AMS Common Exam Part B, Computational Biology Track, January Exam 2007

DO THREE OUT OF FOUR QUESTIONS ONLY.

One question must be based on AMS 535 (questions 1-2) and one question must be based on CSE 549 (questions 3-4). The remaining question can be from either section.

Name:

Questions Based on AMS 535:

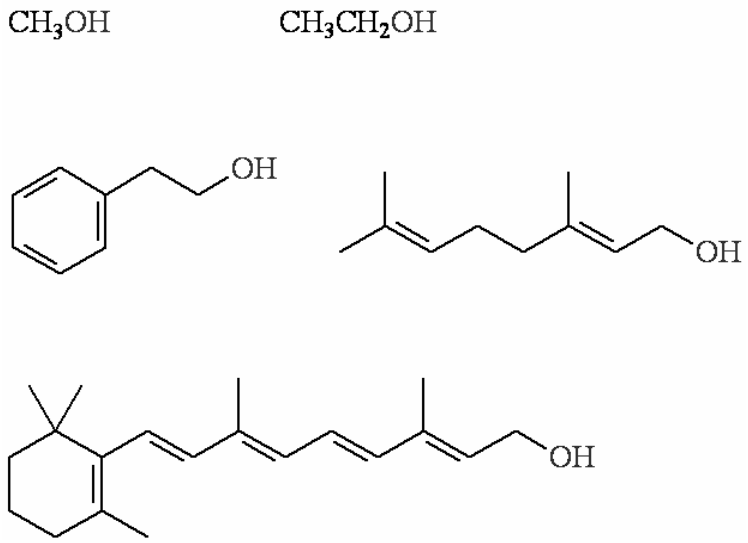
Question 1. Draw a thermodynamic cycle used to determine the relative free energy of hydration ($\Delta\Delta G_{hyd}$) between two molecules A and B. Clearly label all parts of the figure. Write a simple expression which relates how two legs of the cycle (computed using techniques such as free energy perturbation) are equivalent to the difference in the two experimental free energy of hydration values ΔG_{hyd} (A) and ΔG_{hyd} (B)

Question 2. Answer the following questions using 3 letter codes for the amino acids if appropriate. For structural drawings label all atoms and indicate stereo chemistry with hashed or wedged lines.

(a) Electron distributions can be modeled as a collection of "point charges" centered on the atomic nuclei (true or false).

_____

(b) The following molecules have what functional group in common?



_____

(c) How many hydrogen bonds can the common functional group shown above donate to a suitable hydrogen bond acceptor?

_____

(d) Phospholipids are major components of protein structure (true or false)

_____

(e) Write the three letter name of four hydrophilic amino acid side chains.

_____

(f) Which two amino acid side chains are negatively charged under most conditions.

_____

(g) Which two amino acid side chains are positively charged under most conditions.

_____

(h) Draw a simple line representation showing the side chain for tyrosine.

_____

(i) Draw a simple line representation showing the side chain for glutamic acid.

_____

(j) Draw a simple line representation showing the side chain for valine.

_____

(k) There is only one set of partial atomic charges that will yield a net formal charge (i.e. +1, +2, -1, -2, etc. (true or false)?

_____

(l) Electron density in a molecule is always homogenous (true or false)?

_____

(m) Drug development typically takes years, costs upwards of millions of dollars, and is accomplished in various stages. Arrange the following stages in the order commonly performed.

1 clinical trials
2. database screening
3. testing and development
4. target isolation and validation

_____

(n) Carbohydrates contain many hydroxyl groups (true or false)?

_____

(o) Which energy terms in classical Molecular Mechanics force fields are considered the “non-bonded” terms?

_____

Questions Based on CSE 549:

Question 3. Nucleotide sequences are sometimes written in an alphabet with five characters *A, T, G, C,* and *N,* where *N* stands for an unspecified nucleotide (in essence a wildcard).

A sequence with an *N* is referred to as a degenerate string; for example *ATTNG* may correspond to four interpretations:

ATTAG, ATTTG, ATTGG, and *ATTCG*.

In general, a sequence with *k* unspecified nucleotides *N* will have 4^k different interpretations.

Given a non-degenerate string *v* and a degenerate string *w* with *k N*'s, devise an efficient method to find the best interpretation of *w* according to *v*. That is, out of all 4^k possible interpretations of *w*, find *w'* with the minimum alignment score *s(w,w')*

Question 4. Describe how a phylogenic tree on *n* species can be constructed given DNA sequence data about each of the species.