Name: Date:

Final Exam

Biophysics
Bioinformatics degree

1) (1.5 points) Thermodynamic.

Depending of the degree of opening of a channel, a macromolecule could be in three different conformations. One experiment at 25ºC determines the population of these three states to be 60%, 35% and 5%, for opened, semi-opened and closed conformations, respectively. a) Determine the relative energies of these three states assuming a degeneration of one for the three levels.
b) Recompute the relative populations assuming a degeneration of 5 for the semi-open conformation.
c) Compute the temperature at which the population of opened and semi-opened states has the same population assuming the same degenerations of b).

a) The populations of conformations, i=0,1,2 (open, semi-open and close, respectively), with energies e_0, e_1 and e_2 and degenerations of $g_0=g_1=g_2=1$ could be computing using

$$p_i=\frac{g_i e^{-\beta\varepsilon_0}}{\sum_{i=0}^{2} g_i e^{-\beta\varepsilon_i}} \quad \textbf{or} \quad p_i=\frac{g_i e^{-\beta\varepsilon_i}}{q}$$

The lowest level has the reference value of 0.

(Reference) $\varepsilon_0=0$

From p_0/p_1 $\varepsilon_1=RT\ln\left(\frac{p_0 g_1}{p_1 g_0}\right)=1336\,\text{J/mol}$

From p_0/p_2 $\varepsilon_2=RT\ln\left(\frac{p_0 g_2}{p_2 g_0}\right)=6160\,\text{J/mol}$

Alternatively, you can compute first the partition function, q=1.67, and then the energies e_1 and e_2

b) Using the values of denegerations $g_0=1$, $g_1=5$ and $g_2=1$, the previous energies and the expressions:

$$p_i=\frac{g_i e^{-\beta\varepsilon_0}}{\sum_{i=0}^{2} g_i e^{-\beta\varepsilon_i}} \text{ or } \quad p_i=\frac{g_i e^{-\beta\varepsilon_0}}{q}$$

It is obtained the following populations:

$$p_0=0.25 \; p_1=0.73 \; p_2=0.02$$

In this system the partition function is q=4

c) With the equality of populations at the two lowest levels we obtain:

$$p_0 = p_1 \frac{g_0 e^{-\beta\varepsilon_0}}{q} = \frac{g_1 e^{-\beta\varepsilon_1}}{q} \quad e^{-\beta\varepsilon_0} = 5\, e^{-\beta\varepsilon_1} \qquad \beta = \frac{1}{RT} \qquad T = 99.9\,\mathrm{K} = -173.3\,\mathrm{C}$$

2) (1.5 points) Chemical kinetics.

The following data were obtained on the initial rate of isomerization of a compound S catalyzed by an enzyme E:

$[S]_0$ /(mmol dm^{-3})		**1.00**	**2.00**	**3.00**	**4.00**
V_0 /(mmol dm^{-3} s^{-1})	**(a)**	**4.5**	**9.0**	**15.0**	**18.0**
	(b)	**14.8**	**25.0**	**45.0**	**59.7**
	(c)	**58.9**	**120.0**	**180.0**	**238.0**

The enzyme concentrations are (a) 1.0 mmol dm^{-3}, (b) 3.00 mmol dm^{-3}, and (c) 10.0 mmol dm^{-3}. Find the orders of reactions with respect to S and E, and the rate constant. It is not required to use all data of the table.

$$v = k[\mathrm{S}]^{\alpha}[\mathrm{E}]^{\beta}$$

Taking two experiments at concentration of *S* constant, and two experiments at *E* constant, we can obtain the respective exponents using the following expressions:

$$\alpha = \frac{\ln \frac{v_1}{v_2}}{\ln \frac{[\mathrm{S}]_1}{[\mathrm{S}]_2}} \approx 1 \qquad \beta = \frac{\ln \frac{v_1}{v_2}}{\ln \frac{[\mathrm{E}]_1}{[\mathrm{E}]_2}} \approx 1$$

Thus , the reaction rate follows:

$$v = k[\mathrm{S}]^{1}[\mathrm{E}]^{1}$$

The rate constant is calculated taking any initial rate with its corresponding concentrations of S and E: $k = 4500\,\mathrm{L\,mol^{-1}s^{-1}} = 4.5\,\mathrm{dm^3\,mmol^{-1}s^{-1}}$

Name: Date:

3) (1.5 points) Transport. A tetrameric protein of 100 kDa is diffusing in the cytoplasm at 37ºC. i) Discuss how will be the diffusion of the monomers with respect to the tetrameric protein. ii) Discuss how will be the diffusion of the same protein in water at the same temperature.

b) Estimate the time that is required for a initial concentration of 15 mM of glycerol to escape from the interior of a cell of a diameter of 0.5 μm. The permeability through the membrane of the cell of glycerol is of $4 \cdot 10^{-8}$ m/s. Consider a spherical shape for the cell, a null concentration of glycerol outside the cell and approximate that the the initial flux, corresponding to the initial concentration of glycerol, is constant during time.

a) The effect of size and viscosity could be intrepreted through the Stokes-Einstein equation $D = \frac{k_B T}{6 \pi \eta a}$

i) A tetrameri protein has greater size than their corresponding monomers. Thus, the diffusion of monomers will be faster than the corresponding to the tetrameric protein.
ii) The viscosity in the cytoplasm is greatter than in the water. Thus, the diffusion of the protein in water will be faster (greater diffusion coefficient) than in the cytoplasm.

b)

$$\tau = \frac{\text{number of molecules inside}}{\text{amount lost each second}} = \frac{V C_{\square}}{\underset{\text{¿ } p A}{C_{\square} - C_{out}}}$$

$$\tau = \frac{r}{3p} = \frac{0.25 \cdot 10^{-6}\,\text{m}}{3 \cdot 4 \cdot 10^{-8}\,\text{m/s}} = 2\,\text{s}$$

4) (2.5 points) Justify briefly the following sentences (all them are true)

a. Protein stability can be evaluated experimentally by following the denaturation process

Following the denaturation process we can get the relations U/F at different degrees of unfolding, and deduce the corresponding ΔG that can be then extrapolated to obtain protein stability

b. Replacement of positive ionic residues located at the protein surface by negative or neutral polar residues is acceptable

At the surface polar and ionic interactions are made with the solvent, so these changes have little energetic implications.

c. Following enzyme catalysed chemical transformations requires the use of quantum mechanics methods

Only quantum mechanics methods can analyse bond reorganization, so chemical processes. The other methods assume a that bonds are fixed.

d. Macromolecules' stable conformations are thermodynamic energy minima

A "Stable conformation" requires a minimum in the energy landscape, as at minima there is no energy gradient that induce changes.

e. Increasing the yield of protein folding "in vivo" requires to stabilize unfolded structures.

Proteins can be trapped in misfolded conformations or aggregates. A strategy followed "in vivo" by chaperones, is to keep stable the unfolded conformations, and provide "second opportunities" to increase the overall yield.

f. Classic forcefields cannot be used to evaluate whether a protein fold is correct

Classic ff cannot evaluate entropy, so there is no solvation of hydrophobic effect, that drives folding. Statistical potentials are used instead.

g. Monte Carlo simulations is not the preferred approach to relax experimental structures

Relax experimental structures implies to drive them of a energy minimum as close as possible to the original coordinates. MC does random changes that can significantly alter the structure. The preferred method es molecular mechanics.

h. NPT are the recommended ensemble in Molecular Dynamics simulations

NPT keeps constant the number of particles, pressure and temperature. This is the most realistic situation.

Name: Date:

i. A hyperbolic shape in the plot of binding degree against ligand concentration indicates the participation of a protein in the process

A hyperbolic shape implies that beyond some ligand concentration, binding does not increase (Saturation). This behaviour is indicative of a limited number of binding sites, so a protein should participate. Otherwise the increase would be linear

j. The value of the dissociation constant of a protein-ligand complex can be obtained or estimated from the concentration of ligand giving 50% of saturation

The shape of a binding curve is hyperbolic, and can be adjusted to Y = L/(Kd+L). For Y = 0.5 L shoud be Kd

5) (1.5 points) Regarding macromolecular energies

a. (0.5 points) Classify the following energy terms according to their influence in macromolecule stability (favourable (F) / unfavourable (U) / indifferent (I))

i. U Macromolecule conformational entropy

ii. U Electrostatic interaction between main chain atoms of nucleic acids

iii. F Entropy of solvent

iv. I Electrostatic interaction between surface residues in proteins

v. F Van der Waals energies

Justification (not required):

i) Conformation entropy decreases on folding (less available conformations)
ii) Main chains in Nucl. Acids have a strong negative charge, the interaction is repulsive
iii) This is the hydrophobic effect that is the principal driving force for folding
iv) At the surface electrostatics is small due to the effect of solvent
v) VdW interactions are favourable as in improves contacts.

b. (1 point) Outline a procedure to parameterize the following energy terms (refer to equations in the annex when necessary). Indicate which parameters should be obtained and a possible strategy.
GENERAL COMMENT: The question was about obtaining parameters for the forcefield components, not about explaining the equations.

i. Bonded terms (bond distance, bond angles, and torsions)

Parameters to be obtained: Force constats (Kb), equilibrium values (r0), Energy barriers (Vn), periodic parameters (PN, PH).

Typical strategy: Obtain energy values for different molecular geometries (distances, angles, torsion) using quantum mechanics and obtain the parameters by fitting the appropriate equations to such energies.

ii. Hydrophobic solvation energies

They can be deduced from the correlation between energy and ASA
Parameters: Atom coefficients (σi)

Strategy: Obtain experimental ΔΔGsolv (a typical experiment is the transfer between water and an organic solvent), for a list of amino acid like compounds, calculate ASA for each atom and obtain the coefficient by linear regression.

iii. Electrostatic interactions

Using a coulomb law, parameters are the partial charges, and the dielectric constant. Charges can be obtained from quantum mechanics calculations, and dielectric constant is usually fixed by the type of calculation (1 for simulations, Mehler-Solmajer for point calculations, etc.)

iv. Statistical potential based in accessibility data

Statistical potentials are calibrated with known protein structures. This would require to obtain a statistical distribution of residue surfaces of all proteins at PDB, and derive a potential for each residue type based on such distribution.

Name: Date:

6) (1.5 points) We wish to evaluate the influence of several amino acid residues in the stability of a protein-protein complex, by analysing the interaction energy change after mutations to Ala. The following table summarizes calculated changes in interaction and solvation energies due to the modification:

Mutation	**ΔΔGint (Kcal/mol)**	**ΔΔGsolv (Kcal/mol)**
Glu 300 Ala	**10.5**	**-3.0**
Trp 260 Ala	**1.1**	**3.2**
Val 310 Ala	**0.1**	**0.3**
Gln 302 Ala	**1.2**	**-1.1**

a) Order the analysed residues from less to more importance in the complex stability

The global **ΔΔG** for the mutations is the combination of interaction and solvation so:

Mutation	**ΔΔGint (Kcal/mol)**	**ΔΔGsolv (Kcal/mol)**	**ΔΔG (Kcal/mol)**
Glu 300 Ala	10.5	-3.0	7.5
Trp 260 Ala	1.1	3.2	4.3
Val 310 Ala	0.1	0.3	0.4
Gln 302 Ala	1.2	-1.1	0.1

All **ΔΔG** are positive, so all mutations are distabilizing, and the order es

Gln302 < Val310 < Trp260 < Glu300

b) Suggest for each residue the reason behind the observed energy values

Glu 300: A large change is in the interaction energy. Since Glu is a negative amino acid, this suggest that the mutation may break a electrostatic interaction in the complex. Solvation change is negative as the solvation of the Glu side chain is not lost, what confirms that Glu is buried in the interface.

Trp 260: Low change in the interaction energy, that can be attributed to less Vdw interaction (Ala is a lot smaller than Trp). Most of the change comes from solvation, what is compatible with the larger surface of Trp, what in turn indicates that Trp is also buried in the interface.

Val 310: Same explanations as in the case of Trp, but the change from Val to Ala is much smaller.

Gln 302: Gln is polar but neutral, it may lose hydrogen bonds. Solvation effect is positive what indicates the Gln is in the interface (there is a surface change). In any case these effects are small, so the implication of this residue is not relevant.

c) Knowing that the experimental dissociation constant for the unmutated complex is 2.4 nM, evaluate the expected dissociation constants for the mutants.

$$\Delta\Delta G_{WT\to A}=\Delta G_{Ala}-\Delta G_{WT}=RT\ln K_D^{Ala}-RT\ln K_D^{WT}=RT\ln\frac{K_D^{Ala}}{K_D^{WT}}$$

$$K_D^{Ala}=K_D^{WT}e^{\frac{\Delta\Delta G_{WT\to A}}{RT}}$$

$$K_D^{WT}=2.4\,nM\;R=1.987\,\frac{cal}{molK},T=298\,K$$

Mutation	**ΔΔG (Kcal/mol)**	**K_D**
Glu 300 Ala	7.5	0.76 mM ($7.6\ 10^{-4}$ M)
Trp 260 Ala	4.3	3.4 μM ($3.4\ 10^{-6}$ M)
Val 310 Ala	0.4	4.7 nM ($4.7\ 10^{-9}$ M)
Gln 302 Ala	0.1	2.8 nM ($2.8\ 10^{-9}$ M)

Name: Date:

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=8.314 J K^{-1} mol^{-1}
R=0.082 atm L K^{-1} mol^{-1}
N_A=6.022·10^{23} mol^{-1}

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

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

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

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

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

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

Morse law for bond stretching energy

$$E_{pot}=D_M\left(1-e^{-a(r-r_0)}\right)^2$$

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 (length) or bending (angle) energies (Amber parm99 forcefield)

$$E_{pot}=\frac{K_B}{2}(r-r_0)^2$$

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

Periodic law for parameterization of bond torsions (Amber parm99 forcefield)

$$E_{pot}=\frac{V_N}{2}(1+\cos(PN\varphi-PH))$$

V_N: Barrier height (kcal mol-1); PN: Periodicity; PH; phase (degrees); ϕ: dihedral angle (degrees)

Correlation between ΔΔG transfer (oct->wat) and ASA (hydrophobic solvation)

$$\Delta\Delta G=\sum_{AtTypes}\sigma_i ASA_i$$

σ_i: Surface coefficient for atom type i. ASA_i: Surface (ASA) corresponding to cumulated atoms of type i

Coulombic interaction

$$E_{pot} = 332.16 \frac{q_i q_j}{\varepsilon r_{ij}}$$

q_i, q_j: partial charges (e); ε: relative dielectric constant; r_{ij}: distance (Å). Resulting energy is kcal mol^{-1}

Free energy related to dissociation constants

$$\Delta G = RT \ln K_D$$

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)