

# IQB-1

---

Assignment 1 Of IQB

## Instructions to run the scripts

---

All scripts run with the standard command : `python Qx.py -i __inputFile__ -o __outputFile__`

## Question-1

---

For this question, we simply maintain a dictionary of all proteins corresponding to a given sequence of RNA residues. The final output is the protein sequence derived from the RNA sequence. The RNA sequence is built by replacing all the T's by U's in the given file.

## Question-2

---

For this question, we open the input file, split it line-wise and perform linear search on all the lines obtained. If the line contains 'HEADER', we extract the header directly. The same goes for TITLE. For the resolution, as specified by the PDB format, it is always found in REMARK 2. Thus, it is also extracted easily.

## Question-3

---

This was the real meat of assignment. Task is simple :

- Two input sequence are given.
- Align the two using dynamic programming algorithm **Needleman Wunsch algorithm**
- Output the alignment along with a dot plot and sum-matrix (DP Table)

### Penalties

$\Psi$	Score( $\Psi$ )
Match	1
Mismatch	0
Gap	0

$\Psi$  represents possible interaction in alignment

### Recurrence relation

Let `Sum` be our table. `first` and `second` be our sequences. `R_MAX` and `C_MAX` be last row and column of

out table.

Initially,

- $\text{Sum}[i,j] = 1$  ,if  $\text{first}[i] == \text{second}[j]$  .

$\text{Sum}[i,j] = 0$  ,otherwise

After that,

```
v1 = Sum[i+1,j+1]
v2 = -∞
v3 = -∞
for k in range(j+2,C_MAX):
    v2 = max(v2,Sum[i+1,k])
for k in range(i+2,R_MAX):
    v2 = max(v2,Sum[k,j+1])
Sum[i,j] = Sum[i,j] + max(v1,v2,v3)
```