

M2 CHPS

# CHPS1001 – CRTP Docking

COGNE Romain  
CORTINHAL Simão  
BATISTE Quentin  
2024/2025

## Sommaire

|                                           |   |
|-------------------------------------------|---|
| Introduction .....                        | 1 |
| 1. Matériels et méthodes .....            | 1 |
| 1.1. Logiciels et matériels utilisés..... | 1 |
| 1.2. Méthode .....                        | 1 |
| 2. Résultats et interprétations.....      | 2 |
| Conclusion .....                          | 5 |
| Annexes .....                             | 6 |

## Introduction

Le sujet de ce TP est de tout d'abord prendre en main le logiciel d'amarrage moléculaire, puis de réaliser deux étapes : la première est de réaliser des tests d'amarrages entre la molécule de galactose et de la  $\beta$ -galactosidase, qui est la molécule liée à la tolérance au lactose chez l'homme, à l'aide du logiciel Autodock. La deuxième étape consiste à mettre en place un script permettant de réaliser plusieurs tests de Docking plus petit qui ensemble recouvrent l'ensemble de la molécule, qui donne des résultats plus précis qu'un seul test de Docking sur l'intégralité de la molécule.

## 1. Matériels et méthodes

### 1.1. Logiciels et matériels utilisés

Pour réaliser ce TP, nous avons utilisé les logiciels Autodock/AutoGrid, qui permet de visualiser l'amarrage moléculaire, ou « molecular docking ». Cela nous permettra de visualiser l'interaction entre le galactose (ligand) et la  $\beta$ -galactosidase. Aussi, nous avons dû utiliser le logiciel de visualisation VMD pour vérifier le résultat de nos Docking et de comparer l'emplacement de la galactose obtenu après Docking et celui fourni dans le fichier de référence « Complexe\_BetaGal\_Galactose ». Le TP a été réalisé sur les iMac de la salle PMMM du bâtiment 18.

Vous retrouverez également l'intégralité du code python joint à ce Rapport.

### 1.2. Méthode

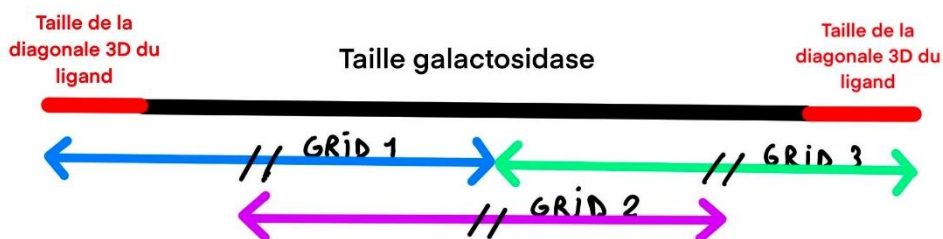
Dans un premier temps, nous avons dû prendre en main le logiciel Autodock en réalisant l'intégralité de l'Étape 1 (et un peu plus...). A la fin de cette étape, Nous avons obtenu un résultat en définissant le nombre de lancement d'algorithme génétique à 10 puis en le modifiant à 50 pour obtenir de meilleurs résultats. Nous avons ensuite pu comparer les résultats avec le fichier de référence « Complexe\_BetaGal\_Galactose » avec VMD pour 10 et 50 lancements plusieurs fois (cf. Figurexx-xx).

Dans un second temps, pour réaliser la 2<sup>ème</sup> et dernière étape du TP, nous avons du réalisé un script python qui permet de découper la molécule en plusieurs petites grilles, plutôt qu'une grosse grille englobant presque l'entièreté de la molécule. En effet, cette méthode de Docking ne permet pas de réaliser des tests précis, pour ce faire, la méthode est de découper cette grille en 27 plus petites grilles (3 \* 3 \* 3) et de réaliser des tests de Docking sur chacune de ces petites grilles.

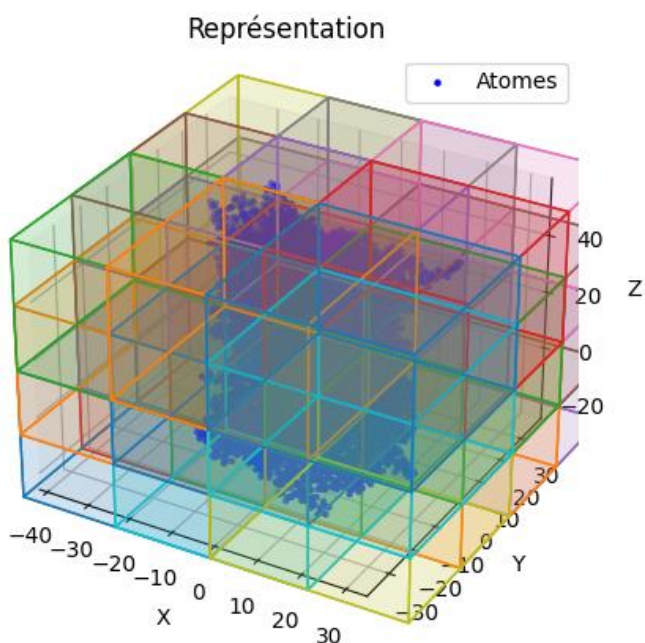
Pour la taille des grilles, on utilise la taille de la galactosidase et la longueur de la diagonale du ligand. Par exemple, pour la hauteur de chaque grille, on fait le calcul suivant :

- $(\text{Hauteur galactosidase} + (2 * \text{diagonale ligand}))/2$

Les grilles seront disposées comme dans le schéma suivant que ce soit en hauteur ou en largeur :



On peut voir ici qu'une des grilles permet un large recouvrement au cas où le ligand se situerait entre les deux grilles.



Pour la parallélisation, chaque thread travaille sur sa grille dans un répertoire différent pour avoir des fichiers map pour chaque thread et réalise le Docking dessus. A la fin, on parcourt les différents fichiers de sorti pour obtenir les meilleurs résultats (ici on vient récupérer les 5 docking où l'énergie de liaison est la plus basse.)

## 2. Résultats et interprétations

Comme dit précédemment, la première étape consistait à prendre en main le logiciel et de faire des premiers tests sur l'ensemble de la molécule. Nous avons tout d'abord réalisé une série de 10 lancements de Docking comme indiqué sur le sujet de TP (cf. Figure 1), puis avons réalisé une série de 50 lancements de Docking (cf. Figure 2) pour obtenir de meilleurs résultats.

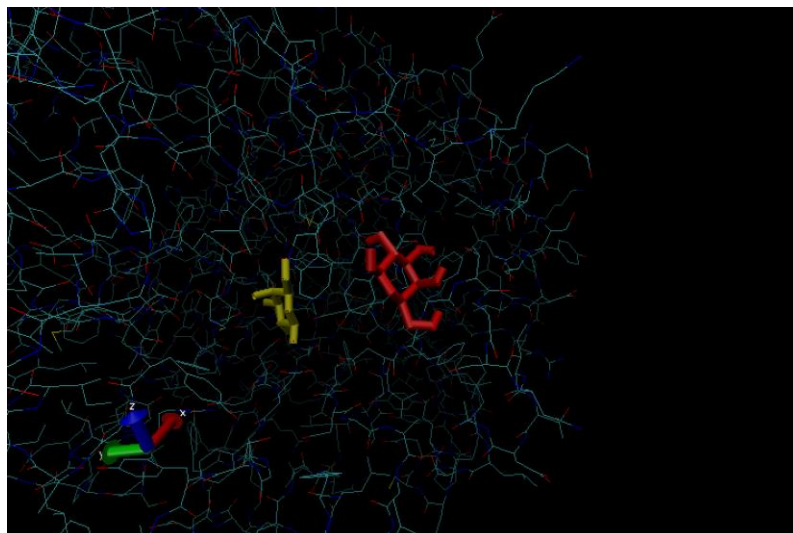


Figure 1 – Visualisation résultats avec VMD (10 lancements) Jaune = Référence, Rouge = le nôtre

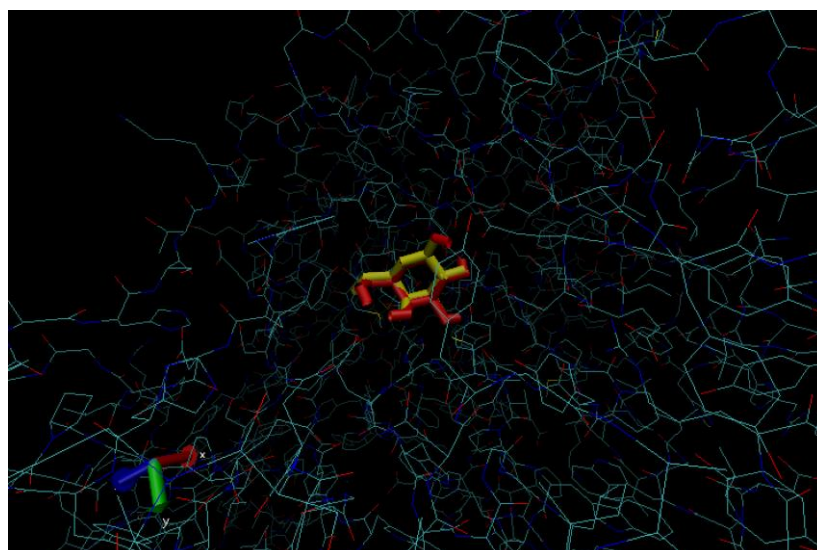


Figure 2 - Visualisation résultats avec VMD (50 lancements) Jaune = Référence, Rouge = le nôtre

Ici, on remarque bien que les résultats obtenus avec 50 lancements sont bien meilleurs que ceux obtenus avec 10 lancements : en effet, le ligand se confond avec celui fourni de référence, comparé à celui de la *Figure 1*. L'énergie de liaison obtenue était également bien plus basse (-4.87 pour 50 lancements, contre -1.95 pour 10 lancements).

Une fois l'étape 1 terminée, nous allons re-effectuer des lancements d'algorithme génétique (10 puis 50), mais en suivant les instructions de l'étape 2, c'est à dire en découpant le cube avec 27 petite grilles plutôt qu'une seule, et effectuer 10, puis 50 lancements dans chacune de ces grilles (comme décrit dans la partie "*Méthode*").

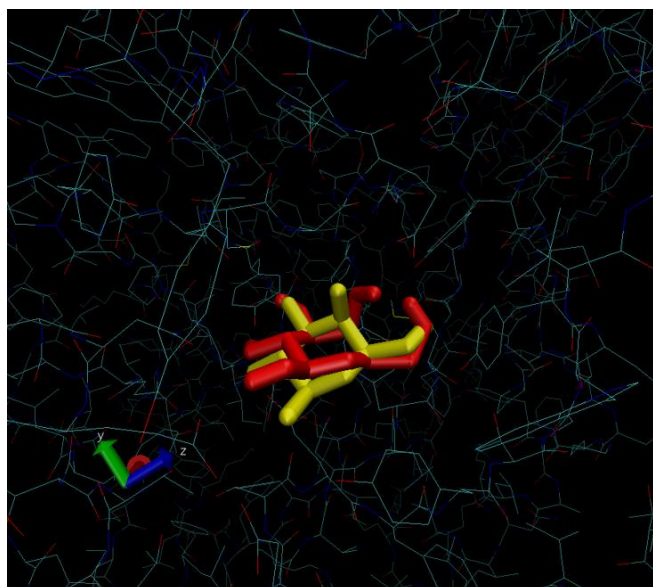


Figure 3 – Visualisation avec VMD (10 lancements, meilleur résultat des 27 grilles) Jaune = Référence, Rouge = le nôtre

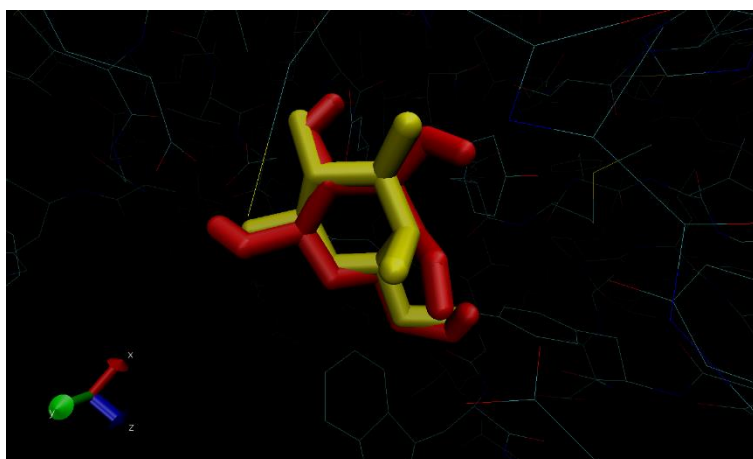


Figure 4 – Visualisation avec VMD (50 lancements, meilleur résultat des 27 grilles) Jaune = Référence, Rouge = le nôtre

En comparant les résultats obtenus dans la partie 2 avec ceux de la partie 1, on remarque qu'ils sont bien meilleurs :

|                  | 10 lancements | 50 lancements |
|------------------|---------------|---------------|
| <b>Méthode 1</b> | -1.95         | -4.87         |
| <b>Méthode 2</b> | -4.53         | -5.26         |

Cette différence s'explique par plusieurs raisons : dans la méthode 1 on se contente de réaliser uniquement 10 ou 50 lancements sur l'ensemble de la molécule, ainsi les tests de Docking se font sur l'ensemble de la molécule (et encore, certaines parties ne sont même pas pris en compte), tandis que dans la deuxième méthode, nous avons réalisé 10, puis 50 lancements pour chacune des grilles, cela fait un total de  $10 * 27 = 270$  lancements puis  $50 * 27 = 1350$  lancements sur l'ensemble de la molécule. De plus, dans la deuxième méthode, nous avons bien fait attention à ce qu'aucune partie de la molécule ne soit oubliée. Finalement, il ne faut pas oublier que chaque test de Docking est aléatoire, ainsi, plus il y a de lancements, plus la probabilité d'obtenir de meilleurs résultats est élevé.

## Conclusion

Pour conclure, ce TP nous a permis de découvrir les techniques de Docking moléculaire à l'aide du logiciel Autodock, ainsi que de comprendre l'importance de la précision dans la méthodologie pour obtenir des résultats fiables. En réalisant des docking à l'aide de plus petites grilles mais plus nombreuses nous avons pu obtenir des résultats nettement améliorés par rapport aux tests réalisés sur l'ensemble de la molécule en une seule grille.

Les résultats montrent une nette amélioration des énergies de liaison, notamment avec 50 lancements par grille, cela valide notre approche d'échantillonnage intensif. Ces améliorations s'expliquent par l'augmentation du nombre total de lancements de l'algorithme génétique, ce qui permet d'explorer davantage de configurations potentielles. Cette expérience met en évidence l'importance des méthodes d'échantillonnage par sous-grilles pour des molécules complexes, et ouvre des perspectives pour optimiser encore cette approche en augmentant le nombre de lancements et en ajustant la taille des grilles en fonction des caractéristiques du ligand et de la molécule cible.

## Annexes

Veillez retrouver l'ensemble du code sur notre dépôt git disponible à l'adresse suivante :  
<https://gitlab-mi.univ-reims.fr/bati0008/tp-docking.git>