

Projets de Bioinformatique Structurale
Responsables : Julien Diharce et Jean-Christophe Gelly

L'objectif de ces projets est de vous familiariser avec les différentes méthodes et outils de Bioinformatique et de Modélisation Moléculaire.

Différentes séquences ont été sélectionnées et elles n'ont pas pour l'instant de structures 3D disponibles à part celles prédites par AlphaFold. Vous choisirez, en binôme, une séquence parmi la liste fournie (une séquence unique par binôme).

Pour chacune de ces séquences, vous devrez :

- ***Partie Génération Modèles***

- 1- Identifier sur **la base d'alignements de séquence** des régions conservées ou ayant co-évolué.
- 2- **Construire un modèle 3D**, en utilisant différents outils dont vous pourrez comparer les résultats (**OmegaFold, ESMFold, RoseTTAFold, DMPFold2, Modeller, RaptorX, SwissModel, I-Tasser, Rosetta** etc...)
Attention, il peut s'agir d'une protéine membranaire, vous devrez donc au préalable vous aider d'outils dédiés.
Vous **comparerez vos résultats à celui d'AlphaFold2 et AlphaFold3**.
- 3- Comparer les résultats et discuter
- 4- Analyser la qualité de votre modèle avec les métriques internes aux outils.
TRES IMPORTANT
- 5- Identifier les **partenaires de votre protéine** en interrogeant la littérature et les base de données STRING ou INTACT.

- ***Partie Etude de la conformation***

Pour chacun des modèles construits et sélectionnés par votre précédente analyse :

1. Réaliser une simulation gros-grain en utilisant le champ de force Martini ou tout-atome avec le champ de force CHARMM36 (à vous de choisir selon la taille du système).
2. Analyser les différents mouvements et le comportement dynamique général de la protéine avec des mesures et outils adaptés.

- ***Partie Docking Protéine-Protéine***

1. Sélectionner un partenaire de la protéine, en privilégiant un dont la structure 3D est connue. Sinon faites le modèle correspondant.
2. Effectuer le docking Prot-Prot avec différents outils (Haddock, ClusPro, Zdock etc..). Comparer éventuellement avec les résultats de AlphaFold2 et 3 (AlphaFold Multimer).
3. Si possible, réaliser une simulation du modèle d'interaction protéine/protéine en simulation gros-grains (champs de force Martini)

Suivi et évaluation du module

Un premier suivi impliquant une présentation orale courte des premières recherches /résultats obtenus par les binômes sera réalisé le **mercredi 11 décembre 2024** la journée à partir de 10h du matin. L'idée est d'ici échanger sur votre plan et vos premières analyses. Prévoyez une présentation courte d'environ 5-10 min.

L'ensemble des résultats devra faire l'objet d'un rapport d'environ 20 pages dans lequel vous présenterez les propriétés biologiques, la question traitée, les outils et méthodes utilisées et bien sur les résultats obtenus en faisant preuve d'esprit critique.

Le rapport pourra être en anglais (recommandé) et sera à rendre sous moodle la veille de la soutenance du projet, **le dimanche 5 janvier 2025 à 23h59. Vous devrez aussi rendre les diapositives de votre présentation orale de votre projet à cette même heure sur un dépôt différent sur moodle.** Ces présentations orales auront lieu le **Lundi 6 Janvier 2025** durant toute la journée. Vous devrez tous être présents ce jour et votre participation lors des séances de questions sera appréciée. Nous serons également très sensibles à la pertinence de la démarche et à son argumentation.

Prenez soin de bien lire la bibliographie et de la fournir.

N'hésitez à passer une journée sur cet aspect qui pourra vous aider grandement par la suite.

Liste des protéines à considérer.

Les protéines surlignées en turquoise sont à choisir obligatoirement, soit 11 protéines. Les autres sont à choisir parmi celles non surlignées. Chaque binôme doit traiter une séquence différente.

Indiquez dans un google sheet, la liste des binômes pour chaque séquence et renvoyez nous cela par mail pour grand maximum la fin de la semaine.

UniProtKB - G7J6M8 (G7J6M8_MEDTR)	UniProtKB Q9GZN6-2 isoform 2 (SLC6A16_Human)
UniProtKB - Q96PL5 (ERMAP_HUMAN)	UniProtKB Q6P1W5 · CA094_HUMAN
UniProtKB - Q93070 (NAR4_HUMAN)	UniProtKB A0A016X1N4_9BILA
UniProtKB - P50895 (BCAM_HUMAN)	UniProtKB A0A2G9UZM2_TELCI
UniProtKB - P0DTC1 (R1A_SARS2) : Region 819-2763 :ATTENTION	Q6UDW7 · Q6UDW7_PLAFA
UniProtKB- P15976 · GATA1_HUMAN	
UniProtKB P16833 · UL116_HCMVA	
UniProtKB: A0A015JVU2 · A0A015JVU2_RHIIW	
UniprotKB :A0A037UPF4 · A0A037UPF4_SPIME	
UniProtKB : A0A059A5L5 · A0A059A5L5_EUCGR	
UniprotKB: A0A063Y959 · A0A063Y959_9GAMM	
UniprotKB: A0A093F6V7 · A0A093F6V7_TYTAL	
UniProtKB - A0A2G9RHE5 (A0A2G9RHE5_LITCT)	
UniProtKB - D6WKK0 (D6WKK0_TRICA)	
UniProtKB - E1ZG21 (E1ZG21_CHLVA)	
UniProtKB - X0K9N7 (X0K9N7_FUSOX)	
UniProtKB - M7T4C5 (M7T4C5_EUTLA)	
UniProtKB - W5MKN1 (W5MKN1_LEPOC)	
UniProtKB - W7X0N2 (W7X0N2_9BURK)	
UniProtKB - A0A1K1M4I5 (A0A1K1M4I5_9BACT)	
UniProtKB - W7IKS8 (W7IKS8_9PSEU)	
UniProtKB - X6IVT1 (X6IVT1_9RHIZ)	
UniProtKB - A0A0W7VPD0	