

# TP MODÉLISATION MOLÉCULAIRE

\*\*\*\*\*Consignes\*\*\*\*\*

- Outils à installer

Autodock tools (disponible pour les différentes plateformes Unix/win/Mac)

<http://mgltools.scripps.edu/downloads>

- Pour toute question/discussion ou problème rencontré, contacter M.Benabderrahmane [mohammed.benabderrahmane@unicaen.fr](mailto:mohammed.benabderrahmane@unicaen.fr)
- Envoyez un compte rendu: NOM\_PRENOM\_M1\_DD.pdf en répondant aux questions du TP à l'adresse ci-dessus avant le 26.02.2021.

\*\*\*\*\*

## I. Découvrir une protéine en partant d'une séquence

On considère la séquence protéique suivante:

GPLGSDELYRQSLEIISRYLREQATGAKDTKPMGRSGATSRKALETLLRRVGDGVQRN  
HETAFQGMLRKLDIKNEDDVKSLSRVMIHVFSDGVTNWGRIVTLISFGAFVAKHLKT  
INQESCIEPLAESITDVLVRTKRDWLVKQRGWDGFVEFFHVEDLEGG

En utilisant l'algorithme BlastP (<https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins>), interrogez la base de données swissprot.

1. De quelle protéine s'agit-il ?
2. Récupérer son numéro d'accèsion (choisir l'accèsion à 100 % d'identité).
3. Interrogez la banque de données (Uniprot) <https://www.uniprot.org/> en utilisant le numéro d'accèsion récupéré.
4. En partant de la fiche Uniprot, faites un résumé sur cette protéine (sa fonction, localisation cellulaire, domaines, modifications post-traductionnelles ... etc).

## II. Protein Data Bank (PDB) et PDBsum

En partant de la fiche Uniprot, aller sur la section "STRUCTURES", choisir la structure 3WIX, et cliquer sur "RCSB PDB".

1. De quel type de complexe s'agit-il ?

2. Quelle méthode a été utilisée pour l'obtention de cette structure ? sa résolution ?  
Quelle différence avec une structure RMN d'une protéine ?
3. La structure obtenue contient-elle des mutations ?

La PDB permet d'avoir un ensemble d'informations relatives à la structure étudiée. La PDBsum permet une vue sur des annotations sur chaque structure de la PDB, tout en fournissant un accès rapide à un ensemble d'analyses intéressantes sur ces structures.

4. À partir du portail de la PDBsum  
(<http://www.ebi.ac.uk/thornton-srv/databases/cgi-bin/pdbsum/GetPage.pl?pdbcode=indx.html>), chercher l'entrée PDB "3WIX".
5. En se basant sur la fiche PDBsum, détaillez le contenu de cette structure ?
6. Quel est le contenu en structures secondaires de cette protéine ?
7. Reportez un schéma de l'interaction entre la protéine et le ligand (Voir rubrique Ligand/LigPlot).
8. Quelles sont les interactions qui entrent en jeu dans ce complexe ? Quelle est leur distance moyenne ?

### III. Amarrage (Docking) moléculaire avec autodock-vina (webina)

Cette partie illustre un docking moléculaire, vu en cours, entre une protéine et un ligand synthétique.

Le docking moléculaire est une étape permettant de filtrer un ensemble de molécules contre une cible biologique (protéine ou autre). Cette analyse nécessite 4 données principales:

- Une structure tridimensionnelle d'une protéine correctement préparée.
- Une définition du site actif (précise si connue, large sinon), par la grille de calcul ("boîte de docking").
- Une structure du ligand à docker correctement préparée.
- Un algorithme de recherche conformationnelle et un optimiseur (avec une fonction de score à optimiser ...).

On utilisera Autodock-vina comme algorithme de docking. Cet outil a été porté sur une interface web (<https://durrantlab.pitt.edu/webina/>).

1. Faites une simple recherche bibliographique sur Autodock vina. Quelle est la fonction de score utilisée ? Quel algorithme d'optimisation ?

Avant de procéder au docking, il faudra préparer correctement la protéine et le ligand à docker.

De plus, pour autodock vina, un type de fichier (PDBQT), incluant les charges, est nécessaire.

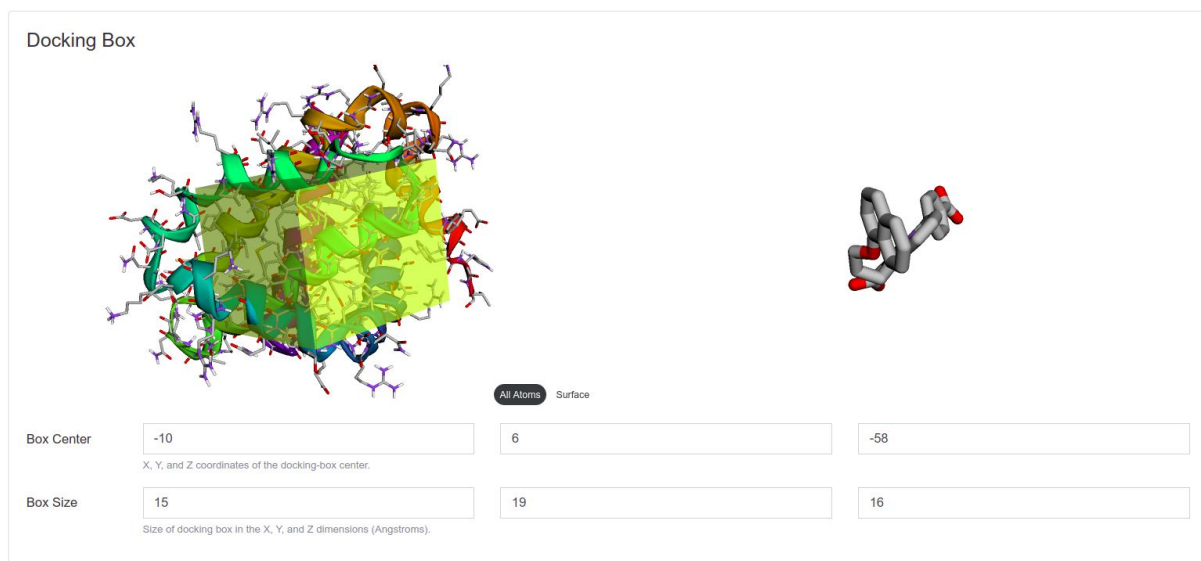
- Préparation de la protéine:
  - ⇒ Lancer AutodockTools et explorer son interface graphique.
  - ⇒ File menu > Read Molecule. et sélectionnez le fichier 3wix\_A.pdb; la structure s'affichera.
  - ⇒ Edit menu > add hydrogens: add polar hydrogens > fix > bond order, and renumber residues including the newly added hydrogen atoms.
  - ⇒ Edit menu > add charges: add Kollman charges. Cela ajoutera les charges partielles pour la protéine.
  - ⇒ Grid menu > Macromolecules > Choose. Sélectionner 3WIX\_A, puis sauvegarder le fichier généré (**3WIX\_A.pdbqt**) dans votre dossier de travail.
- Préparation du ligand:
 

Rq. Vous pouvez supprimer la structure 3WIX\_A. clic droit de la souris puis delete.

  - ⇒ File menu > Read Molecule > Sélectionner ensuite le fichier 3WIX\_lig.pdb
  - ⇒ Edit menu > atoms > assign AD4 type.
  - ⇒ Edit menu > hydrogens > add > polar only
  - ⇒ Edit menu > charges > Compute Gasteiger
  - ⇒ Ligand menu > Input > Choose > select 3WIX\_lig > Select Mol for Autodock4
  - ⇒ Ligand menu > Output > **Save as PDBQT file** (dans votre dossier de travail).
- Aller sur Webina : <https://durrantlab.pitt.edu/webina/>
  - ⇒ Receptor > clic sur Browse puis choisir le fichier 3WIX\_A.pdbqt
  - ⇒ Ligand > clic sur Browse puis choisir 3wix\_lig.pdbqt

Les deux fichiers seront ensuite affichés sur la partie visualisation de l'interface (vous pouvez essayer le mode surface pour voir la poche à viser ...).

  - ⇒ Centrer la boîte de docking autour du résidu F270 (voir image ci-bas)



⇒ Choisir 4 CPU (ou selon votre machine) et 8 pour exhaustiveness.

Webina exécutera le calcul sur votre propre station et non sur un serveur distant ...

Une fois le calcul terminé, le résultat s'affiche sur la rubrique Output. Les différents modes de binding seront affichés.

2. Évaluer qualitativement les différentes poses proposées par le docking (position des groupements fonctionnels, enfouissement au niveau de la poche ...etc).
3. Noter l'énergie d'interaction calculée pour la meilleure pose proposée.
4. Refaire le docking en modifiant la taille de la boîte (réduire par exemple la taille). Quel effet constatez-vous sur le résultat obtenu ?
5. Le complexe protein-ligand correspondant à la meilleure pose est téléchargeable (OUTPUT PDBQT File). Ce complexe peut être utilisé pour des analyses de stabilité par d'autres méthodes de modélisation moléculaire (simulations de dynamique moléculaire ... etc).

Rq. Il est possible d'ajouter le complexe initial (comme point de comparaison) sur Webina avant de lancer le calcul.

#### IV. Evaluation du profil ADME du ligand

Dans un processus de Drug Design, il peut être utile d'évaluer et prévoir le profil ADME d'une molécule afin de guider son optimisation.

⇒ D'abord, il faudra convertir la structure 3D du ligand (fichier 3WIX\_lig.pdb) vers le format SMILES(smi).

⇒ Aller sur : <http://cdb.ics.uci.edu/cgi-bin/BabelWeb.py>

⇒ Upload File et choisir le fichier 3WIX\_lig.pdb

⇒ Input format : PDB

⇒ Output format : SMI absolute.

⇒ Cliquer sur Convert format

⇒ Récupérer la chaîne de caractère générée.

⇒ Aller sur <http://www.swissadme.ch/>

⇒ Coller la chaîne de caractère SMI sur le champ dédié et cliquer sur RUN.

6. Ajoutez au rapport le profil ADME prédit et commentez le caractère physico-chimique de ce ligand (Solubilité, Drug-Likeness, Pharmacocinétique) ?

\*-----\*

#### V. Pour aller plus loin (ressources à consulter)

1. Apprendre un langage de programmation (R, Python, ou autre).

Pour le R par exemple, le livre de Vincent Goulet

[https://cran.r-project.org/doc/contrib/Goulet\\_introduction\\_programmation\\_R.pdf](https://cran.r-project.org/doc/contrib/Goulet_introduction_programmation_R.pdf)

Autres : <https://www.r-bloggers.com/>

2. Bibliothèques pour "l'analyse de données structurales": <http://thegrantlab.org/bio3d/>

3. Apprendre à utiliser un outil de visualisation moléculaire (VMD):  
<https://www.ks.uiuc.edu/Development/Download/download.cgi?PackageName=VMD>
4. Livre : **Molecular** modelling: principles and applications - A. Leach

**Quelques journaux et périodiques à consulter:**

LiveComJ. <https://www.livecomsjournal.org/about>

JCIM. <https://pubs.acs.org/journal/jcisd8>

Physical Chemistry B <https://pubs.acs.org/journal/jpcbfk>

J.Med.Chem. <https://pubs.acs.org/journal/jmcmar>

Annual Review of Biophysics <https://www.annualreviews.org/journal/biophys>