Dr Rout

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

MID SEMESTER EXAMINATION-2015

B. Tech. VI Semester (BI)

COURSE CODE: **10B11BI611** MAX. MARKS: 30

COURSE NAME: ADVANCED CHEMOINFORMATICS

COURSE CREDIT: 04 MAX. TIME: 2 HRS

Note: All questions are compulsory. Provide schematic diagram/figure wherever necessary in Section B & C.

Section A: Each question carries 1 mark. **(1x6=6)**

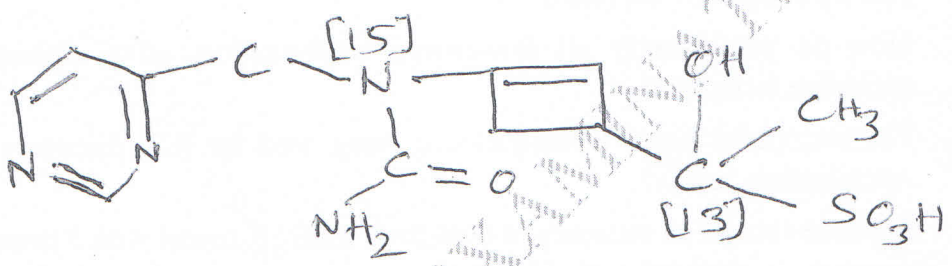
1. What are the functional aspects of chemoinformatics and how these aspects will be useful in drug design/development?
2. How do you identify all low-energy conformations of a molecule using simulated annealing technique?
3. Pharmacophore based techniques are being used for lead discovery in contrast to lead optimization, justify?
4. Suppose you get an accuracy of 80% from a QSAR model with 5 descriptors whereas 90% accuracy is obtained with 50 descriptors using similar method. Out of these two which QSAR model is more appropriate for lead molecule refinement and why?
5. How do you predict solubility of a molecule using quantitative structure-property relationship?
6. Why is internal coordinate system preferred over absolute coordinate one? What are the parameters used in internal coordinate system?

Section B: Each question carries 3 marks. **(3x3=9)**

1. Suppose you have 100 molecules (containing different functional groups) which are active against a particular receptor. How do you decide which groups are to be considered as pharmacophoric groups and why? Suppose 10 molecules contain similar pharmacophoric groups then how do you identify bioactive conformation(s) in these molecules? (1.5+1.5)
2. Suppose 20 molecules active against lungs cancer contain four pharmacophoric groups which are present in each molecule. How do you represent each conformation from each molecule in a pharmacophoric space? How do you employ constrained systematic search technique to identify pharmacophore for these molecules? (1+2)
3. Fifty molecules known to active against a target of Mycobacterium *tuberculosis*. Suppose you have extracted many descriptors from conformational, electronic, quantum mechanical, topological, graph-theoretical, etc. categories. How do you determine which descriptors are important? How do you valid such model and also how do you use this model for better lead identification? (1.5+1.5)

Section C: Each question carries 3 marks. **(3x5=15)**

1. Why do you calculate descriptors of a molecule and how do you prioritize the descriptors derived from a molecule? How do you use the 3D descriptor data for lead discovery and optimization?
2. Suppose you wish to develop a new 3D file format to represent small molecules. What are the components do you need to include in your file format so that it provides better information to researchers? How do you include 2D data so that the same file format can directly be used for faster database searching? (2+1)
3. Take any three diverse 2D chemical molecule representation formats and make a comparative analysis of their advantages and disadvantages? Why is distance metric representation considered as 3D representation? (2.5+0.5)
4. A molecule is given as follows:



Provide SMILES and Beilstein line format for the above molecule? How do you develop a query which uses SMILES of any fragment this molecule as query? (2.5+0.5)

5. You have developed a database which contains all types (1D, 2D & 3D) descriptors of small molecules. How do you calculate similarity between any two molecules using the data available in this database? Why similarity score calculated using Tanimoto coefficient method is so popular? When Tanimoto, Dice and cosine coefficients are equal? (2+0.5+0.5)