

PROJET BISII - Projet Dimères/Tétramères

Contexte

Les homomères consistent en des assemblages protéiques formés de la même sous-unité présente en plusieurs copies (2 copies = homodimère, 3 copies = homotrimère, 4 copies = homotétramère etc). Les homomères sont très fréquents dans la nature, pourtant les mécanismes évolutifs survenant dans leur apparition demeurent encore peu connus. Dans ce projet, pour plus de simplicité, on se focalisera sur l'exemple des homotétramères présenté en Figure 1. Comme vous pouvez le constater, les homotétramères impliquent quatre interfaces différentes réparties en deux types différents (ie, les interfaces bleues identiques entre elles et les interfaces jaunes identiques elles aussi).

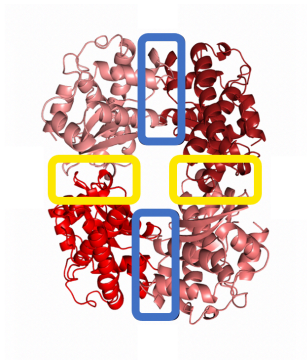


Figure 1 : l'homotétramère (PDB code: 3mds) présente 4 interfaces de deux types différents (bleu et jaune respectivement). Les interfaces bleues sont identiques et impliquent les mêmes résidus, de même que les interfaces jaunes sont identiques entre elles. Chacune des 4 sous-unités du complexe comprend donc 2 sites d'interaction impliqués respectivement dans une des interfaces bleues ou dans une interfaces jaunes. L'homotétramère 3mds peut être vu comme un dimère de dimères.

En 2008, Levy et al, par des analyses de structures quaternaires et de comparaison de séquences, ont proposé un modèle évolutif du chemin ayant donné naissance à ces homotétramères via un processus en deux étapes. En effet, ces derniers ont raisonné qu'il était plus parcimonieux de créer une interface puis une autre que de créer les deux types d'interfaces simultanément. Les auteurs proposent ainsi une première étape consistant d'abord à la formation d'un homodimère via l'apparition d'une première interface (qu'on peut ainsi qualifier d'interface "ancestrale" puisque c'est la première à être apparue - en bleu sur la figure 2). Autrement dit, la protéine de départ en mutant, va acquérir un nouveau site d'interaction lui permettant d'homodimériser. Cette étape conduit à l'obtention d'un premier homodimère. Ensuite, au cours de l'évolution, la protéine en continuant de muter, va pouvoir acquérir un nouveau site d'interaction (en jaune sur la figure 2). Ce site présent en deux copies sur l'homodimère (un site par sous-unité) va permettre à l'homodimère de dimériser à nouveau (ie, former un homotétramère) via deux nouvelles interfaces (qu'on peut qualifier d'interfaces plus "récentes" puisqu'elles sont apparues dans un second temps - en jaune sur la figure 2). Évidemment, le processus peut continuer vers des assemblages de plus en plus complexes ou s'être arrêté avant (on peut retrouver uniquement la forme dimérique chez certaines espèces et la forme tétramérique chez d'autres - Figure 2B). De façon intéressante, les auteurs ont montré par des expériences de spectrométrie de masse que le mécanisme d'assemblage de ces homotétramères passe aussi par un chemin en deux étapes (homodimérisation pour donner un homodimère puis assemblage des deux homodimères pour former l'homotétramère) qui reflète le chemin évolutif ayant donné naissance à ces homotétramères.

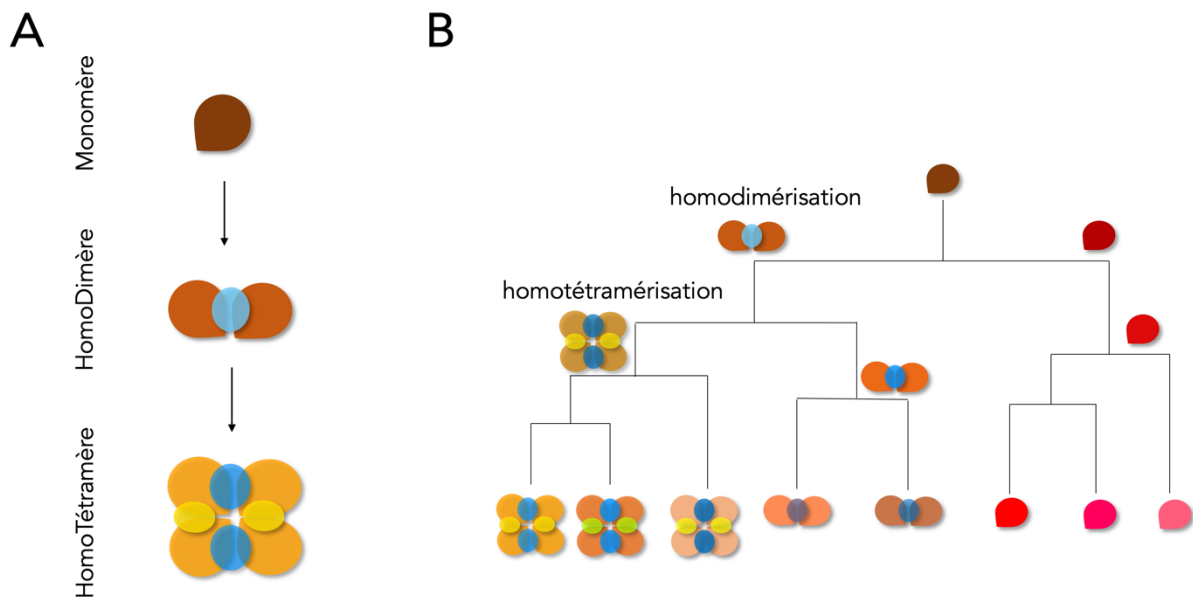


Figure 2 : (A) modèle évolutif de l'émergence d'un homotétramère en 2 étapes. La première étape consiste en l'apparition d'une première interface (bleue) permettant au monomère de départ de dimériser et de former le complexe orange foncé. Plus tard, au cours de l'évolution et suite à des mutations, une nouvelle interface va se créer (jaune) et permettre au dimère de dimériser (càd, de former un dimère de dimère) pour former un homotétramère. Ce dernier est alors constitué de 2 types d'interfaces différents, les interfaces communes au dimère et au tétramère qui sont apparues en premier et donc qualifiées "d'ancestrales" et les interfaces spécifiques du tétramère qui sont apparues dans un second temps et donc dites plus "récentes". (B) exemple de scénario évolutif. L'ancêtre (marron foncé) n'était pas capable de multimériser et existait sous la forme monomérique uniquement. Chez certaines espèces, des mutations ont conduit à l'apparition d'un site d'interaction permettant de dimériser via l'interface bleue (branche gauche de l'arbre) tandis que chez d'autres espèces, la protéine homologue est restée sous forme monomérique (branche droite de l'arbre). Ensuite, dans la branche à l'extrême gauche de l'arbre, des mutations ont conduit à l'apparition d'un nouveau site d'interaction (jaune) permettant au dimère de former un tétramère via les interfaces jaunes. Dans la branche du milieu, la protéine homologue n'a pas évolué de site jaune et ne forme que des dimères. Les gradients de couleur illustrent de façon schématique l'homologie de séquence entre les différentes protéines de cette famille (protéine actuelles (feuilles de l'arbre) ou protéines ancestrales (noeuds de l'arbre)).

Travail demandé

Dans ce projet, on se propose de comparer les propriétés des interfaces "ancestrales" (ou interfaces "communes" car elles sont communes au dimère et au tétramère) avec celles des interfaces plus "récentes" (ou interfaces spécifiques du tétramère) afin de voir si elles présentent des propriétés différentes et ainsi (si différence il y a) de mieux comprendre les déterminants structuraux ayant permis l'émergence des premiers homodimères puis des homotétramères.

Pour cela, nous vous avons préparé un jeu de données constitué de couples de protéines homologues formant soit un homodimère, soit un homotétramère (nous parlerons de dimère et de tétramère pour plus de simplicité). Pour chaque couple d'homologues, nous vous fournissons le fichier PDB de l'homologue formant le dimère et de celui formant le tétramère. Nous avons déjà identifié les résidus appartenant aux interfaces "ancestrales" et "récentes" et reporté l'information dans le bfactor de chaque pdb ce qui vous permettra d'extraire et de manipuler facilement les résidus des interfaces "ancestrales" et "récentes". Le complexe du dimère implique une seule interface - les résidus participant à cette dernière ont une valeur de bfactor égale à "1", les autres résidus ont une valeur de "0". Le complexe du tétramère implique,

quant-à-lui, deux types d'interfaces, les interfaces "ancestrales" et les interfaces "récentes" spécifiques au tétramère. Les résidus participant à la première ont une valeur de bfactor égale à "1" (vous pourrez vérifier sous pymol que ces résidus correspondent globalement à ceux vus aussi dans l'interface du complexe dimérique - "globalement" car la séquence du monomère formant le dimère et celle du monomère formant le tétramère a muté depuis l'ancêtre commun et ne sont pas les mêmes, ainsi une valine dans le monomère formant le dimère peut correspondre à une leucine dans le monomère formant le tétramère). Ceux participant à l'interface "récente" ont une valeur de bfactor de "2" - le reste des résidus a une valeur de "0". A partir de ce jeu de données et des descripteurs que nous avons vus en cours, nous vous demandons d'analyser les propriétés des différents types d'interfaces "ancestrales" VS "récentes". Des analyses intégrant de l'information évolutive sont les bienvenues (on rappelle que ConSurf permet de calculer le niveau de conservation des résidus d'une protéine et reporte cette information dans le bfactor de la protéine d'intérêt). Une piste est de regarder le niveau de conservation des différents types d'interfaces.

Pour aller plus loin : vous pouvez aussi regarder chez les complexes tétramériques, les propriétés des résidus de l'interface récente et les comparer avec les résidus qui leur correspondent (i.e. qui s'alignent avec) dans le dimère (résidus exposés qui n'appartiennent pas à une interface) - sont-ils moins conservés dans le complexe dimérique, sont-ils de type "surface" ?

Rendu

Vous présenterez le contexte de l'étude, votre Matériel et Méthodes et les résultats de votre analyse dans un rapport de 10 pages maximum et lors d'une présentation de 10 minutes.

Détails données

Toutes les données nécessaires à la réalisation du projet sont disponibles ici : <https://drive.google.com/drive/folders/1yXuqzy41sSL8VI9E1dFTJr3CD-tKsZNq?usp=sharing>

- **Dim_Tet_interfaces** : répertoire comprenant pour chaque couple d'homologues, le fichier PDB du monomère formant le dimère et celui du monomère formant le tétramère avec le statut de chaque résidu comme "non-interface", "interface_ancestrale" ou "interface_récente" dans le bfactor (valeurs 0, 1 et 2 respectivement). Pour une manipulation plus aisée des PDB et faciliter l'analyse, nous n'avons stocké qu'un seul monomère pour chaque PDB. Pour voir les complexes dimériques ou tétramériques correspondant, merci de consulter les répertoires Dimers et Tetramers.

- **Dimers** : répertoire comprenant les fichiers PDB de chaque dimère.

- **Tetramers** : répertoire comprenant les fichiers PDB de chaque tétramère.

- **pymol_sessions** : répertoire comprenant pour chaque couple d'homologues une session pymol où le dimère est aligné avec son homologue tétramérique ainsi qu'avec les monomères extraits du dimère et du tétramère et colorés selon le bfactor. Ce répertoire peut être utile pour analyser visuellement quelques cas, peut servir de base pour préparer des figures et illustrer votre rapport et/ou présentation ou tout simplement pour vous approprier les données.

- **couples_dimers_tetramers** : fichier comprenant la liste des couples d'homologues. 1ère colonne : code pdb des dimères / 2nde colonne : code pdb des tétramères.

- Levy_Nat08.pdf : article sur lequel s'appuie le projet.

Références

E Levy, E Boeri Erba, CV Robinson, and SA Teichmann, 2008. Assembly reflects evolution of protein complexes. Nature, 453(7199): 1262–1265. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2658002/>