International Journal of
Molecular Sciences

MDPI

Editorial

Emerging Topics in Protein Crystallography

Zhongzhou Chen [1,*] and Giuseppe Zanotti [2,*]

[1] State Key Laboratory of Animal Biotech Breeding and Beijing Advanced Innovation Center for Food Nutrition and Human Health, College of Biological Sciences, China Agricultural University, Beijing 100193, China
[2] Department of Biomedical Sciences, University of Padova, Via Ugo Bassi 48B, 35131 Padova, Italy
* Correspondence: chenzhongzhou@cau.edu.cn (Z.C.); giuseppe.zanotti@unipd.it (G.Z.)

Protein crystallography is the discipline concerned with the determination of the three-dimensional structure of biological macromolecules in a crystalline state. It took its first steps in the 1950s [1], and today, we can confidently assert that it has reached a phase of complete development [2]. Technically, bio-crystallography has undergone an extraordinary evolution in its more than seventy years of existence, but it could not have made the grand strides that have led it to be a technically mature discipline without the tumultuous parallel development of computing systems and related software [3–6]. Much water has passed under the bridge since computers were the size of cabinets that occupied numerous rooms and received their input data through punched cards; today's tools are a little larger than a tablet and have computing and data storage power unimaginable at the beginning. Crystallography has not only benefited from the developments of computer tools, both hardware and software, but it has also been strongly aided by the development of X-ray sources, particularly synchrotrons [7], X-ray detectors [8], and the improvement in crystallization methods (contribution 1). To give an immediate idea of the progress in the field, we recall that at the time of its foundation, the Protein Data Bank (PDB, http://www.rcsb.org, accessed on 9 May 2024), a database in which the coordinates and data of all resolved three-dimensional structures of macromolecules in the world are deposited, contained 12 structures. Today this number has exceeded 218,000, the vast majority of which (more than 84%) were determined using X-ray diffraction [9].

At this point, we must mention the development of what is defined as cryo-Electron Microscopy (cryo-EM), a technique that allows the determination of the three-dimensional structure of biological macromolecules at nearly atomic level from an ultra-thin frozen layer of the protein solution, without the need for the protein to be present in crystalline form [10–12]. This, in addition to the absence of the phase problem, has opened up the field to the determination of structures of huge macromolecular complexes, almost impossible to obtain in crystalline form and generally present in very small quantities within the cell.

In this context, it is understandable that this Special Issue received only five contributions, of which only two are of a methodological nature. We can confidently state that much of the technology inherent to the crystallographic technique has already been extensively developed in recent decades, and there are few truly significant developments on the horizon. We might expect more in terms of new insights into the structure of macromolecules particularly relevant in the field of human health, although, in this area, we cannot ignore the contribution of artificial intelligence, which has allowed the building of a database containing the prediction of the three-dimensional structures of essentially all of the proteins of numerous genomes [13,14]. Although these are predictions, they appear to be quite reliable and are often sufficient to draw conclusions about the functional and mechanistic aspects of the biological process. The only aspect where predictions are still particularly lacking concerns macromolecular complexes and protein–small-molecule complexes, although major progress is expected in this field in the coming years.

The two methodological articles presented in this Special Issue concern improvements in the structure determination process. Carrozzini et al. (contribution 2) describe REMO22,

check for updates

Citation: Chen, Z.; Zanotti, G. Emerging Topics in Protein Crystallography. *Int. J. Mol. Sci.* **2024**, *25*, 5311. https://doi.org/10.3390/ijms25105311

Received: 29 April 2024
Accepted: 10 May 2024
Published: 13 May 2024

Copyright: © 2024 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/).

a software program that uses prior information to bypass the phase problem using a computational technique called molecular replacement, which involves using an existing or predicted molecular model to determine a new structure based on an approximate existing model (and, of course, the crystal diffraction data of the structure to be determined). REMO22 represents an evolution of the previous software REMO09 and is inserted into a pipeline that also includes additional computational stages, such as phase refinement and automated model building. This pipeline tends to be completely automatic for solving the structures of proteins and nucleic acids, with minimal intervention from the operator. In the other methodological work, Ma et al. (contribution 3) describe the use of an anomalous signal for the correct positioning and orientation of protein fragments containing anomalous scatterers. The technique has been used, in particular, for positioning in the electron density of fragments containing chlorine and sulfur in the non-structural protein 1 (nsp1) of the most recent dangerous pandemic SARS-CoV-2, but the method is generally applicable for any anomalous scatterers. Its use requires a tunable source that can reach the appropriate wavelengths, possibly at low energy in vacuum. In the case in question, the authors, by measuring data at two different energies, were able to identify multiple orientations in various fragments containing chlorine or sulfur.

The other three papers in the Special Issue are of an applied nature. Gao et al. (contribution 4) describe the structural characterization of the enzyme N-Acetyl Sugar Amidotransferase, an enzyme involved in the biosynthesis of lipopolysaccharides in the bacterium *Legionella pneumophila*. The structure, resolved at 2.33 Angstrom resolution, contains a Rossmann-like fold with a PP loop, suggesting that the catalyzed reaction probably requires the conversion of ATP to AMP and PPi. In their work, Shang et al. (contribution 5) solve three complex structures of 6 mA demethylase *Caenorhabditis elegans* NMAD-1A through rational mutations and find a unique "stretch-out" conformation of its DNA binding region. Biochemical and structural studies reveal that it preferentially demethylates 6 mA, an emerging epigenetic mark, in the unpairing regions and binds to the nucleosome with the help of the carboxy-terminal domain and the zinc finger domain.

The last paper, Del Giudice et al. (contribution 6), is not strictly crystallographic, but describes the use of the SAXS (Small-Angle X-ray Scattering) technique in solution to determine the content of disordered structure in the different redox states of the enzyme CP12 from *Arabidopsis Thaliana*, a protein universally distributed in all photosynthetic organisms and involved in regulating the Calvin–Benson cycle of photosynthesis. Considering different levels of disorder, the SAXS data suggest that the reduced form is completely disordered, and that the oxidized form is better described by conformers that include partially ordered portions in some areas, particularly around the sulfide bridge, which coexist with portions that remain disordered.

Funding: This work was supported by the National Key Research and Development Program of China (2023YFE0123700) and the National Natural Science Foundation of China (32071210 and 32371268).

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

List of Contributions:

1. Russo Krauss, I.; Merlino, A.; Vergara, A.; Sica, F. An overview of biological macromolecule crystallization. *Int. J. Mol. Sci.* **2013**, *14*, 11643–11691. https://doi.org/10.3390/ijms140611643.
2. Carrozzini, B.; Cascarano, G.L.; Giacovazzo, C. The Automatic Solution of Macromolecular Crystal Structures via Molecular Replacement Techniques: REMO22 and Its Pipeline. *Int. J. Mol. Sci.* **2023**, *24*, 6070. https://doi.org/10.3390/ijms24076070.
3. Ma, S.; Mykhaylyk, V.; Bowler, M.W.; Pinotsis, N.; Kozielski, F. High-Confidence Placement of Fragments into Electron Density Using Anomalous Diffraction—A Case Study Using Hits Targeting SARS-CoV-2 Non-Structural Protein 1. *Int. J. Mol. Sci.* **2023**, *24*, 11197. https://doi.org/10.3390/ijms241311197.

4. Gao, J.; Xu, W.; Liu, T.; Sun, W.; Wang, N.; Ma, J.; Ge, H. Structural Characterization of an N-Acetyl Sugar Amidotransferase Involved in the Lipopolysaccharide Biosynthesis in Bacteria. *Int. J. Mol. Sci.* **2023**, *24*, 15491. https://doi.org/10.3390/ijms242015491.
5. Shang, G.; Yang, M.; Li, M.; Ma, L.; Liu, Y.; Ma, J.; Chen, Y.; Wang, X.; Fan, S.; Xie, M.; et al. Structural Basis of Nucleic Acid Recognition and 6mA Demethylation by Caenorhabditis elegans NMAD-1A. *Int. J. Mol. Sci.* **2024**, *25*, 686. https://doi.org/10.3390/ijms25020686.
6. Del Giudice, A.; Gurrieri, L.; Galantini, L.; Fanti, S.; Trost, P.; Sparla, F.; Fermani, S. Conformational Disorder Analysis of the Conditionally Disordered Protein CP12 from Arabidopsis thaliana in Its Different Redox States. *Int. J. Mol. Sci.* **2023**, *24*, 9308. https://doi.org/10.3390/ijms24119308.

References

1. Thorp, H.H. Proteins, proteins everywhere. *Science* **2021**, *374*, 1415. [CrossRef] [PubMed]
2. Jaskolski, M.; Dauter, Z.; Wlodawer, A. A brief history of macromolecular crystallography, illustrated by a family tree and its Nobel fruits. *FEBS J.* **2014**, *281*, 3985–4009. [CrossRef] [PubMed]
3. Agirre, J.; Atanasova, M.; Bagdonas, H.; Ballard, C.B.; Baslé, A.; Beilsten-Edmands, J.; Borges, R.J.; Brown, D.G.; Burgos-Mármol, J.J.; Berrisford, J.M.; et al. The CCP4 suite: Integrative software for macromolecular crystallography. *Acta Crystallogr. D Struct. Biol.* **2023**, *79 Pt 6*, 44–461. [CrossRef]
4. Krissinel, E.; Uski, V.; Lebedev, A.; Winn, M.; Ballard, C. Distributed computing for macromolecular crystallography. *Acta Crystallogr. D Struct. Biol.* **2018**, *74 Pt 2*, 143–151. [CrossRef] [PubMed]
5. Adams, P.D.; Afonine, P.V.; Bunkóczi, G.; Chen, V.B.; Echols, N.; Headd, J.J.; Hung, L.W.; Jain, S.; Kapral, G.J.; Grosse Kunstleve, R.W.; et al. The Phenix software for automated determination of macromolecular structures. *Methods* **2011**, *55*, 94–106. [CrossRef] [PubMed]
6. Jain, D.; Lamour, V. Computational tools in protein crystallography. *Methods Mol. Biol.* **2010**, *673*, 129–156. [CrossRef] [PubMed]
7. Dauter, Z.; Jaskolski, M.; Wlodawer, A. Impact of synchrotron radiation on macromolecular crystallography: A personal view. *J. Synchrotron Radiat.* **2010**, *17*, 433–444. [CrossRef] [PubMed]
8. Powell, H.R. X-ray data processing. *Biosci. Rep.* **2017**, *37*, BSR20170227. [CrossRef]
9. 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] [PubMed]
10. Dubochet, J. Cryo-EM—The first thirty years. *J. Microsc.* **2012**, *245*, 221–224. [CrossRef] [PubMed]
11. Stuart, D.I.; Subramaniam, S.; Abrescia, N.G. The democratization of cryo-EM. *Nat. Methods* **2016**, *13*, 607–608. [CrossRef]
12. Benjin, X.; Ling, L. Developments, applications, and prospects of cryo-electron microscopy. *Protein Sci.* **2020**, *29*, 872–882. [CrossRef] [PubMed]
13. 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]
14. 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] [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.