[MS04-P06] Reciprocal space clustering of BORGES-ARCIMBOLDO partial solutions: Practical cases

Claudia Millán[1], Massimo Sammito[1], Kathrin Meindl[1], Iñaki M de Ilarduya[1], Ivan De Marino[1], Isabel Usón[1,2]

[1] *Instituto de Biología Molecular de Barcelona, Spain*
[2] *Institució Catalana de Recerca I Estudis Avançats, Barcelona, Spain*
E-mail: cmncri@ibmb.csic.es

ARCIMBOLDO [2,3,5] is a multisolution ab-initio phasing method for data up to 2.0 Å. It is based on location of small models with PHASER [1] and density modification with SHELXE [4] . In analogy with the paintings from Giuseppe Arcimboldo (1527-1593), who assembled portraits out of vegetables, ARCIMBOLDO assembles structural hypotheses out of ideal secondary structure elements, such as alpha-helices, or composite fragments retrieved from a library. Density modification and autotracing with SHELXE reveals the "portrait" of our protein when the starting hypothesis is close enough. The program BORGES takes its name from the writer Jorge Luis Borges (1899-1986), who described in one of his novels a universe conceived as an infinite library with all possible books, and is used to generate custom libraries extracting composite fragments from the, vast if not infinite, PDB that fit into a given geometrical description. It has been recently successful in solving cases in which the use of single elements of secondary structure within ARCIMBOLDO was not enough. Solved cases comprise all-beta proteins, mixed alpha-beta or all helical proteins. Other kind of prior information can be exploited within this frame, such as the one coming from derivatives or substructures. In either case, initial stages of our multisolution method render a large number of partial solutions with similar figures of merit, that need to be pursued to reach a substructure that is complete enough to be expanded to the full structure. This makes discrimination and combination of promising solutions in early stages a powerful and desirable tool.

Combination of information in reciprocal space depends on finding the correct origin shift. This can be challenging as the phase information provenient from such small fragment is very noisy but exploiting it can be helpful in different ways at various stages of the process. Non-polar spacegroups, with a limited number of possible origin shifts present favorable cases.

A set of test cases and previously unknown structures has been solved recently with BORGES-ARCIMBOLDO, and various analyses were performed in order to get our way through the solution or assess the keys to success and failure. Testing of reciprocal space clustering at different steps and for different purposes was done. Over this time we have learned some interesting and practical issues relating to combination of phase information and identification of correct hypothesis at early stages that will be presented here illustrated through particular cases.

[1] Airlie J. McCoy *et al. J Appl Crystallogr.* 40(4):658–674. (2007)
[2] Dayte D. Rodríguez Martínez *et al*. Nature Methods 6(9):651-653 (2009)
[3] Dayté D. Rodríguez Martínez *et al.* Acta Crystallogr D. 68(Pt 4):336-43. (2012)
[4] George M. Sheldrick. Z. Kristallogr. 217: 644-650 (2002)
[5] Isabel Usón et al. Springer, Dordrecht, The Netherlands, (2013), pp.123-132.

Keywords: computing, macromolecular crystallography, phase determination