- **Title** : Knowledge Discovery in structural databases of small molecules
- **MSc thesis advisor** : Sergei Grudinin, sergei.grudinin@inria.fr, Phone: +33 4 38 78 16 91
- **Group leader** : Stephane Redon, stephane.redon@inria.fr
- **Laboratory** (+working place) : NANO-D, INRIA Rhone-Alpes Research Center

Minatec Campus 17 rue des Martyrs, 38054 Grenoble France, https://team.inria.fr/nano-d/

- **If a PhD is forseen** : yes

Internship presentation :

Structure-based drug design [1] relies on knowledge of the three dimensional structure of the biological target obtained through methods such as X-ray crystallography, which is used as a basis for designing new ligands by applying accepted principles of molecular recognition. The basic assumption underlying structure-based drug discovery is that a good ligand molecule should bind tightly to its target. Thus, one of the most important principles for designing or obtaining potential new ligands is to predict the binding affinity of a certain ligand to its target and use it as a criterion for selection. The binding affinity can be approximated with certain physics-based empirical functions or using knowledge–driven potentials [2, 3].

Internship objectives:

We have recently developed an isotropic (i.e., direction–independent) formulation for the potential function for interactions between proteins [3]. We would like to extend this formulation to a class of anisotropic (i.e., direction–dependent) potential functions.

The goals of the internship are:

• develop a pairwise anisotropic potential function able to describe various types of non-covalent interactions in small molecules, such as (i) halogen bonds, (ii) hydrogen bonds, and (iii) aromatic interactions

• deduce coefficients for the potential from the knowledge-base (i.e., PDBBind data set)

• study the effect of several angular basis sets, as well as several radial basis sets on the quality of the potential

• validate the potential on a set of structures of high resolution extracted from the Protein Data Bank and Cambridge Structural Database and also using a number of standard benchmarks and blind docking competitions (CAPRI, GPCR Dock, CSAR Dock).

Requirements :

We are looking for creative, passionate and hard-working individuals from applied math / computer science background with exceptional talent for computer science and mathematics and interest in computational chemistry / biophysics. Excellent oral, written and interpersonal communication skills are essential (working language will be English – knowledge of French is a plus). Good knowledge of C++ / signal processing / machine learning / orthogonal polynomials will be an asset.

References:

[1] T. L. Blundell, "Structure-based drug design," Nature, (1996), 384, 23.

[2] H. Gohlke, M. Hendlich, and G. Klebe, "Knowledge-based scoring function to predict protein-ligand interactions," Journal of molecular biology, (2000), 295, 337.

[3] P. Popov & S. Grudinin. Knowledge of Native Protein-Protein Interfaces Is Sufficient To Construct Predictive Models for the Selection of Binding Candidates. (2015). J Chem Inf Model, doi: acs.jcim.5b00372.