Nous allons aujourd’hui discuter de nos résultats concernant le projet sur les dimères et tétramères. Le projet portait donc sur les homéomères, qui sont des structures quaternaires protéiques formées d’un groupe de protéines où une même protéine se lie à une ou plusieurs copies d’elle-même. Il s’agit d’un phénomène répandu et pourtant les mécanismes évolutifs conduisant à leur création sont encore méconnus.

L’article de Levy et all, *Assembly reflects evolution of protein complexes*, publié en 2008, présente les résultats d’une expérience impliquant 2000 structures différentes étudiées dans le but de trouver une explication à ce problème. Dans cet article, les auteurs proposent l’hypothèse d’une évolution à temps biparti des tétramères, c’est-à-dire que la création des deux types d’interfaces ancestrales et récentes ne s’est pas faite de façon simultanée, mais au contraire successivement. L’interface récente serait donc apparue dans un deuxième temps.

Nous nous sommes alors demandé si les deux types d’interfaces avaient des propriétés différentes permettant d’apporter du crédit à cette hypothèse de formation successive des interfaces. Pour répondre à cette question, nous décrirons dans un premier temps les données à notre disposition et la façon dont nous les avons exploitées pour montrer leurs propriétés et les analyser. Dans un deuxième temps, nous verrons les résultats des analyses et leur interprétation. Pour finir, nous conclurons en parlant de (…)

Dans ce projet, nous avons eu l’occasion d’étudier un jeu de données comportant des dimères, qui sont des homéomères composés de 2 sous-unités, et des tétramères, qui sont eux composés de 4 sous-unités. Ces structures se présentent sous forme de fichiers PDB où 21 couples d’homéomères associent à un dimère son tétramère homologue. D’autres fichiers PDB reprennent les mêmes structures mais avec un bfactor modifié, qui prend la valeur 0 si l’acide aminé appartient à la surface, 1 si c’est l’interface récente et 2 pour l’interface ancestrale.

La manipulation de ces fichiers nous a permis d’obtenir de nombreuses informations sur les complexes. Tout d’abord, les scores de conservation des acides aminés ont été obtenus pour la chaîne A de chaque dimère et tétramère, sans modification de paramètres. Un score ConSurf est calculé pour chaque acide aminé du fichier PDB donné. Si le score est faible, on sait que l’acide aminé est fortement conservé au cours de l’évolution, tandis que si le score est élevé, cela correspond à une faible conservation. Pour trouver ces scores, l’outil recherche des homologues proches de la séquence donnée et effectue un alignement multiple. Un arbre phylogénétique est alors construit et enfin, les scores sont donnés.

Le langage de programmation Python a été utilisé à plusieurs reprises pour ce projet. Le script colour\_from\_consurf\_scores.py a permis d’obtenir les commandes permettant de représenter les scores ConSurf des homéomères sur le logiciel de visualisation PyMOL. La conservation est indiquée par une coloration allant du bleu si elle est faible au rouge si elle est forte.

Le script Python écrit lors des cours pour générer des fichiers csv a été modifié pour obtenir des fichiers similaires pour chaque monomère. Les fichiers obtenus contiennent des informations sur la taille de la surface et des interfaces, leur composition en acides aminés et le nombre des types d’acides aminés hydrophobes, polaires et chargés. Les scores ConSurf ont également été extraits pour analyse.