

**MASTER « In Silico Drug Design »**  
**2ème année**

**PROPOSITION DE STAGE**  
**Année Universitaire 2021 / 2022**

A envoyer à Mme Pr Camproux  
[anne-claude.camproux@u-paris.fr](mailto:anne-claude.camproux@u-paris.fr)



**Etudiante : Ferial YALA**

**Nom du Responsable du Laboratoire ou de l'Entreprise: Xavier ROUCOU**

Affiliation administrative (CNRS, INSERM, ...) et Numéro d'affiliation de l'unité : Laboratoire Xavier ROUCOU, département de biochimie et de génomique fonctionnelle, faculté de médecine et des sciences de la santé, Université de Sherbrooke

Adresse précise du Laboratoire : 3201 rue Jean-Mignault, Room Z8-2070 | Sherbrooke, Quebec, Canada, J1E 4K8

Nom du Responsable de l'équipe d'accueil (EA) : Pascale LAFRANCE

E-mail : [stage.intl@usherbrooke.ca](mailto:stage.intl@usherbrooke.ca)

---

**Nom du Responsable du stage : Xavier ROUCOU**

Téléphone : 819-821-8000 x 72240

E-mail : [Xavier.Roucou@usherbrooke.ca](mailto:Xavier.Roucou@usherbrooke.ca)

Ecole doctorale de rattachement : Programmes d'études supérieures en Biochimie et génomique fonctionnelle.

Spécialité du stage : Recherche

Indiquez par quelques mots clés, l'orientation scientifique du sujet : Prédiction de structures, AlphaFold, RosettaFold, OpenProt.

---

**Titre du stage :**

Protéogénomique fonctionnelle et structurale en utilisant la ressource bio-informatique OpenProt.

Ce sujet constitue-t-il un premier pas vers un travail de thèse : Oui

---

**Description du sujet (quelques lignes):**

OpenProt est une ressource protéogénomique qui en plus des séquences codantes canoniques présentes dans les bases de données conventionnelles comme Ensembl ou RefSeq, intègre également l'annotation des séquences codantes alternatives. Ces séquences codantes alternatives sont très nombreuses et codent pour plusieurs centaines de milliers de protéines (PMID : 33179748). Le stage consistera à utiliser les données les nouveaux outils de prédiction de structure des protéines comme AlphaFold ou RosettaFold afin de prédire la structure des protéines alternatives et d'insérer ces structures dans OpenProt.

---

## **Bibliographies :**

1. *Leblanc, S., Brunet, M. A., Jacques, J. F., Lekehal, A. M., Duclos, A., Tremblay, A., ... & Roucou, X. (2021). Newfound coding potential of transcripts unveils missing members of human protein communities. bioRxiv, 2020-12.*
  2. *BRUNET, Marie A., LEBLANC, Sebastien, et ROUCOU, Xavier. Reconsidering proteomic diversity with functional investigation of small ORFs and alternative ORFs. Experimental cell research, 2020, vol. 393, no 1, p. 112057.*
  3. *Brunet, M. A., Lucier, J. F., Levesque, M., Leblanc, S., Jacques, J. F., Al-Saedi, H. R., ... & Roucou, X. (2021). OpenProt 2021: deeper functional annotation of the coding potential of eukaryotic genomes. Nucleic Acids Research, 49(D1), D380-D388.*
  4. *Brunet, M. A., Lekehal, A. M., & Roucou, X. (2020). How to Illuminate the Dark Proteome Using the Multi-omic OpenProt Resource. Current Protocols in Bioinformatics, 71(1), e103.*
  5. *Baek, M., DiMaio, F., Anishchenko, I., Dauparas, J., Ovchinnikov, S., Lee, G. R., ... & Baker, D. (2021). Accurate prediction of protein structures and interactions using a three-track neural network. Science, 373(6557), 871-876.*
- 

Retour par e-mail : [anne-claude.camproux@u-paris.fr](mailto:anne-claude.camproux@u-paris.fr)