

## Améliorations du Mémoire - Théo MENARD

## Améliorations proposées pour le mémoire de stage M1 bioinformatique

### 1. Améliorations sémantiques et rédactionnelles

#### A. Orthographe et grammaire

- Corriger les fautes récurrentes :

- "Vu globale" -> "Vue globale"
- "la méthode utilise" -> "la méthode utilisée"
- "script" -> harmoniser en français ou adopter un terme unique (ex. "programme")

#### B. Reformulations et style

- Préférer des tournures académiques :

- "On retrouve huit candidats" -> "Huit candidats ont été identifiés"
- "Certain fragment comporte" -> "Certains fragments comportent"

#### C. Clarté et fluidité

- Raccourcir les phrases longues et complexes.

- Ajouter un ou deux schémas explicatifs, par exemple pour :

- Le pipeline complet (analyse PDB -> génération -> docking -> post-traitement)
- La méthode d'assemblage avec linkers

#### D. Uniformisation

- Harmoniser les temps : présent ou passé composé, mais pas les deux.

- Vocabulaire scientifique constant : "interaction", "conformation", "ligand", etc.

### ### 2. Améliorations techniques du pipeline

#### #### A. Extraction des hotspots

- Intégrer des outils plus robustes pour l'identification des résidus clés :
  - PISA, HotSpot Wizard, FMO (Fragment Molecular Orbital)
  - Modèles pré-entraînés (GNN ou CNN) pour la détection de hotspots

#### #### B. Génération de pseudo-molécules

- Contraindre la géométrie avec minimisation énergétique
  - Utilisation de RDKit (UFF, MMFF) + OpenMM
- Contrôle de la faisabilité chimique (valence, cycles, isoméries)

#### #### C. Docking

- Ajouter une étape de rescoring :
  - MM-GBSA, Prime MM-GBSA
- Tester plusieurs moteurs : Smina, PLANTS, AutoDockFR

#### #### D. Autoencodeur / Génération

- Intégrer un agent DQN ou PPO pour guider la génération
- Test de modèles conditionnels : Conditional VAE, Diffusion Models, SyntaLinker
- Ajouter des contraintes de pharmacocinétique (toxicité, solubilité, etc.)

### ### 3. Perspectives pour la génération d'inhibiteurs d'interfaces PPI

## Améliorations du Mémoire - Théo MENARD

### #### A. Objectifs medium terme

#### 1. Hotspot-centric generation :

- SMILES conditionnes sur la geometrie locale de l'interface

#### 2. Multi-objectif :

- Combiner QED, docking, toxicite, SAS, solubilite, permeabilite, etc.

#### 3. Generation 3D directe :

- a l'aide de DiffDock, EquiBind, GeoDiff

#### 4. Nouvelles banques de fragments :

- ZINC, ChEMBL fragments, DrugBank fragments
- Ajouter des metadonnees : polarite, rigidite, volume

### #### B. Apprentissage supervise sur PPI

#### - Creation de datasets d'entrainement :

- iPPI-DB, 2P2IDB, TIMBAL (positifs)
- Ligands inactifs (negatifs)

#### - Entrainement de modeles discriminants : SVM, GNN, CNN

### #### C. AlphaFold et ESMFold

- Generer des structures en l'absence de donnees cristallographiques
- Integrer la geometrie prevue dans le pipeline de docking ou de generation

### ### 4. Propositions supplementaires d'accompagnement

- Aide a la correction directe du PDF (si souhaite)
- Reecriture des sections cles : introduction, abstract, conclusion, perspectives
- Preparation d'un diaporama de soutenance

## **Améliorations du Mémoire - Théo MENARD**

- Préparation d'un pitch pour publication ou poursuite en M2
- Mise en forme d'un projet GitHub propre avec README, notebooks, figures, etc.