Master II
Internship Subject

Title: Modelling and design of small molecules for antiviral purposes: application to SARS-CoV-2 protease M^{pro} by development of a chemoinformatics search strategy, virtual screening and affinity assessment method.

Proposed by : Arnaud BLONDEL, Dynamics & Drug Design Group (3D), Structural Bioinformatics Unit, UMR3528, Institut Pasteur, 28 rue du Dr Roux, 75015 Paris.
Telephone: 01.40.61.36.93

E-mail address: ablondel@pasteur.fr

Keywords: Structural Bioinformatics // Chemoinformatics and Pharmacophore Filtering // Docking and Screening // Mechanics & Molecular Dynamics // Affinity Assessment

Description of the subject:
Cleavage of the polyproteins encoded by the beta-coronavirus genome by the viral proteases M^{pro} (nsp5) and PL^{pro} (nsp3) is required to produce the virus proteins. Thus, these proteases represent attractive targets for broad-spectrum drug development because they are essential to the virus life cycle and highly conserved among beta-coronaviruses. In particular, the nsp5 of SARS-CoV-2 and SARS-CoV share 95% amino acid sequence similarity. Protease inhibitors have already been shown to be effective in the treatment of HIV and HCV infections, and have become an important component of standard treatments.

Keys were identified in collaboration with N.Naffack and F.Agou (I. Pasteur) using a test on cells carrying a protease-nanoluciferase construct that is quenched by protease cleavage, and thus luminous in the presence of active inhibitor (gain of function). These hits were then validated by a virus infection assay on cells expressing the ACE2 receptor (IP Screening Platform, PF-CCB). These hits were generated by a relatively narrow *in silico* selection (359 compounds) from the campus chemical library (~15,000 compounds, PF-CCB, I. Pasteur). This first relatively promising set will be used to analyze the determinants of ligand anchoring: chemical groups, positions, etc.. This will allow the development of a pre-filtering/screening strategy that will include several channels targeting higher or lower degrees of similarity with the first hits in order to identify the most plausible candidates possible while expanding the chemical space searched to allow the identification of novel molecules. A large library (7.6 million compounds / 42 million tautomers - conformers) will then be used to identify the most promising candidates that will be selected for testing. This project benefits from collaborations within the Pasteur Institute for the realization of the tests and some compounds could be submitted to co-crystallization with the protease to reinforce the approach.

Skills required: Good knowledge of computer tools, unix, programming, etc... Basics and foundations of structural biology and computational chemoinformatics.

Other Information
- This subject is a first step towards a thesis. Focused on the exercise of a Master II, it can, if necessary, be continued according to various applications/developments/opportunities. If necessary, this project and its developments can be discussed upon contact.
- Internship dates: adaptable from the beginning of 2022 until the end of July 2022 (Note: count 2-3 months for the setting up of the contract).

Bibliographic references:

- Chen, K.Y., Krischuns, T., Ortega Varga, L., ... Blondel, A. & Naffakh, N. (in preparation): A highly sensitive cell-based luciferase assay for high-throughput automated screening of SARS-CoV-2 nsp5/3CLpro inhibitors
- US Food and Drug Administration. Coronavirus (COVID-19) Drugs. https://www.fda.gov/drugs/emergency-preparedness-drugs/coronavirus-covid-19-drugs.
- Zhang L, Lin D, Sun X, Curth U, Drosten C, Sauerhering L, et al. Crystal structure of SARS-CoV- 2 main protease provides a basis for design of improved alpha-ketoamide inhibitors. Science. 2020;368(6489):409-12. Epub 2020/03/22. doi: 10.1126/science.abb3405. PubMed PMID: 32198291; PubMed Central PMCID: PMCPMC7164518.
- Choi SG, Olivet J, Cassonnet P, Vidalain PO, Luck K, Lambourne L, et al. Maximizing binary interactome mapping with a minimal number of assays. Nat Commun. 2019;10(1):3907. Epub 2019/08/31. doi: 10.1038/s41467-019-11809-2. PubMed PMID: 31467278; PubMed Central PMCID: PMCPMC6715725.

Other related references:

- Monet, D., Desdouits, N., Nilges, M. and Blondel, A (2019) '*mkgridXf*: consistent identification of plausible binding sites despite the elusive nature of cavities and grooves in protein dynamics' J. Chem. Inf. Model. in 59(8) 3506-3518. https://doi.org/10.1021/acs.jcim.9b00103.
- Blondel, A. (2004): Ensemble Variance in Free Energy Calculations by Thermodynamic Integration: Theory, Optimal "Alchemical" Path, and Practical Solutions. J. Comp. Chem. 25, 985-993. https://doi.org/10.1002/jcc.20025
- Laine E, Goncalves C, Karst JC, Lesnard A, Rault S, Tang W-J, Malliavin TE, Ladant D, Blondel A. Use of allostery to identify inhibitors of calmodulin-induced activation of Bacillus anthracis edema factor. Proc Natl Acad Sci USA. 2010 Jun 22; 107:11277-11282.
- Aguiar Amaral, P., Autheman, D., Dias de Melo, G., Gouault, N., Cupif, J.F., Goyard, S., Dutra, P., Coatnoan, N., Cosson, A., Monet, D., Saul, F., Haouz, A., Uriac, P., Blondel, A. and Minoprio, P. (2018) "Designed mono- and di-covalent inhibitors trap modeled functional motions for Trypanosoma cruzi proline racemase in crystallography" PLOS Neglected Tropical Diseases, https://doi.org/10.1371/journal.pntd.0006853