MEDP-SiteClassifier

Explore All Binding Site Similarities including Interfamily Links at PDB scale

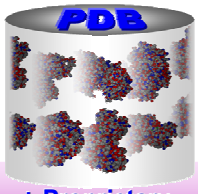
Overview

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



+ Proprietary Protein-Ligand Structures

66432 Proteins
3204 Proteins-DNA/RNA
1288 DNA
832 RNA

35944 Eukaryota
25997 Bacteria
4113 Viruses
2860 Archaea

17178 Homo sapiens
4927 Escherichia coli
2955 Mus musculus
1964 Saccharomyces cerevisiae
1805 Bos taurus
1535 Rattus norvegicus
874 Bacillus subtilis
37707 Other

13311 Hydrolases
9388 Transferases
5887 Oxidoreductases
2687 Lyases
1498 Isomerases
1253 Ligases

1

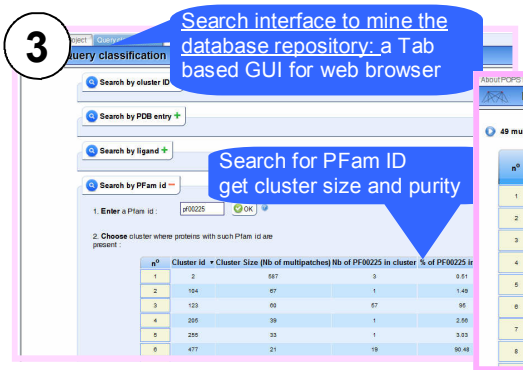
3D comparison of all binding sites, normalizing the MED-SuMo score into a similarity matrix that is clusterized. All data are stored into a relational database.

2

Biostructural Data repository with all detected 3D binding site similarities and biological annotations

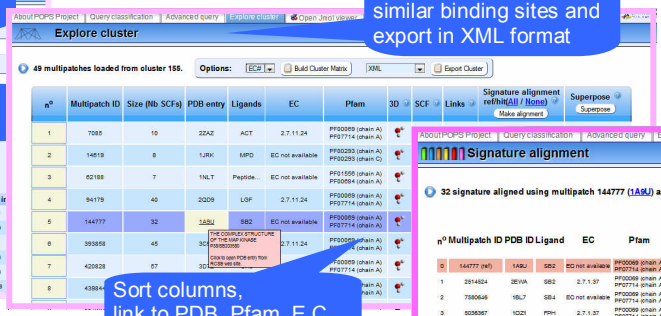
3

Search interface to mine the database repository: a Tab based GUI for web browser



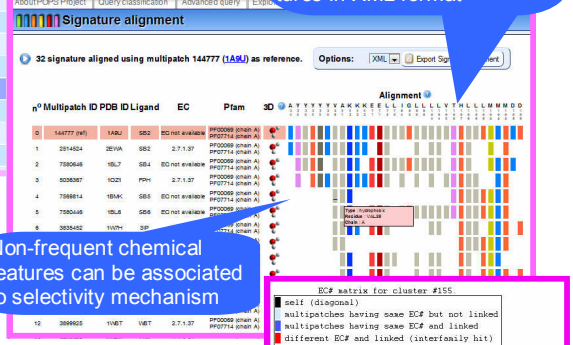
Search for PFam ID get cluster size and purity

Explore your cluster of similar binding sites and export in XML format



Sort columns, link to PDB, Pfam, E.C.

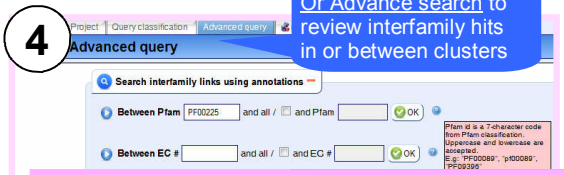
Analyze the 3D superposition in term of conserved 3D chemical features (Hbond donor, acc., aromtic, hydrophobic,…) and export signatures in XML format

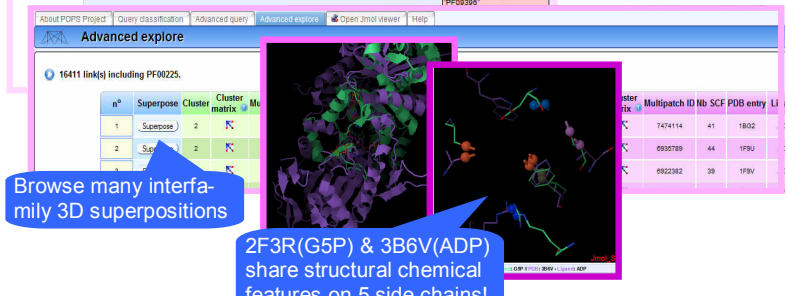


Non-frequent chemical features can be associated to selectivity mechanism

4

Or Advance search to review interfamily hits in or between clusters

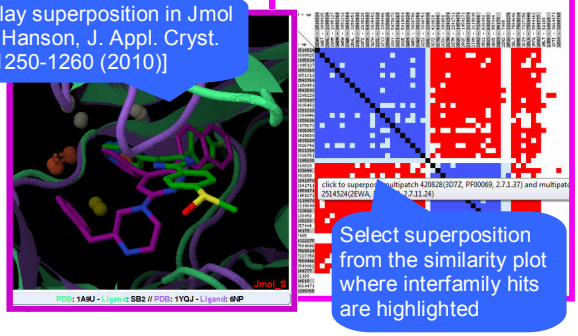




Browse many interfamily 3D superpositions

2F3R(G5P) & 3B6V(ADP) share structural chemical features on 5 side chains!

Display superposition in Jmol [RM Hanson, J. Appl. Cryst. 43, 1250-1260 (2010)]



Select superposition from the similarity plot where interfamily hits are highlighted

MEDP-SiteClassifier for

- **Functional Annotation**
- **Binding Site Comparison**
- **Off target Identification**
- **Drug Repurposing**
- **Scaffold Hopping**

Your Biostrucural Data Repository to Explore all 3D-interaction Local Similarities

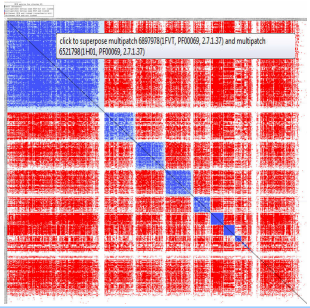
MEDP-SiteClassifier

Explore All Binding Site Similarities including Interfamily Links at PDB scale

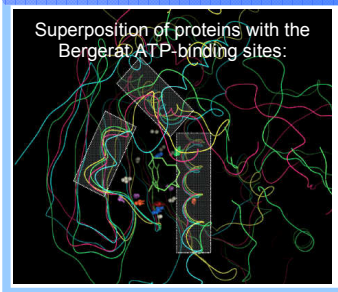
Functional classification on Purine-Binding proteins

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.

From a classification of the PDB (as June 2008): Here cluster#1 on the right includes already 714 kinase binding sites, ranked into subcluster (blue area) according to Enzyme Class numerotation 2.7.1.37, .11.1, .10.1, .1.112, .10.2, .22,22, .11.11.



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 α-ketoacid dehydrogenase kinase C (BCK).

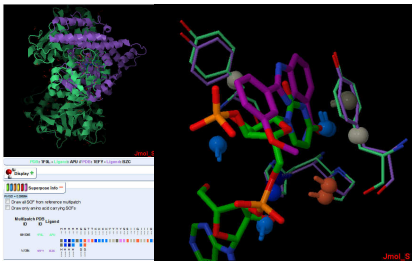
The constituent families are quite different but their ATP binding sites appear quite alike. MED-SMA detects five different clusters in a two minute job on a four CPU machine.

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

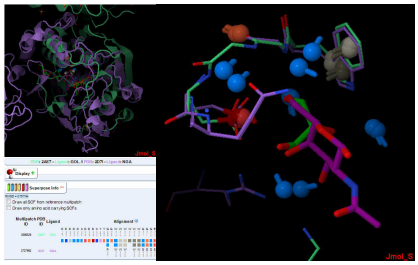
Interestingly, the proteins which can bind radicicol are represented in cluster n°4 (superpimposition in the figure)

Doppelt-Azeroual O, Moriaud F, Delfaud F and de Brevern AG "Analysis of HSP90-related folds with MED-SuMo classification approach", Drug Design Development and Therapy, 3:59-72

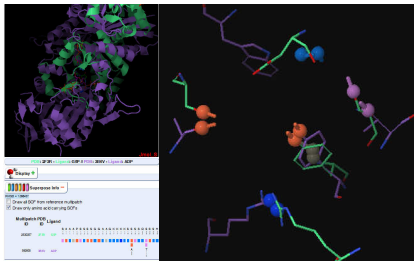
Interfamily-based Local similarities within the PDB



Interfamily hit (15% seq. Id.) 1F0L (blocks protein synthesis by transfer of ADP-ribose from NAD to a diphthamide residue of EF-2) and 1EFY (poly ADP-ribose polymerase catalyses the covalent attachment of ADP-ribose units from NAD+ to itself and to a limited number of other DNA binding proteins)



Interfamily hit (32% seq. Id.) 2D7I (human UDP-GalNAc: polypeptide alpha-N-acetylGalactosaminyltransferase pp-GalNAc-T10) and 2AE7 (human M340H-Beta1,4-Galactosyltransferase-I in Complex with Pentasaccharide)



Interfamily hit (35% seq. Id.) 2F3R (GMP kinase) and 3B6V (Motor domain of human kinesin family).

Color codes of 3D shared chemical features: *Red=HBond acceptor, blue= HBond donor, darkblue= positive charge, magenta=hydroxyl, grey=hydrophobic.*

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
- Webpages with sorting capabilities & Jmol 3D viewer
- Export clusters in CSV and XML file format with signatures of shared 3D-structural chemical features
- Enlarge PDB with your proprietary biostructures
- Versatile Relational Database architecture including biological annotations (Pfam, EC, ...)
- Optimized for multicore and multinode Linux servers

Request further information about MEDP-SiteClassifier today !

MEDIT

www.medit.fr

MEDIT France: 2 rue du Belvédère, 91120 Palaiseau, Tel +33 (0)1 6014 8743

MEDIT US: 7985 Dunbrook Rd., San Diego CA 92126, Tel +1 (858) 342 6807

info@medit.fr

www.medit.fr

Browse all detected 3D local binding site similarities, look at selective chemical features, superpose biostructures