PREGUNTES QUE ENS PODRIEN ENTRAR + COM RESOLDRE

1) Do you think the structure is correct? Can you prove it? show an image of the energies that prove it and name it p28g.png

We have to open PROSA and compare.

2) Do you think the protein PROBLEM can work as a tetramer (or dimer)? show an image that can prove it (name it p28i.png)

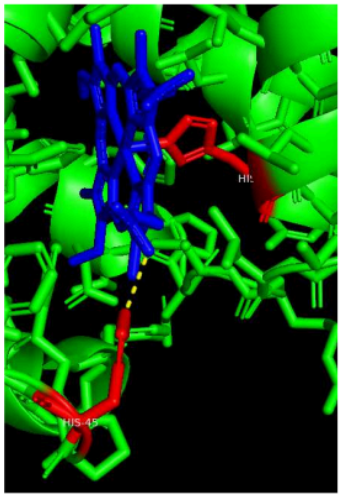
- fetch the hemoglobin
- remove waters → remove resn hoh
- rename the tetramer chains (hemoA, hemoB, hemoC and hemoD)
  select hemoA, chain A
- Now we will superimpose the target monomer on the chain A of the tetrameric:
  super target, hemoA
- Save the target monomer in this new set of coordinates as a new PDB file:
  save target_A.pdb, target
- Repeat this same procedure for chains B, C and D.
- Open all the target pdb files generated to see the tetrameric form

It should work as a tetramer, because the chains are not clashing!
It should NOT work as a dimer, because the chains are clashing!

3) Even if it was a monomer, do you think it will work with the function you selected in answer "a"? Use a sequence alignment (ie. you can reuse p28b.aln) to show the conservation non-conservation of functional residues. Mark the functionally important amino-acids (conserved or non-conserved) in the alignment with the symbol # at the bottom.

We must find the aa that are essential for the correct functioning of our target protein. Thus, we need to find an homologous PDB structure that contains the substrate in the active site. Then, we just find the residues that are interacting with the substrate and search them in the alignment.

The residues that are interacting with the hemo group are HIS45 and HIS87



```
target          vlspaDKTNVKAAWGKVGAHAGEYGAEALERMFLSFPTTKTYFPHF-DLS----HGSAQV
2B7H_1|Chains   vlspaDKTNIKSTWDKIGGHAGDYGGEALDRTFQSFPTTKTYFPHF-DLS----PGSAQV
1FHJ_1|Chains   vlspaDKTNIKSTWDKIGGHAGDYGGEALDRTFQSFPTTKTYFPHF-DLS----PGSAQV
2QLS_1|Chains   vlspaDKTNIKSTWDKIGGHAGDYGGEALDRTFQSFPTTKTYFPHF-DLS----PGSAQV

target          KGHGKKVADALTNAVAHV--D-DMPNALSALSDLHAHKLRVD--DPVNDKLLSHCL-lvtl
2B7H_1|Chains   KAHGKKVADALTTAVAHL--D-DLPGALSALSDL--HAYKLrVDPVNFKLLSHCLLvtla
1FHJ_1|Chains   KAHGKKVADALTTAVAHL--D-DLPGALSALSDL--HAYKLrVDPVNFKLLSHCLLvtla
2QLS_1|Chains   KAHGKKVADALTTAVAHL--D-DLPGALSALSDL--HAYKLrVDPVNFKLLSHCLLvtla
```

4) This protein can bind ligands such as drugs or hormones. Find an homologous structure that contains one of these ligands. Name it ligand_template.pdb. (practical 3)

Ir mirando templates y ver en el pdb. El que hemos usado anteriormente ya contiene lo que queremos: 3d24

5) Locate the amino acids you identified in the previous question in the sequence alignment you created in question d. Are these amino acids conserved or not? Explain why you think this is happening?

6) Do you think that the interaction that you modeled in the previous question could happen in nature? Select some amino acids from SH2 that you think are important for the interaction between the two proteins. Show images of these amino acids to proof your answers. Name them aa_picture1.png, aa_picture2.png, and so on. (taking images with pymol)

Raones en funció de les seves propietats, cadenes i dependent de si hi ha clashes.

7) Search for the amino acids that you selected in the previous question in the MSA you created in question d. Are these amino acids conserved? Discuss the degree of conservation of these amino acids and why do you think this is happening.

8) This protein can bind ligands such as drugs or hormones. Find an homologous structure that contains one of these ligands. Name it ligand_template.pdb. Use the structure you obained to reconstruct the interaction between our protein of interest and the ligand included in that structure.

9) We want to study the L253R mutation in SH2. Model the structure of SH2 containing this mutation. Be aware that the structure that you have is not complete, therefore the first amino acid corresponds with position 232. Name your model as mutant_model.pdb.

Alignement entre el original y el mutado, target el SH2B.fa mutado manualmente

Practice 4 - Build a model:

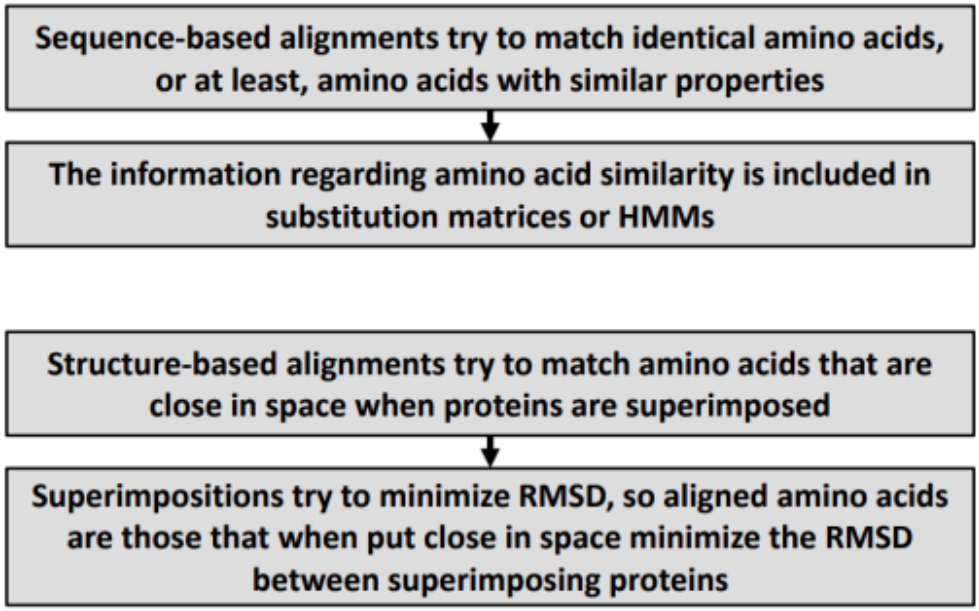
1. Find templates: We have done it already with blast.
2. Build a protein model. We need :
   - Target file (target.fa) →yes
   - Script file (modeling.py): contains the executing commands for MODELLER →yes (needs to be modified)
   - Alignment file: contains the alignment between the target and the template/s in PIR format →no

Alignment file:

```
cat target.fa > target_template.fa
cat template.fasta >> target_template.fa (template found before →download its FASTA)
clustalw2 target_template.fa
perl /shared/PERL/aconvertMod2.pl -in c -out p<target_template.aln>target_template.pir
```

mod10.5 modeling.py

10) Explain what are the differences between the alignment you obtained in question d (hmm_alignment.aln) and the alignment you obtained in question f (structural_alignment.aln).


Sequence-based alignments try to match identical amino acids, or at least, amino acids with similar properties
The information regarding amino acid similarity is included in substitution matrices or HMMs
Structure-based alignments try to match amino acids that are close in space when proteins are superimposed
Superimpositions try to minimize RMSD, so aligned amino acids are those that when put close in space minimize the RMSD between superimposing proteins