MS11-1-3 Evolutionary covariance to assess the biological relevance of putative homo-oligomeric interfaces #MS11-1-3

J.J. Burgos-Mármol [1], R.M. Keegan [2], E. Krissinel [2], D.J. Rigden [1]
[1]ISMIB, University of Liverpool - Liverpool (United Kingdom), [2]UKRI-STFC, Rutherford Appleton Laboratory, Research Complex at Harwell - Didcot (United Kingdom)

Abstract
Evolutionary covariance is a phenomenon that occurs due to natural selection of mutations that tend to preserve the structure and, consequently, the function of a protein. When a mutation appears in a protein, the interaction with the surrounding residues can be affected, modifying the structure and possibly its functionality in a detrimental way. However, if a second compensating mutation occurs so that the contact is preserved, the functionality may not be adversely affected and the new variant could pass down to new individuals. By studying a protein's homologues, these pair correlations that signal structural contacts can be found.
This principle has been employed in the last decades to develop software tools that are able to predict contacts within a protein chain by looking at the sequence alone, with accuracy related to the number of homologues available. While more recent tools, using machine learning algorithms, have been trained mostly to predict intramolecular contacts, previous programs with little or no training to distinguish intramolecular contacts from those within oligomer interfaces can be used to infer intermolecular contacts.
We present project PISACov, a tool that exploits first-generation contact-prediction programs as well as novel algorithms to evaluate interfaces in homo-oligomeric proteins. To this end, the workflow includes widely used tools to produce sequence alignments and contact predictions that are finally put together to assess the interface. The application involves the use of PISA, a CCP4 program that determines the stability of the interfaces observed in a crystal structure by calculating their solvation free-energy. With this method, PISACov acts a an independent source of information to determine which of the crystal interfaces are likely biologically relevant. Other uses beyond validation of PISA results are possible. For instance, alternative computationally generated interfaces (e.g. from AlphaFold) can be evaluated by PISACov to be assessed with this approach.