# M.E.D.I.T.

Molecular Extended Distribution in Information Technology

**YOUR FAVORITE PARTNER FOR *in silico* DRUG DISCOVERY**

# From Target based drug design to innovative cheminformatics tools

Fabrice Moriaud CSO, François Delfaud, CEO

– Chemoinformatics in Europe: Research and Teaching – 1 June 2006

-1- **MED-SuMo** → how to compare active sites ; affinity and selectivity assessment

-2- Cheminformatics Applications: fragment based drug design

-3- Demo

-1- **MED-SuMo** → how to compare active sites ; affinity and selectivity assessment

-2- Cheminformatics Applications: fragment based drug design

-3- Demo

# 1) MED-SuMo ⇒ a CNRS/IBCP collaboration

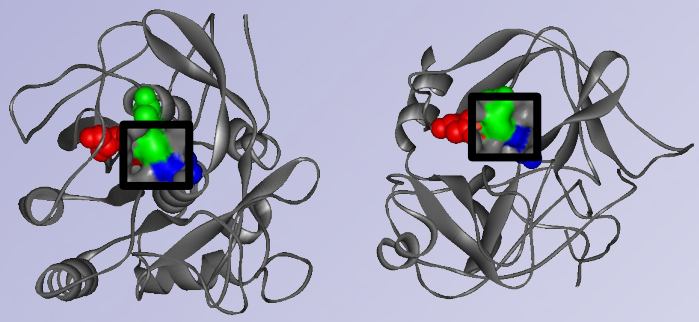
Pôle Bio-Informatique Lyonnais

SuMo

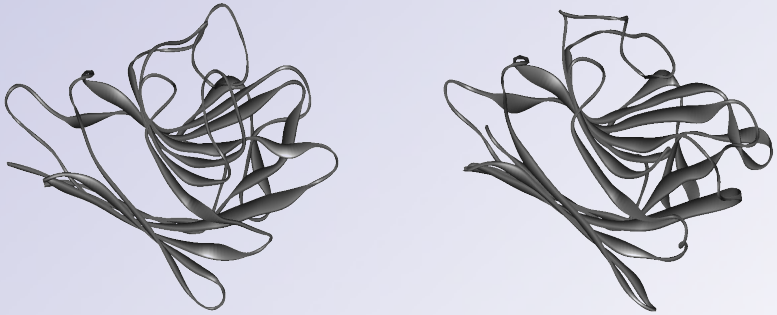
Search for similar 3D sites in proteins

Version 4.4-Boom

- Technology to compare surfaces and detect similarities between macromolecular structure
  - VERY fast: compare a binding site *vs* PDB sites: 10-30 minutes on a desktop PC
  - Comparison at the level of chemical groups (pharmacophoric points)
  - Patented technology
  - Detect convergent or divergent evolution



- Local similarities
- But different backbone (2 serine proteases)

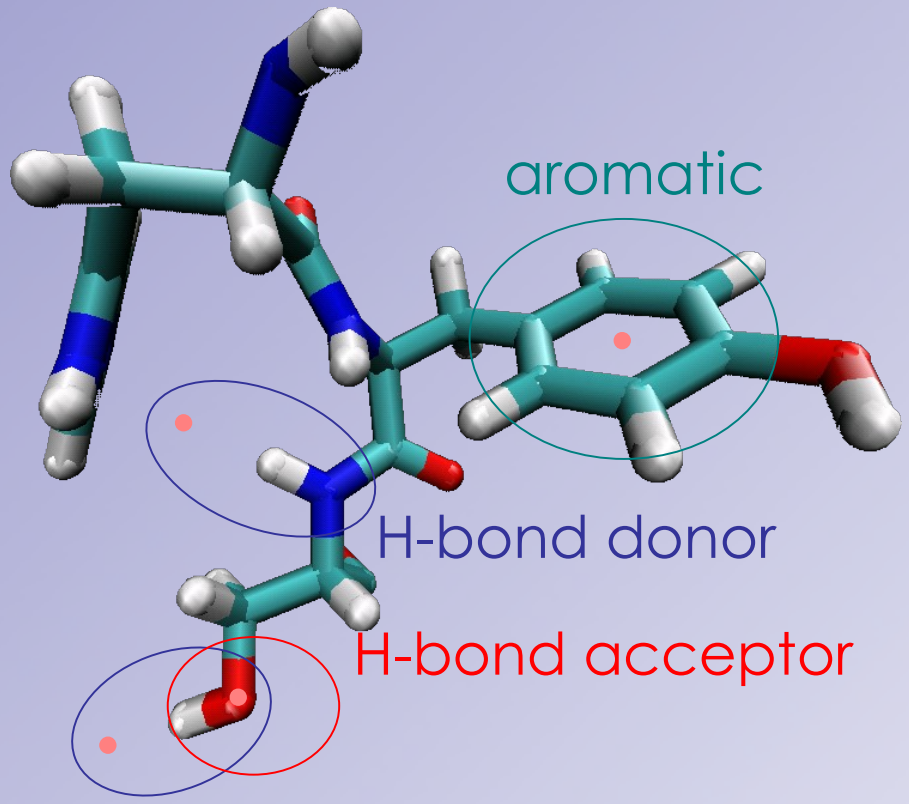


Similar backbone and sequence

- But different or antagonist functions

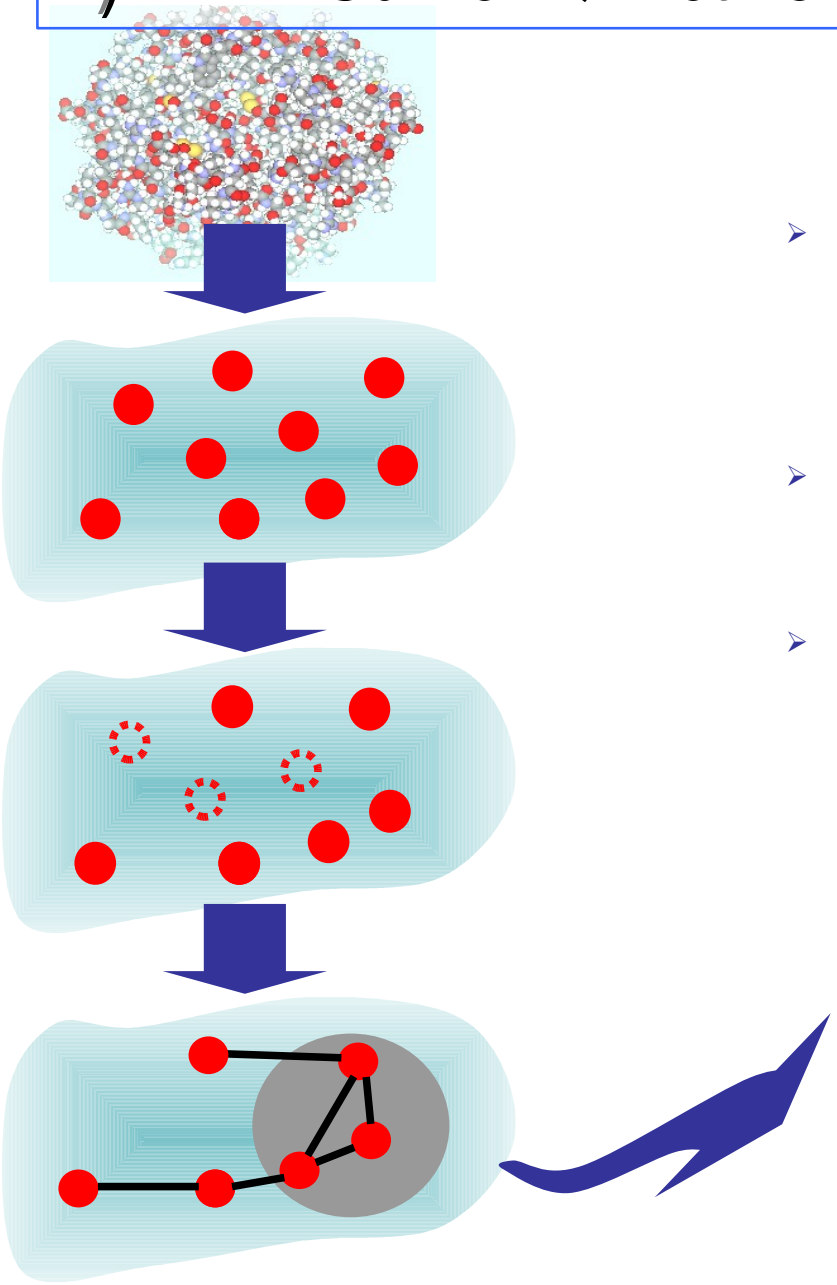
(Lectins 2PEL & 1IOA)

# 1) MED-SuMo ⇒ Chemical groups : definition

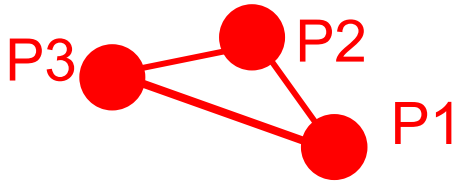



- Chemical group involve into binding interaction
  - set of atoms
  - depend of the local environment
- Position:
  - physical position
  - functional position
- Proper geometry

# 1) MED-SuMo ⇒ Active site: Conversion into Triplets



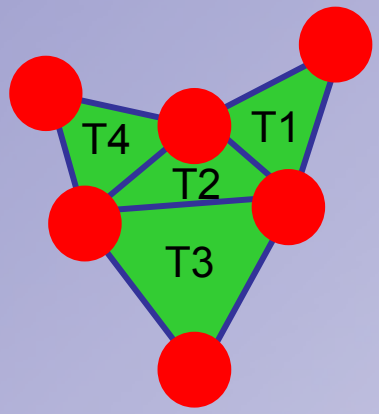
- Identification of MED-SuMo chemical groups
- Selection of chemical groups
- Generation of triplets

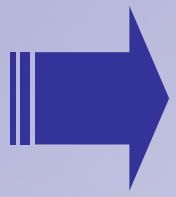


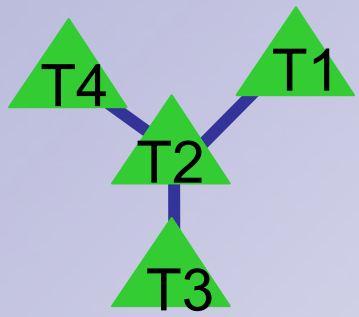

Formation of triplets

# 1) MED-SuMo ⇒ Molecules: Internal representation

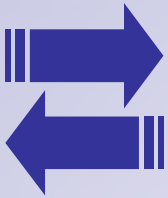
Chemical Groups






Graph mapping triplets of chemical groups



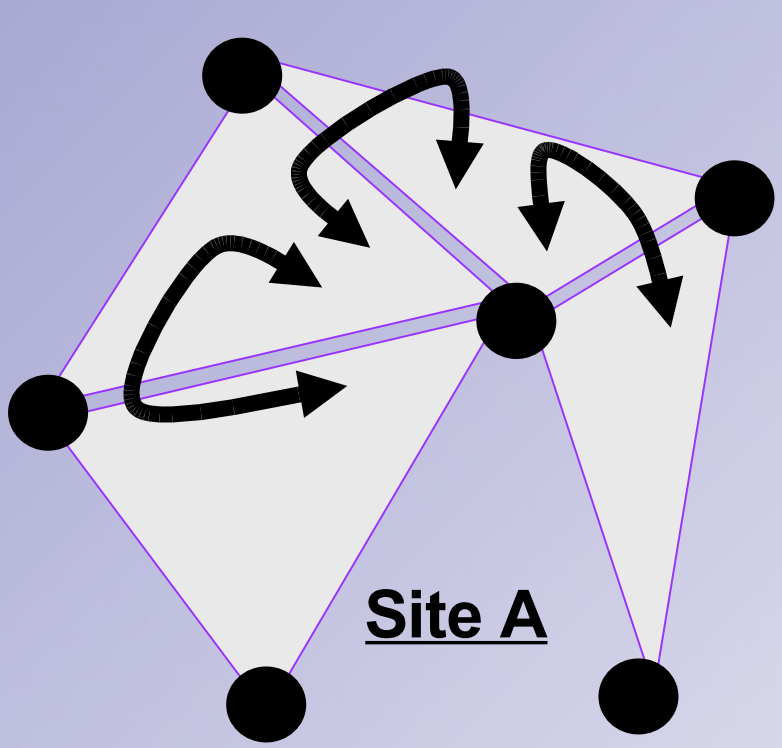
Database

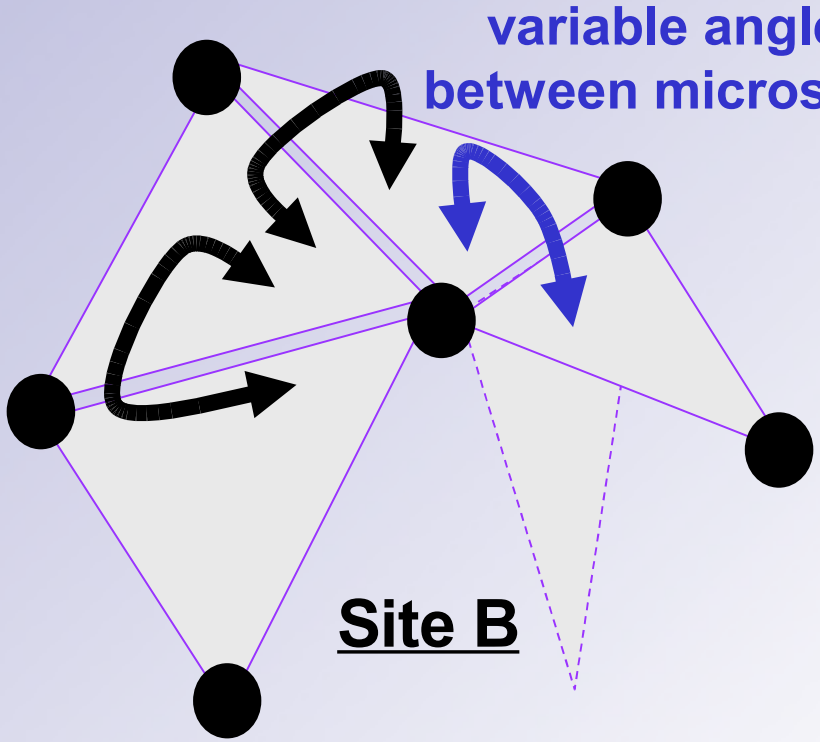
Comparison

- Same type of chemical groups
- Similar distances (between chemical groups)
- Similar burying level (atomic density)
- Similar orientation of chemical groups
- Similar shape of the local environment

# 1) MED-SuMo ⇒ Dealing with ligands flexibility

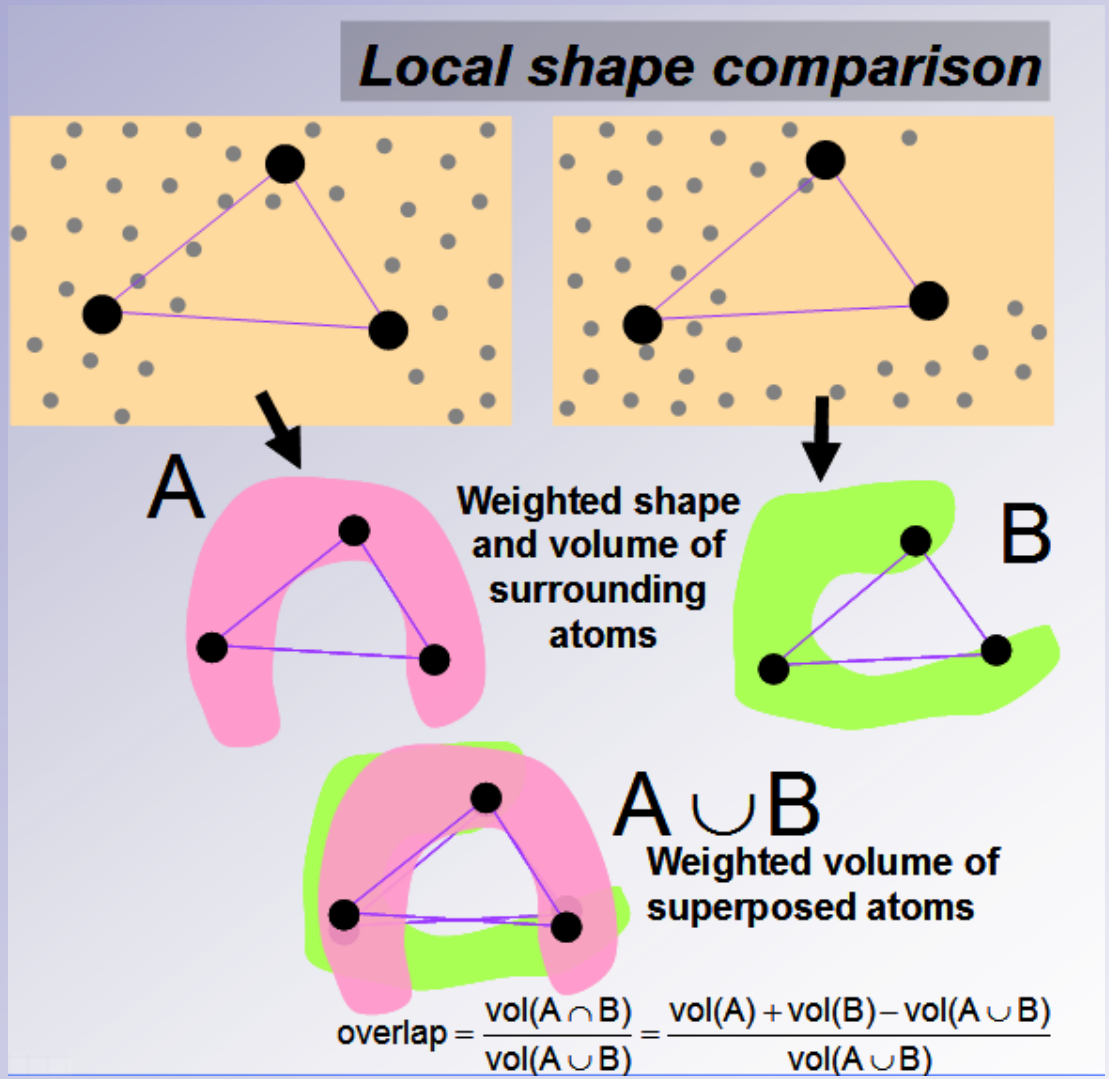
## Non superposable sites … but same ligand



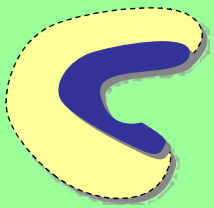
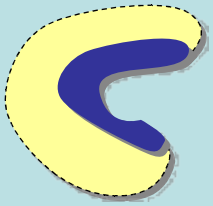




## Not a rigid superposition

# 1) MED-SuMo ⇒ Local shape comparison



# 1) MED-SuMo ⇒ Multiple applications

| Data Base / Query | Site  | Full  |
|---|---|---|
| Site  | • Site comparison for drug-design process<br>• Ligand prediction | • Enhance site characterization<br>• Activity and selectivity assesment |
| Full  | • Site detection | • Pre-characterization for protein-protein docking<br>• Template search for homology modeling |

# 1) MED-SuMo ⇒ Application to Kinases: Classifications from Naumann & Al.

3)

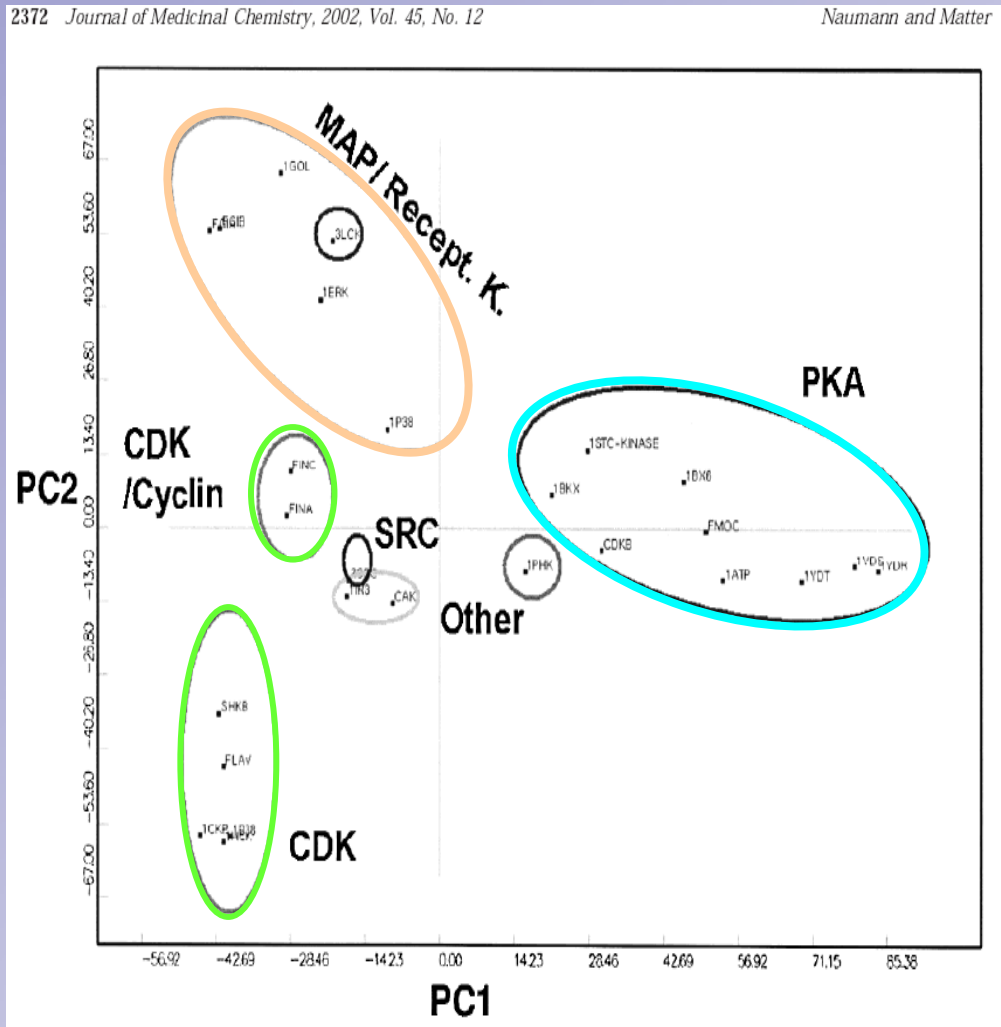



**Figure 4.** PCA score plot for analysis using all three different GRID probes (N1, O, and DRY) for this final analysis. The plot (target family landscape) illustrates the differences of several kinase families in the chemometrical space.

| Famille | Type | Structures | EC Number |
|---|---|---|---|
| Ser/Thr kinase | PKAc | 1ATP | 2.7.1.37 |
| | | 1BKX | 2.7.1.37 |
| | | 1BX6 | 2.7.1.37 |
| | | 1CDK | 2.7.1.37 |
| | | 1STC | 2.7.1.37 |
| | | 1YDR | 2.7.1.37 |
| | | 1YDS | 2.7.1.37 |
| | | 1YDT | 2.7.1.37 |
| | | 1FMO | 2.7.1.37 |
| | CDK2 | 1HCK | 2.7.1.37 |
| | | 1CKP | 2.7.1.37 |
| | | FINA | 2.7.1.- |
| | | FINC | 2.7.1.- |
| | | 1B38 | 2.7.1.37 |
| | Map Kinases | 1P38 | 2.7.1.- |
| | | 1GOL | 2.7.1.- |
| | | 1ERK | 2.7.1.- |
| | Caseine kinase | 1CSN | 2.7.1.- |
| Tyr kinase | Phosphorylase Kinase | 1PHK | 2.7.1.38 |
| | SRC | 2SRC | 2.7.1.112 |
| | SRC | 3LCK | 2.7.1.112 |
| | Recepteur de facteur de croissance | FGIA | 2.7.1.112 |
| | | FGIB | 2.7.1.112 |
| | Recepteur à l'insuline | 1IR3 | 2.7.1.112 |

# 1) MED-SuMo ⇒ Application to Kinases: MED-SuMo score to classify protein sites

## PKAc group (9 structures) – Ranking according to MED-SuMo score:

| run | 1ATP | 1BKX | 1CDK | 1BX6 | 1STC | 1YDT | 1YDR | 1YDS | 1FMO |
|---|---|---|---|---|---|---|---|---|---|
| | 1ATP | 1BKX | 1CDK | 1BX6 | 1STC | 1YDT | 1YDR | 1YDS | 1FMO |
| | 1CDK | 1FMO | 1FMO | 1ATP | 1FMO | 1YDR | 1YDT | 1YDR | 1CDK |
| | 1FMO | 1CDK | 1ATP | 1CDK | 1CDK | 1ATP | 1CDK | 1FMO | 1ATP |
| | 1YDT | 1YDT | 1YDR | 1YDR | 1YDT | 1CDK | 1FMO | 1ATP | 1YDR |
| | 1YDS | 1ATP | 1YDT | 1FMO | 1YDR | 1YDS | 1YDS | 1YDT | 1YDS |
| | 1BX6 | 1YDS | 1BX6 | 1YDT | 1PHK | 1BX6 | 1ATP | 1CDK | 1BX6 |
| | 1YDR | 1YDR | 1YDS | 1YDS | 1YDS | 1FMO | 1BX6 | 1BX6 | 1BKX |
| | 1BKX | 1BX6 | 1BKX | 1BKX | 1BX6 | 1BKX | 1PHK | 1BKX | 1YDT |
| | 1PHK | 1STC | 1STC | 1STC | 1ATP | 1PHK | 1BKX | 1PHK | 1STC |
| | 1STC | 1PHK | 1PHK | FGIA | 1BKX | 1STC | 1STC | 1STC | 1PHK |
| | 1CSN | 1CSN | 1CSN | 1HCK | 1HCK | 1CSN | 1CSN | FGIA | 1CSN |
| Structures | FGIA | FINC | FGIA | 1PHK | 1CKP | FINC | 1ERK | FGIB | 1HCK |
| | FINC | 1HCK | FINC | 1CSN | 1B38 | 1HCK | FINC | FINC | FINA |

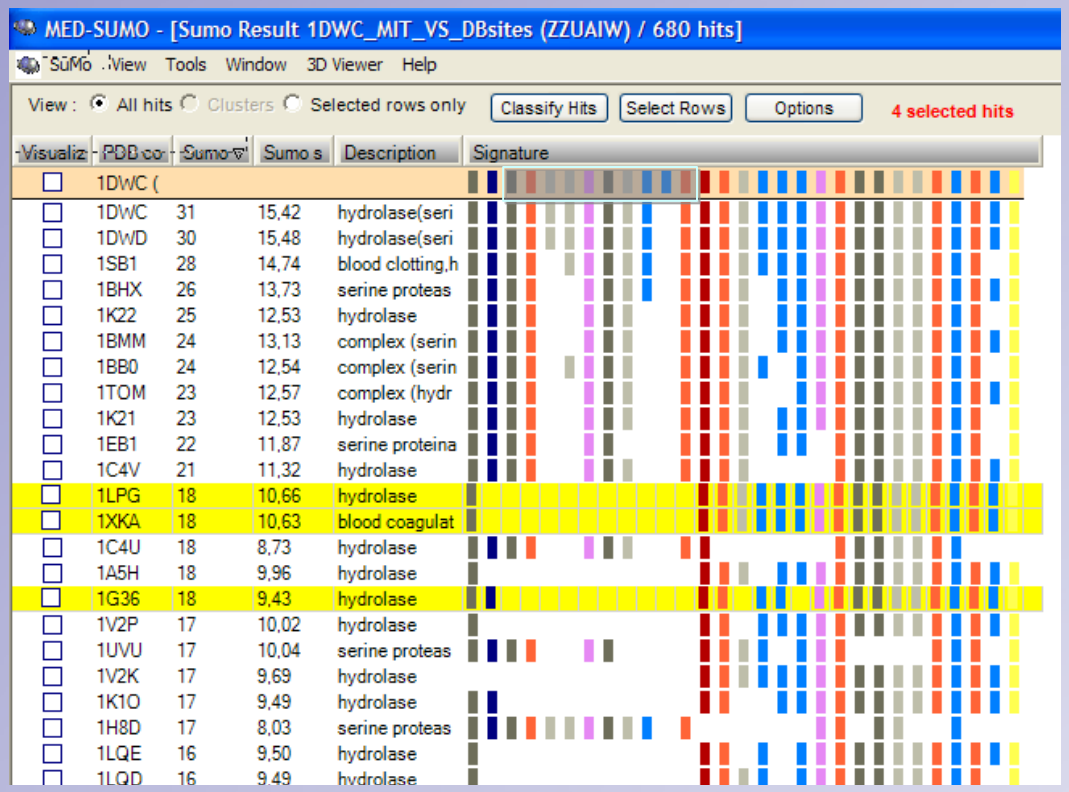
## CDK2 group (4 structures) – Ranking according to MED-SuMo score:

| run | 1HCK | 1CKP | 1FINA | FINC | 1B38 |
|---|---|---|---|---|---|
| | 1HCK | 1CKP | FINA | FINC | 1B38 |
| | 1B38 | 1HCK | FINC | FINA | 1HCK |
| | 1CKP | 1B38 | 1HCK | 1CKP | 1CKP |
| | FINA | FINC | 1CKP | 1HCK | FINA |
| | FINC | FINA | 1B38 | 1B38 | FINC |
| | 2SRC | FGIA | 2SRC | 3LCK | FGIB |
| | 1STC | FGIB | 3LCK | 2SRC | FGIA |
| | 1GOL | 1STC | 1YDR | 1YDR | 2SRC |
| | FGIA | 2SRC | FGIB | 1YDT | 1STC |

**Legends**

- CDK2
- MAP kinase
- SRC
- PKAc

# 1) MED-SuMo ⇒ Query = α-thrombine (1DWC) vs Sites

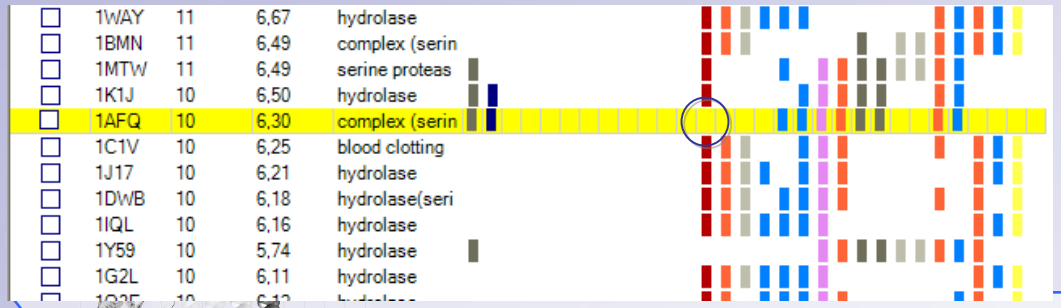



- A 60 loop (9 residues) specific to alpha-thrombin
- ASP189 -> SER189 for chymotrypsin

} Factor Xa

} Trypsin

Activity and Selectivity assessment




} Chymotrypsin

-1- **MED-SuMo** → how to compare active sites ; affinity and selectivity assessment

-2- Cheminformatics Applications: fragment based drug design

-3- Demo

# 2) Applications ⇒ Hybridation of ligands

Hybridation of known ligands bound to a common target can be described as a two steps protocol [1]

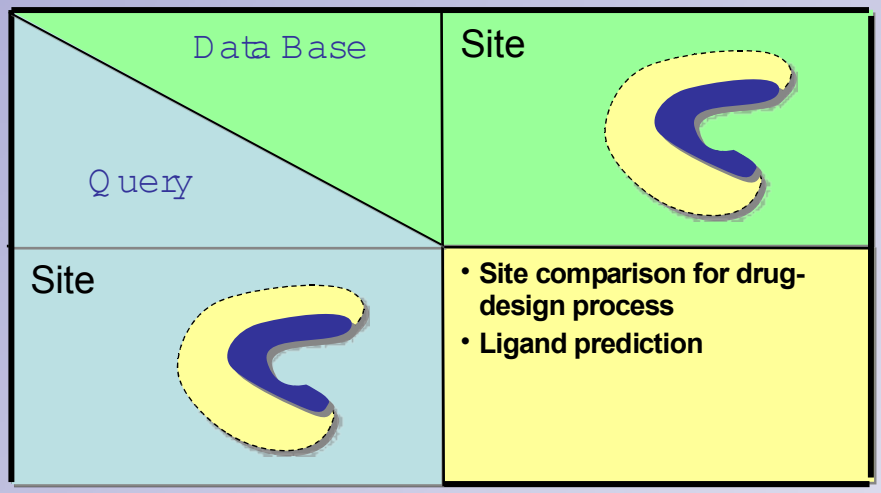
- "By hand" overlay of the p38 MAP kinases 1DI9, 1A9U and 1BMK
- Ligands hybridation with a matching bond algorithm

MED-SuMo fragment based drug design: common & related target

- Overlay: Site vs Site
- Fragments generation + ligands hybridation within a user friendly Graphical User Interface dedicated to the end-user: the chemist, the biologist, the molecular modeler

1 - Pierce AC, Rao G, Bemis GW "BREED: Generating novel inhibitors through hybridization of known ligands. Application to CDK2, p38, and HIV protease." J Med Chem. 2004 May 20;47(11):2768-75

# 2) Applications ⇒ Site vs Site mode to compare ligands

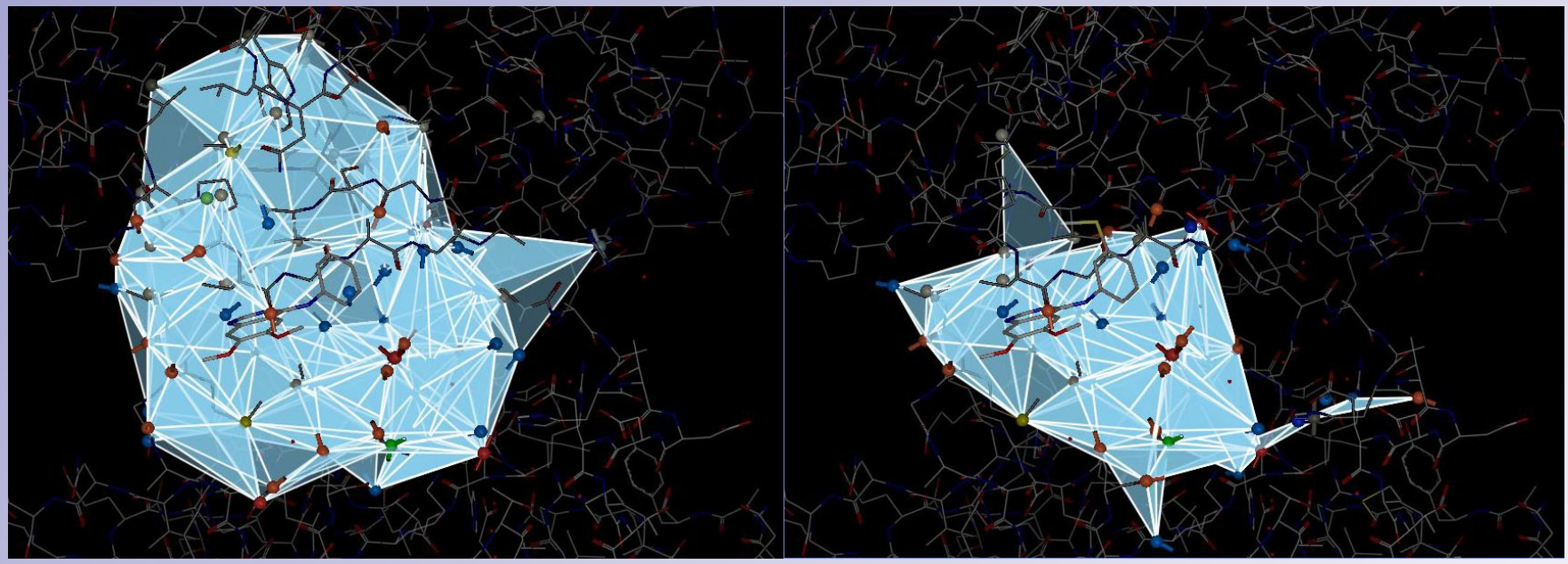



## Site are defined by the neighborhood of the ligand:

➔ SuMo objects are likely to be involved in binding

➔ A different binding mode is described as a different site

# 2) Applications ⇒ Optimize the Site query



(1) Large: 92 SuMo objects, 752 triplets
Diversity of Hits is higher
But false positive rate is higher

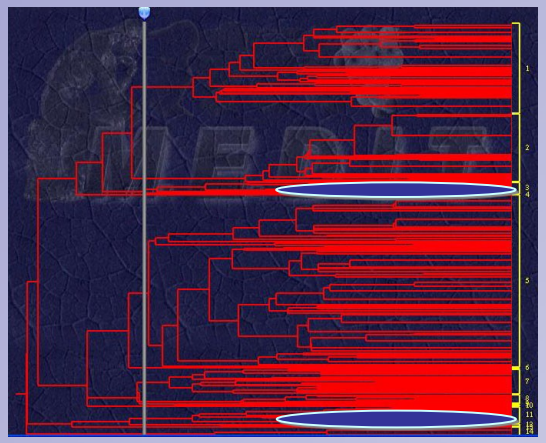
(2) Optimized: 59 objects, 379 triplets

Easier to analyse and to report results

Results of (1) are browsed to identify key residues: most common and/or in the best hits

THR106, LEU108, MET109, ASP150, SER154, ASN155, LEU167, ASP168, PHE169, HIS174

# 2) Applications ⇒ Clustering of 831active sites,check homogenity, resize clusters

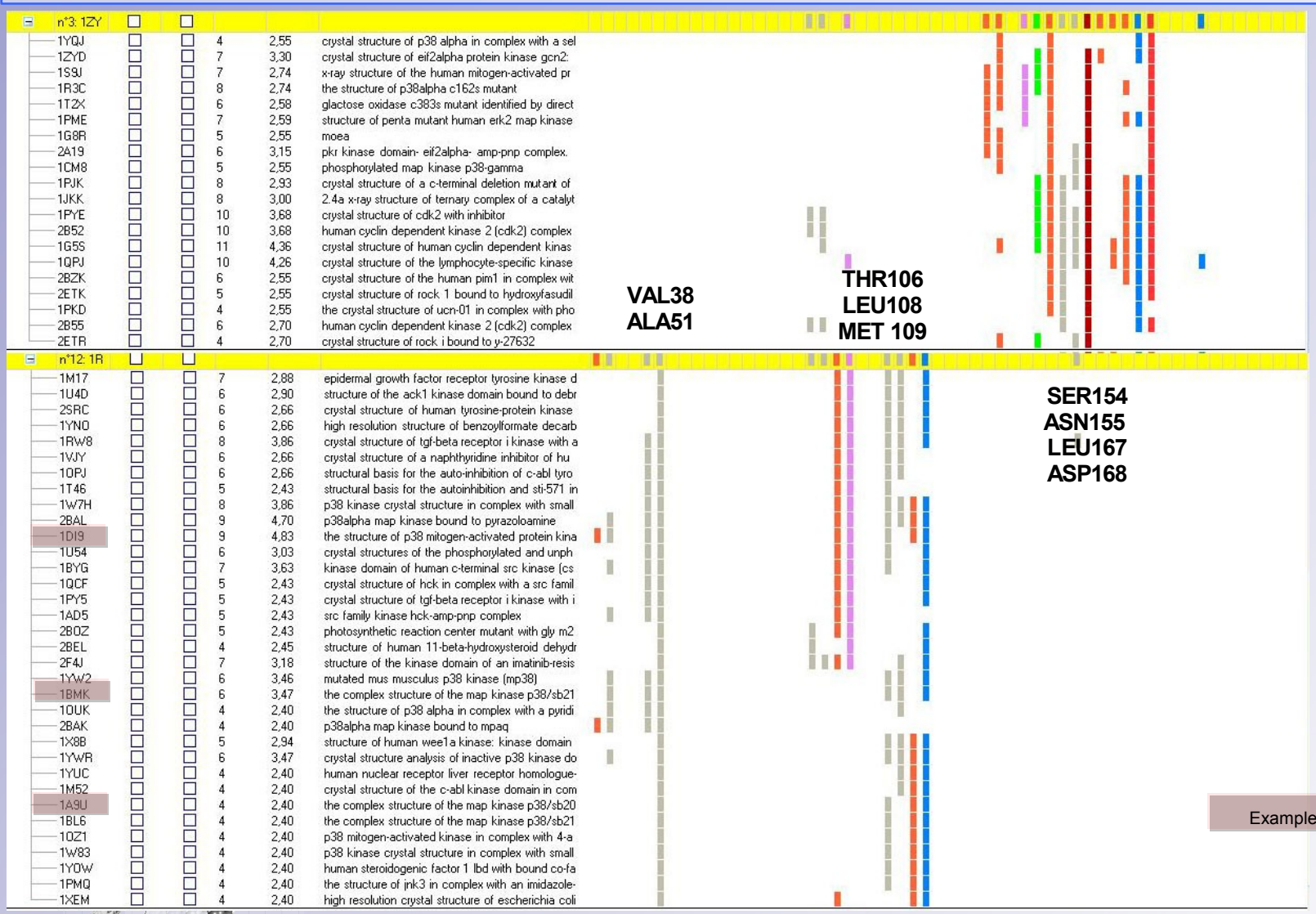


2 small and distants clusters contain kinases


20 Hits in cluster 3

34 Hits in cluster 12

# 2) Applications ⇒ Analysis of the clusters

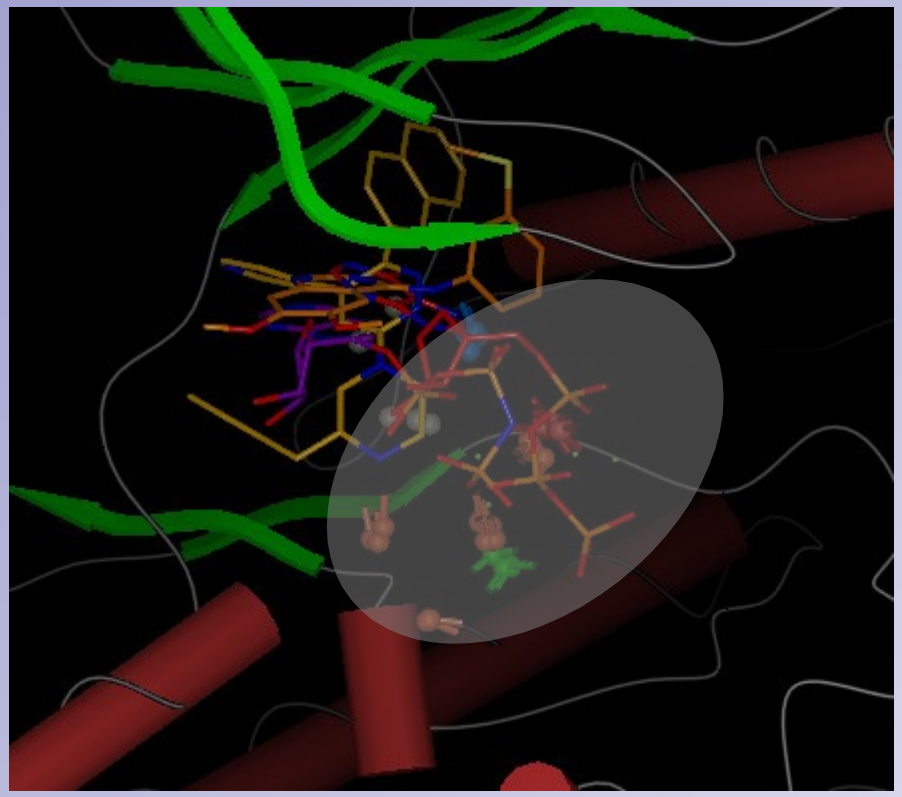



20 Hits in cluster 3

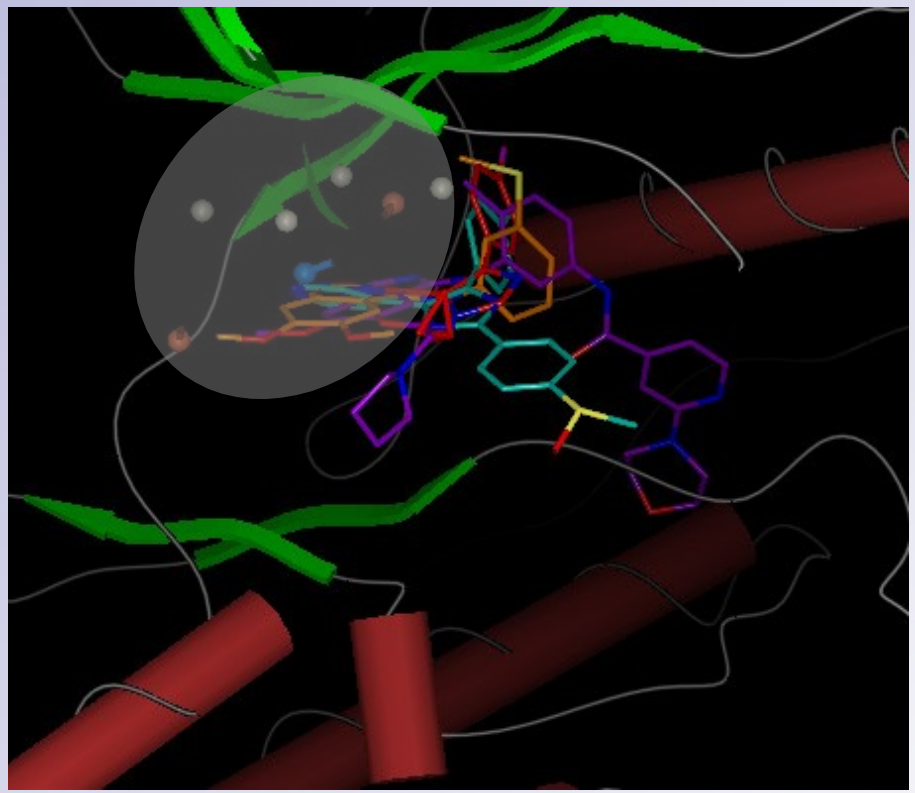
34 Hits in cluster 12

Examples from Pierce et al.

# 2) Applications ⟹ Cluster 3 and Cluster 12 correspond to 2 subpockets in the binding pocket



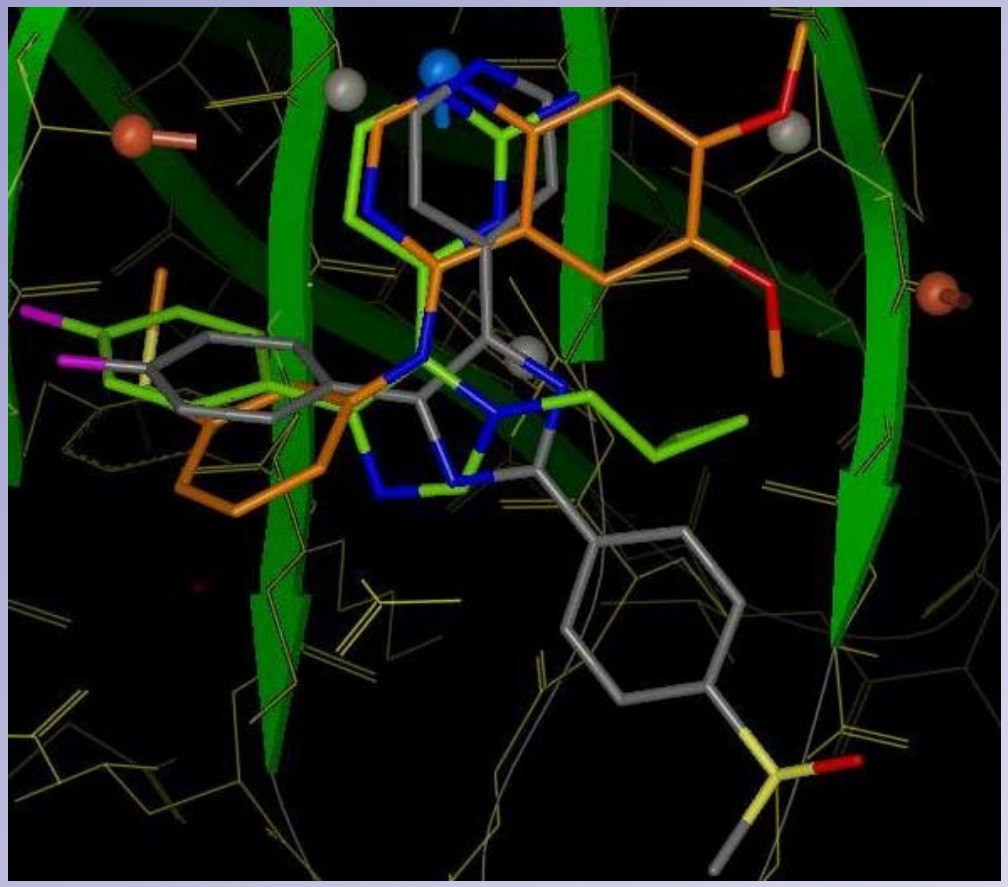
Ligands from Cluster 3

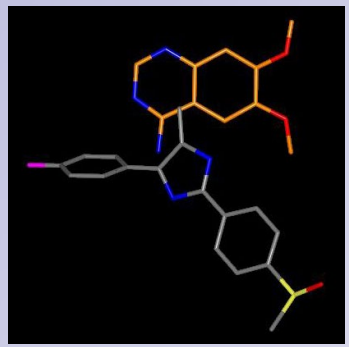


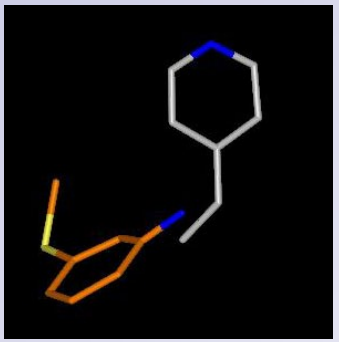
Ligands from Cluster 12

Fragments can be extracted in each subpockets

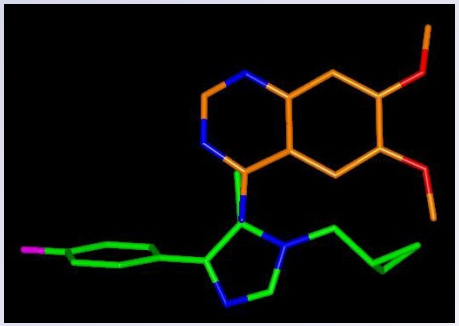
# 2) Applications ⇒ Ligands overlay and hybridation







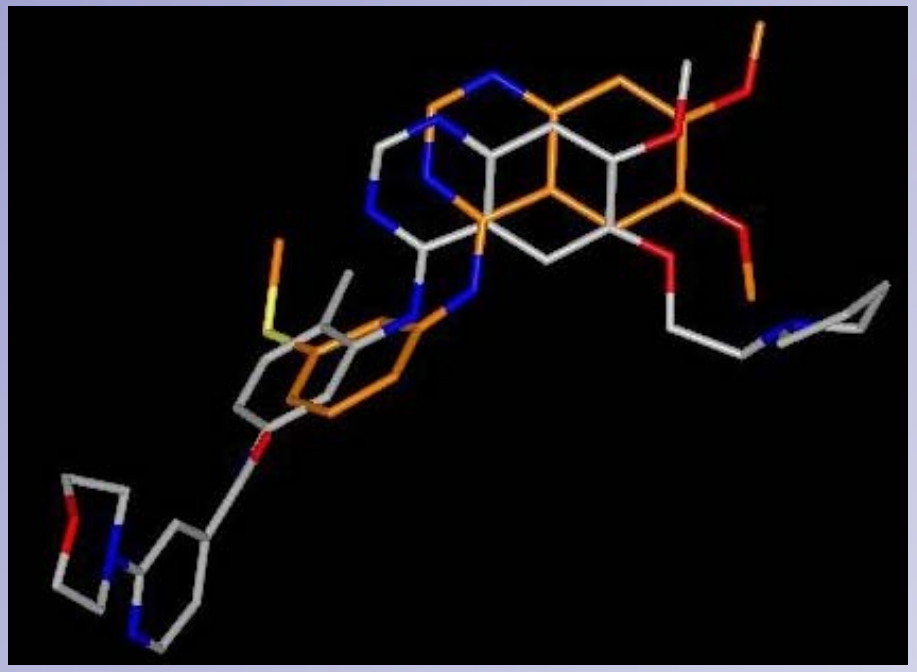
Couples of fragments with matching bonds (automatized in Pierce et al.)
Here: extracted from Cluster 12

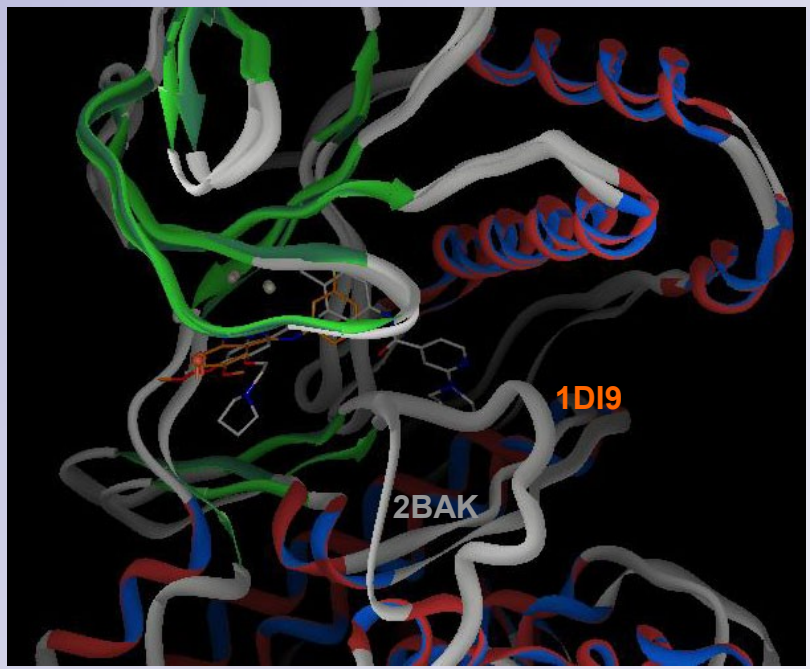


**Legend**: protein=p38 MAP kinase (1DI9)

| | |
|---|---|
| 1DI9 p38 ligand | 4-[3-METHYLSULFANYLANILINO]-6,7-DIMETHOXYQUINAZOLINE |
| 1A9U p38 ligand | 4-[5-(4-FLUORO-PHENYL)-2-(4-METHANESULFINYL- PHENYL)-3H-IMIDAZOL-4-YL]-PYRIDINE |
| 1BMK p38 ligand | 4-(FLUOROPHENYL)-1-CYCLOPROPYLMETHYL-5-(2- AMINO-4-PYRIMIDINYL)IMIDAZOLE |

# 2) Applications ⇒ Common target and steric constraints






Common target (p38 MAP kinase) : all fragments are kept

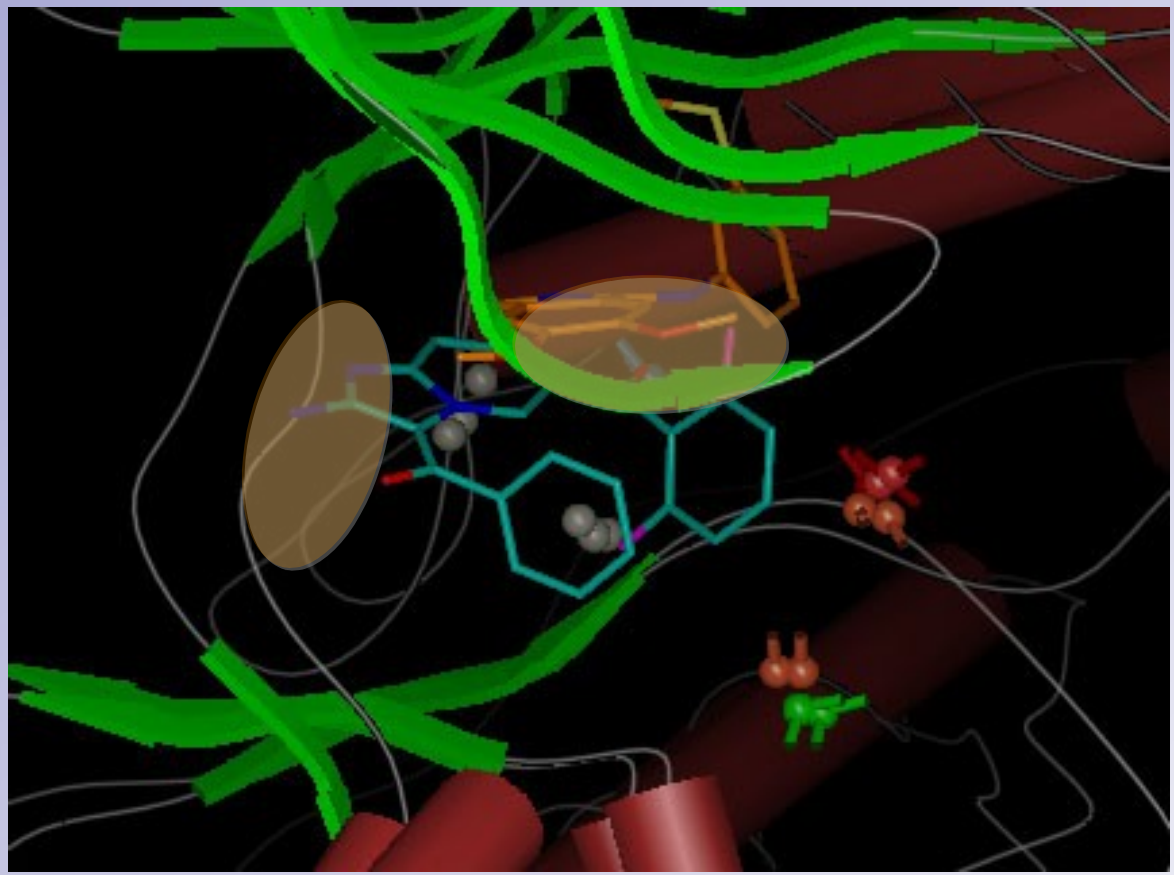
➔ Related target (other kinase): rank fragments

**Legend**:

**1DI9 p38 ligand** **4-[3-METHYLSULFANYLANILINO]-6,7-DIMETHOXYQUINAZOLINE**

2BAK p38 ligand MPAQ = (S)-[(3-{[(6R)-7-METHOXY-6-(2-PYRROLIDIN- 1-YLETHOXY)-5,6-DIHYDROQUINAZOLIN-4-YL]AMINO}- 4-METHYLPHENYL)AMINO](2-MORPHOLIN-4-YLPYRIDIN- 4-YL)METHANOL

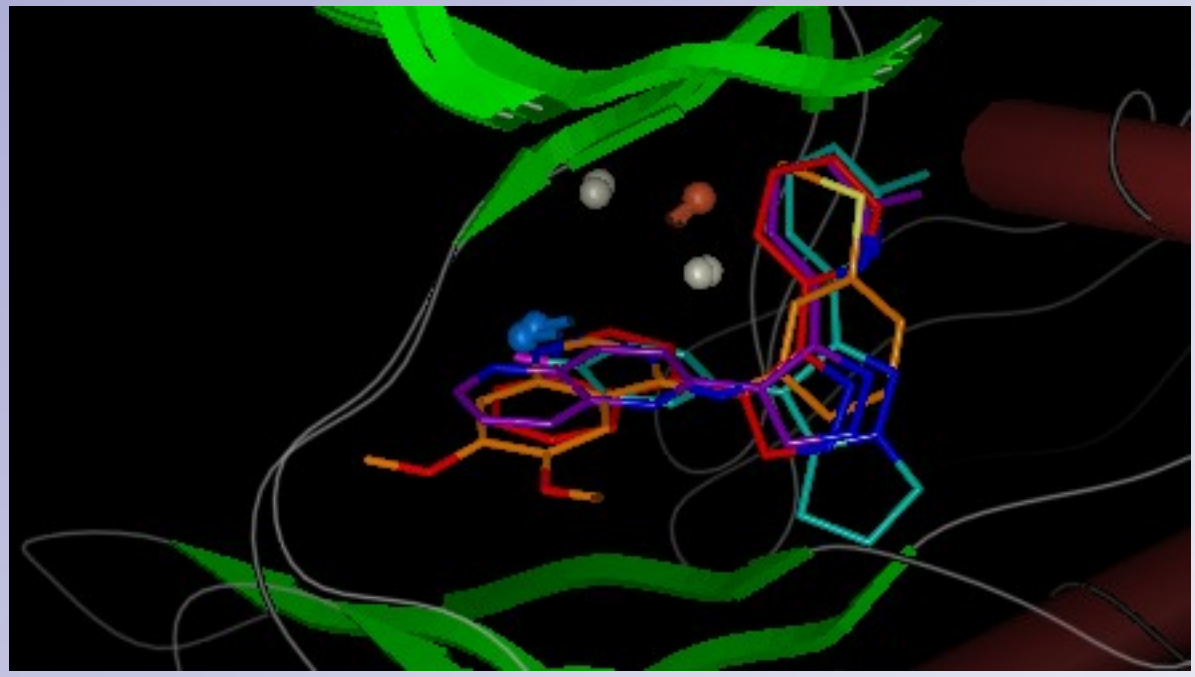
# 2) Applications ⇒ CDK2 Hits



Cluster 3 contains 5 CDK2 Hits (none in cluster 12) : 2B55, 2B52, 1G5S, 1PYE and 1PKD

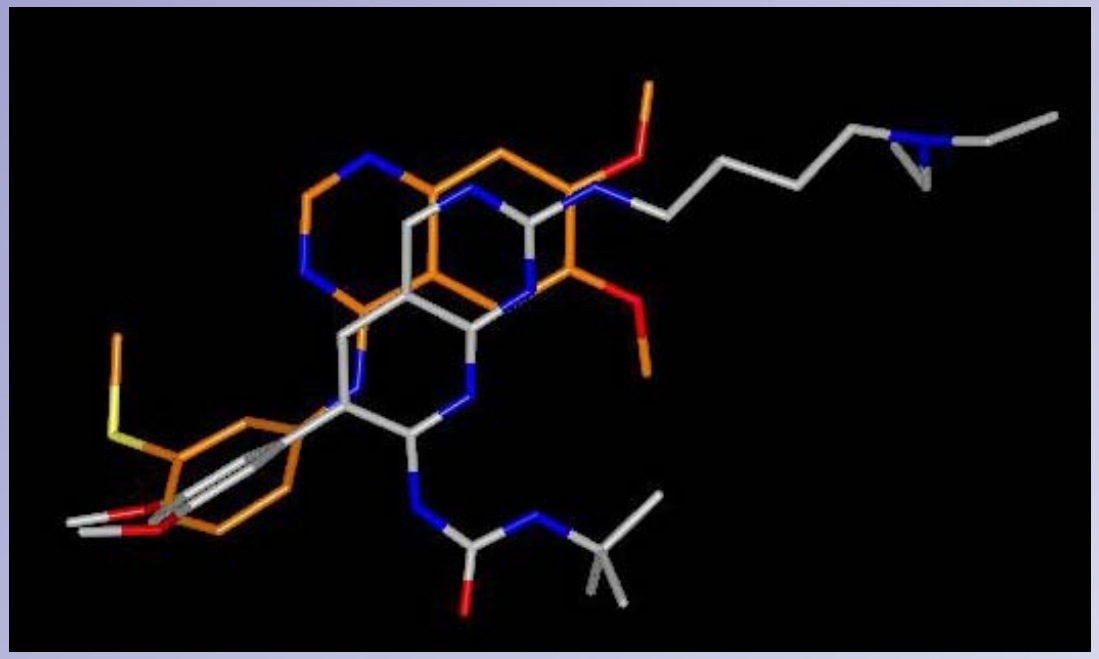
Example of a  case of picking up a fragment in a similar active site

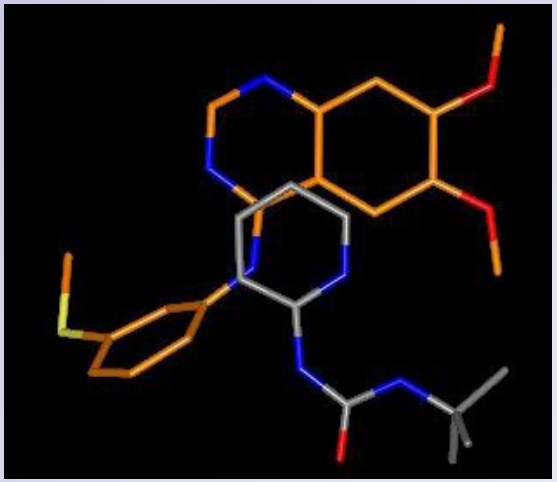
# 2) Applications ⇒ TGFβ Hits



Cluster 12 contains 3 TGFβ Hits (none in cluster 3) : 1VJY, 1RW8 and 1PY5
Example of an easy case of picking up a fragment in a similar active site

# 2) Applications ⇒ Case of matching bonds in cycles

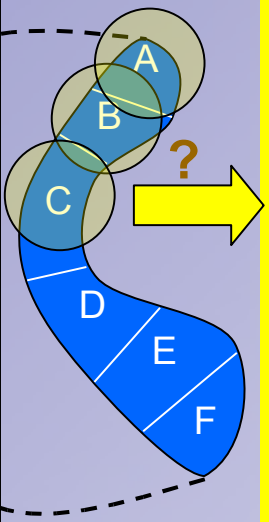


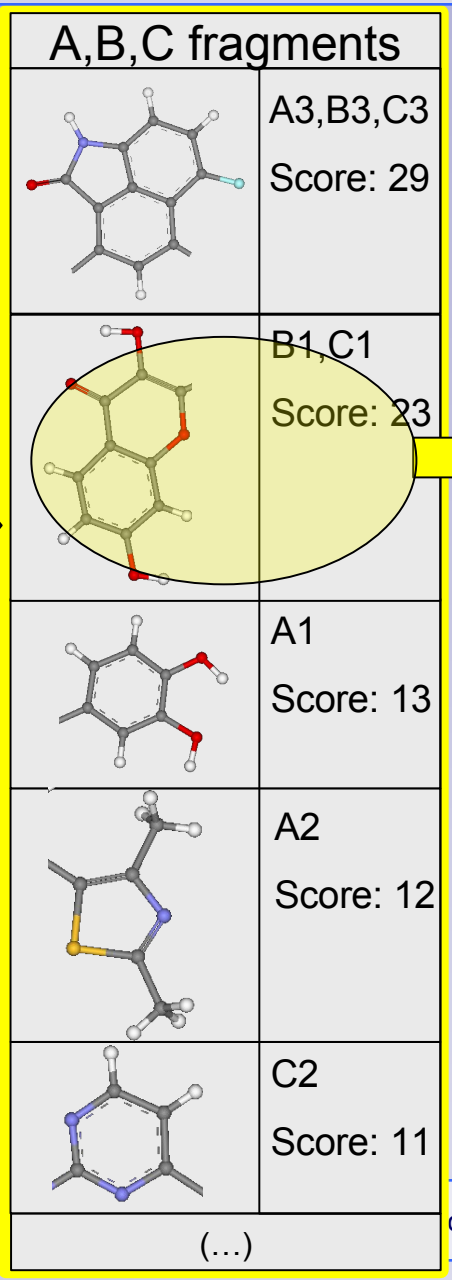


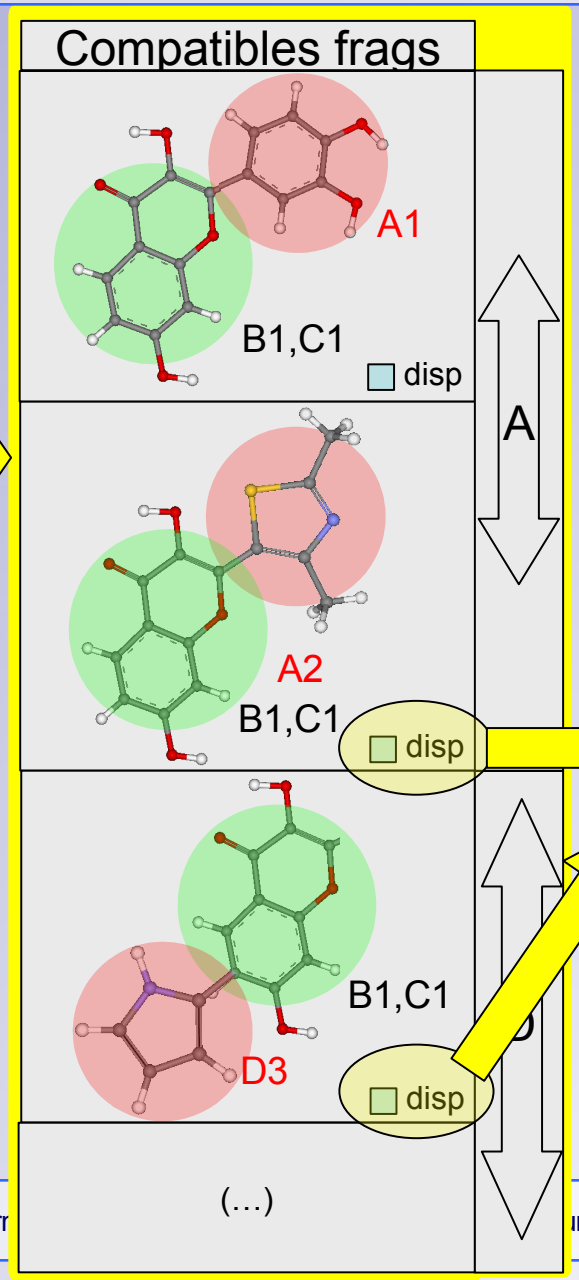
Complicated overlay cases are probably best hybridized « by hand » directy by the end-user

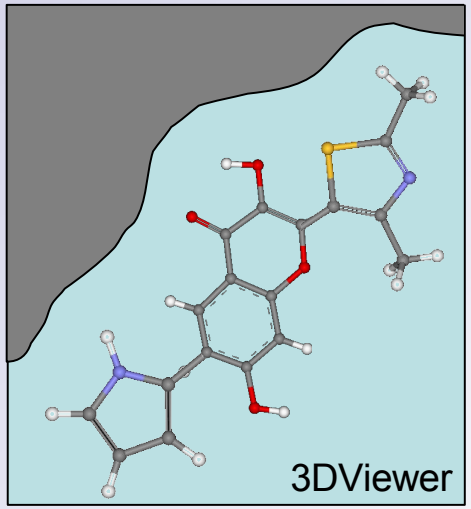
**Legend**:
**1DI9 p38 ligand** **4-[3-METHYLSULFANYLANILINO]-6,7-DIMETHOXYQUINAZOLINE**
**2BAK p38 ligand** **CRYSTAL STRUCTURE OF THE TYROSINE KINASE DOMAIN OF FGF RECEPTOR 1 IN COMPLEX WITH 1-TERT-BUTYL-3-[6-(3,5-DIMETHOXY-PHENYL)- 2-(4-DIETHYLAMINO-BUTYLAMINO)-PYRIDO[2,3- D]PYRIMIDIN-7-YL]-UREA**

# 2) Applications ⇒ Fragment Hybridization - chemist expert tool



-1- **MED-SuMo** → how to compare active sites ; affinity and selectivity assessment

-2- Cheminformatics Applications: fragment based drug design

-3- Demo

# 3) Demo ⇒ Steps

- Launch query:
  - Launch query with 1DI9 active site vs Active sites DB
  - Others modes
- Results analysis:
  - Open results
  - Browse results : protein and ligands


***FASTER TO BETTER MOLECULAR MODELING PREDICTIONS***

# THANKS

**To contact M.E.D.I.T. :** **2 Rue du Belvédère, Palaiseau, France**

**Tel: +33 (0)6 855 66 722**

http://www.medit.fr **&** info@medit.fr