



Contribution ID: 41

Type: Talk

The Small Angle Scattering Biological Data Bank (SASBDB) - more than a FAIR repository

Wednesday, 18 October 2023 09:30 (30 minutes)

The Small Angle Scattering Biological Data Bank (SASBDB) at EMBL is an open-access, freely available, and searchable resource for the structural-biological/biophysical and biotechnology communities and is the only curated repository for small angle X-ray (SAXS) and neutron scattering (SANS) data. SASBDB was established in 2014 for the deposition SAS projects focusing on the structure(s) of biological macromolecules in solution and related biomaterials. From the outset, the data bank adheres to an open access philosophy in keeping with Findability, Accessibility, Interoperability, and Reusability (FAIR) data principles. In addition, working with partners across the scientific community, SASBDB and EMBL are involved in developing and establishing data standards, validation, and reporting protocols. The data bank has grown significantly and is integrated into numerous resources, including UniProt, the Protein Ensemble Databank, PDB-Dev, Disprot, the 3D-Beacons Network, as well as being mirrored through the Yorodumi PDB/EMDB/SASBDB browser at the PDBj and is part of the wwPDB OneDep system. Several journals require the deposition of SAXS or SANS data for biological solution scattering into SASBDB, and it has become a recommended databank for the deposition of these projects. SASBDB also acts as an “experimental hub” for certain ELIXIR communities, and is moving toward integration with other European and international data re-sources, including the Worldwide Protein Data Bank (wwPDB), Biological Magnetic Resonance Data Bank (BMRB) and Electron Microscopy Databank (EMBD).

The provision of SASBDB as a data resource at EMBL and coordinating the integration of SASBDB across larger life science data infrastructures is a key objective that relies on adaptability, communication, and responsiveness with the needs of the diverse SAS community. New and exciting opportunities to evolve the databank toward housing high-end scattering experiments include devising tools for the deposition of anisotropic data, biomaterials data, neutron scattering, time-resolved data and other ‘4D-structural biology’ projects. This requires developing common and interoperable representations of data, metadata conventions, standards, and workflows in consultation with existing core data resources (e.g., the PDBe, the BioImage Archive) as well as the ELIXIR IDP and Bioschemas communities, the SAS validation task force (SASvtf) and the canSAS initiative. Dialogue and collaboration with large-scale instrument facilities is also crucial to achieve the overall objectives of SASBDB to make it ‘more than a data repository’ and an integral component of community-based objectives for SAS standards and reporting that are as FAIR as possible.

Presenter: JEFFRIES, Cy (EMBL Hamburg)

Session Classification: Open access / databases / collaborations between facilities