeneous nucleation proceeds, when the protein molecules themselves combine. During heterogeneous nucleation, a proto-crystal is formed, due to the adsorption of molecules onto the solid particles present in the solution or introduced into it; remaining for some time on a solid surface, the adsorbed molecules form a nucleus [41,46]. The optimal conditions for the growth of a protein crystal differ significantly from the conditions required for nucleation. The most ordered crystal grows at the minimum degree of supersaturation, when the process proceeds at a low rate. In order for the supersaturation level not to reach the labile zone, the rate of increase in the precipitant concentration must also be sufficiently low. At low supersaturation and a low growth-rate, molecules that have been improperly integrated into the crystal lattice can dissociate and reattach in the correct orientation. Such different conditions of nucleation and growth are one of the features of crystallization of macromolecules. As a result, it is possible to provide optimal conditions for growth by carrying out both stages (nucleation and growth) separately, for example, by introducing separately obtained nuclei (seeds) into a slightly supersaturated solution [45].

2. Protein-Crystallization Techniques

Currently, a number of methods for protein crystallization have been developed and are widely used. Most protein crystals have been, and are still, grown by solvent vapor diffusion [47]. The advantage of this method is its simplicity and economy. Crystallization

takes place in a hermetically sealed cell containing an undiluted precipitant solution. Water from a drop with a mixture of protein and a precipitant, where the concentration of the precipitant is lower, is distilled into the reservoir solution until the partial vapor pressure over the drop and the surface of the solution is equal. Due to the increase in the concentration of the precipitant and protein, the solution in the droplet becomes supersaturated, and at a certain stage, crystals or an amorphous precipitate appear in it. The method is carried out in two versions—the drop can be “hanging” or “sitting” [48]. Another widely used method is free diffusion through the liquid–liquid interface [49]. In this method, a protein solution is carefully layered onto a precipitant solution in a narrow test tube. Due to the significantly higher diffusivity of the salt compared to the protein, in the first stages of mixing the concentration of the precipitant increases to a greater extent than the concentration of the protein. The concentration of the precipitant is selected in such a way that a larger number of nuclei formed in the first stages of mixing dissolve, and a limited number of large crystals grow from a small number of the remaining ones. A variant of this method can be considered as the dialysis method, where the precipitant solution diffuses into the protein solution through the dialysis film, so the protein concentration remains constant [50]. The dialysis membrane increases the likelihood of nucleation, by serving as a substrate for epitaxial growth. A very simple and convenient method of crystallization is under a layer of paraffin oil [50]. Another not very common but very fast method for obtaining crystals is crystallization during protein precipitation in an ultracentrifuge [51]. This method is applicable to proteins of sufficiently large molecular weight. A solution containing protein and a low concentration of a precipitant is placed in a centrifuge tube and centrifuged for 20–40 h at speeds at which the protein slowly sediments to the bottom of the tube. Under the action of acceleration, active transport of the protein to the crystallization zone occurs. As the protein concentration at the bottom of the tube approaches the protein concentration in the crystal, nucleation occurs and then crystal growth occurs. At the same time, the level of supersaturation of the precipitant remains low, and directional acceleration, which promotes a certain orientation of protein molecules, facilitates crystallization. To prevent the crystals from dissolving after the centrifuge is stopped, they must be transferred to a solution with a high concentration of precipitant, the composition of which is selected empirically. Spanish researchers proposed carrying out crystallization through a layer of gel using the method of counterdiffusion in a capillary [52]. The protein solution was placed in an X-ray capillary, one end of which was closed, and the other end was immersed in agarose gel in a plastic box. The precipitant solution was applied to the agarose gel. Slow diffusion of the precipitant through the gel layer led to the formation of a precipitant concentration-gradient in the capillary, and crystals grew at different distances from the capillary inlet, under different conditions. As a result, the counterdiffusion method allows for the testing of several growth conditions within a single capillary.

Currently, there are still no rational approaches for choosing the nature and composition of the precipitant to obtain a protein in the crystalline state. It is not clear what kind of precipitant and in the presence of which additives, or at what pH values and at what temperature this protein will form a crystalline, rather than amorphous, precipitate. As a result, the least predictable is the initial stage—obtaining a protein in a crystalline form. Precipitant selection is sometimes aided by knowledge of protein behavior and properties, but in general, crystallization conditions are screened using commercial equipment and commercial crystallization-reagent kits. A number of companies offer a large number of sets and various crystallization devices for setting up crystallization. Each kit typically contains 50 or 96 precipitant solutions, which include a pH buffer, the precipitant itself, and low-molecular-weight additives. During screening, a protein solution in a certain proportion is mixed with a solution of a precipitant, and the formed precipitate is examined under a microscope at certain intervals. Nevertheless, researchers are making attempts to study the mechanisms of the formation and growth of protein crystals. The growth mechanisms were studied by electron- and atomic-force microscopy, as well as by interfer-

ometry, using Michelson and Mach-Zehnder interferometers [53–55]. It was shown that macromolecular crystals grow using the same mechanisms as crystals of other molecules; however, during the growth of protein and virus crystals, a new mechanism, unknown for small molecules was discovered—growth by direct addition and subsequent development of whole three-dimensional nuclei. In addition, a number of attempts were made to study the structure of precrystallization solutions by small-angle X-ray scattering [56,57].

Despite well-developed methods of crystallization and an understanding of the general patterns of growth of protein crystals, a number of proteins cannot be crystallized or only poorly diffracting crystals can be obtained. In such cases, it is advisable to use protein-engineering methods to increase the probability of the formation of additional intermolecular contacts in the crystal lattice [58]. To create new intermolecular contacts, site-directed mutagenesis replaces individual amino-acid residues on the surface of a protein molecule. If it is necessary to increase the solubility of the protein, some hydrophobic residues on the surface can be replaced with polar ones [59]. It has been noted that proteins have entropy surface-protection: the presence of charged lysine and glutamic-acid residues on the surface prevents the formation of nonspecific aggregates and precipitation [60]. Replacing them with alanine or other small residues reduces the surface conformational-entropy. In this way, crystallization conditions can be optimized. New intermolecular contacts can also be created by the chemical modification of individual amino-acid residues, for example, by the acetylation of lysine residues [61]

3. Peculiarities of the Growth of Protein Crystals under Weightless Conditions

One way to improve the quality of protein crystals is to grow them under weightless conditions. Gravity-related convection currents are absent under weightless conditions. At an acceleration of 10^{-5} – 10^6 g, the transport of growth units to the growing crystal is carried out mainly through diffusion [62]. Under microgravity conditions, the protein-depleted zone, which is formed around the crystal when molecules are included in the crystal lattice, expands and stabilizes. Within this zone, a stable protein-concentration-gradient is established, where transport is carried out by diffusion [63]. The absence of sedimentation and the spherical shape of the diffusion field favor the growth of isometric crystals. Due to the low diffusivity of macromolecules during diffusion mass-transfer, growth proceeds slowly, and the rate is controlled by surface kinetics. Slow growth at a low level of supersaturation promotes the attachment of molecules in the optimal orientation, and increases the ordering of the crystal. The fundamentally uncontrolled process of growth under weightless conditions becomes partially regulated. Along with the concentration gradient of the protein around the crystal, the concentration gradient of impurities is also established. Since impurities are often aggregates that have a lower diffusivity than the original macromolecules, and their incorporation into the crystal is reduced. Thus, the concentration gradient is a diffusion filter that protects the crystal from the inclusion of impurities [64]. Therefore, for macromolecules having a low diffusion-coefficient, weightlessness is a suitable environment for the growth of high-quality crystals. Experiments on the crystallization of macromolecules under weightless conditions have been carried out since the mid-1970s, and are still ongoing. The results of numerous experiments have shown the advantages of weightlessness as a medium for growing high-quality crystals, and have made a significant contribution to understanding the laws governing the growth of macromolecular crystals. Compared to the ground-based control crystals, crystals grown in zero gravity had a larger size, a more perfect structure, and a smaller mosaic-value which is also significant for neutron crystallography [65–67]. In addition, they were not growing at the sticky surfaces between the vial and the protein solution, so the difficulty in picking them up could be avoided.

4. Crystallization of Membrane Proteins

Crystallization of membrane proteins is another challenge for crystallographers. Membrane proteins include proteins that are embedded in or associated with the cell membrane

or the membrane of a cell organelle. Approximately 25% of all proteins are membrane proteins [68]. Membrane proteins are important elements of the cell membrane, responsible for a number of critical functions for the life of the cell. For example, a number of membrane proteins are ion channels. Such proteins carry out the transport of ions into the cell and into the intercellular space. For viruses, membrane glycoproteins are extremely important, and are responsible for the penetration of the virus into the cell. In addition, such glycoproteins are often antigenic determinants of viruses, and their spatial structures can be used to design vaccines [9]. Membrane-protein molecules contain on their surface, along with hydrophilic regions, hydrophobic regions embedded in the biological membrane. Membrane proteins often function in combination with other proteins or ensembles of molecules. Significant difficulties in working with membrane proteins are associated with their isolation: they must be transferred into solution without damaging or changing the unit involved in crystallization. The dissolution of membrane proteins occurs when detergents are added. Detergents belonging to the group of amphiphilic molecules cover the hydrophobic part of the protein surface hidden in the membrane. After converting a membrane protein or an ensemble of proteins into a soluble state, they are treated as with ordinary proteins, reducing their solubility by adding precipitants with respect to crystallization by vapor diffusion (either by sitting- or hanging-drop). The nature of the detergent and its concentration are important parameters. Detergents with large micelles tend to prevent crystal contacts by reducing protein hydrophilic-regions. In addition, high detergent-concentration leads to many unbound detergent micelles precluding the crystallization process and leading to detergent crystals or phase separation in the crystallization drops. A special method has also been developed for the crystallization of membrane proteins: crystallization in the lipid mesophase (or lipid cubic phase) [69]. The lipid cubic phase is composed of two components, one being the lipid and the other the protein solution. The percent of lipid has to be around 60% for successful crystallization. In this method, the membrane protein in complex with the detergent moves freely through the solvent channels, diffusing slowly into the bicontinuous bilayer of the cubic phase. (A solution of the membrane protein in a specially selected detergent is mixed with an equal volume of an aqueous solution of a precipitant. The composition and ratio of solutions are selected in such a way that the so-called cubic phase is formed, which is a combination of continuous channels of an aqueous solvent with a system of continuous channels of a protein-detergent solution. In this case, only the hydrophilic areas are in contact with the aqueous phase, while the hydrophobic part is hidden inside the detergent phase.) Crystal growth occurs at the phase interface. For simultaneous screening and optimization of crystallization conditions in the case of membrane proteins, it is advisable to use microfluidic devices [70]. The microfluidic technique makes it possible to use nanoliter amounts of the drug for a complete cycle of experiments, which is especially valuable for membrane proteins. In addition, convection flows are minimized in microfluidic devices and diffusion mass-transfer predominates.

5. Application of Computational Methods for the Study of Crystals and Structures of Proteins

It should be noted that in recent years, in connection with the development of computer technology, computational methods have been widely used to study the properties of protein crystals, as well as protein structures. Therefore, there are works on modeling using the method of molecular dynamics of crystals of various proteins [71–76]. To study the process of crystal formation, various computational methods are also used, including molecular dynamics. For example, an attempt was made to investigate the dependence of the stability of a protein crystal on the concentration of ions, as well as the charges on the surface amino-acid residues [77]. There are studies that model the process of protein-crystal formation [78]. In addition, using molecular dynamics and computational techniques that allow one to estimate the change in free energy during interaction, including that of protein molecules, precrystallization solutions of a number of proteins were studied [79,80]. It was mentioned earlier that more than 200 thousand spatial structures of proteins and their

complexes with ligands have now been determined, using experimental methods. This amount of structural information can be used to determine the structure of a protein, using computational methods. One of the most well-known and widely used computational methods for determining the structure of a protein is the so-called homology modeling. It is known that many proteins have typical motifs of spatial organization. It has been empirically established that if the sequences of two proteins are more than 30% identical to each other, then the proteins are almost certainly related, and the degree of evolutionary divergence is not yet so great that their structures lose commonality. These observations form the basis of a spatial-structure prediction technique called homology-based modeling. The homology-modeling process includes several steps, the main ones being the search for a structural template and the construction of an amino-acid alignment. The decisive factor determining the quality of the resulting models is the degree of sequence homology between the protein being modeled and the template. High identity means that the evolutionary divergence of both proteins from a common ancestor occurred not so long ago that these proteins lost their structural commonality [81–84]. Currently, there are many programs for modeling the spatial structures of proteins using homology [85–88]. Despite being widely used among researchers, the homology-modeling method has a relatively low accuracy, even at the level of the main-chain and secondary-structure elements. Modeling moving loops by this method is a separate challenge. Amino-acid side chains in a protein model defined by homology are often misconfigured. Recently, machine-learning methods have been used to model the spatial structure of proteins. One of these tools is Alpha Fold. Alpha-fold2 is an artificial intelligence system developed by DeepMind [89], which predicts the spatial structures of proteins, based on the amino-acid sequence [90]. This artificial neural network is based on the so-called “attention” approach. Attention is a way to tell the network what to pay more attention to; that is, to predict the probability of a particular outcome, depending on the state of the neurons and the input data [91]. Identification of important factors is carried out through the method of backpropagation of the error [92]. It should be noted that previous successes in structural biology and, above all, in protein crystallography, made it possible to obtain more than 170 thousand spatial structures of proteins, on which alpha-fold2 was trained [93]. A number of works have shown that the accuracy of predicting the spatial structures of proteins in most cases is not inferior to experimentally obtained structures [94], and in some cases gives the same quality as the crystal structure obtained at a resolution of 1.1 Å [94]. This fact opens up significant opportunities for an approach based on artificial neural networks for medicine, including overcoming the consequences of future pandemics [95]. Alpha-fold2 has also been shown to be one of the important tools for IDP research [96].

6. Conclusions

Even now, obtaining a crystal suitable for structural study is the least predictable stage, and the one on which the time of the entire study depends. The complexity of protein crystallization is explained both by the structural features of the biological macromolecules themselves and by the features of the crystals built from such molecules.

At present, in connection with the improvement of sources and detectors of synchrotron radiation, a complete structural study can be carried out with crystals as small as 10 µm for classical single-crystal experiment and 2–5 µm for serial microcrystallography. However, even relatively small but highly ordered protein crystals are not always easy to obtain. Growing protein crystals of high diffraction-quality is still the main obstacle to the establishment of the spatial structure of the protein.

The unique properties of the protein molecules are factors that impede crystallization. The process of protein crystallization itself is also characterized by a number of features which still do not allow placing it under full control. Therefore, despite significant advances in technology and in understanding the process of crystal growth, macromolecular crystallization continues to be an empirical science of rational trial-and-error, guided by the results of previous attempts.

Author Contributions: Writing—original draft preparation, V.T.; writing—review and editing, V.S.; funding acquisition, V.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Russian Ministry of Science and Education within the framework of the Federal Scientific and Technical Program for the Development of Synchrotron and Neutron Research and Research Infrastructure for 2019–2027 (Agreement No. 075-15-2021-1355 (12 October 2021)).

Conflicts of Interest: The authors declare no conflict of interest.

References

- Murray, R.F.; Harper, H.W.; Granner, D.K.; Mayes, P.A.; Rodwell, V.W. *Harper's Illustrated Biochemistry*; Lange Medical Books/McGraw-Hill: New York, NY, USA, 2006.
- Bairoch, A.T. The ENZYME Database in 2000. *Nucleic Acids Res.* **2000**, *28*, 304–305. [CrossRef] [PubMed]
- Erickson, H.P. Evolution of the cytoskeleton. *Bioessays* **2007**, *29*, 668–677. [CrossRef] [PubMed]
- Vale, R.D. The molecular motor toolbox for intracellular transport. *Cell* **2003**, *112*, 467–480. [CrossRef] [PubMed]
- Hartman, M.A.; Spudich, J.A. The myosin superfamily at a glance. *J. Cell Sci.* **2012**, *125*, 1627–1632. [CrossRef] [PubMed]
- Cohen, S.; Bigazzi, P.E.; Yoshida, T. Similarities of T cell function in cell-mediated immunity and antibody production. *Cell. Immunol.* **1974**, *12*, 150–159. [CrossRef]
- Weed, R.I.; Reed, C.F.; Berg, G. Is hemoglobin an essential structural component of human erythrocyte membranes? *J. Clin. Investig.* **1963**, *42*, 581–588. [CrossRef]
- Araf, Y.; Moin, A.T.; Timofeev, V.I.; Faruqi, N.A.; Saiara, S.A.; Ahmed, N.; Parvez, M.S.; Rahaman, T.I.; Sarkar, B.; Ullah, M.A.; et al. Immunoinformatic Design of a Multivalent Peptide Vaccine Against Mucormycosis: Targeting FTR1 Protein of Major Causative Fungi. *Front. Immunol.* **2022**, *13*, 863234. [CrossRef]
- Abass, O.A.; Timofeev, V.I.; Sarkar, B.; Onobun, D.O.; Ogunsola, S.O.; Aiyenuro, A.E.; Aborode, A.T.; Aigboje, A.E.; Omobolanle, B.N.; Imolele, A.G.; et al. Abiodun Immunoinformatics analysis to design novel epitope based vaccine candidate targeting the glycoprotein and nucleoprotein of Lassa mammarenavirus (LASMV) using strains from Nigeria. *J. Biomol. Struct. Dyn.* **2021**, *40*, 7283–7302. [CrossRef]
- Tollenaere, J.P. The role of structure-based ligand design and molecular modelling in drug discovery. *Pharm. World Sci.* **1996**, *18*, 56–62. [CrossRef]
- Guner, O.F. *Pharmacophore Perception, Development, and Use in Drug Design*; International University Line: La Jolla, CA, USA, 2000.
- Mauser, H.; Guba, W. Recent developments in de novo design and scaffold hopping. *Curr. Opin. Drug Discov. Dev.* **2008**, *3*, 365–374.
- Zhang, L.; Lin, D.; Sun, X.; Curth, U.; Drosten, C.; Sauerhering, L.; Becker, S.; Rox, K.; Hilgenfeld, R. Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α -ketoamide inhibitors. *Science* **2020**, *368*, 409–412. [CrossRef]
- Korendovych, I.V. Rational and Semirational Protein Design. *Methods Mol Biol.* **2018**, *1685*, 15–23. [PubMed]
- Goncharuk, M.V.; Baleeva, N.S.; Nolde, D.E.; Gavrikov, A.S.; Mishin, A.V.; Mishin, A.S.; Sosorev, A.Y.; Arseniev, A.S.; Goncharuk, S.A.; Borshchevskiy, V.I.; et al. Structure-based rational design of an enhanced fluorogen-activating protein for fluorogens based on GFP chromophore. *Commun. Biol.* **2022**, *5*, 706. [CrossRef] [PubMed]
- Ghislieri, D.; Green, A.P.; Pontini, M.; Willies, S.C.; Rowles, I.; Frank, A.; Grogan, G.; Turner, N.J. Engineering an enantioselective amine oxidase for the synthesis of pharmaceutical building blocks and alkaloid natural products. *J. Am. Chem. Soc.* **2013**, *135*, 10863–10869. [CrossRef] [PubMed]
- Rotticci, D.; Rotticci-Mulder, J.C.; Denman, S.; Norin, T.; Hult, K. Improved enantioselectivity of a lipase by rational protein engineering. *ChemBioChem* **2001**, *2*, 766–770. [CrossRef]
- The RCSB Protein Data Bank. Available online: <https://www.rcsb.org/> (accessed on 4 December 2022).
- Spronk, C.A.; Nabuurs, S.B.; Krieger, E.; Vriend, G.; Vuister, G.W. Validation of protein structures derived by NMR spectroscopy. *Prog. Nucl. Magn. Reson. Spectrosc.* **2004**, *45*, 315–337. [CrossRef]
- Svergun, D.I.; Koch, M.H. Small-angle scattering studies of biological macromolecules in solution. *Rep. Prog. Phys.* **2003**, *66*, 1735–1782. [CrossRef]
- Hunefeld, F. *Die Chemismus in der Thierischen Organization*; Brockhaus: Leipzig, Germany, 1840; pp. 158–163.
- Funke, O. Über das milzvenenblut. *Z. Rat. Med.* **1851**, *1*, 172–218.
- Hoppe-Seyler, F. Über die oxydation in lebendem blute. *Med.-Chem Untersuch Lab.* **1866**, *1*, 133–140.
- Tulinsky, A. Chapter 35. The Protein Structure Project, 1950–1959: First Concerted Effort of a Protein Structure Determination in the U.S. In *Annual Reports in Medicinal Chemistry*; Elsevier: Amsterdam, The Netherlands, 1996; Volume 31, pp. 357–366.
- Kendrew, J.C.; Bodo, G.; Dintzis, H.M.; Parrish, R.G.; Wyckoff, H.; Phillips, D.C. A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. *Nature* **1958**, *181*, 662–666. [CrossRef]
- Bank, P.D. Protein Data Bank. *Nat. New Biol.* **1971**, *233*, 233.
- Liang, M.; Williams, G.J.; Messerschmidt, M.; Seibert, M.M.; Montanez, P.A.; Hayes, M.; Milathianaki, D.; Aquila, A.; Hunter, M.S.; Koglin, J.E.; et al. The Coherent X-ray Imaging instrument at the Linac Coherent Light Source. *J. Synchrotron Rad.* **2015**, *22*, 514–519. [CrossRef]

28. Milne, C.J.; Schietinger, T.; Aiba, M.; Alarcon, A.; Alex, J.; Anghel, A.; Arsov, V.; Beard, C.; Bettoni, S.; Bopp, M.; et al. SwissFEL: The Swiss X-ray Free Electron Laser. *Appl. Sci.* **2017**, *7*, 720. [[CrossRef](#)]
29. Pedrini, B.; Martiel, I. Available online: <https://www.psi.ch/swissfel/internal-reports> (accessed on 3 July 2017).
30. Mehrabi, P.; Müller-Werkmeister, H.M.; Leimkohl, J.P.; Schikora, H.; Ninkovic, J.; Krivokuca, S.; Andriček, L.; Epp, S.W.; Sherrell, D.; Owen, R.L.; et al. The HARE chip for efficient time-resolved serial synchrotron crystallography. *J. Synchrotron Radiat.* **2020**, *27*, Pt 2, 360–370. [[CrossRef](#)] [[PubMed](#)]
31. Mehrabi, P.; Schulz, E.C.; Agthe, M.; Horrell, S.; Bourenkov, G.; von Stetten, D.; Leimkohl, J.P.; Schikora, H.; Schneider, T.R.; Pearson, A.R.; et al. Liquid application method for time-resolved analyses by serial synchrotron crystallography. *Nat. Methods* **2019**, *16*, 979–982. [[CrossRef](#)]
32. Tenboer, J.; Basu, S.; Zatsepin, N.; Pande, K.; Milathianaki, D.; Frank, M.; Hunter, M.; Boutet, S.; Williams, G.J.; Koglin, J.E.; et al. Time-resolved serial crystallography captures high-resolution intermediates of photoactive yellow protein. *Science (N. Y.)* **2014**, *346*, 1242–1246. [[CrossRef](#)] [[PubMed](#)]
33. Ihee, H.; Rajagopal, S.; Srajer, V.; Pahl, R.; Anderson, S.; Schmidt, M.; Schotte, F.; Anfinrud, P.A.; Wulff, M.; Moffat, K. Visualizing reaction pathways in photoactive yellow protein from nanoseconds to seconds. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 7145–7150. [[CrossRef](#)]
34. Schotte, F.; Lim, M.; Jackson, T.A.; Smirnov, A.V.; Soman, J.; Olson, J.S.; Phillips, G.N., Jr.; Wulff, M.; Anfinrud, P.A. Watching a protein as it functions with 150-ps time-resolved x-ray crystallography. *Science (N. Y.)* **2003**, *300*, 1944–1947. [[CrossRef](#)]
35. Ahn, S.; Kim, K.H.; Kim, Y.; Kim, J.; Ihee, H. Protein tertiary structural changes visualized by time-resolved X-ray solution scattering. *J. Phys. Chem. B* **2009**, *113*, 13131–13133.
36. Frauenfelder, H.; Chen, G.; Berendzen, J.; Fenimore, P.W.; Jansson, H.; McMahon, B.H.; Strope, I.R.; Swenson, J.; Young, R.D. A unified model of protein dynamics. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 5129–5134. [[CrossRef](#)]
37. Moeglich, A.; Moffat, K. Engineered photoreceptors as novel optogenetic tools. *Photochem. Photobiol. Sci.* **2010**, *9*, 1286–1300. [[CrossRef](#)] [[PubMed](#)]
38. Suga, M.; Shimada, A.; Akita, F.; Shen, J.R.; Tosha, T.; Sugimoto, H. Time-resolved studies of metalloproteins using X-ray free electron laser radiation at SACLA. *Biochimica et biophysica acta. Gen. Subj.* **2020**, *1864*, 129466. [[CrossRef](#)] [[PubMed](#)]
39. Giegé, R. A historical perspective on protein crystallization from 1840 to the present day. *FEBS J.* **2013**, *280*, 6456–6497. [[CrossRef](#)] [[PubMed](#)]
40. McPherson, A.; Gavira, J.A. Introduction to protein crystallization. *Acta Crystallogr. Sect. F Struct. Biol. Commun.* **2014**, *70* Pt 1, 2–20. [[CrossRef](#)]
41. McPherson, A. A brief history of protein crystal growth. *J. Cryst. Growth* **1991**, *110*, 1–10. [[CrossRef](#)]
42. Kundrot, C.E. Which strategy for a protein crystallization project. *Cell. Mol. Life Sci.* **2014**, *61*, 524–536. [[CrossRef](#)]
43. Bugg, C.E. The Future of Protein Crystal Growth. *J. Cryst. Growth* **1986**, *76*, 535–544. [[CrossRef](#)]
44. McPherson, A. Current approaches to macromolecular crystallization. *Eur. J. Biochem.* **1990**, *189*, 1–23. [[CrossRef](#)] [[PubMed](#)]
45. Rupp, B. *Biomolecular Crystallography: Principles, Practice, and Application to Structural Biology*; Garland Science: New York, NY, USA, 2009; p. 800.
46. Chayen, N.E.; Saridakis, E.; Sear, R.P. Experiment and theory for heterogeneous nucleation of protein crystals in a porous medium. *Prot. Natl. Acad. Sci. USA* **2006**, *103*, 597–601. [[CrossRef](#)]
47. Davies, D.R.; Segal, D.M. [25] Protein crystallization: Micro techniques involving vapor diffusion. *Methods Enzymol.* **1971**, *22*, 266–269.
48. Wlodawer, A.; Hodgson, K.O. Crystallization and crystal data of monellin. *Prot. Natl. Acad. Sci. USA* **1975**, *72*, 398–399. [[CrossRef](#)]
49. Salemme, R.R. A free interface diffusion technique for the crystallization of proteins for X-ray crystallography. *Arch. Biochem. Biophys.* **1972**, *2*, 533–539. [[CrossRef](#)] [[PubMed](#)]
50. Zeppezauer, M.; Eclund, H.; Zeppezauer, E. Micro diffusion cells for the growth of single protein crystals by means of equilibrium dialysis. *Arch. Biochem. Biophys.* **1968**, *126*, 564–573. [[CrossRef](#)]
51. Chayen, N.E. The role of oil in macromolecular crystallisation. *Structure* **1997**, *5*, 1269–1274. [[CrossRef](#)] [[PubMed](#)]
52. Garcia_Ruiz, J.M.; Moreno, A. Investigations on protein crystal growth by the gel acupuncture method. *Acta Cryst.* **1994**, *50*, 484–490.
53. Durbin, S.D.; Feher, G.J. Studies of crystal growth mechanisms of proteins by electron microscopy. *J. Mol. Biol.* **1990**, *212*, 763–774. [[CrossRef](#)]
54. Malkin, A.J.; Kuznetsov, Y.G.; Glantz, W.; McPherson, A.J. Atomic force microscopy studies of surface morphology and growth kinetics in thaumatin. *Phys. Chem.* **1996**, *100*, 11736–11743. [[CrossRef](#)]
55. Shlichta, P.J. Feasibility of mapping solution properties during the growth of protein crystals. *J. Cryst. Growth* **1986**, *76*, 656–662. [[CrossRef](#)]
56. Kovalchuk, M.V.; Blagov, A.E.; Dyakova, Y.A.; Gruzinov, A.Y.; Marchenkova, M.A.; Peters, G.S.; Pisarevsky, Y.V.; Timofeev, V.I.; Volkov, V.V. Investigation of the Initial Crystallization Stage in Lysozyme Solutions by Small-Angle X-ray Scattering. *Cryst. Growth Des.* **2016**, *16*, 1792–1797. [[CrossRef](#)]
57. Marchenkova, M.A.; Konarev, P.V.; Kordonskaya, Y.V.; Ilina, K.B.; Pisarevsky, Y.V.; Soldatov, A.V.; Timofeev, V.I.; Kovalchuk, M.V. The Role of Cations and Anions in the Formation of Crystallization Oligomers in Protein Solutions as Revealed by Combination of Small-Angle X-ray Scattering and Molecular Dynamics. *Crystals* **2022**, *12*, 751. [[CrossRef](#)]

58. Lawson, D.M.; Artymiuk, P.J.; Yewdall, S.I.; Smoth, J.M.; Livingstone, J.C.; Treffry, A.; Levi, S.; Arosio, P.; Cesareni, G. and Thomas, C.D. Solving the structure of human H ferritin by genetically engineering intermolecular crystal contacts. *Nature* **1991**, *349*, 541–544. [[CrossRef](#)] [[PubMed](#)]
59. Trevino, S.R.; Scholtz, J.M.; Pace, C.N. Measuring and increasing protein solubility. *J. Pharm. Sci.* **2008**, *97*, 4155–4166. [[CrossRef](#)] [[PubMed](#)]
60. Deriwenda, Z.S.; Vekilov, P. Entropy and surface engineering in protein crystallization. *Acta Cryst.* **2006**, *52*, 116–124.
61. Rayment, I.; Rypniewski, W.R.; Schmidt-Bäse, K.; Smith, R.; Tomchick, D.R.; Benning, M.M.; Winkelmann, D.A.; Wesenberg, G.; Holden, H.M. Three-dimensional structure of myosin subfragment-1: A molecular motor. *Science* **1993**, *261*, 50–58. [[CrossRef](#)] [[PubMed](#)]
62. Ramachandran, N.; Baugher, C.R.; Naumann, R.J. Modeling Flows and Transport in Protein Crystal Growth. *Microgravity Sci. Technol.* **1995**, *VIII/3*, 170–179.
63. Otalora, F.; Novella, M.L.; Gavira, J.A.; Thomas, B.R.; Garcia-Ruiz, J.M. Experimental evidence for the stability of the depletion zone around a growing protein crystal under microgravity. *Acta Cryst.* **2001**, *D57*, 412. [[CrossRef](#)] [[PubMed](#)]
64. Chernov, A.A.; Garcia-Ruiz, J.M.; Thomas, B.R. Visualization of the impurity depletion zone surrounding apoferritin crystals growing in gel with holoferritin dimer impurity. *J. Cryst. Growth* **2001**, *232*, 184. [[CrossRef](#)]
65. Snell, E.H.; Helliwell, J.R. Macromolecular crystallization in microgravity. *Rep. Prog. Phys.* **2005**, *68*, 799–853. [[CrossRef](#)]
66. Eistrikh-Heller, P.A.; Rubinsky, S.V.; Samygina, V.R.; Gabdulkhakov, A.G.; Kovalchuk, M.V.; Mironov, A.S. Crystallization in Microgravity and the Atomic-Resolution Structure of Uridine Phosphorylase from *Vibrio cholerae*. *Crystallogr. Rep.* **2021**, *66*, 777–785. [[CrossRef](#)]
67. Boyko, K.M.; Timofeev, V.I.; Samygina, V.R.; Kuranova, I.P.; Popov, V.O.; Koval'Chuk, M.V. Protein crystallization under microgravity conditions. Analysis of the results of Russian experiments performed on the International Space Station in 2005–2015. *Crystallogr. Rep.* **2016**, *61*, 718–729. [[CrossRef](#)]
68. Stevens, T.J.; Arkin, I.T. Do more complex organisms have a greater proportion of membrane protein in their genomes? *Proteins* **2000**, *39*, 417–420. [[CrossRef](#)]
69. Caffrey, M. A lipid's eye view of membrane protein crystallization in mesophases. *Curr. Opin. Struct. Biol.* **2000**, *10*, 486–497. [[CrossRef](#)] [[PubMed](#)]
70. Li, L.; Mustafi, D.; Fu, Q.; Tereshko, V.; Chen, D.L.; Tice, J.D.; Ismagilov, R.F. Nanoliter microfluidic hybrid method for simultaneous screening and optimization validated with crystallization of membrane proteins. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 19243–19248. [[CrossRef](#)]
71. Meinhold, L.; Smith, J.C. Protein Dynamics from X-Ray Crystallography: Anisotropic, Global Motion in Diffuse Scattering Patterns. *Proteins* **2007**, *66*, 941–953. [[CrossRef](#)]
72. Meinhold, L.; Merzel, F.; Smith, C.J. Lattice dynamics of a protein crystal. *Phys. Rev. Lett.* **2007**, *99*, 138101. [[CrossRef](#)] [[PubMed](#)]
73. Cerutti, D.S.; Le Trong, I.; Stenkamp, R.E.; Lybrand, T.P. Simulations of a protein crystal: Explicit treatment of crystallization conditions links theory and experiment in the streptavidin-biotin complex. *Biochemistry* **2008**, *47*, 12065–12077. [[CrossRef](#)] [[PubMed](#)]
74. Cerutti, D.S.; Le Trong, I.; Stenkamp, R.E.; Lybrand, T.P. Dynamics of the streptavidin-biotin complex in solution and in its crystal lattice: Distinct behavior revealed by molecular simulations. *J. Phys. Chem. B* **2009**, *113*, 6971–6985. [[CrossRef](#)]
75. Schnieders, M.J.; Fenn, T.D.; Pande, V.S.; Brunger, A.T. Polarizable atomic multipole refinement: Application to peptide crystals. *Acta Crystallogr. D* **2009**, *65*, 952–965. [[CrossRef](#)]
76. Cerutti, D.S.; Freddolino, P.L.; Duke, R.E.; Case, D.A. Simulations of a protein crystal with a high resolution X-ray structure: Evaluation of force fields and water models. *J. Phys. Chem. B* **2010**, *114*, 12811–12824. [[CrossRef](#)]
77. Kuzmanic, A.; Zagrovic, B. Dependence of Protein Crystal Stability on Residue Charge States and Ion Content of Crystal Solvent. *Biophys. J.* **2014**, *106*, 677–686. [[CrossRef](#)]
78. Taudt, A.; Arnold, A.; Pleiss, J. Simulation of protein association: Kinetic pathways towards crystal contacts. *Phys. Rev. E* **2015**, *91*, 033311. [[CrossRef](#)] [[PubMed](#)]
79. Kordonskaya, Y.V.; Timofeev, V.I.; Dyakova, Y.A.; Marchenkova, M.A.; Pisarevsky, Y.V.; Kovalchuk, M.V. The Role of Cations of the Precipitant in the Interaction of Protein Molecules in the Lysozyme Oligomers in Crystallization Solutions. *Crystals* **2021**, *11*, 1534. [[CrossRef](#)]
80. Kordonskaya, Y.V.; Timofeev, V.I.; Dyakova, Y.A.; Marchenkova, M.A.; Pisarevsky, Y.V.; Kovalchuk, M.V. Free Energy Change during the Formation of Crystalline Contact between Lysozyme Monomers under Different Physical and Chemical Conditions. *Crystals* **2021**, *11*, 1121. [[CrossRef](#)]
81. Chothia, C.; Lesk, A.M. The relation between the divergence of sequence and structure in proteins. *EMBO J.* **1986**, *5*, 823–826. [[CrossRef](#)]
82. Marti-Renom, M.A.; Stuart, A.C.; Fiser, A.; Sanchez, R.; Melo, F.; Sali, A. Comparative protein structure modeling of genes and genomes. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 291–325. [[CrossRef](#)]
83. Baker, D.; Sali, A. Protein structure prediction and structural genomics. *Science* **2001**, *294*, 93–96. [[CrossRef](#)]
84. Zhang, Y. Progress and challenges in protein structure prediction. *Curr. Opin. Struct. Biol.* **2008**, *18*, 342–348. [[CrossRef](#)]
85. Fiser, A.; Sali, A. Modeller: Generation and refinement of homology-based protein structure models. *Macromol. Crystallogr. Part D Meth. Enzymol. Methods Enzymol.* **2003**, *374*, 461–491.

86. Cooper, S.; Khatib, F.; Treuille, A.; Barbero, J.; Lee, J.; Beenen, M.; Leaver-Fay, A.; Baker, D. and Popović, Z. Predicting protein structures with a multiplayer online game. *Nature* **2010**, *466*, 756–760. [[CrossRef](#)]
87. Jian, P.; Xu, J. RaptorX: Exploiting structure information for protein alignment by statistical inference. *Proteins* **2011**, *79* (Suppl. 10), 161–171.
88. 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](#)]
89. Available online: <https://deepmind.com>.
90. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596*, 583–589. [[CrossRef](#)] [[PubMed](#)]
91. Niu, Z.; Zhong, G.; Yu, H. A review on the attention mechanism of deep learning. *Neurocomputing* **2021**, *452*, 48–62. [[CrossRef](#)]
92. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [[CrossRef](#)]
93. 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](#)]
94. Zweckstetter, M. NMR hawk-eyed view of AlphaFold2 structures. *Protein Sci.* **2021**, *30*, 2333–2337. [[CrossRef](#)] [[PubMed](#)]
95. Higgins, M.K. Can We AlphaFold Our Way Out of the Next Pandemic? *J. Mol. Biol.* **2021**, *433*, 167093. [[CrossRef](#)]
96. Ruff, K.M.; Pappu, R.V. AlphaFold and Implications for Intrinsically Disordered Proteins. *J. Mol. Biol.* **2021**, *433*, 167208. [[CrossRef](#)] [[PubMed](#)]

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.