

M2 BIBS. Module ABA.
Séance 3 : mise en pratique
Daniel Gautheret

Première partie : S'orienter dans une structure 3D d'ARN

Objectifs

A l'aide d'un logiciel de visualisation moléculaire, être en mesure de reconnaître les principaux éléments structuraux d'un acide nucléique, notamment:

- Formes d'hélice et sillons
- Paires Watson-Crick et Wobble
- Interactions non canoniques (paires non WC, triplets, interactions base-squelette)

a) Récupération de structure PDB

- Connectez-vous sur votre station Unix avec votre nom et votre mot de passe.
- Rendez-vous sur le site de la protein databank. Récupérez la séquence `1bna.pdb` (séquence complète avec coordonnées).
- Visualisez ce fichier avec la commande `more 1bna.pdb`. Encodage des coordonnées ?

b) Rasmol

- Chargez dans rasmol la structure `1bna.pdb`, avec la commande: `rasmol 1bna.pdb`
- Familiarisez-vous avec les rotations, translations et zoom (souris, avec combinaisons de boutons différentes et de la touche "shift").
- Passez en mode stereo et tentez une visualisation stereo. Si vous n'y parvenez pas, revenez en mode normal.
- Cliquez sur un atome et visualisez ses informations dans la fenêtre de dialogue: atome, nucléotide, brin. Ex: **Atom: C1* 136 Group: T 7 Chain:A** signifie atome C1' No 136, nucléotide T7, brin A.
- Apprenez à sélectionner les atomes dans la fenêtre de dialogue avec la commande `Select`.

Notes sur la sélection: lorsqu'il n'y a qu'une chaîne `select x` sélectionne le nucléotide (ou aa) No X et `select x-y` sélectionne les nt de x à y.

Lorsqu'il y a plusieurs chaînes, la sélection doit spécifier la chaîne: `select rxa` sélectionne le résidu de type r, numéro x de la chaîne a (par exemple: `select T7A`).

Pour sélectionner un atome particulier, spécifier le nom de l'atome séparé du résidu par un point, par exemple `select T7A.N1` pour le N1 du résidu T7 du brin A.

Les sélections peuvent être annulées avec la commande `select all`

- Sélectionnez un groupe d'atomes et colorez-le en rouge avec la commande `color red`.

- Sélectionnez plusieurs nucléotides consécutifs, puis essayez les options du menu "display". L'option "spacefill" est particulièrement utile pour appréhender le volume effectif des atomes.
- L'option "export" permet de produire un fichier graphique à partir de la vue courante. Sauvegardez la vue courante en format POSTSCRIPT. Visualisez le fichier sauvé avec `gs`

c) B-DNA

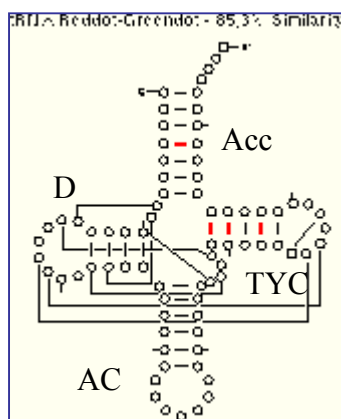
Toujours sur la structure 1bna...

- Repérez une paire A-T et une paire G-C. Colorez-les en jaune.
- Repérez deux donneurs de ponts H dans le sillon majeur. Colorez ces atomes en rouge
- Repérez deux accepteurs de ponts H dans le sillon mineur. Colorez ces atomes en bleu
- Affichez la molécule en "space-filling", puis sauvegardez une vue sous forme de fichiers POSTSCRIPT montrant les atomes colorés. Nommez le fichier 1.ps.
NOTE: les hétéroatomes, comme les molécules d'eau et les ligands, peuvent être supprimés avec l'option "Heteroatom" du menu "Display"

d) A-DNA

- Récupérez sur la pdb une structure de double hélice d'ADN de type A (recherche par mot-clé « A-DNA »)
- Ouvrez avec Rasmol le fichier que vous venez de récupérer. Colorez l'ensemble des atomes en blanc.
- Repérez deux donneurs de ponts H dans le sillon majeur. Colorez ces atomes en rouge
- Repérez deux accepteurs de ponts H dans le sillon mineur. Colorez ces atomes en bleu
- Représentez la molécule en "space-filling" et sauvez le graphique. Nommez le fichier 2.ps.

e) tRNA



- Trouvez sur le serveur pdb le fichier de coordonnées du tRNA Phe de levure 1EVV.pdb. Téléchargez-le et ouvrez le fichier avec Rasmol. Colorez l'ensemble des atomes en blanc.
- A l'aide de la structure secondaire ci-jointe, repérez et colorez les éléments suivants: La tige acceptrice, la tige "D", la boucle de l'anticodon.
- Sauvez le graphique, en nommant le fichier 3.ps. puis recolorez toute la molécule en blanc.
- Repérez dans la tige acceptrice: une paire G-C et une paire G-U. Colorez-les
- Colorez le U du U-turn de la boucle de l'anticodon.
- Repérez et colorez au moins deux triplets de bases.
- Sauvez le graphique en nommant le fichier 4.ps.

- Option "expert": recherchez au moins deux interactions non-Watson-Crick entre bases, et deux interactions entre squelette phosphodiester et bases. Colorez-les. Sauvez le graphique en nommant le fichier 5.ps.

Deuxième partie : Prédiction de structure secondaire d'ARN

a) MFOLD

- Récupérez la séquence mir.fasta dans ~gauthere/TD
- Rendez-vous sur le serveur Web MFOLD de Michael Zuker
- Repliez la séquence d'ARN
- Visualisez le dotplot d'énergie
- Visualisez la structure secondaire optimale
- Notez l'énergie de la structure

b) Intérêt d'un repliement ?

A l'aide d'un générateur de séquences aléatoire tel que :

<http://www.faculty.ucr.edu/~mmaduro/random.htm>

Imaginez et réalisez un protocole permettant de déterminer si l'énergie de repliement de cette séquence de miRNA précurseur est significative.

Troisième partie : Recherche d'ARNnc dans un génome

a) RFAM.

Rfam is a collection of multiple sequence alignments and statistical models covering many common non-coding RNA families. The statistical models used in RFAM are called covariance models. They use both sequence and structure information.

RFAM pipeline:

Seed alignment (expert-made) -> covariance model -> Infernal software -> full alignment

1. Use “browse RFAM” to obtain seed alignment for miRNA MiR1 family.

The screenshot shows the RFAM website interface in a Mozilla Firefox browser. The page title is "Rfam: seed alignment for mir-1". The URL is "http://www.sanger.ac.uk/cgi-bin/Rfam/getalignment.pl?acc=RF00103&type=seed&format=link". The page content includes the RFAM logo, the Wellcome Trust Sanger Institute logo, and a navigation menu with options: Home, Keyword Search, Sequence Search, Browse Rfam, Genomes, ftp, Help, miRNA, and mir-1 family. Below the menu, the text "seed alignment for mir-1" is displayed. The main content area shows a table of sequences from various species, including Dro. mel., Mus. mus., Hom. sap., and Cae. ele., with their corresponding seed alignments. The alignment is shown in a color-coded format with red and blue highlights. The page also includes navigation links like "Next" and "Previous" for each sequence.

Check secondary structure symbols.

Both secondary structure and alignment are necessary to build covariance model.

Note: distinguish miRNAs families in RFAM and the miRNA collection in the miRNA registry/miRBase.

- RFAM miRNAs: contains only miRNAs that can be aligned into families.
- MiRBase: all miRNAs. No alignment, no covariance model.

2. Use “Browse” to obtain seed alignment for tRNA.

Lots of sequences and little variation except in variable loop. Excellent signature. This is why tRNAs are so easy to find. Specificity and sensitivity are very high.

3. Use Keyword search to obtain seed alignment for Group I intron.

Highly variable structure. Poor signature. Very hard to identify. Either lots of false positives (specificity low) or lots of false negatives (sensitivity low).

4. Go to the Ensembl Web site. Retrieve the human sequence of coordinates:

8:65453473..65455125 (select Homo sapiens+range+export+fasta).

Which RFAM sequence is there?

