Biophysics 2020-21
Bioinformatics degree
Final Exam. Part I

1. (1.5 points) Thermodynamics.
It has been seen that a transmembrane protein could make conformational changes to adopt an active or inactive states. The estimated free energy for the active state of the protein is -200.5 kJ/mol with a degeneration of one. The degeneration of the inactive state is 2 with a free energy of -198.1 kJ/mol.
a) If a cell has around of 10^5 of those transmembrane proteins, calculate the number of active and inactive proteins at 37ºC.
b) Calculate the temperature at which the proportions of both states are the same.
c) Discuss which proportions will be obtained at 0K and at infinity temperature.

a) $p_j = \frac{g_j e^{-\beta\varepsilon_j}}{\sum_{i=0}^{2} g_i e^{-\beta\varepsilon_i}}$ with g_0= 1, g_1=2
E_0=-200.5 10^3 j/mol E_1= -198.1 10^3 J/mol

Or making relative energies to the more stable state (E_0=0 and E_1=2400 J/mol) we obtain the same results:

$$p_{act} = p_0 = \frac{g_0 e^{-\beta\varepsilon_0}}{\sum_{i=0}^{2} g_i e^{-\beta\varepsilon_i}} = 0.559$$
$$p_{inact} = p_1 = \frac{g_1 e^{-\beta\varepsilon_1}}{\sum_{i=0}^{2} g_i e^{-\beta\varepsilon_i}} = 0.441$$

$$N_{act} = 10^5 p_{act} = 55912 \quad N_{act} = 10^5 p_{inact} = 44087$$

b) $p_0 = p_1 \quad g_0 e^{-\beta\varepsilon_0} = g_1 e^{-\beta\varepsilon_1} \quad \beta = \frac{1}{RT}$

$$T_{eq} = \frac{-E_{Act} + E_{inact}}{R\ln(2)} = 416K$$

c) At 0K, $p_0 = 1$, all the population is in the most stable level
At infinite temperature, the population is distributed according to the degeneration of levels.
$p_0 = g_0/(g_0 + g_1) = 1/3$
$$p_1 = g_1/(g_0 + g_1) = 2/3$$

2) (1.5 points) Chemical kinetics.

It is studied a reaction A+B→ P using different initial concentrations of the both reactants at several temperatures. A summary of the experiments is provided in the table.
a) Determine the partial order of A and B, and the rate constant at 25C
b) If we have an initial concentration of $[A]_0$ of 0.04 mol dm^{-3} and $[B]_0$ of 0.4 mol dm^{-3}, calculate the concentration of [A] at 1 minute.
c) Determine the activation energy of this reaction.

Experiment	$[A]_0$ / (mol dm^{-3})	$[B]_0$ / (mol dm^{-3})	v_0 /(mol dm^{-3} s^{-1})	Temperature/ºC
1	0.21	0.10	0.45	25
2	0.42	0.10	1.81	25
3	0.42	0.15	1.83	25
4	0.21	0.10	0.93	50

a) $v = k[\mathrm{A}]^{\alpha}[\mathrm{B}]^{\beta}$

$$\alpha = \frac{ln\frac{v_1}{v_2}}{ln\frac{[\mathrm{A}]_1}{[\mathrm{A}]_2}} \approx 2 \qquad \beta = \frac{ln\frac{v_2}{v_3}}{ln\frac{[\mathrm{B}]_2}{[\mathrm{B}]_3}} \approx 0$$

$$v = k[\mathrm{A}]^2$$

For example, from the data of first experiment we can obtain the value of the constant:

$$k_{25} = 10.2\mathrm{Lmol^{-1}s^{-1}}$$

b) It is a second order reaction (We don't need the concentration of B):

$$\frac{1}{A} = \frac{1}{A_0} + kt \;\; A = \frac{1}{\frac{1}{A_0}+kt} = 0.00156M$$

c) $k = Ae^{\frac{-Ea}{RT}}$

It is required the constants for two temperatures. Thus, first calculate the constant at 50C with the experiment 4.

$$v_4 = k_{50C}[\mathrm{A}_0]^2 \quad k_{50} = 21.1\mathrm{Lmol^{-1}s^{-1}}$$

Making a system with both temperatures (numbers in kelvins):

$$E_a = R * ln\,(k_{25}/k_{50})/(1/T_{50} - 1/T_{25}) = 23259 J/mol$$

3) (1.5 points) Transport.

a) Considering that the viscosity of water is about 10^{-3} N s m^{-2} estimate the diffusion coefficient of a protein of 3.5 nm of radius in water at 25ºC.

b) One experiment determined that the diffusion of this protein inside a particular type of cells is three times smaller. Estimate the viscosity inside this type of cells.

c) If this cell has 53 micrometers in diameter, estimate the time required to travel this distance.

a) $a = 3.5nm = 3.5 \cdot 10^{-9} m$

$$D = \frac{k_B T}{6\pi\eta a} = 6.24 \cdot 10^{-11}\, m^2/s$$

b) $\eta = \frac{k_B T}{6\pi\eta a(D/3)} = 3 \cdot 10^{-3} N\, s/m^2$

c) $D_{cell} = D/3 \quad t = \frac{R^2}{6D_{cell}} = 22.5s$

Biophysics 2020-21
Bioinformatics degree
Final Exam. Part II

4) (2 points) Indicate whether the following sentences are **true or false**. Justify briefly the answer.

a) Solvation energy in water cannot be calculated using classical forcefields

True. Solvation energy contains entropic terms, and entropy cannot be calculated from an FF

b) Monte Carlo simulations are limited to the conformational space between energy barriers

False. MC simulations do random conformational changes, so they are not limited by energy barriers

c) Flexible residues like Gly decrease protein stability

True. They increase the degrees of freedom of the unfolded form, and then the loss in conformational entropy on folding is larger.

d) Hydrogen bonds do not contribute to the overall stability in nucleic acid double helices

True. Hydrogen bonds have the same energy in the double helix or when they are made with water in the unfolded chain. Most stability comes from stacking interactions (hydrophobic + Vdw) between base-pair steps.

e) Statistical potentials are needed to identify properly folded structures

True. Normal FF cannot distinguish properly folded structures. Statistical potentials do so by comparing with known protein structures.

f) Protein folding is driven by entropic contributions

True. Protein folding is energetically driven by the hydrophobic effect that comes from the gain in entropy of the solvent.

g) NVT ensembles can be obtained by molecular dynamics simulations at constant pressure

False. NVT corresponds to constant VOLUME and TEMP. Constant pressure simulations are used for the NPT ensemble.

h) Calculation of binding energies requires only the calculation of interaction energies between interacting molecules.

False. Also, solvation energies of both the complex and the unbound components is relevant.

i) When analyzing reaction rates, the concentration of substrate known as K_M corresponds to a fully saturated protein

False. K_M corresponds to the concentration that gives 50% of saturation.

j) Chaperones accelerate protein folding "in vivo" by catalyzing the isomerization of Pro residues.

False. Chaperones stabilize unfolded structures avoiding mis-folding.

5) (1 point) Identify related terms in the two columns

a. Solvent entropy	A. Hooke equation
b. Bond angles	B. Generalized-Born equation
c. Van der Waals energy	C. Hydrophobic effect
d. Electrostatic solvation energy	D. Quantum mechanics calculation
e. Bond breaking energy barrier	E. 12-6 Potential

a-C
b-A
c-E
d-B
e-D

6) (1 point) Bond length energy can be evaluated using both Morse and Hooke equations. The energy profile of a C-C single bond can be estimated (Morse law) using D_M=520 kcal mol^{-1}, b_0 = 1.52 Å, and a = 0.55 $Å^{-1}$.
Evaluate the error that implies using Hooke's law instead, for the following C-C distances:
a) 1.0 Å, b) 1.5 Å, c) 2 Å, and d) 5 Å

Parameters for the Hooke equation should be deduced considering that it is the 2nd order polynomial approximation of the Morse equation (Solved already in Exercise 2.1.3). This can be done using the 2nd order Taylor series.

$$f(b) = f(b_0) + f'(b_0)(b - b_0) + f''(b_0)\frac{(b - b_0)^2}{2!}$$

$f(b_0) = 0$ and since b_0 is a minimum $f'(b_0) = 0$

$$f(b) = f''(b_0)\frac{(b - b_0)^2}{2!}$$

What identifies as K_{Bond} as $f''(b_0)$, then

$$E'_{pot} = 2aD_M\left(e^{-a(b-b_0)} - e^{-2a(b-b_0)}\right)$$

$$E''_{pot} = 2a^2D_M\left(2e^{-2a(b-b_0)} - e^{-a(b-b_0)}\right)$$

$$K_{Bond} = E''_{pot}(b_0) = 2a^2D_M = 314.6\ Kcalmol^{-1}Å^{-2}$$

Using this parameter:

Dist (A)	*Morse (Kcal/mol)*	*Hooke (Kcal/mol)*	*%Error*
1.0	*57.004*	*42.534*	*25.38%*
1.5	*0.064*	*0.063*	*1.09%*
2.0	*27.995*	*36.242*	*29.46%*
5.0	*377.923*	*1904.966*	*404.06%*

Alternatively, as 1,5 Å is very close to the minimum (1,52 Å) it would be possible, although less correct, to estimate KBond assuming that there is no error at 1.5 Å , in that case K_{Bond} = 318,08 Kcal mol^{-1} $Å^{-2}$ that is not far from the true value.

REMINDER: KBond is different from kB (the Boltzmann constant)!!!!, only names are similar, mixing the two is a clear indication you have not understood anything!

7) (1.5 points) We wish to evaluate the influence of several amino acid residues in the stability of a protein-protein complex (R-L). We have access to some mutagenesis experiment results where residues of R protein have been mutated to Ala, and the dissociation constant of the complex has been evaluated. The dissociation constant of the unmutated R-L complex es 2.4 nM.

At the same time, we have estimated theoretically the changes in interaction energy (Electrostatic + VdW).

Results are summarized in the following table:

Mutation	Complex K_D	ΔΔGint (Kcal/mol)
Glu 300 Ala	0.76 mM	10.5
Trp 360 Ala	3.4 µM	1.1
Val 310 Ala	4.7 nM	0.1
Gln 302 Ala	2.8 nM	1.2

a) Evaluate the global ΔΔGwt→A and the contribution of the solvation to this value.

Using $\Delta G = RTlnK_D$ we can obtain the $\Delta G_{binding}$ for each protein, including WT. Then the differences with WT give $\Delta\Delta G_{WT\rightarrow A.}$ This can be done as follows.

$$\Delta\Delta G_{WT\rightarrow A} = \Delta G_A - \Delta G_{WT} = RTlnK_{D(A)} - RTlnK_{D(WT)} = RTln\frac{K_{D(A)}}{K_{D(WT)}}$$

NOTE that both K_D 's should be expressed in the same UNITS!!

$\Delta\Delta G_{WT\rightarrow A}$ is the sum of the interaction component and the solvation component, then the solvation component is just the difference.

$$\Delta\Delta G^{solv}_{WT\rightarrow A} = \Delta\Delta G_{WT\rightarrow A} - \Delta\Delta G^{int}_{WT\rightarrow A}$$

Protein	K_D (M)	$\Delta G_{binding}$ *(Kcal/mol)*	$\Delta\Delta G_{WT\rightarrow A}$ *(Kcal/mol)*	$\Delta\Delta G^{solv}_{WT\rightarrow A}$ (Kcal/mol)
WT	2,40E-09	-11,75		
Glu - Ala	7,60E-04	-4,25	7,50	-3,00
Trp - Ala	3,42E-06	-7,45	4,30	3,20
Val - Ala	4,72E-09	-11,35	0,40	0,30
Gln - Ala	2,84E-09	-11,65	0,10	-1,10

WARNING: Energy units depend on which value of the R constant is used, anyone is fine but ALWAYS indicate the UNITS, and DO NOT MIX them!!!! Note that int energies were expressed in Kcal/mol

b) Indicate whether the following sentences are True or False, justifying briefly (with reference to the energy values obtained).
   a. Glu 300 is involved in an electrostatic interaction in the interface

True. Glu 330 Ala show a large $\Delta\Delta G_{WT\rightarrow A}$ that comes mostly from ΔΔGint. Since Glu is charged is should be the loss of an electrostatic interaction.

b. Trp 360 is a central residue in the protein-protein interface

True. Most of the change appears as a solvation contribution. Since Trp is a large hydrophobic, Trp should be buried in the interface after binding (ASA loss)

c. Val 310 is a central residue in the protein-protein interface

False. Val is also hydrophobic, but the solvation energy contribution is much lower, so Val is not buried, as that change in surface is not much relevant.

d. Gln 302 is involved in an electrostatic interaction in the interface

False. Gln is neutral, so it cannot be involved in an electrostatic interaction.

e. Solvation energies are only relevant for polar residues

False. Solvation is important for all residues.

c) *Indicate* briefly how ΔΔGint can be evaluated.

Interaction energies can be obtained using the standard electrostatic and vdw equations, evaluated between all pairs of atoms in the two components of the complex.
The WT→A comes either from the difference of those values when removing the contribution the appropriate atoms of the side chain or repeating the calculation after replacing the residue.

d) Evaluate the degree of complex formation of wild-type and each mutant at a concentration of free L of 2.4 nM.

The degree of complex formation corresponds to the "saturation degree" Y = RL/Rtot
WARNING: Again UNITS are relevant, K_D and L should be expressed in same units. Reulst should show that higher K_D correspond to less affinity, end then less saturation.

	K_D (nM)	Y
WT	2.40	0.50
Glu - Ala	7.60E+05	3.16E-06
Trp - Ala	3.42E+03	7.01E-04
Val - Ala	4.72	0.34
Gln - Ala	2.84	0.46

Additional data and equations:

M(H)=1 g mol^{-1}; M(C)=12 g mol^{-1}; M(O)=16 g mol^{-1}; M(N)=14 g mol^{-1}
k_B=1.3806488·10^{-23} J K^{-1}
R=1.987 cal K^{-1} mol^{-1} = 8.314 J K^{-1} mol^{-1} = 0.082 atm L K^{-1} mol^{-1}
N_A=6.022·10^{23} mol^{-1}

$$v^{mp} = \sqrt{(2\,RT/M)}$$

$$\overline{v} = \sqrt{(8\,RT/(\pi M))}$$

$$v^{rms} = \sqrt{(3\,RT/M)}$$

$$f(v) = 4\pi\left(\frac{m}{2\pi k_B T}\right)^{3/2} v^2 e^{-mv^2/(2k_B T)}$$

$$\frac{1}{[A]_0 - [B]_0} ln\frac{[B]_0[A]}{[A]_0[B]} = kt$$

$$D = \frac{k_B T}{6\pi\eta a}$$

Morse law for bond stretching energy

$$E_{pot} = D_M\left(1 - e^{-a(b-b_0)}\right)^2$$

D_M: Well depth (kcal mol^{-1}), b_0: bond length at energy minimum (Å); a: well width parameter ($Å^{-2}$); b: bond length (Å)

Hooke law for bond stretching energy (as used in the Amber parm99 forcefield)

$$E_{pot} = \frac{K_{Bond}}{2}(b - b_0)^2$$

K_{Bond}: Force constant (kcal mol^{-1} $Å^{-2}$), b_0: equilibrium bond length (Å); b: bond length (Å)

Free energy related to dissociation constants

$$\Delta G = RTlnK_D$$

ΔG: Process free energy (units: energy/mol, depending on the R value used). R (gas constant), T: Temperature (K); K_D: Dissociation constant (Concentrations in M)

Saturation degree in a simple binding process

$$Y = \frac{L}{K_D + L}$$

Y: Saturation degree, K_D: Dissociation constant, L: Concentration of free ligand