MANIPAL INSTITUTE OF TECHNOLOGY
MANIPAL
(A constituent unit of MAHE, Manipal)

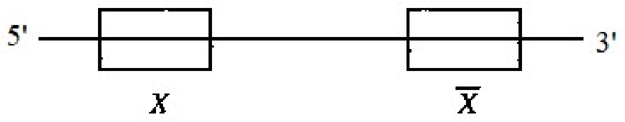
V SEMESTER B.TECH. END SEMESTER ONLINE EXAMINATIONS
JANUARY 2021

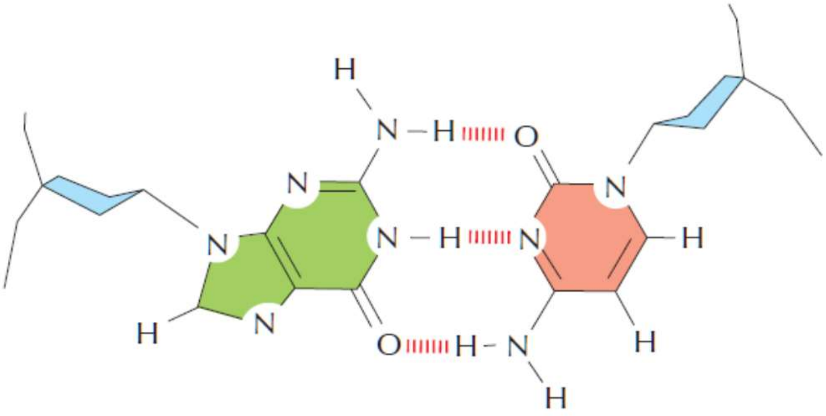
SUBJECT: [BIO 3151] Bioinformatics

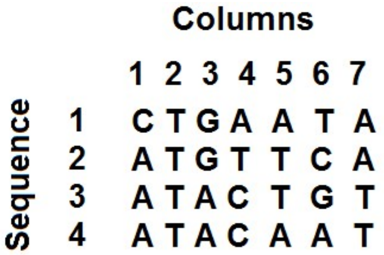
Date of Exam: **28-01-2021** Time of Exam: **2-5pm** Max. Marks: **50**

Instructions to Candidates:

- ❖ Answer ALL the questions & missing data may be suitably assumed

1A.	In general, the catalytic site of most proteins contain amino acids such as histidine. Find out whether the histidine is required to catalyze a reaction in a protonated or unprotonated state? Out of these two forms which form of histidine is required for the enzyme activity?	**2**
1B.	Which sequence is better to identify a distantly-related genes based on homology? Is it better to search for a short conserved sequence or a long sequence with more conserved patterns? Explain	**3**
1C.	The inverted repeat of the form ‘X’ is given below. Replace ‘X’ with the sequence “AGCAT” and show the inverted repeats in both the strands of DNA  Using dynamic programming approach find a global match between the two segments: ***X*** and $\overline{X}$ derived from the given strand with 5’ to 3’ orientation. Use the following scores: gap penalty of -2, match score of +2 and a mismatch score of -1	**5**
2A.	Which of the following description is **the best** for defining a protein domain? **i)** The region of the cell in which a protein functions **ii)** An evolutionarily conserved sequence of amino acids **iii)** A discrete structural and functional unit within a protein **iv)** A group of proteins that share related functions **Note:** Also explain why other given options are not suitable.	**3**

2B.	The structure and function of a protein may be affected by phosphorylation. Identify which of the following statements are correct in this context and explain the reason for which they may affect the protein conformation or function. **i)** Charge attraction between the phosphate group and positively charged amino acids cause a conformational change in the protein. **ii)** Charge repulsion between the phosphate group and negatively charged amino acids cause a conformational change in the protein. **iii)** The phosphate group creates a binding site for another protein	**3**
2C.	Describe about fragment assembly and elaborate the complexities involved with a special focus on repetitive sequences. How do you handle them?	**4**
3A.	Which secondary structure is adapted by most proteins found in the lipid membrane (bilayer), Is it α-helices or β-barrels or disordered chains? Explain your choice with an example	**3**
3B.	Which sequence is better to align, DNA or protein? Justify with examples	**3**
3C.	Distinguish homology modeling and fold recognition methods. Which is more efficient in protein structure prediction? Why?	**4**
4A.	How would you identify the major groove side and minor groove side of the given base pair? Also indicate the name of the nitrogenous base pair given below  **Note:** Instead of re-drawing the structure, only indicate which site is found at the TOP, and at the BOTTOM of the diagram?	**3**

4B.	Examine the following RNA sequence and check whether they form stem and loop structure? If it forms, then highlight those bases and draw their secondary structure **ACCCAAAGGCUCUUUCAGAGCCACCCAC**	**3**
4C.	In order to develop a gene finding algorithm that should recognize the protein coding genes from a eukaryotic genome, list out the important features that distinguish coding sequences from the non-coding DNA and also the untranslated sequences.	**4**
5A.	The only horizontal gene transfer that might have occurred in animals is from the mitochondrial genome to the nuclear genome - contemplate. Also add whether can we use mitochondrial genome to construct the universal ‘tree of life’?	**3**
5B.	Explain why the prevalence of horizontal gene transfer is more in single-celled organisms than in multicellular organisms?	**2**
5C.	Explain the ‘Maximum Parsimony’ approach for the given multiple sequence alignment. Identify the informative columns and explain the processing steps to get an optimal MP-tree  **Columns** **Sequence** 1 2 3 4 5 6 7 1 C T G A A T A 2 A T G T T C A 3 A T A C T G T 4 A T A C A A T	**5**