

TP de Visualisation de Structures Protéiques

A. Fichiers de coordonnées 3D : format PDB

- (1) Aller sur le site de la Protein Data Bank (<http://www.pdb.org>) et rechercher les entrées 1CLL et 1CFC. Quel est le point commun entre ces deux structures ?
- (2) Télécharger les fichiers de coordonnées 1CLL.pdb et 1CFC.pdb et les ouvrir avec un éditeur de texte, pour rechercher les informations suivantes, en notant les champs (TITLE, COMPND, SOURCE, EXPDTA...) qui les contiennent.

	1CLL	1CFC
Nom de la molécule		
Nombre de chaînes		
Publication associée		
Méthode expérimentale		
Résolution		
Nombre de conformères		
Classe structurale		

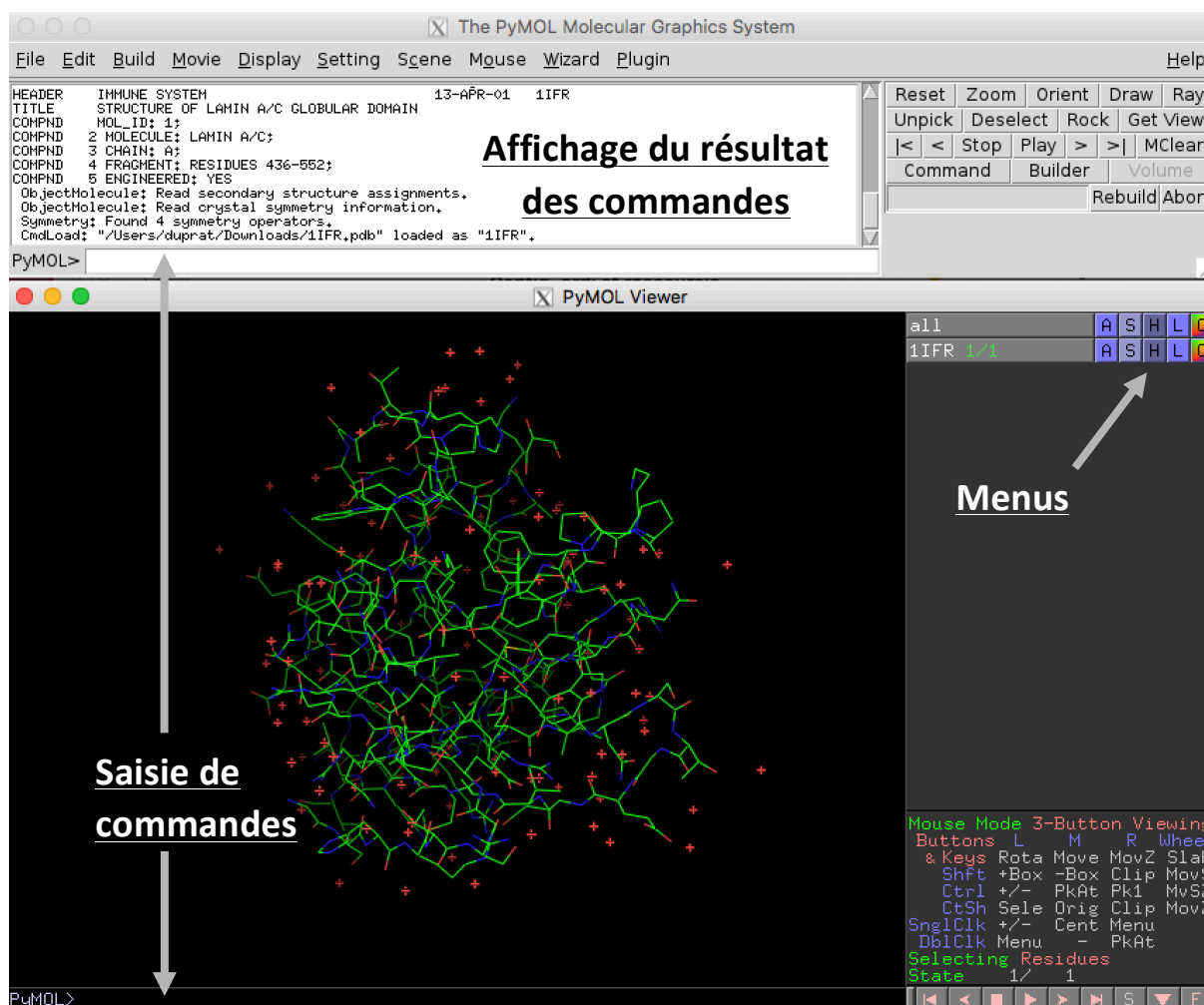
- Comparer le contenu des champs DBREF, SEQRES. Quelles informations fournissent-ils et en quoi diffèrent-elles ?
- Quel type d'information est fourni par les champs ATOM et HETATM ? Quelle est la différence entre ces champs ?
- Quelles informations sont fournies par les champs SITE et HETNAM ?

B. PyMOL : guide d'utilisation et premiers pas

- (1) Ouvrir le fichier 1CLL.pdb avec PyMOL.

L'objet 1CLL est chargé. PyMOL propose deux modalités d'action : par des commandes interactives (syntaxe PyMOL, décrite sur la page Wiki <http://pymolwiki.org/index.php/Category:Commands>) ou par les menus associés à chaque objet : **A** (Action), **S** (Show), **H** (Hide), **L** (Labels), **C** (Colors).

- Afficher la séquence en acides aminés (code 3 lettres, puis code 1 lettre) de chaque chaîne : Onglet *Display* > *Sequence* (sélectionner le *Mode* adéquat).
- Sélectionner un résidu au hasard en cliquant dessus. Que se passe-t-il au niveau de l'affichage de la séquence ? Que se passe-t-il au niveau du menu à droite ?
- Sélectionner les « O » rouges à la fin de la séquence. Les représenter sous forme de sphères : **S** (show) > *as* > *spheres*
A quoi correspondent-ils ?



Pour appliquer un mouvement à la molécule :

- **rotation** : clic gauche en déplaçant la souris (toutes directions)
- **zoom** : clic droit en déplaçant la souris verticalement

Quelques actions par Menu :

S (show) ou H (hide) >

- **cartoon** : représentation **simplifiée** des **structures secondaires** (telle qu'assignées dans le fichier PDB)
- **ribbon** : représentation **simplifiée** de la **chaîne principale** (pas de structure secondaire, pas de chaîne latérale)
- **sticks / lines** : représentation des **liaisons covalentes** (code couleur dans : C > by atom) ; on peut masquer les chaînes latérales (H > side chains)
- **spheres** : représentation des **atomes** comme des sphères pleines (rayon de van der Waals)
- **surface** : représentation de la **surface moléculaire**, obtenue en faisant rouler une molécule d'eau (représentée par une sphère de rayon 1.4 Angströms) sur la représentation des rayons de van der Waals ; cette surface prend en compte le recouvrement des creux par la sphère représentant le solvant.

C (Color) > by Chain (Chaîne polypeptidique), by SS (Structures Secondaires), ...

A (Action) >

- Visualisation des liaisons H entre atomes de la chaîne principale :
find > polar contacts > just intra-main chain
- Représentation du potentiel électrostatique de surface :
generate > vacuum electrostatics > protein contact potential (local)
- Superposition de deux objets (mobile, référence) :
align [mobile] > to molecule > [choix référence]

Pour accéder à l'aide sur une commande :

help [commande]

- (2) Exécuter les commandes ci-dessous et décrire leur action (pensez à consulter les affichages dans la boîte de dialogue) :

```
select sel1, object 1CLL and chain A and resn Met
select sel2, /1CLL//A/Met
select sel3, /1CLL//A/M
select segment, object 1CLL and chain A and resi 100-110
show sticks, /1CLL//A/82 or /1CLL//A/86
distance /1CLL//A/82/OE2, /1CLL//A/86/NH1
distance d, /1CLL//A/*/O*, /1CLL//A/*/N*, 3.2, mode=2
```

- (3) Ouvrir également dans PyMOL (c'est-à-dire sans fermer l'application) le fichier 1CFC.pdb : *File > Open > [Adresse fichier]*

Note : chaque objet chargé peut être affiché ou masqué en cliquant sur le nom correspondant.

Exécuter les commandes ci-dessous et décrire leur action :

```
align /1CLL//A//CA, /1CFC//A//CA, object=ali1
align /1CFC//A//CA, /1CLL//A//CA, object=ali2
align /1CFC//A//CA, /1CLL//A//CA, cutoff=1.5, object=ali3
align /1CFC//A//CA, /1CLL//A//CA, mobile_state=7, object=ali4
```

Que permet de faire l'option `mobile_state` ? Visualiser les deux structures superposées en faisant varier l'état de 1CFC (en bas à droite).

```
phi_psi /1CLL//A/104
get_dihedral /1CLL//A/103/C, /1CLL//A/104/N, /1CLL//A/104/CA,
/1UFG//A/104/C
get_dihedral /1CLL//A/104/N, /1CLL//A/104/CA, /1CLL//A/104/C,
/1CLL//A/105/N
```

Comparer le résultat de ces trois commandes. Conclure.

Application 1 : Analyse des interactions stabilisatrices de structures secondaires (PDB : 1IFR et 1CLL)

Pour chacune des structures 1IFR et 1CLL, sélectionner un fragment de la protéine ayant une structure secondaire bien définie (feuillet β ou hélice α) et visualiser les liaisons hydrogènes qui la stabilisent.

Application 2 : Analyse d'un site de liaison à un ion (PDB : 1CLL)

Dans la structure 1CLL, il y a 4 « résidus » nommés CA. Que représentent-ils ? Les sélectionner et les représenter sous forme de sphères.

Identifier les résidus qui les coordonnent avec la commande :

```
select sites, byres (resname CA around 3.5)
```

Les représenter sous forme de bâtons (sticks).

Comparer les quatre sites de liaison. Sont-ils différents, et si oui, en quoi ? Calculer les distances de coordination correspondantes.

Application 3 : Analyse d'une interface protéine-protéine (PDB : 2BBM)

Décrire cette structure : que contient-elle ? Comment a-t-elle été résolue ? Colorer les différentes chaînes de différentes couleurs et les représenter en cartoon : sont-elles en contact ?

Pour chaque chaîne, créer 1 objet en utilisant la commande `copy to object` (dans le Menu Action). Visualiser les surfaces moléculaires de ces objets : sont-elles en contact ?

Utiliser le script Python `color_by_restype.py` pour colorer les résidus en fonction de leur type. Les résidus à l'interface sont-ils différents des autres résidus à la surface de la protéine ?

Application 4 : Analyse de la dynamique structurale d'une protéine (PDB : 1CFC)

Que représentent les différents états de la structure 1CFC ? Avec la commande `intra_rms`, calculer la déviation standard (RMSD) de chaque état par rapport au premier. Lequel en est le plus éloigné ?

Utiliser le script `rmsf_states.py` pour calculer les fluctuations atomiques de chaque résidu de la protéine sur l'ensemble des états :

```
run rmsf_states.py
```

```
rmsf_states 1CFC, byres=1, reference_state=1
```

Les valeurs de fluctuations sont maintenant stockées dans la colonne des facteurs B de la structure. Pour les visualiser, utiliser la commande `spectrum`. Quelles sont les régions les plus flexibles de la protéine ?