

## **Final Project Proposal**

Proteins are the object of interest of many researchers because they perform many functions in living organisms. For the same reason, it is practical to analyze the coding regions of the genome by means of bioinformatics tools, such as ExPASy, to get an idea of what protein could be coded by the sequenced DNA.

The purpose of the d2p project is to provide a tool that allows users to upload a DNA sequence that will be translated to the amino acid sequence and predict the rate of protein degradation based on the N-end rule. The tool will try to find the top 5 homologous proteins in the human genome using BLAST. Finally, the user will be offered to save the result in the database.

Further details of the d2p interface and functions are described below:

1. The tool's main page will contain a form with a text area and a file upload field. The user can upload the sequence either way: it will be possible to paste the sequence into the text area or upload a file in FASTA format. Whenever both fields are used, the file will be parsed, and the text area will be ignored. Additionally, the page will contain another form that the user can use to retrieve the analysis saved previously.
2. The tool will parse the DNA sequence into three amino acid sequences based on three possible reading frames and employ Bio.Blast.NCBIXML module upon submission in order to retrieve the potential homologs of each reading frame in the human genome.
3. On the result page, the user will see the original DNA sequence, the computed amino acid sequence, and the top 5 homologs, if any were found. In addition, the user can save this result in the database and receive an ID to access this information.

The following technologies will be used throughout the implementation of the proposed project:

1. The UNIX OS and filesystem for the placement and organization of files within a web application and relational database schema design.
2. Python CGI programming for implementation of the logic and results retrieval.
3. CGI and HTML templates along with the HTML5 document markup and page styling with CSS to create the frontend.
4. mysql.connector python module to connect to MySQL and MySQL queries to save and get information from the database.
5. Javascript and JQuery for some interactive elements, such as showing a small pop-up with brief information about the amino acid upon clicking on its letter.