**Structure Integration with Function, Taxonomy and Sequences (SIFTS)**

Ganesh Jawahar Swaminathan[a],Sameer Velankar[a], Phil McNeil[a], Antonio Suarez[a], Daniel Barrell[b], Jules Jacobsen[b], Kim Henrick[a], [a]*Macromolecular Structure Database.* [b]*UnProtein Resource, European Bioinformatics Institute, EMBL Outstation, Hinxton, Cambridge CB10 1SD, Unite*d *Kingdom*. E-mail: msd@ebi.ac.uk

One of the major obstacles to the improved integration of structural databases such as MSD (http://www.ebi.ac.uk/msd) and sequence databases like UniProt is the absence of up to date and well maintained mapping between corresponding entries. MSD and UniProt have worked closely to clean up the taxonomy and sequence cross-reference, a vital step in integrating sequence family databases such as Pfam and Interpro with structure-oriented databases (SCOP and CATH). This information has been made available to the eFamily group (http://www.efamily.org.uk/), and now forms the basis of the regular interchange of information between the member databases (MSD, Uniprot, Pfam, Interpro, SCOP and CATH). This work was carried out under the SIFTS initiative (http://www.ebi.ac.uk/msd-srv/docs/sifts/). As well as the domain boundary definitions, the schema also allows any associated sequence or structural alignments to be encapsulated in XML. Since the *e*Family schema is complex, an API has been developed for easy integration into BioPerl. For example, the API allows a Pfam alignment wrapped up in the *e*Family XML to be returned as a BioPerl alignment object and *vice versa*. The use of such API's in the development of production quality webservices will also be discussed.
**Keywords: data integration, webservices, data exchange**