Name:

BE.011/2772J
Spring 2004
QUIZ I
February 23, 2004

You have 1 hour for this exam.

CLOSED BOOK
1 page notes allowed

1 (10points)	
2 (5 points)	
3 (25 points)	
4 (60 points)	
total (100 points)	

$k = 1.38 \times 10^{-23}$ J/K

Continued on next page....

Name:

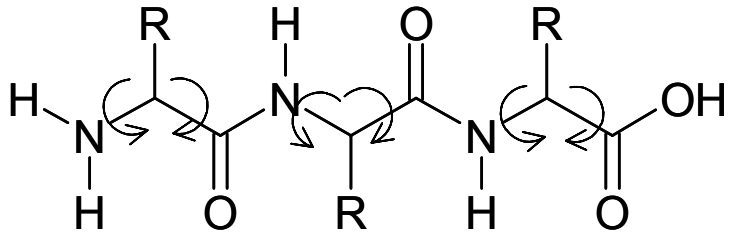
1.) (10 points) **ELVIS is everywhere**

a.) Given 20 naturally occurring amino acids, what is the probability that the amino acid sequence ELVIS occurs in a stretch of a protein sequence?

b.) What is the probability if the order of the amino acids did not matter, i.e., VLSEI, etc.?

2.) (5 points) **Protein folding.**

a.) A protein is a linear chain of amino acids. The amino acid has two torsional angles than can vary around the α carbon, ψ and ϕ. Due to sterics, ψ and ϕ have 3 possible configurations apiece, yielding $3 \times 3 = 9$ possible configurations per amino acid. Pictured below is a three residue peptide:



If a protein has n =100 amino acids, how many different configurations are possible?

Name:

3.) (25 points) **Conditional probabilities of the genetic code**

a) A codon is a sequence of 3 nucleotides that specifies a particular amino acid. How many codons are possible out of the 4 nucleotides?

b) Using the table for the genetic code below, what is the joint probability of obtaining a G in the second position (G_2) and A in the first (A_1)?

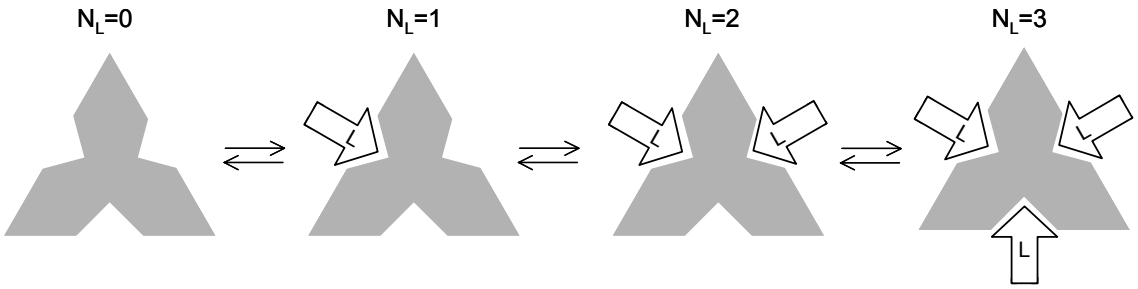
First Letter	Second Letter								Third Letter
	U		C		A		G		
U	UUU	Phe	UCU	Ser	UAU	Tyr	UGU	Cys	U
	UUC	Phe	UCC	Ser	UAC	Tyr	UGC	Cys	C
	UUA	Leu	UCA	Ser	UAA	Stop	UGA	Stop	A
	UUG	Leu	UCG	Ser	UAG	Stop	UGG	Trp	G
C	CUU	Leu	CCU	Pro	CAU	His	CGU	Arg	U
	CUC	Leu	CCC	Pro	CAC	His	CGC	Arg	C
	CUA	Leu	CCA	Pro	CAA	Gln	CGA	Arg	A
	CUG	Leu	CCG	Pro	CAG	Gln	CGG	Arg	G
A	AUU	Ile	ACU	Thr	AAU	Asn	AGU	Ser	U
	AUC	Ile	ACC	Thr	AAC	Asn	AGC	Ser	C
	AUA	Ile	ACA	Thr	AAA	Lys	AGA	Arg	A
	AUG	Met	ACG	Thr	AAG	Lys	AGG	Arg	G
G	GUU	Val	GCU	Ala	GAU	Asp	GGU	Gly	U
	GUC	Val	GCC	Ala	GAC	Asp	GGC	Gly	C
	GUA	Val	GCA	Ala	GAA	Glu	GGA	Gly	A
	GUG	Val	GCG	Ala	GAG	Glu	GGG	Gly	G

c) Calculate the degree of correlation for obtaining a Serine given G_2. Is it negatively/ positively/not correlated, or mutually exclusive?

Name:

4.) (60 points) **Protein ligand binding**

A protein has M = 3 sites for binding a ligand. The sites are indistinguishable from each other, as are the ligands.



a) Express the number of ways N_L ligands can be arranged in M sites, W (M, N_L).

b)Calculate the multiplicity and also the entropy for the following states. N_L= 1 means that one ligand is bound, N_L=0 means no ligands are bound, etc. You may leave the entropy in terms of the Boltzmann constant, k.

(i) N_L =0

(ii) N_L = 1

(iii) N_L = 2

(iv) N_L = 3

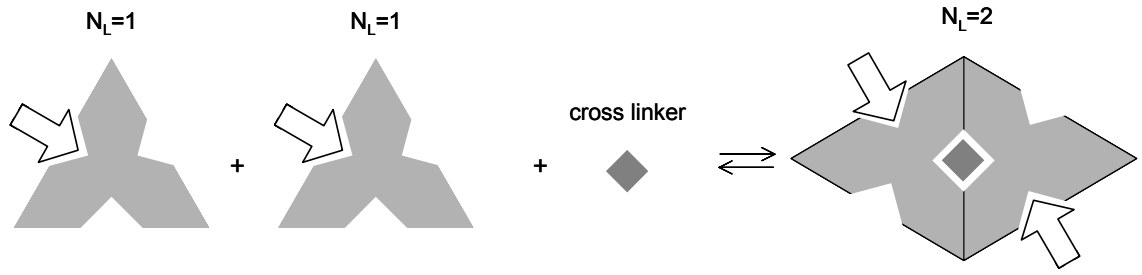
Name:

c) Plot the probability distribution as a function of the number of ligands, $p(N_L)$.

d) Calculate the variance of the distribution.

Name:

e) Upon addition of a cross linker, the protein can dimerize to form the following dimer (pictured below). The dimer can bind up to 4 ligands.



What is the entropy of the dimeric protein where N_L =2 (pictured)? You may leave in terms of k.

f) Compare the entropy of two proteins as monomers (left) with N_L=1 with the entropy of the dimer (right). What is the change in entropy, ΔS, going from monomeric to dimeric, in terms of k? Based on your calculation, is the system more likely to be in the monomeric or dimeric form?