STUDENT INTERNSHIP

Topic: **Computational identification of off-target proteins for drug candidates**
Duration: 3 to 6 months start on February
Location: Oncodesign HQ – Dijon
Benefits: Monthly indemnity + meal Ticket

Our Company

OPM is a technological company specialized in precision medicine. OPM's mission is to bring innovative therapeutic and diagnostic solutions to treat therapeutic resistance and metastasis evolution. The patient is at the center of our reflection, of our unique innovative model, and our investments. For OPM "our collective success is paramount", there can be no value creation without exchange, without dialogue. The value creation resulting for us from reciprocity, i.e. balanced and fair exchanges at all levels, whether between internal collaborators, or with our partners, therapists, patients, experts and investors.

Context

Obtaining the structure of a protein is a challenge: experimental methods such as x-rays are expensive, laborious and it is not always possible to crystallize the protein. On the other hand, considering that proteins can be composed of hundreds of amino acids, generating an algorithm capable of predicting the structure of a protein is a rather a complex task. Proteins play a fundamental role in living beings and are the main target for therapeutic molecules. Thus, the folding problem (how to obtain the structure of a protein from its sequence) has occupied the minds of researchers for most of the 20th century.

AlphaFold2 won the main competition for protein structure prediction (CASP14) in 2020. AlphaFold2's predictions were considered to be almost at the level of those determined experimentally. DeepMind has recently made both the code and the model available on GitHub as open source, allowing the community to be able to use the model both for the prediction of structures from an amino acid sequence and to incorporate it into other models for other applications.

The possibility of accurately predicting the structure of a protein opens up different applications (from the possibility of designing new enzymes for the food industry, to nanotechnology for medicine). The pharmaceutical industry and especially the drug discovery field is a domain where the greatest effects are expected. Indeed, most approved drugs are small molecules and biologics that interact with a protein. Typically for small molecules, having identified a target of interest and the structure of a protein, molecular modeling could be used to design virtual compounds that can bind to the protein's active site and modulate its function.

To date, only 10 % of drug candidates make it through the clinical trial stages and reach the market. The main reason for the failure of clinical trials is safety. While in general a drug candidate is selected to have high affinity for its target, it could potentially bind to other targets (off-targets) resulting in secondary effects. Indeed, drugs often have off-targets that lead to unwanted effects that can be serious. This is why it is important to identify potential off-targets before a molecule enters the clinical trial phase, which can cost up to a billion dollars.

One of the most important class of target is the kinases protein family. Kinase is an enzyme that catalyzes the transfer of a phosphate group to a specific substrate. This mechanism has different functions and it involved in fundamental process, which can be often dysregulated in cancer. There are known 500 kinases, while they can be expressed in different sites they have a similar structure with an ATP binding domain. OPM has developed a class of molecules that are highly specific for kinases, flat molecules called macrocycles and has developed a specific platform called Nanocyclix®. In most cases, off-targets are proteins that display similarity in the region of the ATP-binding domain to the active site of the protein of interest. The aim of this training is to develop an algorithm that can identify potential off-target sites and develop a similarity scoring function.

The internship candidate will use an algorithm based on AlphaFold2 to be able to identify potential off-target proteins.

Missions & activities of the internship

- Build an algorithm based on AlphaFold2 source code to generate embedding representation of protein kinases active sites.
- Testing potential other algorithms as Rosetta Fold
- Identity potential candidates for off-target in a case study
- Modeling of the active site and the interaction of small molecules

Keywords
Python, structural biology, deep learning

Student expected background/Knowledge
M2 or last year of engineer school with specialty/knowledge in Computer Science / Bioinformatics / Structural Biology/Statistics Biology with knowledge in programming (R / Python).
Docking knowledge, working with computer clusters are a plus

References
1. Jumper et al., 2021 **Highly accurate protein structure prediction with AlphaFold**. Nature
2. Baek et al; 2021. **Accurate prediction of protein structures and interactions using a three-track neural network**. Science

How apply?
Contact: Thierry Billoué – Chief Human Resources Officer – Oncodesign Precision Medicine
Send your application (resume & motivation letter) under ref "ComputID" to tbilloue@oncodesign.com