

Methods for Protein Crystallization

Subjects: **Crystallography**

Contributor: Vladimir Timofeev , Valeriya Samygina

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. 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.

protein crystallography

X-ray

nucleation

1. Introduction

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 researchers to elucidate complicated virus replication and evasion mechanisms, and create more effective and safe vaccines based on peptides [8][9]. As a separate topic, people 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][11][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][15][16][17].

The first protein crystal was obtained in 1840 [18]. In 1851, a method was described for obtaining such crystals from erythrocytes [19]. This protein was later named hemoglobin [20]. The first diffraction pattern from a hemoglobin crystal was obtained in 1934 [21]. The first spatial structure of a protein was obtained for myoglobin in 1958 [22]. In October 1971, the Protein Data Bank (PDB) appeared. At first it had seven protein structures [23]. 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 [24][25][26][27]. Such experiments are realized in the fourth generation synchrotrons or FELs [28][29]. 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 [30][31][32][33][34][35]. The importance of anomalous dispersion in the development of protein crystallography should also be noted [21].

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 [36]. 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” [37]. Another widely used method is free diffusion through the liquid–liquid interface [38]. 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 [39]. 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 [39]. Another not very common but very fast method for obtaining crystals is crystallization during protein precipitation in an ultracentrifuge [40]. 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 [41]. 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 interferometry, using Michelson and Mach-Zehnder interferometers [42][43][44]. 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 [45][46].

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 [47]. 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 [48]. 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 [49]. 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 [50].

References

1. 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.
2. Bairoch, A.T. The ENZYME Database in 2000. *Nucleic Acids Res.* 2000, 28, 304–305.
3. Erickson, H.P. Evolution of the cytoskeleton. *Bioessays* 2007, 29, 668–677.
4. Vale, R.D. The molecular motor toolbox for intracellular transport. *Cell* 2003, 112, 467–480.
5. Hartman, M.A.; Spudich, J.A. The myosin superfamily at a glance. *J. Cell Sci.* 2012, 125, 1627–1632.
6. 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.

7. 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.
8. Araf, Y.; Moin, A.T.; Timofeev, V.I.; Faruqui, 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.
9. 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.
10. Tollenaere, J.P. The role of structure-based ligand design and molecular modelling in drug discovery. *Pharm. World Sci.* 1996, 18, 56–62.
11. Guner, O.F. Pharmacophore Perception, Development, and Use in Drug Design; International University Line: La Jolla, CA, USA, 2000.
12. Mauser, H.; Guba, W. Recent developments in de novo design and scaffold hopping. *Curr. Opin. Drug Discov. Dev.* 2008, 3, 365–374.
13. 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.
14. Korendovych, I.V. Rational and Semirational Protein Design. *Methods Mol Biol.* 2018, 1685, 15–23.
15. 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.
16. 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.
17. 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.
18. Funke, O. Über das milzvenenblut. *Z. Rat. Med.* 1851, 1, 172–218.
19. Hoppe-Seyler, F. Über die oxydation in lebendem blute. *Med.-Chem Untersuch Lab.* 1866, 1, 133–140.

20. 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.
21. 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.
22. Bank, P.D. Protein Data Bank. *Nat. New Biol.* 1971, 233, 233.
23. 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.
24. 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.
25. Pedrini, B.; Martiel, I. Available online: <https://www.psi.ch/swissfel/internal-reports> (accessed on 3 July 2017).
26. 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.
27. 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.
28. 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.
29. 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.
30. 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.
31. 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.
32. Frauenfelder, H.; Chen, G.; Berendzen, J.; Fenimore, P.W.; Jansson, H.; McMahon, B.H.; Stroer, I.R.; Swenson, J.; Young, R.D. A unified model of protein dynamics. *Proc. Natl Acad. Sci. USA* 2009, 106, 5129–5134.

33. Moeglich, A.; Moffat, K. Engineered photoreceptors as novel optogenetic tools. *Photochem. Photobiol. Sci.* 2010, 9, 1286–1300.
34. 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.
35. Giegé, R. A historical perspective on protein crystallization from 1840 to the present day. *FEBS J.* 2013, 280, 6456–6497.
36. Davies, D.R.; Segal, D.M. Protein crystallization: Micro techniques involving vapor diffusion. *Methods Enzymol.* 1971, 22, 266–269.
37. Wlodawer, A.; Hodgson, K.O. Crystallization and crystal data of monellin. *Prot. Natl. Acad. Sci. USA* 1975, 72, 398–399.
38. Salemme, R.R. A free interface diffusion technique for the crystallization of proteins for X-ray crystallography. *Arch. Biochem. Biophys.* 1972, 2, 533–539.
39. 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.
40. Chayen, N.E. The role of oil in macromolecular crystallisation. *Structure* 1997, 5, 1269–1274.
41. Garcia_Ruiz, J.M.; Moreno, A. Investigations on protein crystal growth by the gel acupuncture method. *Acta Cryst.* 1994, 50, 484–490.
42. Durbin, S.D.; Feher, G.J. Studies of crystal growth mechanisms of proteins by electron microscopy. *J. Mol. Biol.* 1990, 212, 763–774.
43. 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.
44. Shlichta, P.J. Feasibility of mapping solution properties during the growth of protein crystals. *J. Cryst. Growth* 1986, 76, 656–662.
45. 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.
46. 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.
47. 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.
48. Trevino, S.R.; Scholtz, J.M.; Pace, C.N. Measuring and increasing protein solubility. *J. Pharm. Sci.* 2008, 97, 4155–4166.
49. Deriwenda, Z.S.; Vekilov, P. Entropy and surface engineering in protein crystallization. *Acta Cryst.* 2006, 52, 116–124.
50. 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.
-

Retrieved from <https://encyclopedia.pub/entry/history/show/90977>