

# Prédiction de la structure des protéines.

Projets M2 IGF

Adrien Melquiond - Julien Maupetit

Décembre 2007

## Remarques générales :

Vous rédigerez un rapport synthétique (maximum 6 pages) en privilégiant les points biologiques et méthodologiques importants ainsi que la discussion. Vous décrirez assez succinctement les programmes utilisés, sauf si les paramétrages choisis étaient différents de leurs valeurs standards ou si vous avez fait appel à des méthodes originales.

Pour les serveurs de modélisation qui requièrent une adresse académique, vous disposez d'un compte à Paris 7 (login : *m2\_igf@paris7.jussieu.fr*; passwd : *bioinfo*); la messagerie est accessible depuis le site <https://paris7.jussieu.fr>. Attention à bien différencier vos soumissions sur les différents serveurs, en précisant par exemple votre nom dans le titre associé à la requête.

## 1 Construction d'un modèle structural pour une séquence protéique inconnue

L'objet de cette première partie est de proposer un modèle structural pour une séquence protéique d'intérêt (seq1.fasta).

Réalisez dans un premier temps une recherche d'homologues sur les banques de votre choix. Y-a-t'il un domaine conservé dans cette séquence, de quelle nature, que peut-il vous apprendre ? Vous prédirez la structure de cette protéine soit par modélisation comparative, soit par enfilage<sup>1</sup> (*threading*) ou par modélisation *ab initio*.

Afin d'affiner votre prédiction, vous disposez de données expérimentales obtenues par diffusion dynamique de lumière (DLS) et par dichroïsme circulaire (CD) sur cette protéine :

- ▷ Une analyse par dichroïsme circulaire de notre protéine fournit le spectre d'absorption suivant (Fig 1). Quelles conclusions pouvez-vous en tirer ? Des expériences complémentaires (Cryo EM) montrent qu'elle contient cinq éléments de structure secondaire : quatre forment un coeur hydrophobe et le dernier est flexible.

---

<sup>1</sup>3DPSSM semble connaître quelques soucis de maintenance, vous utiliserez donc **Fugue** en plus de **Phyre** et **Genthrader**

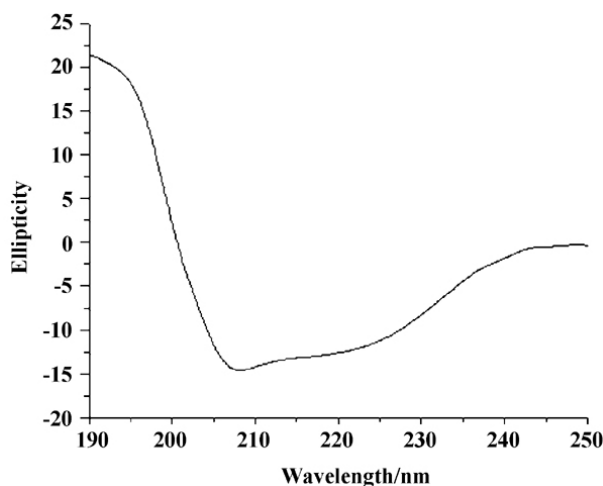


Fig. 1: Far-UV CD-spectrum of the protein X.

- ▷ Le rayon hydrodynamique ( $R_H$ ) d'une molécule est calculé à partir de son coefficient de diffusion ( $D$ ) en utilisant la relation de Stocke-Einstein (1) où  $k$  est la constante de Boltzmann,  $T$  la température et  $\eta$  la viscosité du milieu.

$$R_H = \frac{kT}{6\pi\eta D} \quad (1)$$

Cette mesure nous renseigne sur la taille effective de la molécule détectée à partir de son mouvement brownien, elle est liée à la taille de la molécule **solvatée** mais également à sa dynamique pendant la diffusion. La mesure de  $R_H$  permet une bonne approximation du rayon de gyration de la molécule ( $R_g \simeq R_H * 0.774$ ). Pour notre étude, le rayon hydrodynamique mesuré par DLS vaut approximativement 21.5 Å. Utilisez le programme `rgyr.pl` du package MMTSB pour calculer le rayon de gyration des différents modèles générés.

Si le modèle n'est pas fourni ou si certaines régions n'ont pas été reconstruites à cause d'un score trop faible (boucles en général), vous devrez reconstruire le modèle à l'aide de l'option **Alignement Interface** de SWISS\_MODEL. Pour cela, récupérez l'alignement réalisé qui correspond à votre modèle d'intérêt et si besoin transformez-le au format approprié grâce à l'outil `clustalw_convert`. L'utilisation du mode **Alignement Interface** de SWISS\_MODEL est très simple, n'oubliez pas de rechercher le code `ExpDB` associé à votre *template* (jeter un oeil au tutoriel).

Une fois votre modèle généré et optimisé, n'oubliez pas de procéder à son évaluation à l'aide des programmes vus en TP. Vous pouvez également regarder si les outils `El Nemo` ou `Pints` vous apportent de l'information supplémentaire sur cette structure.

**Remarque :** il est possible par la résolution parfaite du problème d'obtenir un modèle qui dévie de 1.3 Å par rapport à la solution (RMSD calculé par alignement structural sur 78.5 % de la séquence).

## **2 Construction d'un modèle structural pour deux phospholipases**

Vous avez deux séquences de deux sous-unités de phospholipases A2 appartenant au groupe XII. Avec les méthodes vues en cours, et similairement au premier projet, réalisez un modèle de ces deux sous-unités.