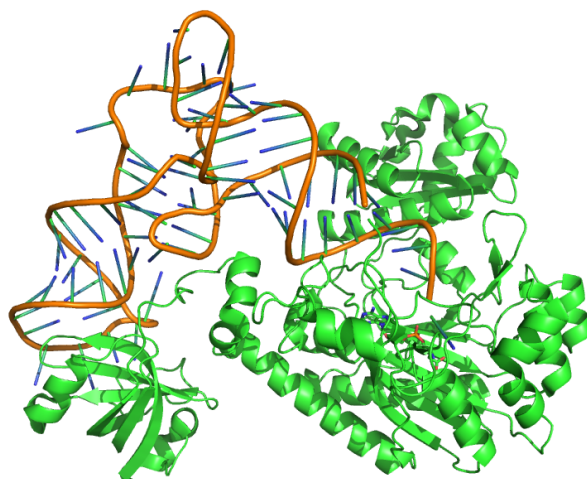# *In silico* directed mutagenesis of the aspartyl-tRNA synthetase

The objective of the TP is to study by molecular modeling the specific recognition between the aspartyl-tRNA synthetase and its substrate Asp. We will try to evaluate the specificity by comparing the binding of the ligands Asp and Asn. We will then seek to identify and model mutations in the active site that could favor the binding of Asn instead of Asp. This is a first step towards an engineering of the genetic code.

Introduction

Aminoacyl-tRNA synthetases (aaRS) are an enzyme family implicated in protein synthesis. They are involved in translation by allowing the binding of an amino acid to its transfer RNA. They are very specific of the amino acid concerned and corresponding transfer RNA. Thus there exists one for each amino acid.

We will be interested particularly in the aspartyl-tRNA synthetase (AspRS), the goal being to perform point mutations of this enzyme in order to reduce its affinity for its natural ligand aspartate and favor its binding to asparagine.

For that, we will consider the problem in terms of protein sequences and structures. The study includes three steps :

— Study of the sequences of some aaRS, including AspRS, in order to highlight the distinctive features of the AspRS.
— Inspection of the AspRS structure in order to identify amino acids in the active site that would be good candidates for a mutagenesis.
— Mutagenesis and affinity calculations by molecular modeling.

This analysis will lead to propose judicious mutations of the active site, allowing to modify the AspRS specificity by favoring Asn binding over Asp.

Protocol

A   Analysis of aaRS sequences

1. Retrieve the E. coli AspRS sequence in the UniProt (`http://www.uniprot.org`) bank

2. Obtain homologous sequences : BLAST search

The E. coli AspRS has three domains : the tRNA anticodon binding domain, the catalytic site domain, and a third one inserted into the catalytic site domain.

Launch a BLAST search. What types of proteins are found ?

3. Identify important residues : make a multiple sequence alignment

— Make a multiple alignment of the previously obtained sequences.
— Try different formats for display and coloring by amino acid.
— Repeat the operation by refining the choice of sequences.

Identify strongly conserved regions that can correspond to the active site. Choose some positions that seem distinctive of the AspRS and of Asp binding.

Which strategy have you employed ? Which mutations do you propose to modify the AspRS affinity for aspartate and asparagine ?

B   Structural analysis : inspection of the AspRS structure

— Visualize with PyMOL the structure of AspRS and its ligand.
— Locate the active site region, with the help of the sequence analysis.
— Examine the active site to refine the choice of amino acids to mutate.

With the informations previously obtained, propose judicious mutations to modify the AspRS affinity for aspartate and asparagine. We will try to test several of them in the next step of modeling.

Does the active site examination lead you to modify your mutation propositions made from the sequences ?

Can we use the structure to verify the sequence alignment ?

C   Molecular modeling study

This is the most ambitious and complex part of the TP. There are two steps :

— Estimate the AspRS affinity difference for Asp and Asn.
— Mutate one or more residues in the active site and estimate again the affinity difference.

We will follow this protocol with the XPLOR program :

1. Inspect the files available :

asprs.seq	sequence of the AspRS protein
asprs.pdb	experimental structure of the AspRS protein
asprs.xplor.pdb	experimental structure of the AspRS protein formatted for XPLOR
asp.xplor.pdb	structure of the Asp ligand formatted for XPLOR
amber.rtf	topology file for XPLOR
isolated_aa.rtf	additional topology file for isolated aa
amber.prm	parameter file for XPLOR
build.inp	model building of the protein:ligand complex
minimize.inp	energy minimization of the complex
energy.inp	energy calculation of the complex
run.sh	script to drive the calculations

2. Compare the files **asprs.xplor.pdb** and **asprs.pdb**

   The PDB files must comply with a particular format to be readable by XPLOR. The segment name has to be written on 4 characters in columns 73-76. We also note that the 3-letter code of histidines has been changed from HIS to HIE. There indeed exists 3 possible protonation states for histidines and it is necessary to tell XPLOR which state is chosen among HID, HIE, or HIP (see the topology file **amber.rtf** for the definition of these states).

   Does the HIE protonation state chosen for all histidines seem reasonable to you? If in doubt, try other protonation states and evaluate the impact on the results.

3. Build a model of the AspRS:Asp complex with XPLOR

   `xplor < build.inp > build.out`

4. Minimize the energy of the complex to improve its geometry

   `xplor < minimize.inp > minimize.out`

5. Estimate the energy of the AspRS:Asp complex, and then of each partner alone

   `xplor < energy.inp > energy.out`

   What is the affinity of the AspRS protein for the Asp ligand?

6. Edit the PDB file of the ligand **asp.xplor.pdb** to change Asp into Asn. We will simply replace one of the carboxylate oxygens of the Asp sidechain by a nitrogen (corresponding to the Asn NH2 group). It will then be easy with XPLOR to position the two missing hydrogens.

   What is the affinity of the AspRS for Asn?

   Experimentally, the wild-type enzyme binds Asp considerably stronger than Asn, with an association free energy difference of more than 7 kcal/mol. Do you find the same tendency?

7. Mutagenesis of the AspRS : choose a mutation among the candidates previously identified.

   A simple mutation (for example Asp→Asn or Gln→Glu) can be realized by editing the PDB file, as explained for the ligand.

   A more complex mutation can be performed with the SCWRL program. The mutation choice (for example R10K) is done by replacing in the **asprs.seq** file the one-letter code of the native amino acid in lowercase by the code of the amino acid chosen for the mutation in uppercase (for example replacing the lowercase "r" in position 10 by an uppercase "K"). We then launch the SCWRL program as follows :

   `scwrl -s asprs.seq -i asprs.wt.pdb -o asprs.pdb > scwrl.out`

   Compare the mutated structure obtained **asprs.pdb** with the native structure **asprs.wt.pdb**. It is recommended to work in a separate folder for each mutant.

8. Perform the affinity calculations for the mutated enzyme.

   In order to use the structure mutated by SCWRL with XPLOR, it must be ensured that it is correctly formatted. For that, we will use the pdb2xplor program as follows :

   `pdb2xplor asprs.pdb A PROT > asprs.xplor.pdb`

   What are the affinities for Asp and Asn obtained with the mutated protein?

   Have you succeeded in inverting the specificity?

   Interpret structurally the effect of the mutations.

9. Which improvements could we bring to the model or the protocol?

10. Was the AspRS the most judicious target for this engineering?