

TP modélisation par homologie

Partie SWISS-Model

1 SWISS-MODEL

1.1 Créer un compte sur SWISS-MODEL

1. Aller sur <http://swissmodel.expasy.org/>
2. Cliquer sur « Start Modelling »
3. Cliquer en bas sur *create an account*
4. Entrer votre adresse email
5. Chercher sur votre compte email le mot de passe pour SWISS-MODEL
6. Cliquer en bas sur *log in* et entrer votre email et mot de passe. Si vous êtes bien enregistré, votre adresse email apparait en haut à droite.
7. En haut il y a quatre menus:
 - a. *Modelling*: (comprends plusieurs étapes de modélisation en même temps)
 1. *myWorkspace*: la liste de vos calculs
 2. *Automated Mode*: modélisation automatisée
 3. *Alignment Mode*: modélisation à partir d'un alignement multiple fournit par vous
 4. *DeepView Project Mode*: modélisation à partir d'un alignement corrigé manuellement avec *DeepView (=Swiss-PdbViewer)*
 - b. *Tools*: (étapes de modélisation une par une)
 1. *Template Identification*: Recherche de templates
 2. *Domain Annotation*: Annotation de la séquence (domaines, structure secondaire, zones désordonnées ou trans-membranaires)
 3. *Structure Assessment*: Evaluation des modèles obtenues
 4. *Template Library*: Informations sur un template (Résolution X-ray, liens vers banques de structures: RCSB PDB, PDBsum, PDBe, PISA et liens vers SCOP et CATH)
 - c. *Repository*: (banques de modèles générés par SWISS-MODEL)
 1. *Query by Sequence*: recherche par séquence
 2. *Query by AC*: la même chose
 3. *Search by full text*: recherche par nom biologique
 - d. *Documentation*: aide en ligne
 1. *Examples*: Un tutoriel de SWISS-MODEL sur un exemple
 2. *Help* : Documentation brève de SWISS-MODEL

3. *Repository* : Documentation sur les modèles déposés sur SWISS-MODEL
4. *Visualisation*: Une liste de logiciels locaux de visualisation de structures
5. *Protein Structure Course* (attention : pas mis-à-jour)
6. *References* : Articles en lien direct ou indirect avec SWISS-MODEL

1.2 Tutoriel de SWISS-MODEL

On trouve sous *Documentation* → *Examples* un tutoriel pour les fonctions de base de SWISS-MODEL. On suivra ici ce tutoriel avec quelques questions supplémentaires :

1. Le tutoriel de SWISS-MODEL propose de créer un modèle par homologie de la protéine « Superoxide Dismutase [Cu-Zn] protein of *Drosophila melanogaster* (SoD, Uniprot AC: P61851) ». La séquence en acides aminés de la protéine est répertoriée sur www.uniprot.org sous le code P61851. Aller sur www.uniprot.org et y entrer ce code. Chercher la séquence sur la page de résultat.
 2. Quelle est la fonction de cette protéine ?
 3. Quels sont les cofacteurs ?
 4. Dans la rubrique « Interaction » chercher la forme oligomérique (monomère, dimère, ...) de la protéine.
5. Une fois loggé sur SWISS-MODEL, aller sur *Modelling* → *myWorkspace* si ce n'est pas déjà le cas.
6. Dans « Target Sequence » entrer le code de la protéine en question. Comparer rapidement la séquence qui s'affiche avec celle que vous avez trouvée sur uniprot.
7. Cliquer sur  Target pour ouvrir les options de coloration et tester les différentes options. Garder à la fin la coloration « Clustal ».
8. Cliquer sur « Search For Templates » pour trouver des structures 3D homologues à notre séquence. Cela peut prendre un peu de temps, en attendant vous pouvez poursuivre les questions suivantes :
9. Avec l'aide du cours de la semaine dernière chercher à quelles propriétés physico-chimiques correspondent les couleurs avec la coloration du type « Clustal » ?
10. Revenons sur SWISS-MODEL qui devrait avoir fourni une liste de templates possibles. Expliquer les titres des colonnes du tableau (Name, Title, Coverage, ...).
11. Les templates qu'on sélectionne en cochant la case à gauche sont superposés en 3D dans le viewer à droite. Tourner les structures avec la souris en tenant le bouton gauche de la souris. La molette de la souris permet de zoomer. Sélectionner les quatre premiers templates, qu'est-ce qu'on peut dire sur la similitude structurale ?
12. Cocher la case d'un seul template et regarder le détail d'un template en cliquant sur le bouton de flèche à droite dans le tableau (agrandir la fenêtre du browser au maximum).

13. Analyser l'alignement des deux séquences, chercher les substitutions et les insertions/deletions.
14. Cliquer sur  Targe et changer l'option « Secondary Structure » en comparant les résultats sur l'alignement des séquences et la structure 3D.
15. Parmi les quatre premiers templates, lequel contient le même nombre et type de ligands que notre target (attention c'est un homodimer) ?
16. Sélectionner plusieurs templates avec le même nombre et type de ligands que notre target dans la liste, puis regarder les alignements dans l'onglet « Alignment of Selected Templates ».
17. Ouvrir l'onglet « Sequence Similarity ». Chaque cercle représente un template et la distance entre deux cercles est proportionnel à l'identité de séquence entre les deux templates correspondants. La séquence du target est représentée par un point rouge. En positionnant la souris sur un cercle, la structure 3D et les infos du template sont affichés directement. Si on clique sur un cercle, cela sélectionne ou désélectionne le template correspondant.
Qu'est-ce qu'on peut dire sur la similitude des structures 3D dans un amas de cercles ? Et en comparant des cercles éloignés ?
18. D'après votre étude sur les propriétés des templates (insertions/deletions, ligands, forme oligomérique), sélectionner un ou deux templates pour construire un modèle.
Ensuite ouvrir le détail du template choisit puis choisir si le modèle doit être un homo-dimer ou un monomère et ensuite appuyer sur « Build Model ».
19. En attendant que le modèle se construit, lisez dans *Documentation* → *Help* les deux dernières parties : « *Display of modelling results* » et « *SwissModel Template Library (SMTL)* »
20. Analyser la qualité du modèle généré.
21. Cliquer en haut à gauche sur « Model 01 », puis sauvegarder et regarder les sorties « Model Report », « PDB File » et « Modelling Logs ». En regardant « Modelling Logs » est-ce que il y avait des problèmes à résoudre lors de la génération de la structure ?

2 AlkD exemple

Comme décrit dans le tutoriel de SWISS-MODEL publié dans Nature Protocols (2009, voir fichier 2009_BordoliSchwede_NatureProtocols.pdf dans le dossier biblio), on essaye de modéliser la *bacterial methylpurine-DNA glycosylase* (AlkD) qui est une enzyme de réparation de l'ADN. AlkD appartient à la superfamille des glycosylases d'ADN. Son code uniprot est : **Q2PAD8**

Regarder aussi la première partie du Matériel&Méthodes de l'article 2007-Dalhus-NAR_AlkD.pdf

2.1 Annotation de la séquence

22. Chercher l'entrée **Q2PAD8** sur uniprot, puis sauvegarder la séquence en format FASTA en cliquant sur le bouton « Format ».
23. Dans SWISS-MODEL aller dans *Tools* → *Domain Annotation*. On trouve ici des liens vers des serveurs de prédiction des propriétés de la séquence du target.

2.2.3 Comparaison

1. Comparer les résultats obtenues avec SWISS-MODEL et avec BLAST (NCBI) directement

A remplir

2. Comparer les résultats par rapport aux listes fournies par InterPro et Pfam

A remplir

3. Comparer le template 3BVS avec 2B6C au niveau séquence et structure. Pour cela utiliser le service *3D Similarity* de la RCSB PDB (voir voir fichier TP_PSF_NV505_serveurs.odt). Dans le *Comparison Tool* choisir par exemple *External: TM-Align*

Quel est le RMSD entre les deux structures ?

A remplir

2.2.4 Choix du template

1. Quel template choisiriez vous ?

A remplir

2. Pourquoi ?

A remplir

3. Quel template choisiriez vous, si vous seriez en 2007 ?

A remplir

2.3 Modélisation

1. Dans SWISS-MODEL créer un modèle avec le template PDB-ID = **2B6C** et Chain: **A**

2.4 Évaluation

2.4.1 SWISS-MODEL, ProQ2, IntFold2

1. Inspecter les différentes parties de la page de résultats. Cliquer sur ? pour obtenir de l'aide.
2. Dans SWISS-MODEL aller sur *Tools* → *Structure Assessment*. Cette page contient des liens sur des serveurs de vérification des modèles.
3. Tester ProQ2, qui a été bien évalué lors du concours CASP 10.
4. Tester IntFold2, qui a aussi été bien évalué lors du concours CASP 10 :

http://www.reading.ac.uk/bioinf/IntFOLD/IntFOLD2_form.html

2.4.2 Comparaison de structure avec PyMol

1. Vous pouvez aussi utiliser en local le logiciel PyMol. PyMol s'utilise le plus facilement avec des commandes:
 - a. Ouvrir une fenêtre Terminal et se déplacer (**cd**) dans le répertoire où se trouve les fichiers PDB des modèles (à renommer: model_1.pdb, model_2.pdb, ...) ainsi que le template 2b6c_A.pdb et la référence 3BVS.pdb.
 - b. Lancer PyMol avec la commande: **pymol**
 - c. Supposons que le fichier PDB d'un modèle se nomme « model_1.pdb », taper les commandes suivantes dans pymol dans la fenêtre noir en bas à gauche:
 1. **load model_1.pdb**
 2. **load 2b6c_A.pdb**
 3. **super model_1, 2b6c_A** (utiliser la touche **Tab** pour compléter les noms)

Appuyer sur **Echap** pour voir les résultats de la commande.

Quel est le RMSD entre le modèle et son template ?

A remplir

4. **hide everything**
5. **show ribbon**

Comparer le squelette peptidique entre les deux structures

6. Dans la deuxième fenêtre de PyMol avec les menus, choisir *Display* → *Sequence*

On peut aussi changer la représentation des résidus: *Display* → *Sequence Mode*

7. En sélectionnant des résidus dans les deux séquences on peut afficher leur chaînes latérales via le menu local (click droit => *show* => *sticks*). Ceci est surtout intéressant pour les positions où les acides aminés ne sont pas les mêmes.

Avec le bouton droit de la souris on peut zoomer et avec le bouton du milieu (=la molette) on peut déplacer la protéine

8. Fermer/Réouvrir pymol et répéter les mêmes commandes pour les autres modèles. Comparer les modèles entre eux aussi, ainsi que contre la structure de référence 3BVS.pdb
9. Est-ce que vous retrouvez les résultats décrits dans l'article biblio/3BVS.pdf (en jaune sur la deuxième page) et biblio/3BVS_supp.pdf (figure S1) ?
10. Comment est-ce qu'on pourrait expliquer les différences décrites dans biblio/3BVS.pdf (en jaune sur la deuxième page) ?