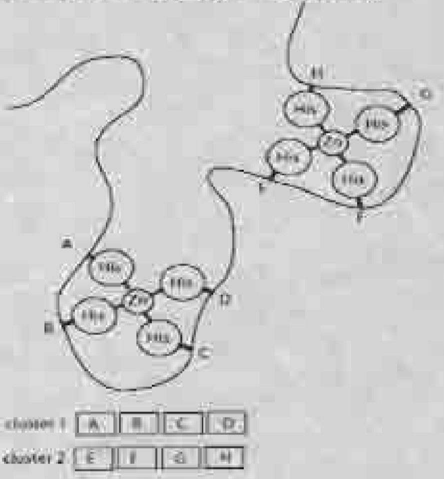
QUESTION 1. (20 points total)

Question 1(A) (10 points)



(A) Shown above is a schematic drawing of a protein chain.
The protein chain contains 8 histidine residues (labeled A, B, C, ... H), and the properly folded protein contains two clusters of 4 histidine residues that each coordinate a zinc ion, as shown in the figure. The precise order of the histidines coordinating the zinc is required for activity.
An experiment analogous to the Anfinsen experiment is carried out, in which the protein is (i) unfolded and refolded in the absence of zinc, and then zinc is added, and (ii) the protein is unfolded, zinc is added, and then the protein is refolded.
In case (i), 100% of the original protein activity is recovered. In case (ii) only a fraction of the protein activity is recovered.

Using the same reasoning that we used in the case of the Anfinsen experiment, estimate the fraction of protein activity that is recovered in case (ii).
Hint: Assume that in the unfolded state all possible combinations of His residues in sets of four can coordinate the zinc atoms. Look at the labels in the boxes denoted cluster 1 and cluster 2 below the drawing, and calculate the number of all equivalent rearrangements of the histidine residues that give two zinc clusters. Be sure to correct for overcounting due to permutations. Use this to estimate the fractional activity, by assuming that only one arrangement corresponds to active protein.

$\boxed{1}\ \boxed{2}\ \boxed{3}\ \boxed{4}$ cluster 1

$\boxed{5}\ \boxed{6}\ \boxed{7}\ \boxed{8}$ cluster 2

Number of ways of picking 8 histinines = 8!
Within cluster 1 if we count all permutations as equivalent then we overcount by 4!. Likewise

(PAGE LEFT PURPOSELY BLANK)

for cluster 2 we also overcount by 4!
Also, the vertical ordering is irrelevant, so we have to divide by 2.

Hence: Number of distinct his arrangements

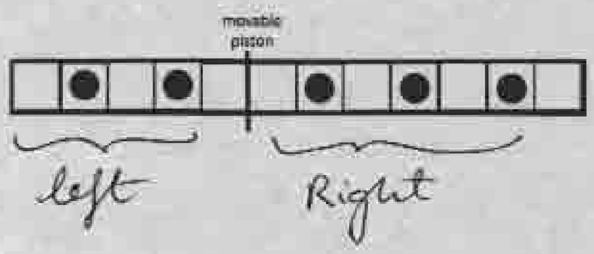
$$= \frac{8!}{2 \times 4! \times 4!} = 35.$$

Hence, for the scenario where the zinc is added to the unfolded protein only 1 in 35 configurations of the histidine correspond to active protein ⇒ $0.0286 \times 100 = 2.9\%$

QUESTION 1(B) (10 points)

Shown below is a simple model for gas expansion. A chamber is divided into two parts, separated by a movable piston. The chamber contains 5 molecules, 3 on the right and two on the left.

Based on the principle of maximum multiplicity (second law of thermodynamics), will the piston move spontaneously to the right by one grid point? Explain how you work out the answer.



$$W_{left} = \frac{5!}{3! \times 2!} = \frac{120}{6 \times 2} = 10$$

$$W_{right} = \frac{7!}{3! \times 4!} = \frac{5040}{6 \times 24} = 35$$

Total multiplicity = $10 \times 35 = 350$

move piston to right by 1 grid point.

$$\text{Now } W_{left} = \frac{6!}{2!\ 4!} = \frac{720}{2 \times 24} = 15$$

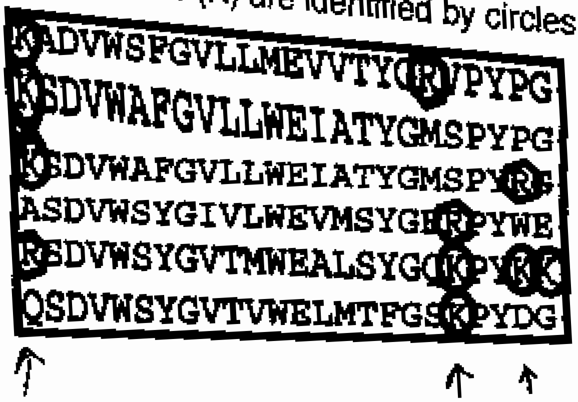
$$\text{and } W_{right} = \frac{6!}{3!\ 3!} = \frac{720}{6 \times 6} = 20$$

Total multiplicity = $15 \times 20 = 300$

Since the multiplicity goes down the piston will not move to the right.

QUESTION 2. (20 points)

Consider the block of 6 aligned sequences, each with 24 residues, shown in the diagram below. All the arginine (R) residues in the block are identified by hexagons, and all the lysines (K) are identified by circles.



```
KADVWSFGVLLMEVVTYGRVPYPG
KSDVWAFGVLLWEIATYGMSPYPG
KSDVWAFGVLLWEIATYGMSPYRG
ASDVWSYGIVLWEVMSYGERPYWE
RSDVWSYGVTMWEALSYGGKPYKK
QSDVWSYGVTVWELMTFGSKPYDG
```

↑ ↑ ↑

QUESTION 2(A) (5 points) What is the probability of finding a K-R pair by random chance, given the sequence composition of the sequence block? That is, if the given sequence block were randomized by completely shuffling all the residues, what would be the probability of finding a K and an R in the same column in the shuffled sequence? Explain how your work out your answer.

There are 7 lysines and 4 arginines.

Probability of finding a lysine anywhere $= \frac{1}{6 \times 24} = \frac{1}{144} \times 7$

Probability of finding R anywhere $= \frac{1}{144} \times 4$

Probability of finding an R in one position and K in another position in the same column $= \frac{4 \times 7}{144^2}$. Since the order does not matter, required probability $= \frac{2 \times 4 \times 7}{144^2} = 2.7 \times 10^{-3}$

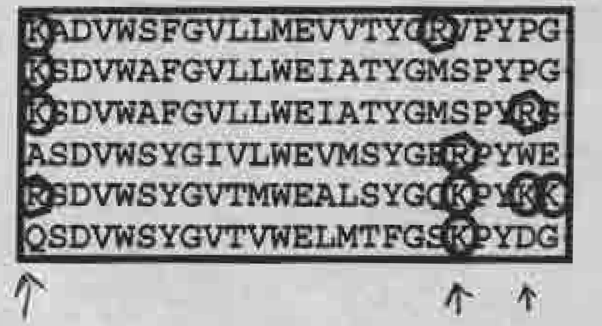
QUESTION 2(B) (5 points) How many K-R pairs (in the same column) are actually found in the given sequence block? Explain how your work out your answer.

There are 3 columns with both K and R. In each there is only 1 R, which is paired with each K.

Total pairs $= 3 + 2 + 1 = 6$

QUESTION 2. (20 points)

Consider the block of 6 aligned sequences, each with 24 residues, shown in the diagram below. All the arginine (R) residues in the block are identified by hexagons, and all the lysines (K) are identified by circles.



KADVWSFGVLLMEVVTYGRVPYPG
KSDVWAFGVLLWEIATYGMSPYPG
KSDVWAFGVLLWEIATYGMSPYRS
ASDVWSYGIVLWEVMSYGERPYWE
RSDVWSYGVTMWEALSYGCKPYKK
QSDVWSYGVTVWELMTFGSKPYDG

QUESTION 2(A) (5 points) What is the probability of finding a K-R pair by random chance, given the sequence composition of the sequence block? That is, if the given sequence block were randomized by completely shuffling all the residues, what would be the probability of finding a K and an R in the same column in the shuffled sequence? Explain how your work out your answer.

There are 7 lysines and 4 arginines.

Probability of finding a lysine anywhere = $\frac{1}{6\times24} = \frac{1}{144}\times7$

Probability of finding R anywhere = $\frac{1}{144}\times 4$

Probability of finding an R in one position and K in another position in the same column = $\frac{4\times7}{144^2}$. Since the order does not matter, required probability = $\frac{2\times4\times7}{144^2} = 2.7\times10^{-3}$

QUESTION 2(B) (5 points) How many K-R pairs (in the same column) are actually found in the given sequence block? Explain how your work out your answer.

There are 3 columns with both K and R. In each there is only 1 R, which is paired with each K.

Total pairs = 3 + 2 + 1 = 6

QUESTION 2 CONTINUED.

QUESTION 2(C) (5 points) What is the total number of amino-acid pairs, without regard to sequence, in the sequence block?

Since there are 6 positions in each column there are $\frac{6 \times 5}{2} = 15$ pairs in each column. There are 24 columns, giving up a total of $15 \times 24 = 360$ pairs.

QUESTION 2(D) (5 points) Are K-R pairs found with greater or lesser frequency than expected based on random chance in this sequence block? Explain how you worked out the answer.

observed frequency $= \frac{6}{360} = 1.7 \times 10^{-2}$

Expected random frequency $= 2.7 \times 10^{-3}$

Hence K-R pairs are found more frequently than random chance.

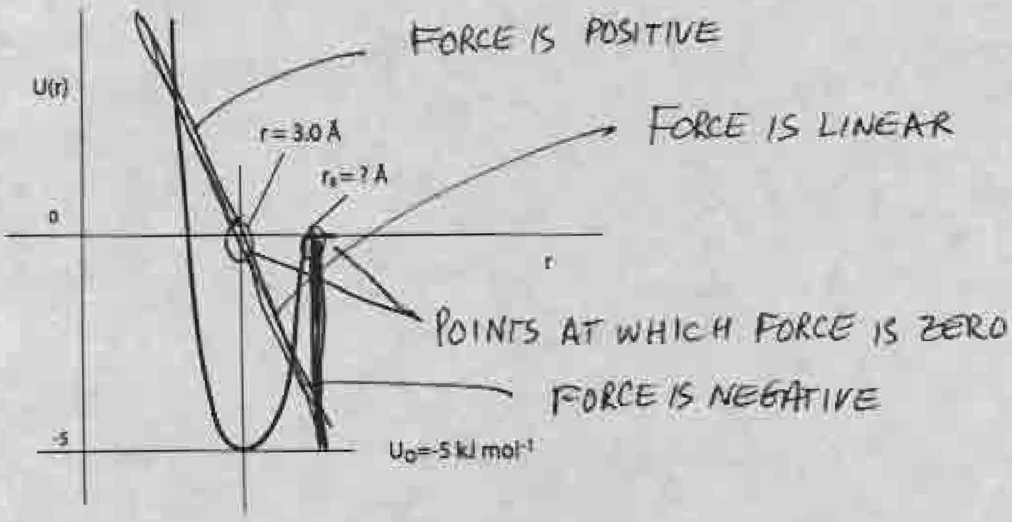
QUESTION 3. (20 Points)
A system consists of two atoms. The energy of interaction, U(r), between two atoms is given by:

$U(r)=U_o+\frac{k}{2}(r\overset{-}{\square}r_o)^2$ when $r < r_b$ Å,

and $U(r)=0$ when $r > r_b$ Å,
where r_b is the interatomic separation at which the energy is zero and the bond breaks.

k is the force constant: $k = 10$ kJ mol^{-1} Å^{-2},
r_o is the optimal interaction distance: r_o=3.0 Å,
and U_o is the stabilization energy: U_o=-5 kJ mol^{-1}.

The schematic diagram below shows the dependence of the energy on inter-atomic separation:



QUESTION 3(A) (7 points) Indicate schematically the force, F, as a function of r on the diagram above. Ignore the discontinuity in energy at $r = r_b$.

Hint: F = -dU/dr

QUESTION 3(B) (8 points) What is the value of the distance r_B in Å?

When $r = r_B$ $\quad U(r) = -5 + \frac{10}{2}(r_o - r_B)^2 = 0$

$\Rightarrow \frac{10}{2}(r-r_B)^2 = 5$

$\Rightarrow (r_o-r_B)^2 = 1$

$\Rightarrow r_o - r_B = \pm 1 \quad \Rightarrow r_B = r_o + 1 = 4$ Å

QUESTION 3 (C) (5 points) What is the force required to break the bond between the two atoms? Report the value of the force in Newtons mol^{-1}.

Hint: $\boxed{J = N\,m^{-1}}$ $J = N m$

$$F = -\frac{dU}{dr} = -\frac{d}{dr}\left(U_0 + \frac{k}{2}(r-r_0)^2\right)$$

$$= -\frac{k}{2}\,2\,(r-r_0)$$

$$= -k\,(r-r_0)$$

When the bond breaks $r = r_B$ and $r - r_0 = 1$

So the magnitude of the force $= k = 10\,kJ\,mol^{-1}\,Å^{-1}$

$$10\ kJ\ mol^{-1}\ Å^{-1} = 10^3\ J\ mol^{-1}\ Å^{-1}$$

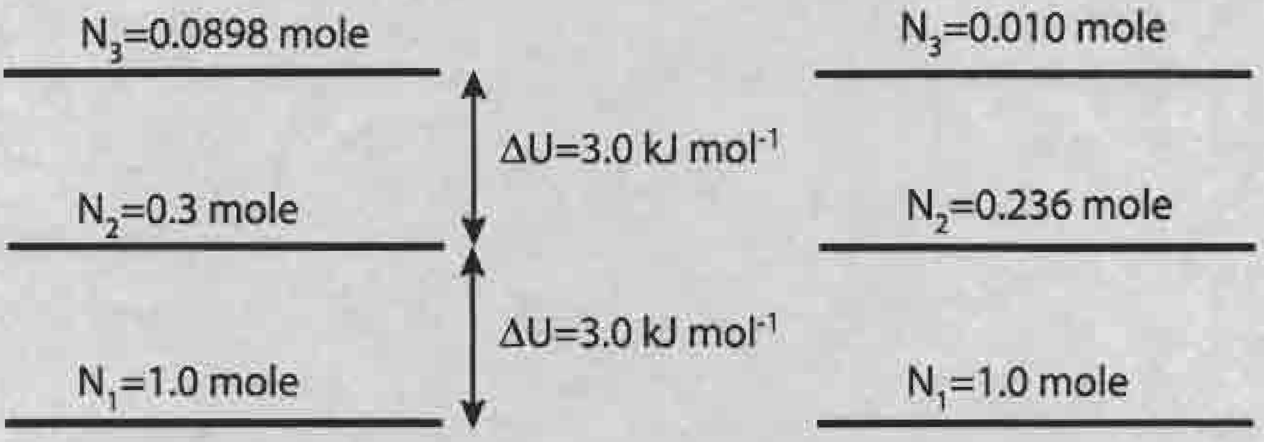
$$= \frac{10^3}{[illegible]}\ J\ Å^{-1}\ mol^{-1}$$

$$= 10^3 \times 10^{10}\ J\ m^{-1}\ mol^{-1}$$

$$= 10^{13}\ N\ mol^{-1}$$

QUESTION 4. 20 Points

(A) (10 Points) A molecular system consists of 3 energy levels per molecule. Consider the two energy distributions, shown in the diagram below. The number of molecules in each energy level are shown in the diagram.



One of these distributions corresponds to a Boltzmann distribution, and one does not. Indicate which one is the Boltzmann distribution, and explain completely your reasoning.

Since the energy gap between the levels is the Same in all cases, the ratios of molecules ~~is~~ in adjacent levels should be the same. For the distribution at the left:

$\frac{N_2}{N_1} = 0.3 \qquad \frac{N_3}{N_2} = 0.2993 \approx 0.3$

For the distribution on the right,

$\frac{N_2}{N_1} = 0.236 \qquad \frac{N_3}{N_2} = 0.0423$

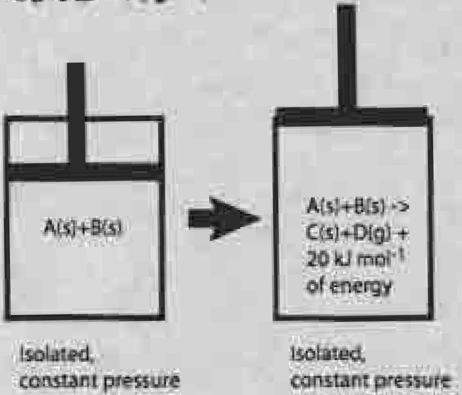
Hence the distribution on the right is not a Boltzmann distribution.

QUESTION 4, CONTINUED

Consider the following chemical reaction:

A (solid) + B (solid) -> C(solid) + D (gas) $\overset{+}{\cancel{\Delta U = -}}$20 kJ mol^{-1} (at 300K and 1 atm pressure)

In other words, the energy of 1 mole each of A and B is higher by 20 kJ mol^{-1} than the energy of 1 mole each of C and D, at 300 K and 1 atm pressure. Assume that there are no other molecules present



QUESTION 4(B) (2 points) One mole each of A and B are completely converted to 1 mole each of C and D in an isolated system at constant pressure, initially at 300 K. The system exchanges no heat with the surroundings, so the temperature of the system increases during the reaction. At the end of the reaction, the temperature of the system is higher by 10° C. What is the heat capacity of the system (C+D) at constant pressure, in units of kJ mol^{-1} K^{-1}?

$$C_p = \frac{\text{Energy delivered}}{\text{rise in temperature}} = \frac{20 \text{ kJ mol}^{-1}}{10 K} = 2 \text{ kJ mol}^{-1} K^{-1}$$

QUESTION 4(C) (6 points) What is the change in energy of the system upon completion of the reaction, in units of kJ mol^{-1}? Assume ideal gas behavior and negligible volume changes for the solids.

Hint: Calculate the work done by using the ideal gas law, $PV_1 = n_1RT_1$ and $PV_2 = n_2RT_2$, where V_1, V_2 are the initial and final volumes, T_1, T_2 are the initial and final temperatures, n_1, n_2 are the initial and final number of moles of the gas, P is the pressure and R is the gas constant.
The gas constant is 8.314 J K^{-1} mol^{-1}.

Work done during expansion under constant pressure is $-P\Delta V$. Since $PV_1 = n_1RT_1$
and $PV_2 = n_2RT_2$
$$-P\Delta V = -(n_2RT_2 - n_1RT_1)$$
We assume that the change in volume is due entirely to the creation of D. So initially $n_1 = 0$, and finally $n_2 = 1$. So $-P\Delta V = -RT_2$ $T_2 = 300 + 10 = 310K$.

QUESTION 4 (D) (2 points)

What is the enthalpy change of the system upon completion of the reaction, in units of kJ mol^{-1}?

QC - continued

So $-RT_2 = -8.314 \times 310 = -2577\ J\,mol^{-1}$

$= -2.58\ kJ\,mol^{-1}$

This amount of energy goes towards doing work. ~~so $\Delta U = 20\,kJm^{-1} - 2.58\,kJm^{-1}$~~

~~$= 17.42\ kJ\,mol^{-1}$~~

~~QD. The enthalpy change~~

Use the 1st law:

$$dq = dU - dw$$

or $q = \Delta U - w$

In this case, $q = 0$, so $\Delta U = w$

$= -2.58\ kJ\,mol^{-1}$

(The total energy decreases, because the system does work).

(D) Enthalpy:

$$\Delta H = \Delta U + P\Delta V$$

$$= -2.58 + 2.58 = 0$$

Recall that enthalpy is the heat transferred under constant pressure, and there is no heat transferred here.

QUESTION 5 (20 points)
Indicate whether the following statements are true or false. 2 points for each correct answer. -1 point for each wrong answer. You do not need to answer all the questions to get the maximum number of points (20 points maximum, 0 points minimum).

(i) In contrast to ribonuclease, the majority of proteins in the cell cannot fold without the assistance of proteins known as molecular chaperones. This means that the thermodynamic hypothesis of protein folding does not apply to most proteins.
TRUE / [✓ FALSE]

(ii) Urea denatures proteins because it mimics an amino acid sidechain. TRUE / [✓ FALSE]

(iii) Cysteine is much more strongly conserved than leucine because it has a unique structural role. [✓ TRUE] / FALSE

(iv) Other than the histidine residue that links to the heme group, essentially every other residue in the globins is varied in one globin or the other. [✓ TRUE] / FALSE

(v) Two proteins that share more than 40% sequence identity over a 100 residue stretch are likely to have the same three-dimensional fold [✓ TRUE] / FALSE

(vi) Genetic selection experiments using the protein lambda repressor showed that residues in the hydrophobic core are sometimes replaced by lysine in functional variants of the protein. TRUE / [✓ FALSE]

(vii) The number of distinct folds in proteins is likely to be essentially unlimited, because secondary structural elements can pack against each other in continuously varying ways. TRUE / [✓ FALSE]

(viii) The carboxyl terminal end of the helix dipole in the Rossmann fold points towards the phosphate groups of the bound nucleotide. TRUE / [✓ FALSE]

(ix) Protein domains can be assembled together in many different ways because surface sidechains can be mutated easily without losing protein stability [✓ TRUE] / FALSE

(x) Electrostatic interactions in water are attenuated approximately 10-fold in water and 2-4 fold inside the interior of proteins. TRUE / [✓ FALSE]

(xi) If water did not attenuate the strengths of hydrogen bonds then protein, DNA and RNA molecules would be held so tightly together that biological function would not be possible [✓ TRUE] / FALSE

(xii) The force required to break the strongest non-covalent interactions in proteins is in the range of approximately 100 piconewtons [✓ TRUE] / FALSE

-END OF EXAM