

TD5: Prediction of protein 3D structures using molecular modelisation

You will need the following programs:

- a sequence editor
- 3D structure viewer (PyMOL)

We will test an hypothesis about the role of a SNP on a gene of interest.

Here is the gene sequence:

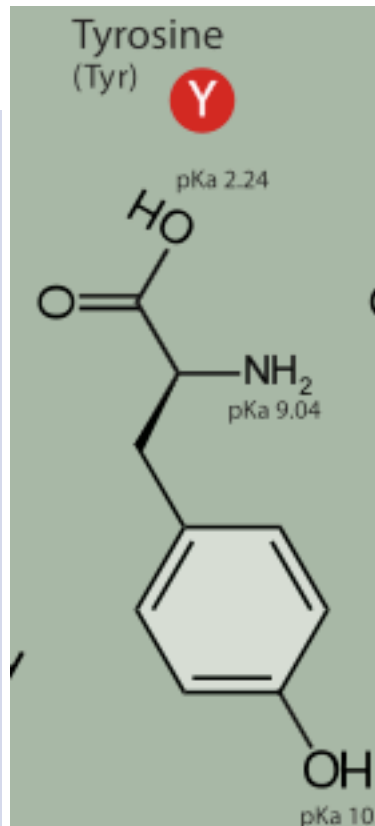
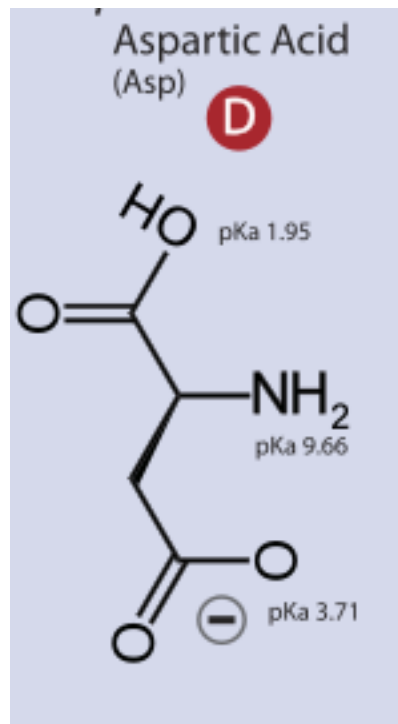
>TDR

```
MGRGDHLLMKNSNAAAAA AVNGGGTS LDAALRPLVGSDGWDYCIYWRLS
PDQRFLEMTGFCCSSELEAQVSALLDLPSIPLDSSSIGMHAQALLSNQP
IWQSSSEEEADGGGGAKTRLLVPVAGGLVELFASRYMAEEQQMAELVMA
QCGGGGAGDDGGGQAWPPPETPSFQWDGGADAQRLMYGGSSLNLFDAAAA
DDDPFLGGGGGDAVGDEAAAAGAWPYAGMAVSEPSVAVAQEQQMQHAAGGG
VAESGSEGRKLLHGGDPEDDGDGEGRS GGAKRQQCKNLEAERKRRKKLN GH
LYKLRS LVPNITKM DRASILGDAIDYIVGLQKQVKELQDELEDNHVHHKP
PDVLIDHPPPASLVGLDNDDASPPNSHQQQPPLAVSGSSSRRSNKDPAMT
DDKVGGGGGGGHRMEPQLEVRQVQGNELFVQVLWEHKPGGFVRLMDAMNA
LGLEVINVNVT TYKTLVLNVFRVMVRDSEVAVQADRVRDSLLEVTRETY P
GVWPSPQEEDDAKFDGGDGGQAAAAAAAAGGEHYHDEVGGGYHQHLHYLA
FD
```

in red, look at the SNP: D in the wild crop becomes Y in the mutant crop.
ITKM DRASIL becomes ITKYRASIL

A residue charged and hydrophilic is replaced by a neutral and hydrophobic one.

What could be the impact of this punctual mutation on the protein function:



How to get the sequence if interest

A simple copy and paste from this document or on the website of the training
Sequence_TD5

Searching data about this sequenc:

Search similar proteins in the Expasy database Swiss-Prot (<http://www.uniprot.org>).

Sequence similarity analysis using BLASTP

Select the tab *Blast* (BLASTP)

Copysefrequency or load the file and slee

BLAST

Filter byⁱ

- Reviewed (1)
- Swiss-Prot
- Unreviewed (99)
- TrEMBL
- Proteomes (63)
- Organisms
- A. thaliana (1)
- Rice (2)
- Rice (2)
- ORYGL (1)
- SETIT (5)
- Other organisms
- Go

Map To

- UniProtKB
- UniRef
- UniParc

View by

- Taxonomy
- Text version
- XML version

Demo

- Help video

Overview

Show all 100

Entry	Protein names	Match hit	Identity
Q6YUS3	BHLH protein-like (Oryza sativa subsp. japonica)	100	100.0%
B8AGK6	Putative uncharacterized protein (Oryza sativa subsp. indica)	200	99.0%
T2B0H9	Uncharacterized protein (Oryza sativa subsp. indica)	300	99.0%
I1NWN6	Uncharacterized protein (Oryza glaberrima)	400	97.0%

Alignments

Columns BLAST Align Download Add to basket

Entry	Alignment overview	Info	Status	Caution
Q6YUS3	Q6YUS3_ORYSJ - BHLH protein-like - Oryza sativa sub... - View alignment	E-value: 0.0 Score: 2,911 Ident.: 100.0%		
B8AGK6	B8AGK6_ORYSI - Putative uncharacterized protein - Oryza sativa sub... - View alignment	E-value: 0.0 Score: 2,882 Ident.: 99.0%		
T2B0H9	T2B0H9_ORYSI - Uncharacterized protein - Oryza sativa sub... - View alignment	E-value: 0.0 Score: 2,874 Ident.: 99.0%		

ctionner

It is possible to select target database (use standard settings)

Launch Blast (can last from 2 to 20 minutes ...) then look at resulting *hits*:

What are resulting proteins (function, molecular weight ..? .)

What is the largest alignment coverage and the largest identity percentage obtained? Is this the same hit ?

Example of request result:

Select Q6YUS3

Explore different types of information

Open pages in new tabs

Look Particularly at

Protein Attributes

General Annotations

In the item "Cross-references":

Launch a request ModBase

Take care of DataSet Information

Open links PFAM and Protein Model Portal

PFAM

See information

Protein Model Portal

Open HLH link (new tab)

Appreciate the quality of the model

Is there a structure of this protein =Is there a PDB access number for this protein ? If so, look for the reference on the PDB (<http://rcsb.org/pdb/>).

Is it an experimentally determined structure (X-ray crystallography...) if yes, you can go home.

If it is a theoretical model, you can try to do a better one.

Modelisation

Edit the sequence to only conserve the spotted part : region of 250-360

Sequence_TD5_eDIT

atome2

<http://atome.cbs.cnrs.fr/AT2B/meta.html>

enter a title

Paste the sequence

Leave email

Click all boxes of Search for homologous ...

Click on submit ... take a cafe

The next step is to validate one or more templates to start the

Warning: Access to @TOME2 server is restricted to academic use only !

This website is free and open to all academic users and there is no login requirement.

This website is optimized for Firefox.

[Documentation](#)

Embedded tools

*HHsearch
Fugue, PS3
Psbiblast*

Psipred, P-Sea

*Tito, Scwrl
Modeller*

*QMean
Verify3D
Eval23D*

*T-Coffee
Clustalw
Muscle
Jalview*

*MaxCluster
Profit, Matt*

*Plants
MedusaScore
XScore
DSX*

*Pat
JMol
Babel
Fconv
...*

Title :	<input type="text" value="TD5_EDIT"/>
Primary Sequence:	<div><p>[One letter code] eg: MPSHRNSNLKFCTVCASNNNRSMESHKVLQEAGYNVSSYGTGSAVRLPGLSIDKPNVYSFGT PYNDIYNDLLSQSADRYKSNGLLQMLDRNRRLKKAPEKWQESTKVFDFVFTCEERCDFDAVCE DLMNRGGKLNKIVHVINVDIKDDDENAKIGSKAILELADMLNDKIEQCEKDDIPFEDCIMDI LTEWQSSHSQPLSLYAPSY</p><div><div>AAGGGVAESGSEGRKLHGGDPEDDGDGEGRSGGAKRQCKNLEAERKRRKKLNG HLYKLRLVNPITKMDRASILGDAIDYIVGLQKQVKELQDELEDNHVHHKPPDVLIDH PPP</div></div></div>
Email :	<input type="text" value="lamotte@supagro.inra.fr"/>

☒ **Search for Homologous Sequences & Comparative Modeling**

Select tools :

☒ Psi-Blast (PDB)

☐ Options...

☒ HHSearch (PDB)

☐ Options...

☒ Fugue (Homstrad)

☐ Options...

☒ SP3

☐ Options...

Global Options:

☒ Hide redundant structural alignments (>98%) in Atome selection

☐ Ligands Selection & Complexes Prediction by Comparative Docking

☐ Modules

Ref: Pons & Labesse - Nucleic Acids Research, Web Server Issue- 2009

Generation of the 3D model

Download the model and the template

Visualization:

Start PyMOL and load model:

Is the general structure coherent ?

Try several types of representation

Rename the model

Load the template (1HLO.pdb)

Align *align PROT1, prot2*

explore PyMOL functionalities, and if necessary, consult [PyMOL Wiki](http://www.pymolwiki.org/index.php/Main_Page)
(http://www.pymolwiki.org/index.php/Main_Page)

Applications:.

Regarding the result model, what applications can, or can't, be tried ?

PMB

10.1007 / s11103-013-0166-5

Molecular Breeding

10.1007 / s11032-013-9972-3

*If you have any questions concerning this TD or that thematic research
contact lamotte@supagro.inra.fr us.*

TD predbyed by Frederic de Lamotte

NB

It is important to keep in mind that the degree of similarities / differences between amino acids are much more complex

