**Overview**

**MEDP-SiteClassifier** is a powerful biostructural data repository for molecular biologists and medicinal chemists to mine any local 3D binding site similarities at a PDB scale.

**MEDP-SiteClassifier** uses a 3 step procedure with (1) MED-SuMo technology to compare all pairs of binding site by looking at 3D shared structural chemical features (HBond donor, acceptor, Charges, Hydrophobic, ...) on a large set of biostructures up to the whole PDB, (2) clustering the whole set of detected binding sites into a relational database, (3) providing advanced data mining interface to analyze clusters and interfamily links in term of conserved SCF (Structural Chemical Feature) and 3D superposition. Multiple binding pockets are automatically assigned.

With **MEDP-SiteClassifier**, any of your protein structure of interest are automatically projected toward all 3D local similarities. Applications include (1) Functional Annotation, (2) Binding site characterization, (3) Off target Identification, (4) Drug repurposing, and (5) Scaffold hopping.

**MEDP-SiteClassifier features**

- **Biostructural Data repository** with all detected 3D binding site similarities and biological annotations.
- **Search interface to mine the database repository**: a Tab based GUI for web browser.
- **Explore your cluster of similar binding sites and export in XML format**.
- **Analyze the 3D superposition in term of conserved 3D chemical features (HBond donor, acc., aromatic, hydrophobic,...) and export signatures in XML format**.
- **Non-frequent chemical features can be associated to selectivity mechanism**.
- **Select superposition from the similarity plot where interfamily hits are highlighted**.
- **Display superposition in Jmol** [RM Hanson, J. Appl. Cryst. 43, 1250-1260 (2010)]

**Your Biostructural Data Repository to Explore all 3D-interaction Local Similarities**
**Functional classification on Purine-Binding proteins**

From Doppelt-Azeroual O, Delfaud F, Moriaud F and de Brevern AG Prot. Sci., 19(4), 847–867 (2010): 2229 selected protein structures containing 2322 purine binding sites were selected from the PDB (as May 2009) by looking at ligands containing either adenosine or guanosine: A*P, NAD, G*P. With the selected clustering parameters, 247 clusters were identified comprising 2115 binding sites. A Shannon Entropy for each cluster was expressed to measure the purity regards to the 442 collected different protein functions. Result analysis shows biological uniformed clusters and heterogeneous family that is directly resulting from a MEDP-SiteClassifier skill to merge subpockets together.

**Analysis of HSP90 protein families**

Classification of 146 binding sites of protein with the Bergerat ATP-binding fold are from different families; 78 are from HSP90, 38 from topoisomerase/MutL, 26 are from histidine kinase, and four are from a-ketoacid dehydrogenase kinase C (BCK).

The constituent families are quite different but their ATP binding sites appear quite alike. MED-SMA detects five different clusters in a two minute job on a four CPU machine.

The classification is detecting similarities of binding modes which are relevant for Drug Design application, rather than pocket similarities nor ligand similarities.

Interestingly, the proteins which can bind radicicol are represented in cluster n°4 (superimposition in the figure)


**Interfamily-based Local similarities within the PDB**

**Summary**

- Access all pre-calculated clusters of similar local binding sites in PDB with MED-SuMo technology
- Qualify frequent & non-frequent 3D chemical features
- Search for specific interfamily hits (defined by Pfam or E.C.) or explore interactive similarity matrix
- Designed for (1)Functional Annotation, (2)Binding site characterization, (3)Off target Identification, (4)Drug repurposing, (5)Scaffold hopping

**Request further information about MEDP-SiteClassifier today!**