

A fragment-based computational protocol at PDB scale - Application to lead-optimization of DFG-out kinase inhibitors

F. Moriaud¹, T. Henry¹, S.A. Adcock¹, A.M. Vorotyntsev¹, L. Martin¹, O. Doppelt^{1,2}, A.G. De Brevern², F. Delfaud¹

¹MEDIT, Palaiseau, France ²EBGM University Paris 7, Paris, France

Fragment-based drug design has emerged in the last decade and has become an established paradigm at many pharmaceutical companies. This exciting field has been recently reviewed [1]. Obtaining structural information on the fragment complexed to the protein target is a key step and also a major limitation to the number and types of target that are amenable to fragment-based approaches. Therefore, computational methods are needed to mine efficiently all the available 3D structures of ligands complexed to proteins, both treated as a whole and as smaller fragments to increase the likelihood of fragment hopping from one target to another.

MED-SuMo [2,3], a target based drug design tool, offers a procedure to adequately characterize the protein binding site. This tool is based on the identification of local shape and 3D Surface Chemical Features similarities in the target binding site with other proteins (with their co-crystallized ligands). MED-SuMo uses the binding site of the target as a query to search either the whole Protein Data Bank (or any corporate protein structure databank) for all the binding sites that display a local match with the query. This valuable information can then be used to identify which residues of the binding site are potentially important for ligand binding affinity and selectivity. As the similar binding sites are overlaid, the co-crystallized ligands are aligned and are therefore a starting material for ligand hybridization. Among the hundreds of overlaid binding sites generated by MED-SuMo, we found the protein-ligand complexes overlaid by Pierce et al. [4] as their starting material for ligand hybridization in the BREED method. Interestingly, they found relevant combinations of ligands starting from only a few protein-ligand complexes structures and we believe that the output of MED-SuMo is a very promising input for automatic methods like BREED.

In this work, we've worked on a fragment database derived from the PDB: each pdb file is converted into one or more pdb files containing a single fragment as ligand. Fragments are converted to MED-Portions which are fragments annotated with protein 3D environnement and dummy bonds. We've used MED-SuMo to query and mine the Protein's Surface Chemical Functions surrounding MED-Portions, seeking similarities with the kinase of interest (i.e. Vegfr DFGout, pdb code 2oh4, ligand code GIG) and collecting a library of 1129 unique MED-Portions positioned in the vegfr's active site and annotated with the counts of contacts and h-bonds. MED-Portions can be used to design novel ligand scaffolds (lead generation) or to optimize attachments on a fixed scaffold (lead optimization). Here we present the optimization of a substructure (i.e. phenylamide) of the GIG ligand to find others DFGout ligands. The 3D hybridisation in 5 iterations of the phenylamide moiety with 1129 fragments suggested by our MED-Hybridiser protocol leads to 22824 molecules. In this list, we identified 3585 different scaffolds, 298 are in PubChem, 46 in the PDB attesting of the diversity and quality of those generated molecules. 25 are marked as active on protein kinase in PubChem bioassay.

1. Hajduk PJ, Greer J. "A decade of fragment-based drug design: strategic advances and lessons learned" *Nat Rev Drug Discov.* **2007** Mar; 6(3):211-9.
2. Jambon M, Imberty A, Deléage G, Geourjon C "A new bioinformatic approach to detect common 3D sites in protein structures" *PROTEINS* **2003** Structure, Function, and Genetics 52:137-145
3. Jambon M, Andrieu O, Combet C, Deléage G, Delfaud F, Geourjon C « The SuMo server : 3D search for protein functional sites" *Bioinformatics.* **2005** Vol 21, n°20, 3929-3930
4. Pierce AC, 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