Keynote Lecture

KN08

Macromolecular Ab Initio Phasing Enforcing Secondary and Tertiary Structure

I. Usón[1]

[1]*ICREA at IBMB-CSIC, Department of Structural Biology, Barcelona, Spain*

Ab Initio phasing of macromolecular structures, from the native intensities alone with no experimental phase information or previous particular structural knowledge has been the object of a long quest, limited by two main barriers: structure size and resolution of the data. Beyond the first successful atomic resolution scenario, current approaches have been developed, exploiting alternative constraints to atomicity, through use of the Patterson function, density modification and data extrapolation [1]. Our own approach relies on the combination of locating model fragments such as polyalanine alpha-helices with the program PHASER [2] and density modification with the program SHELXE [3]. Given the difficulties in discriminating correctly positioned fragments, many putative groups of fragments have to be tested in parallel, thus calculations are performed in a grid or supercomputer. The method has been called after the Italian painter Arcimboldo, who used to compose portraits out of fruits and vegetables. In the case of our program, most collections of fragments remain a "still-life", but some are correct enough for density modification to reveal the protein's true portrait (http://chango.ibmb.csic.es/ARCIMBOLDO). Beyond alpha-helices, any unknown structure should contain fragments already seen in the PDB, but how to retrieve and exploit this information? Our program BORGES identifies, retrieves and exploits this geometric information The PDB database contains a vast amount of information and for any unknown structure, given small enough fragments (e.g. two helices or three strands in a particular disposition), similar models (<0.5Å rmsd) are bound to occur in some of the deposited entries. In analogy to Borges' "Library of Babel" that enclosed books with all random combinations of letters and therefore held any possible book, the information required to phase a structure through fragment search and density modification should already be present somewhere in the PDB. The more so as unlike "Borges library", the PDB is non-random, containing in all sort of structural contexts only meaningful structural units. In addition, to bootstrap phasing our method requires small sentences instead of complete volumes, that is, a small fraction of perfect mainchain rather than a complete description of the structure (http://chango.ibmb.csic.es/BORGES). Using these methods, a number of unknown macromolecules with a few thousand atoms and resolutions around 2 Å have been solved.

[1] *Sheldrick, G.M., Hauptman, H.A., Weeks, C.M., Miller, R. & Usón, I. International Tables for Macromolecular Crystallography vol. F, (eds., M.G. Rossmann and E. Arnold) 333–345 (Boston, 2001); Burla, M.C., Carrozzini, B., Cascarano, G.L., Giacovazzo, C., P,* ***[2]*** *McCoy, A.J. Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. J. Appl. Crystallogr. 40, 658–674 (2007).,* ***[3]*** *A. Thorn and G.M. Sheldrick. Acta Crystallogr. D69, 2251-2256 (2013)*

Keywords: ab initio phasing, macromolecular, supercomputing