Name: _____

MCB 100A/ChemC130 Midterm 1 2009

Please write your name on the first page.

1. Find the letter below that best matches the following statements. Use a letter only once. **(20 pts.)**

A. hydrogen bonds
B. most important structural restraints determined using NMR
C. wide and shallow in RNA duplex
D. flips backbone into the structure
E. reverses chain direction in 4 residues
F. stacking
G. Watson-Crick base pairs display different H-bonding pattern
H. C3' endo sugar pucker
I. measure of quality of a sequence alignment
J. genetic code
K. measure of quality of a NMR structure
L. all tRNAs
M. depends on temperature, denaturant concentration and buried surface area
N. ionic interaction
O. often recognized by hydrogen bonds from side chains in a helix
P. fold
Q. extended protein strand
R. measure of quality of a crystal structure

i) _ E-value
ii) _ through-space interactions
iii) _ requires faithful charging of tRNAs with the correct amino acid
iv) _ hydrophobic effect
v) _ falls off as 1/distance
vi) __ cloverleaf secondary structure
vii) __ β-turn
viii) __ metal-ion core
ix) __ shared by homologous proteins
x) __ major groove

2a. List three different advantages of a DNA genome compared to a RNA genome. **(9 pts.)**

Name: _____

2b. The A-minor motif occurs frequently in folded RNA structures. What is the general role of an A-minor motif? **(5 pts.)**

2c. List two reasons that DNA strands could not form an A-minor motif. **(4 pts.)**

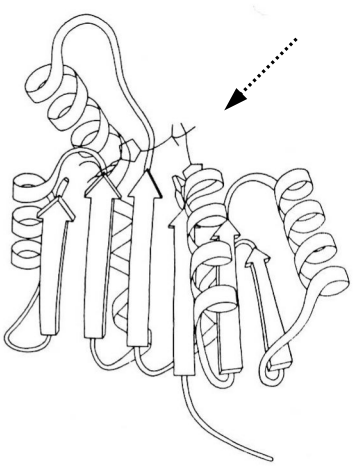
3. The sequence of part of a helix in the folded core of your favorite protein (YFP) is: RMELLKAAIEGD.

3a. Propose a conservative amino acid substitution that you would expect to only minimally disrupt the stability or structure of the protein. Draw the chemical structure of the starting amino acid and the new residue, and give two reasons why this mutation is likely to preserve the folded structure. **(6 pts.)**

3b. The program BLAST assigns a score to define quantitatively the degree of conservation of each identity and each amino acid substitution. What is the name of the table of such scores? What do the numbers in the table represent? **(6 pts.)**

Name: _____

4. Here is Jane Richardson's "ribbon diagram" showing the arrangement of secondary structures and loops in alcohol dehydrogenase (ADH):



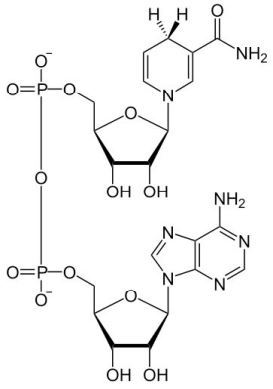
4a. What is the importance of secondary structure to protein folding? **(5 pts.)**

4b. What type of β-sheet is found in ADH? **(3 pts.)**

4c. What is the fold class of ADH? **(3 pts.)**

4d. Why do the helices tend to align with the strands of the sheet? **(5 pts.)**

4e. The stick drawing (dashed arrow) in the ribbon diagram shows the bound cofactor, NADH (chemical structure on the right). What feature of helices, such as the one second-from-the-right in the ribbon diagram, promotes binding of NADH to the structure? **(4 pts.)**



4f. Do you think RNA polymerase could make NADH? Why or why not? **(4 pts.)**

Name: _____

5. Lisa is studying a 32 kDa, 296-amino-acid terpene synthase (TPS) required for virulence of *M. tuberculosis.* The enzyme produces a novel diterpene product *in vivo*, so she knows it's active.

The best hit in a BLAST search against the sequences of proteins in the Protein Data Bank is:

```
CopII coat protein     Length=748
Expect = 0.13
 Identities = 16/42 (38%), Positives = 24/42 (57%), Gaps = 6/42 (14%)

Query  7    KEFLDLPLVSVAEIVRCRGPKVSVFPFDG---TRRWFHLECN  45
            K+ + LP+V+ + IVRCR  +  + PF      RRW   +CN
Sbjct  70   KDLVQLPVVTSSTIVRCRSCRTYINPFVSFLDQRRW---KCN  108
```

5a. What can Lisa conclude about the fold of TPS? Briefly explain your answer. **(8 pts.)**

5b. Lisa compared the TPS sequence to the database of non-redundant sequences and found the match on the next page with a "hypothetical protein" from a single-cell eukaryote called *Dictyostelium discoideum*.

She showed the results to her colleague Larry who said, "Aha, that means that the two proteins have the same fold and the same function!" Lisa says, "Not so fast. The sequence identity is pretty low." Larry said, "OK. What other computational technique could we use to test if both proteins are terpene synthases?" What would you advise Larry? What method should they use, and what result would they expect find if both proteins have the same enzymatic function? **(6 pts.)**

```
Expect = 2e-32
 Identities = 89/320 (27%), Positives = 158/320 (49%), Gaps = 35/320 (10%)

Query  7    KEFLDLPLVSVAEIVRCR--GPKVSVFPFDGTRRWFHLE----------------------  43
            +EF  L    +++I+  R       V+ +DGTRR + +E
Sbjct  11   QEFNKLTDNEISKIINSRLNNCNTMVYAYDGTRRSYLIENTISKLQTNGIHNNKCKFTGK  70

Query  44   ---CNPQYDDYQQAALRQSIRILKMLFEHGIETVISPIFSDDLLDRGDRYI---VQALEG  97
                   YDDY + A+ + +  L M+F+HGI+T++ P++   L DRG  Y+   ++ L G
Sbjct  71   DEKTTIDYDDYCKTAISKLLFDLVMMFKHGIKTIVYPMWFCTLEDRGPEYLPKFIKYLSG  130
```

Name: _____

```
Query  98   MALLANDEEILSFYKEHEVHVLFYGDYKKRLPSTAQGAAVVKSFDDLTISTSSNTEHRLC  157
            +  L  +E ++  YKE  + V+FYG+Y K L        ++++F+ +   T  N  H +
Sbjct  131  LKALLENETLVKLYKECGIRVIFYGEYIKLL-ERGNDPILLETFNKIMELTKDNISHTIL  189

Query  158  FGVFGNDAAESVAQFSISWNETHGKPPTRREIIEGYYGEYVDKADMFIGFGRFSTFDFPL  217
            FG    + ++++ + SI + E +   PT+ ++I+ YYG  VD+   ++GF RFST   P+
Sbjct  190  FGTTIQEPSQTIIENSIDFFEKYNYRPTKNQLIKKYYGVDVDQVSFYLGFDRFSTDGRPI  249

Query  218  LSS--GKTSLYFTVAPSYYMTETTLRRILYDHIYLRHFRPKPDYSAMSADQLNVLRNRYR  275
              S  G   LY+T++P  Y ++   R++L+D +Y R      +Y     D + +++  Y
Sbjct  250  YISDKGNEDLYYTISPHSYFSKINFRKVLFDKLYCRSNTNAKEYELKLTD-IEMMKEFYE  308

Query  276  AQPDRVFGVGCV--HDGIWF  293
                 V G+G V  H   W+
Sbjct  309  NNSTNVMGLGNVNPHGNYWY  328
```

5c. Lisa decides to determine the structure of the *M. tuberculosis* TPS. What <u>experimental</u> technique should she use and what is the most important reason she should use that technique? **(6 pts.)**

5d. Once she generates a finished set of atomic coordinates using her method of choice, what are two criteria she could use to make sure the structural coordinates contain no serious errors? **(6 pts.)**