

**BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE, PILANI**  
**Application of Statistics and Computers in Biology/Pharmacy (BIO/PHA G510)**  
**FIRST SEMESTER 2017 – 2018**

**MID-SEMESTER EXAM (CLOSED BOOK)**

**Full marks: 50      DATE: 14.10.17      DURATION: 90 Mins.**

- **Answer to the point**
- **Answer together all parts of same question.**

**1. a)** “Sequence data of biological molecules are much more abundant than structural data though structural data is much more useful than sequence information” Explain this statement. **[2]**

**b)** What would be your inference to the following BLAST result? **[3]**

```
Score = 92.0 bits (228), Expect = 7e-20,
Identities = 28/153 (18%), Positives = 54/153 (35%), Gaps = 21/153 (13%)

Query: 14  AFEVKDAKGRTVSLEKYKGVSLVNVASDCQLTDRNYLGLKELHKEFGPSHFSVLAFFC 73
          AF + D  G+TVS   +GKV+L+      C          + +  ++  +F VLA
Sbjct: 10  AFSLPDLHGKTVSNADLQGKVTTLINFWFSPCGCVSEMPKIIKTANDYKNKNFQVLAVA- 68

Query: 74  NQFGESEPRPSKEVESFARKNYGVTFPIFHKIKILGSEGEPAFRFLVDSSKKEPRWNF-W 132
          P + V +  K+YG+ F +          +          +          +
Sbjct: 69  -----QPIDPIESVRQYV-KDYGLPFTVM-----YDADKAVGQAFGTQVYPT 109

Query: 133  KYLVNPEGQVVKFWRPEEPIEVIRPDIAALVRQ 165
          L+ +G+++K + E      +  +I  R
Sbjct: 110  SVLIGKKGEILKTYVGEPDFGKLYQEIDTAWRN 142
```

**c)** Explain the fact that in BLOSUM 62 substitution matrix Trp to Trp match score is given as +11 whereas Ala to Ala match score is only +4. **[3]**

**d)** Explain why nucleic acid scoring system is much simpler than amino acid scoring system. **[2]**

**2. a)** A pair of sequences is given to two students and they came up with following two alignments. What would be possible reason for two different alignments? **[2]**

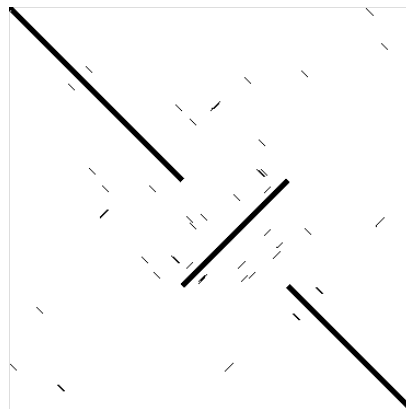
Alignment-1

```
AGATAGAAACTGATATATA
AGA-A-A-ACAGAG-T---
```

Alignment-2

```
AGATAGAAACTGATATATA
AG----AAAACAGAGT-----
```

**b)** Write the characteristic feature of following DOT plot. **[2]**



c) Draw the DOT plot of following palindromic DNA sequence and show that the characteristic feature of palindromic sequence is visible from plot. [4]

5'- T A T C G A T A -3'

3'- A T A G C T A T -5'