Université Libre de Bruxelles

Biophysics II

BIOL-F-459

Task 3 - Modeling of the structure of the dimers

Authors:
Hélène Libion
Charlotte Nachtegael

May 2016

1 What is ClusPro ?

ClusPro is a automated, web-based program for protein-protein docking. The program first performs a rigid-body docking. They select the 1000 rotation/translation combinations that have the lowest score. Then they cluster the structures with a greedy clustering based on a value of rmsd of 9 angström. They define a center of cluster as the structure with the most neighbours within a rmsd of 9 angström. This structure and its neighbours are taken out of the set and defined as a cluster. The operation is repeated until all the structures were clustered.

The clusters are ranked in function of the size of the cluster.

You can also adjust the coefficient weights of the energy function, so that you can privilege what you consider the most important factors for the docking, such as the van der waals or electrostatic interactions for example.

2 Test of performance

To test the performance of Cluspro we first submitted the structure of the chain A of 3MQ1 to Cluspro. Then we compared the results obtained with the experimental structure of the chain A and B of the dimer which is available on PDB.

For this, we took the ten first models given by Cluspro and firstly submitted it to PDBEFold with the experimental structure to superimpose them. We looked at the the percent of matched SSEs (%sse) which indicates the percent of secondary structure of 3MQ1 chain that was identified in the model. As we can see in the figure 2, it is 100% only for the model 3. Indeed the model 0, 1, 2, 4, 5, 6 and 9 have a %sse of 50%, the model 7 is 67%, the model 8 is 83%. So that means that only the chain of 3MQ1 was perfectly identified in the model 3. So we made the hypothesis that it is the best model because it covers the most the chains of 3MQ1. Furthermore, against PDEBEFold, the RMSD of the model 3 is 1.96$\mathring{A}$. So the structure is closed to the experimental one.

To confirm our hypothesis we used Pymol to superimpose all the models with the experimental structure. It confirms that the model 3 is the closest as we can see on the figure 1.

We also used Pymol to compute the RMS (see figure 2). The RMS is the Root Mean Square. As the RMSD it measures the accuracy of a model to 3MQ1. The

lowest it is the closest the structure of the model is to 3MQ1. As we can see the lowest RMS is 1.970 Åfor the model 3.

Cluspro proposed as the closest model the model 0 which is the most populated model. But we saw by our analysis that the closest is actually the model with the lowest energy structure, model 3 in our case.

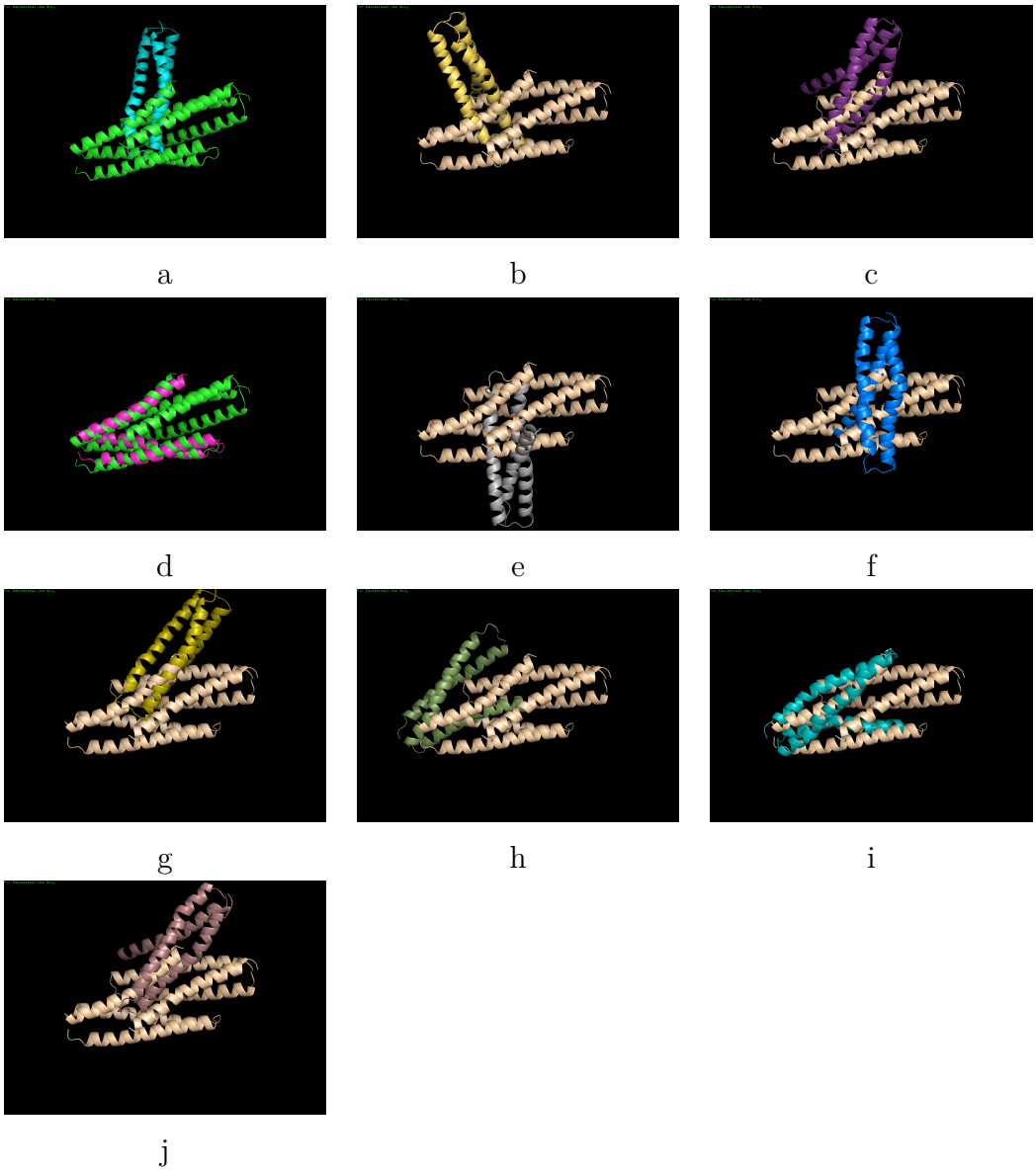


Figure 1: Superimposition with Pymol

So the model 3 is the closest to the experimental structure. more native and not the most populated.

Model	%sse	RMS
0	50%	14.261Å
1	50%	10.679Å
2	50%	16.524Å
3	100%	1.970Å
4	50%	12.112Å
5	50%	15.671Å
6	50%	16.043Å
7	67%	12.981Å
8	83%	12.483Å
9	50%	18.622Å

Figure 2: Tabular of the results obtained with PDBEFold and Pymol

3 Der f 5 based on Der p 5

During the previous task, we modelled Der f 5 using both the alleged dimer Der p 5 and the monomer Blo t 5. We observed better normalized DOPE scores when we used Blo t 5 as template compared to Der p 5. This could mean that Der f 5 has a tendency to be more a monomer than a dimer.

We ran ClusPro using the best model obtained when using Der p 5 as template to see if we would be able to obtain a dimer similar to the native one of Der p 5. We then selected the structure of the biggest cluster and the structure of the cluster with the lowest energy. We superimposed the structure with the one of Der p 5 the aid of PDBEFold.

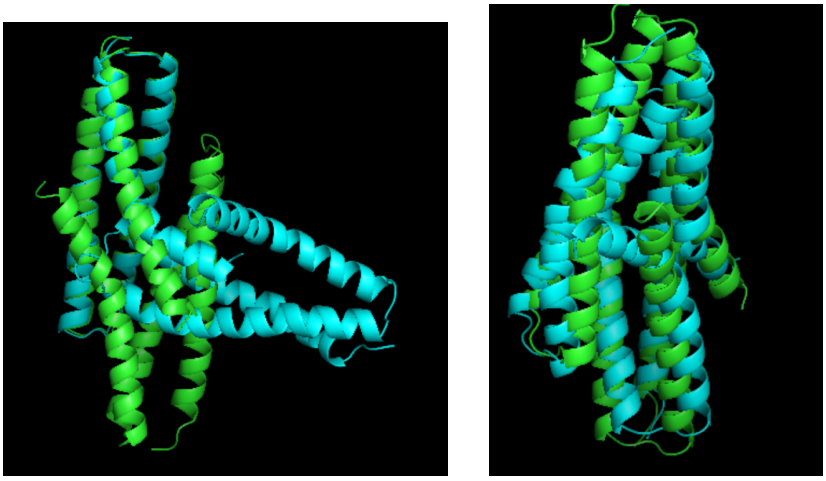


Figure 3: Superimposition of the model in blue of the biggest cluster (left) and the cluster with the lowest energy (right) with the native structure of the dimer Der p 5 in green.

The structure of the biggest cluster does not approach at all the native structure. However, the structure of the cluster with the lowest energy is a very good approximation of the native structure, with a RMSD of 3.595 (Fig 3).

4 Blo t 5 based on Der p 5

Blo t 5 is a monomer. During the previous task, we built a model of Blo t 5 based on the dimer Der p 5. We confirmed the structural difference of an open helix between the native structure of Blo t 5 and the model built. We wanted to see if

the structural difference make the dimerization possible.

In this case, the biggest cluster was also the cluster with the lowest energy. When superimposed with the native structure of Der p 5, we confirmed the structure were very close, with a RMSD of 4.947. So we can imagine that this open helix is a main component of the dimerization (Fig 4).

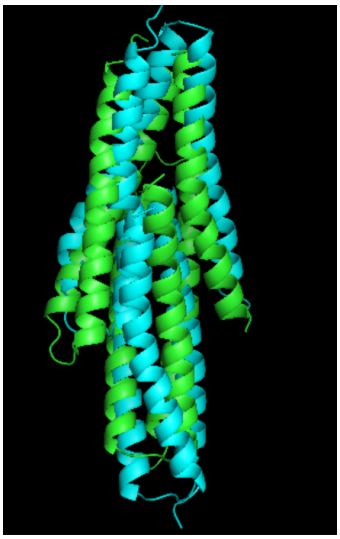


Figure 4: Superimposition of the model of the biggest cluster and lowest energy in blue and the native structure of the dimer of Der p 5 in green

5 Comparison between Der f 5 and Blo t 5 ClusPro simulation

We could observe differences between the scores obtained on ClusPro for the two models (Fig 5). The scores obtained for the clusters for Der f 5 were better than the ones obtained for Blo t 5. This could mean the very nature of monomer of Blo t 5, despite the fact the structure was based on the dimer Der p 5, is somewhat conserved.

Cluster	Members	Representative	Weighted Score
0	160	Center	-1045.6
		Lowest Energy	-1045.6
1	136	Center	-1031.2
		Lowest Energy	-1058.4
2	106	Center	-956.7
		Lowest Energy	-1036.1
3	90	Center	-915.5
		Lowest Energy	-1018.9
4	51	Center	-1143.0
		Lowest Energy	-1143.0
5	43	Center	-867.7
		Lowest Energy	-1000.9
6	42	Center	-970.0
		Lowest Energy	-970.0
7	40	Center	-924.6
		Lowest Energy	-968.9
8	40	Center	-954.7
		Lowest Energy	-1008.2
9	31	Center	-887.5
		Lowest Energy	-975.9
10	29	Center	-896.7
		Lowest Energy	-931.8

Cluster	Members	Representative	Weighted Score
0	105	Center	-830.8
		Lowest Energy	-1045.4
1	98	Center	-714.6
		Lowest Energy	-1007.1
2	65	Center	-713.3
		Lowest Energy	-841.6
3	59	Center	-726.4
		Lowest Energy	-862.2
4	51	Center	-705.3
		Lowest Energy	-771.1
5	48	Center	-743.7
		Lowest Energy	-758.5
6	37	Center	-778.4
		Lowest Energy	-899.6
7	35	Center	-770.5
		Lowest Energy	-849.4
8	31	Center	-700.0
		Lowest Energy	-855.8
9	30	Center	-719.7
		Lowest Energy	-878.8
10	30	Center	-714.8
		Lowest Energy	-905.6

Figure 5: Scores of the top 10 clusters obtained with the model of Der f 5 (left) and Blo t 5 (right)

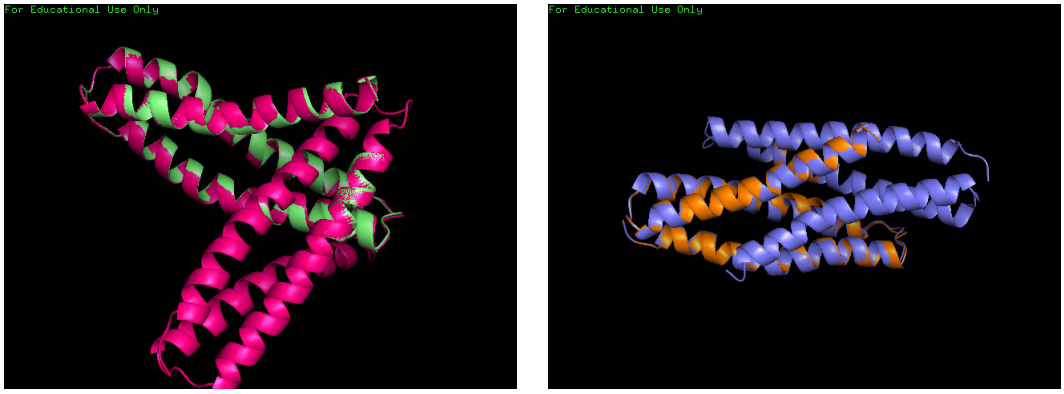
6 Mutated Der p 5

We mutated during the previous task the protein Der p 5 two valine residues at position 85 and 88 into alanine to delete the valine zipper. This was done in the aim to see if the valine zipper was essential to the dimerization.

We submitted the mutated protein to Cluspro and compared the clusters obtained with those of 3MQ1 discussed in the section **??**. We chose to firstly compare the most populated clusters of both models (model 0 in both structure) and secondly the one with the lowest energy structure in both models (model 3 in both structure).

We first submitted them to PDBEFold to align them. We could see that the results are the same in both cases. We saw that the percent of matched SSEs (%sse) was 100%. This means that all the secondary structure of 3MQ1 was identified in the mutated protein. We also saw that the RMSD is 0.38$\mathring{A}$ and that the sequence identity (%seq) is about 98%. That means that the structure of the mutated protein is very close to the structure of 3MQ1 and that the chains are almost identical. As the Q-score (0.98), the P-score (17.0) and the Z-score (12 .3)

are high, we can tell that the match is significant.



a b

Figure 6: Superimposition of the structure of the most populated clusters (a) and the clusters with the lowest energy structure (b)

Then we used Pymol to display the superimposition of the two structures. The figure 6a show the superimposition of the most populated clusters of 3MQ1 (pink) and the mutated protein (green). As we can see, the two structure are very closed. One of the chain is identical so we can only see one color. We observed the same think in the figure 6b which shows the superimposition of the clusters with the lowest energy structure of 3MQ1 (purple) and the mutated protein (orange).

So we can see that the mutation of valine into alanine doesn't change the predicted structure of the dimer and the computed stability of the dimer. We can tell that the valine zipper is not essential to the dimerization.