

Rapport de projet

Bertrand BOUVAREL, Alexandre LECOEUR

Janvier 2019

1 Introduction

Avant la mise en place de la dynamique moléculaire, une étape cruciale de la modélisation consiste à minimiser l'énergie de la molécule étudiée. En effet, si la stabilité d'une molécule dépend de son énergie et, plus celle-ci sera élevée, moins la molécule sera stable, ce qui peut provoquer une "explosion" lors du lancement de la dynamique causée par des interactions qui n'ont pas lieu d'être. Il est donc aisé de comprendre l'importance de la minimisation de l'énergie d'une molécule.

Une étape de cette minimisation de l'énergie concerne les interactions électrostatiques comprises dans les énergies non liées. En effet, la séquence des peptides présente des acides aminés tels que les histidines, les asparagines et les glutamines qui dans le cas des deux dernières en bout de chaînes latérales des NH_2 et des O pouvant se positionner de façon inverse, ce qui, dans le cas des interactions électrostatiques, peut induire des énergies très différentes entre deux configurations différentes.

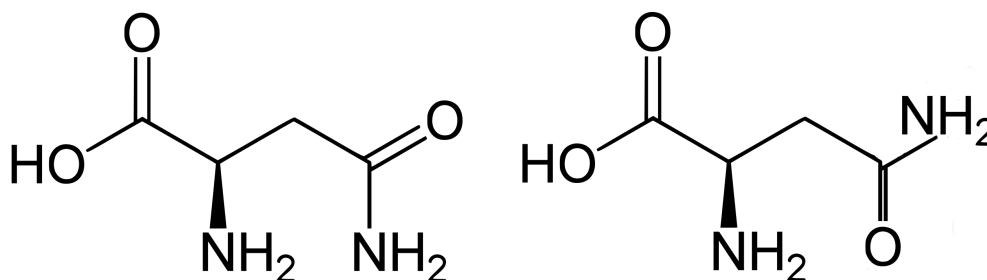


FIGURE 1 – Flip d'une asparagine

À partir d'un fichier pdb contenant la séquence d'une protéine et d'un fichier topologique contenant les charges partielles des atomes de l'ensemble des acides aminés, notre programme a donc pour but dans un premier temps de collecter les acides aminés susceptibles de réaliser un flip, de calculer l'énergie électrostatique avec les atomes lourds se trouvant à moins de 6Å, de comparer ces énergies à l'énergie électrostatique du flip correspondant et ainsi choisir la conformation fournissant la plus faible énergie afin de minimiser l'énergie électrostatique et sortir le fichier pdb avec les modifications réalisées, ainsi qu'un affichage dans le terminal

du nombre de flip réalisées pour chacun des deux acides aminés par rapport à leur nombre total.

2 Méthode

Pour ce programme, nous n'avons quasiment pas utilisé de modules externes, seul le module math nous a été nécessaire, notamment pour obtenir des valeurs telle que Pi ou encore réaliser des calculs mathématiques telles que des racines carrés de façon direct.

Pour l'obtention des acides aminé, nous avons utiliser le format pdb qui, comme il reste constant peut importe la protéine concerné, nous permet de récupérer les acides aminé en fonction de la composition de la ligne (exemple : ATOM en ébut de ligne, etc...)

Listes des fonctions créées :

- calc_distance3D(A, B) : évalue la distance entre les atomes A et B (en ångström) en utilisant les coordonnées x, y et z des chaque atomes

- trouve_charge(resi, atome, file) : recherche et renvoi la charge partielle de l'atome d'un résidu

- trouve ASN_OD1_ND2_GLN_OE1_NE2(nom_pdb) : recherche et renvoi une liste des atomes d'azotes et d'oxygènes impliqués dans les flips des asparagines et des glutamines présent dans la séquence de la protéines

- trouve_atom_N_O(nom_pdb) : cette fonction recherche dans le fichier pdb et renvoi une liste des lignes des atomes lourd présents à moins de 6 Å de l'atome transmis à la fonction, dont de chaque atome de la liste récupéré dans la fonction précédente

- def liste_to_coord(liste) : cette fonction utilise la liste de la fonction précédente et renvoi une liste de coordonnées des atomes concernés

- def fct_calcul_energie(nom_residue, nom_atome_O, liste_residue_O, nom_atome_N, liste_residue_N) : cette fonction permet à partir de l'ensemble des information obtenues avec les fonctions précédentes pour évaluer l'énergie électrostatique des azotes et des oxygènes, et ainsi les comparé à l'énergie électrostatique de la version flip afin de minimiser l'énergie électrostatique de chaque acides aminés concerné. Pour évaluer cette énergie, nous utilisons la fonction mathématique suivante :

- def ecrire_fichier_pdb_avec_flip(fichier_pdb, fichier_ecriture, coord_dans_ordre_ASN, coord_dans_ordre_GLN) : Cette fonction va copier le fichier pdb donné et le réécrire dans le fichier pdb nommé fichier_ecriture. Il va faire le switch des coordonnées si il ya besoin.

$$V_{elec} = \sum_{i=1}^{N-1} \sum_{j=i+1}^N \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}}$$

avec N le nombre d'atome à moins de 6 Å, q les charges partielles, ϵ_0 la permittivité du vide (= 1) et r les distance entre les atomes

3 Application

La comparaison de notre outils avec un outils existant est assez difficile car notre outils demande certaines améliorations, comme la prise en compte de l'encombrement stérique ainsi que le calcul d'autres énergies comme les énergies de Van der Waals par exemple. En résulte lors de la comparaison par rapport à un programme extérieur un nombre de flip plus important du côté de notre programme dû au fait que lors du calcul de l'énergie, nous n'utilisons que l'énergie électrostatique, alors que d'autre programmes utilisent l'énergie de Van der Waals qui, dans certain cas de flip, augmente l'énergie non liante et donc défavorise le dit flip.

Mais il réalise parfaitement ce pourquoi il a été créer, c'est à dire faire un flip des N et O de l'asparagine et de la glutamine en fonction de l'énergie électrostatique la plus faible.

Cela est illustrer par les deux images ci dessous qui représente l'Asparagine 68 de la protéine C, avant et après le flip.

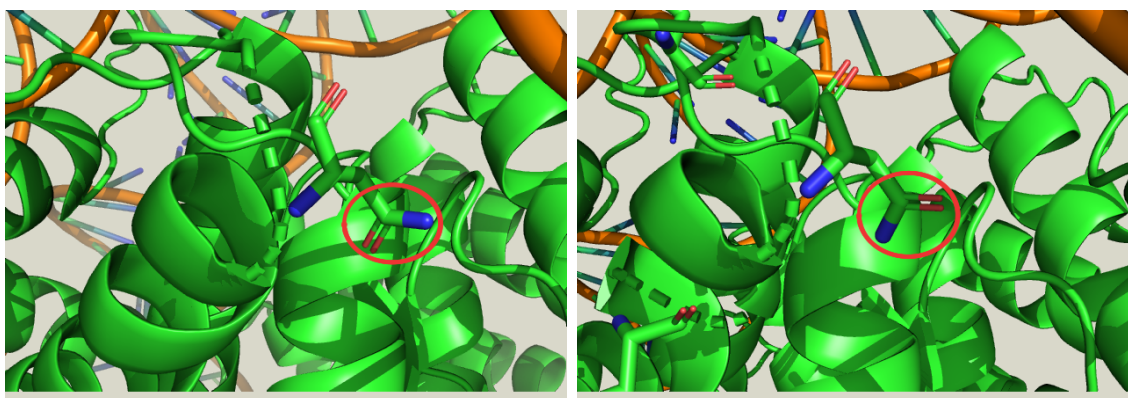


FIGURE 2 – Visualisation du flip de l'asparagine 68 de la protéine C avec pymol

4 Conclusion

Avec ce projet, nous avons donc pu créer un programme qui, à partir du calcul de l'énergie électrostatique, permet la modification des la structure 3D de la protéine en réalisant un flip pour les acides aminés Asparagines et Glutamines quand cela permet la réduction de l'énergie électrostatique avec l'obtention d'un fichier pdb modifié. Deux axes principaux d'amélioration du programme peuvent être envisagés pour assurer un résultat optimal.

Tout d'abord inclure l'histidine, en effet, nous avons traiter les cas de flip de la glutamine et de l'asparagine qui sont relativement "simple" impliquant uniquement des atomes, l'azote et l'oxygène, facilement accessible via le fichier pdb. Contrairement à ces acides aminés, le flip de l'histidine implique une modification des doubles liaisons du cycle de l'histidine.

Un autre moyen d'amélioration de ce programme est le calcul de d'autre énergie pour juger de l'avantage d'un flip. En effet, nous considérons ici uniquement l'énergie électrostatique pour le flip. Cependant, inverser les positions d'un NH2 et d'un O peut induire d'autre changements que les interactions électrostatique. On pourrait imaginer le cas d'un inversion

induisant un encombrement stérique et induire de gros problèmes lors du lancement de la dynamique moléculaire. Cet encombrement stérique pourrait être prévu et donc évité en calculant l'énergie de Van der Waals qui dans ce cas, indiquerait que le flip n'est pas avantageux pour la minimisation de l'énergie de la protéine. Calculer l'ensemble des énergies non liées permet de confirmer ou infirmer l'utilité d'un flip et représenterait donc un axe majeur d'amélioration de notre programme.