Automated evaluation of quaternary structures from protein crystal structures

Jose M. Duarte[1,2], Spencer Bliven[1,4], Aleix Lafita[1], Guido Capitani[1], Stephen K. Burley[2,3]

1. Paul Scherrer Institute, Villigen, Switzerland
2. RCSB Protein Data Bank, SDSC, UCSD, San Diego, CA, USA
3. RCSB Protein Data Bank, Rutgers University, Piscataway, NJ, USA
4. NCBI, NIH, Bethesda MD, USA

Crystallography is the most powerful technique for generating atomic level structures of proteins and other biological macromolecules. However, it does not always yield definitive insights into the quaternary structures of biological macromolecules. In order to provide better tools for determining the most likely quaternary structure in proteins, we have developed the new EPPIC 3 method. It uses evolutionary considerations as the ultimate arbiters of the biological relevance of interfaces and assemblies, thereby offering a complementary approach versus other available methods that rely on thermodynamic considerations. EPPIC 3 extends our previous Evolutionary Protein-Protein Interface Classifier (EPPIC), by going beyond classifying pairwise interfaces. It identifies all possible topologically valid assemblies present in a protein crystal and provides predictions as to likely quaternary structures. Pairwise interface classifications are based on two evolutionary scores and a single geometric score that are in turn combined into a final score. This approach was trained against a large dataset of known biologically relevant and crystal interfaces. Assembly enumeration is achieved by representing the crystal lattice as a periodic graph. Finding valid assemblies is then reduced to the problem of finding subgraphs complying to a set of rules, which guarantees closed assemblies (Point Group symmetries) and isomorphism in the assembly composition and connectivity throughout the crystal.
The software is accessible through an easy to use web graphical interface at http://www.eppic-web.org. The graphical interface is designed to aid the crystallographer in interpreting putative quaternary structures using 2D and 3D graphical tools that operate within the browser.