What biocrystallogenesis tells us – What is needed in the future

R. Giegé

Architecture et réactivité de l'ARN, Université de Strasbourg, CNRS, IBMC, 15 rue René Descartes, France,

r.giege@ibmc-cnrs.unistra.fr

Protein crystallization dates back to the 19th century, but became only recently a mature science thanks to strong *interdisciplinar* efforts based on physics, chemistry, biology, and associated technologies. Thus, many crystallization methods adapted to proteins were developed, but only a few are systematically used by structural biologists [1, 2]. Nevertheless, the number of successful crystallizations increased tremendously in the last two decades, leading to ~10^5 X-ray structures, covering alas only a restricted part of the *macromolecular diversity* and *3D-space* across the tree of life. This relies on idiosyncratic features of biomacromolecules making their crystallization mostly unpredictable, despite it is now well established that biocrystallogenesis is governed by the same physical rules than crystal growth of inorganic molecules. In another perspective, protein crystallization is a *self-assembly* process of particles of more or less defined chemical structure and conformation, with crystals being *supramolecular assemblies*. Empirical observations have shown that crystallizations can occur in the bulk of complex fluids, even *in vivo* for assemblies as large as ribosomes [3], and that purity favours the process [2]. Life also uses *self-assembly* and *supramolecular* processes leading mostly to *transient*- and fewer to *stable* complexes. An integrated view of supramolecularity implies that entities crystallizing or participating in biochemical processes require *determinants* and *antideterminants*, that favour or disfavour *correct* or *incorrect* associations. Moreover, and as a result of *evolution*, biomacromolecules coexisting in a given biological context show a proper balance between features favouring or disfavouring such associations. If this balance is broken, cellular *disorders* / *diseases* may occur. Understanding these phenomena is a challenge for the future.

Altogether, these considerations trace the future of biocrystallogenesis research. Crystals covering the missing gaps in structural genomics have to be grown, especially for hydrophobic proteins and macromolecular complexes. Structural *dynamics* ("*4D-structures*" with time as the 4th parameter), conformational *plasticity*, and *allosteric* phenomena underlying biological processes should be better documented. For that, strategies combining X-ray crystallography with alternative technologies, including cryo-EM and computational tools, are needed [4]. Thus, crystallogenesis research remains essential in this quest. Amongst others, analysis of structures of apo- and liganded proteins captured *in crystallo* under different packing / solvent environments will help to understand their structural and functional plasticity. Knowledge of large panels of *thermodynamic compatible* structures, together with precise analysis of packing contacts and contacts within oligomers, will be crucial to decipher the chemical rules governing macromolecule *self-assembly*. This should moderate the recurrent idea claiming that atypical packing polymorphs are useless 'artefacts'. Understanding of crystal growth *in vivo* and under *crowding* conditions is anticipated, and rational-based engineering of biomacromolecule crystallization becomes possible. *Self-assembly rules* will also be the guide for engineering novel biomaterials and, in a wider perspective, their understanding will open the route towards supramolecular biology [5].

1. C. Sauter, B. Lorber, A. McPherson & R. Giegé, in *International Tables for Crystallography*, edited by E. Arnold, D. Himmel & M. Rossmann (Chichester, John Wiley & Sons), volume F, 2012, pp. 99–121.

2. R. Giegé, *FEBS. J.*, **280**, (2013), 6456.

3. A. Yonath, *Angew. Chem. Int. Ed; Engl.*, **49**, (2010), 4341.

4. R. Giegé, C. Sauter, *HSFP J.*, **4**, (2010), 109.

5. J.-M. Lehn, *Science*, **295**, (2002), 2400.