Oral presentation

Teaching crystallography on real data in life sciences and beyond

S. Gražulis[1*], A. Merkys[1], A. Vaitkus[1], A. Grybauskas[1]

[1)] Vilnius University Institute of Biotechnology, Life Sciences Center, Vilnius University, Saulėtekio al. 7, LT-10257 Vilnius, Lithuania

[*)] correspondence address: grazulis@ibt.lt

Crystallography is applied to many disciplines of natural sciences: materials research, solid state physics, structural biology, engineering – all use results obtained from crystallographic research as an essential part of their toolbox. A great number of excellent textbooks can be recommended [1]. The XXI century however gives us additional opportunities and challenges: a large amount of crystallographic data is accessible at our fingertips, and a lot of Free/Libre Open Source Software is available to process this data. The students thus get the opportunity to have a hands-on experience on working with real data for their classes, do data analysis and even make small "discoveries" of their own. At the same time, doing computations or writing software of their own requires from students dedication and broad set of prerequisite knowledge.

On the examples of the use of the Crystallography Open Database [2] and the Protein Data Bank [3], and the personal experience of teaching crystallography and bioinformatics in the Vilnius University I would like to showcase how crystallographic databases can be used for university courses. I will share the experience of using high level memory safe languages, such as Perl and, more recently, Ada, to teach students the concepts of crystallography, bio- and cheminformatics through practical software construction and application to large data volumes. The emphasis will be in the assignments that are not just "point-and-click" exercises or the use of pre-programmed "black boxes"; I will show how students can implement mathematical abstractions from the basic computer math and apply their implementations to complete databases. I will provide some examples of the assignments and the problems that the students solved. I would also like to discuss difficulties that await us on this road – steep learning curve, the "devil in the details" and the importance of prior education.

[1] Helliwell, J. R. (2021) How should we teach crystallography? A review of teaching books' contents pages. *Crystallography Reviews* 27(3-4), 135-145. Informa UK Limited. DOI: https://doi.org/10.1080/0889311X.2021.1978080.

[2] Gražulis, S. *et al.* (2012) Crystallography Open Database (COD): an open-access collection of crystal structures and platform for world-wide collaboration. *Nucleic Acids Research* 40, D420- D427. DOI: https://doi.org/10.1093/nar/gkr900.

[3] Burley, S. K. *et al.* (2018) Protein Data Bank: the single global archive for 3D macromolecular structure data. Nucleic Acids Research 47(D1), D520-D528. Oxford University Press (OUP). DOI: https://doi.org/10.1093/nar/gky949.