Structure Prediction Module-worksheet

Predict 3D protein structure using ESMFold

1. What is the predicted structure of your proto-gene as determined by ESMFold? (insert picture below). Is ESMFold confident in its prediction?

2. Using Foldseek, Fill in the information about the top scoring target (ie the target with the highest TM-score)

   Target name:
   Probability:
   Sequence identity:
   TM-score:
   Picture of target and proto-gene alignment:

3. Do you find the structure alignment between your proto-gene and the top scoring target convincing? Why or why not?

Using the CATH website answer questions 4-8:

4. What is the CATH protein fold hierarchy for the top scoring Foldseek target?

5. What is the largest percentage of species that this protein fold has been found in?

6. What cellular component has the largest proportion of annotations for this protein fold?

7. What molecular function has the largest proportion of annotations for this protein fold?

8. What biological process has the largest proportion of annotations for this protein fold?