Name: Date:

Biophysics 2021-22
Bioinformatics degree
Final Exam. Part I

1) (1.5 points) Thermodynamics.

One polysaccharide could adopt three different conformations (A, B and C). At 42ºC degrees, the populations of A, B and C were determined to be 80%, 15% and 5%, respectively. Assuming that the degeneration of the conformation B is 2, and the degeneration of both A and C is 1, determine the populations of these three conformations of the polysaccharide at 2ºC. Indicate clearly the values of the energetic levels with appropriated units.

$p_j = \frac{g_j e^{-\beta\varepsilon_j}}{\sum_{i=0}^{2} g_i e^{-\beta\varepsilon_i}}$ with g_A= 1, g_B=2, g_C=1 and $\beta = \frac{1}{RT}$

The most stable level is the energetic reference: E_A= 0 J/mol

$$q = exp\,(-E_A/(RT))/p_A = 1/0.80 = 1.25$$

$$E_B = -RTln(p_B * q/g_B) = 6202\,J/mol$$

$$E_C = -RTln(p_C * q/g_C) = 7264\,J/mol$$

Calculation of the new partition function, *q*, at 2ºC considering 275K and the previous energies.

$$q = \sum g_j\, exp\big(-E_j/(RT)\big) = 1.17$$

and calculate the new proportions with $p_i = g_i exp\,(-E_i/(RT))/q$

$$p_0 = 0.85; p_1 = 0.11; p_2 = 0.036$$

Name: Date:

2) (1.5 points) Chemical kinetics

A beta blocker drug was administrated to a patient. The blood plasma of this patient was analyzed to determine the remaining content of the drug as a function of time. The following table of drug concentration in blood plasma at different times at 37ºC were determined:

t/min:	30	60	120
c/(ng/ml):	699	622	413

a) Determine the kinetic law for the elimination process of this beta blocker. Determine the kinetic constant.
b) Determine at which time the concentration of the drug should be 10 ng/ml.

Because the order it is not given, the data should be use to identify the most appropriated order.

First, it could be explored the order zero ($C=C_0-kt$), with the following expressions for each time:

$C_{30}=C_0-kt_{30}$
$C_{60}=C_0-kt_{60}$
$C_{120}=C_0-kt_{120}$

Combining first and second equation, for example, subtracting the second to the first, a first estimation of the kinetic constant is k=2.5 ng/ml min
Combining first and third equation, for example, subtracting the third to the first, a second estimation of the kinetic constant is k=3.2 ng/ml min
Different values of the kinetic constant are obtained from different values. Thus, we will check the first order.

Application of the data of the table to the first order (($\ln(C)=\ln(C_0)-kt$) gives the following expressions.

$\ln(C_{30})=\ln(C_0)-kt_{30}$
$\ln(C_{60})=\ln(C_0)-kt_{60}$
$\ln(C_{120})=\ln(C_0)-kt_{120}$

Combining first and second equation, the value of the constant is k=0.004 min^{-1}
Combining first and third equation, the value of the constant is k=0.006 min^{-1}
Less variation is obtained for first order kinetics. Thus, we can estimate the constant of this first order reaction with the average value k=0.005 min^{-1}

b) First we have to determine the value of C_0 from any of the previous equations. For example:

$$C_0 = C_{30} exp(-0.005 min^{-1} 30 min) = 812\, ng/ml$$

$$time = ln\left(C_0/C_{10ng/ml}\right)/k = 879 min$$

Name: Date:

3) (1.5 points) Transport.

a) Considering that the viscosity of water is about 10^{-3} N·s·m^{-2} and the viscosity inside the cytoplasm is $3.4 \cdot 10^{-3}$ N·s·m^{-2} at 25ºC. Estimate the diffusion coefficient of a protein of 3.8 nm of radius in water and inside the cytoplasm.

$D = \frac{k_B T}{6\pi\eta a}$ using the values of magnitudes, and considering that

$$10^9 nm = 1m$$
$$1J = 1Nm$$
$$10^9 nm = 1m$$
$$D_{cit} = 1.710^{-11}\, m^2/s$$

b) Determine the average time required for the protein of previous section (a) to travel a distance of 100 micrometers when it is immersed in water or in the cytoplasm.

$$R = 100\mu m = 10^{-4} m$$

$$t_{water} = \frac{R^2}{6D_{water}} = 29s$$

$$t_{cit} = \frac{R^2}{6D_{cit}} = 99s$$

c) In a cell there are 0.01M of glucose. If outside the cell there are a concentration of 0.001M of glucose, determine the flux knowing that the permeability of glucose is 10^{-3} nm/s. Give the flux using the units of mol, m and s.

$$j = -p(C_{ins} - C_{out})$$

Considering:

$$1M = mol/L$$
$$1L = 1m^3$$
$$j = -910^{-12} mol m^{-2} s^{-1}$$

Additional space for questions 1) , 2) or 3):

Name: Date:

Biophysics 2021-22
Bioinformatics degree
Final Exam. Part II

4) (2.5 points) Indicate and justify briefly whether the following sentences are **true or false**

a. The addition of denaturants or the increase in temperature allow to evaluate the stability of protein-protein complexes.

True. The addition of any compound (a denaturant) or a change in conditions (e.g. Temperature) that leads to the unfolding of the protein can be used to follow the process, obtaining the apparent ΔG at each experimental point (from U/F). In the case of protein-protein complexes it can be expected that the dissociation of the complex is the initial event followed later by the denaturation of the complex components (A-B → A + B → A^U + B^U)

b. A larger content of buried hydrophobic residues increases the protein stability

True. A larger content of buried hydrophobic residues imply a larger hydrophobic effect (related to the hydrophobic solvation energy of such residues) and therefore a larger protein stability.

c. Classic forcefields can be used to evaluate kinetic rate constants.

False. The evaluation of kinetic rate constants requires to analyze processes where the molecular structure changes (breaking or forming covalent bonds) or are severely distorted (transition states). Classic forcefields require a fixed molecular structure and equilibrium conformations and cannot be used to analyze chemical transformations (Quantum mechanics is required)

d. The protein folding process implies the exploration of all available conformations for the macromolecule and the selection of the most stable one.

False. The time required to explore and select ALL available conformations for a protein chain is astronomically large. Protein folding follows a preferred kinetic path from the unfolded to folded conformation. (This reason also justifies that the protein could eventually be trapped in incorrect folding structures)

e. Chaperones base their action in accelerating the folding process.

False. Chaperones stabilize unfolded structures to assure that they fold properly even at a lower rate.

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

True. Classic forcefields cannot differentiate between folded and unfolded structures as both are chemically correct. Proper evaluation would require including entropy. The only practical way to evaluate folding are statistical potentials.

g. Molecular mechanics is a required initial step in all simulations

True. Molecular mechanics assures that the structures are a mathematical minimum within the forcefield in use. It is required to avoid numerical instabilities in the simulation.

h. The NPT ensemble is not realistic, as all biological processes take place at constant volume.

False. Biological processes take place a constant pressure and temperature, so NPT is actually the most realistic.

Name: Date:

i. The evaluation of binding energies in macromolecular complexes require the evaluation of entropic energy components.

True. An essential component of binding energies are solvation energies that depends on solvent entropy. Also, conformational restrictions on complex formation can alter the conformational entropy of the complex components

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.

True. In a hyperbolic behaviour like protein-ligand binding the saturation degree follows Y = L/(Kd+L) which gives 0.5 when L = Kd.

5) (1 p) Classify the following energy terms according to their influence in macromolecule stability (favourable (F) / unfavourable (U) / indifferent (I))

i. *U/I* Interaction energies between polar residues and solvent
ii. *I* Hydrogen bonds between of nucleic acids bases
iii. *F* Entropy of solvent
iv. *I* Hydrogen bonds between protein residues
v. *F* Van der Waals energies

Justification (not required):
i: Interaction energies between polar residues and solvent are maximized in the unfolded protein. Any lost of these interactions on folding is unfavourable (U). However, as most of these residues remain in the surface, or form hydrogen bonds in the inside, the overall effect can be small (I).
ii, iv: Hydrogen bonds are always of best quality in the unfolded state, therefore they cannot contribute to the stability (I)
iii: Solvent entropy is the responsible of the hydrophobic effect, that is one of the major contributors to stability
v: Vdw contribution is small but always favourable, it defined the close packing of side chains in the folded structure.

6) (2 p) The analysis using FOLDX of some sequence variants on SARS-Cov-2 Spike protein from the Omicron strain regarding to the binding energy between RBD domain and ACE2 has given the following results. (Note: ΔΔG corresponds to $\Delta G_{Mut} - \Delta G_{Wt}$):

Variant	ΔΔGvdw (kcal/mol)	ΔΔGelec (kcal/mol)	ΔΔGsolv (kcal/mol)
Lys417Asn	0.56	2.23	-1.71
Glu484Ala	0	0	-0.23
Gln493Lys	-0.44	0.84	-0.20
Gly496Ser	-0.12	-0.85	0.35
Gln498Arg	-1.63	1.75	-0.24
Asn501Tyr	4.94	0.61	-2.08
Tyr505His	0.14	1.05	1.02

Name: Date:

a) From the data provided, which would be the order of stability of the complex ACE2-RBD

b) Calculate the ratio of binding constants with respect to the complex made with WT Spike

Total ΔΔGbind for each individual mutant is the sum of all energy components.components. The Ratio between binding constants comes from

$$\Delta\Delta G_{bind} = \Delta G_{mut} - \Delta G_{wt} = -RT \ln K_{Bind}^{Mut} + RT \ln K_{Bind}^{WT} = -RT \ln \frac{K_{Bind}^{Mut}}{K_{Bind}^{WT}} = +RT \ln \frac{K_D^{Mut}}{K_D^{WT}}$$

Note that the ratios deduced from K_{bind} are the inverse from ratios deduced from K_D. Both were correct as long there were correctly indicated.

Variant	$\Delta\Delta G_{vdw}$ (Kcal/mol)	$\Delta\Delta G_{elec}$ (Kcal/mol)	$\Delta\Delta G_{solv}$ (Kcal/mol)	$\Delta\Delta Gbind_v$ (Kcal/mol)	$\frac{K_{Bind}^{Mut}}{K_{Bind}^{WT}}$	$\frac{K_D^{Mut}}{K_D^{WT}}$	Reasons
Lys417Asn	0.56	2.23	-1.71	1.08	0.16	6.20	b,c
Glu484Ala	0	0	-0.23	-0.23	1.47	0.68	c
Gln493Lys	-0.44	0.84	-0.20	0.2	0.71	1.40	b,d
Gly496Ser	-0.12	-0.85	0.35	-0.62	2.85	0.35	d
Gln498Arg	-1.63	1.75	-0.24	-0.12	1.22	0.82	b,c,d
Asn501Tyr	4.94	0.61	-2.08	3.47	$2.85\ 10^{-3}$	350.8	a,b,c
Tyr505His	0.14	1.05	1.02	2.21	$2.39\ 10^{-2}$	41.78	b,e

The order of COMPLEX STABILITY from most to least stable is

Gly486Ser > Glu484Ala > Gln498Arg > Gln493Lys > Lys417Asn > Tyr505His > Asn501Tyr

*Note: Always use the correct value of the R constant **considering the units**. Be aware that R = 1.987 **cal/K.mol** or 0.001987 **Kcal/K.mol**. R can be also used in J/K.mol but in that case energies should be converted t**o** KJ/mol*

c) For each variant choose from the following options, reasons that help to explain the change

Note that this was not just a theoretical exercise, it should be solved considering the energy values shown in the table.

a. Vdw collision due to increased sidechain

Can be recognized by a high, and positive $\Delta\Delta G_{vdw}$. In this case only Asn501Tyr.

b. Loss of an electrostatic interaction/Hbond

Recognized by a positive $\Delta\Delta G_{elec}$. Appears in Lys417Asn (loss of electrostatic interaction) and also in less extent (compatible with Hbond pattern change) in Gln493Lys, Gln408Arg, Asn501Tyr, Tyr505His

c. Increased hydrophobic effect

Recognized by a negative (favourable) $\Delta\Delta G_{solv}$ but only where the major component of the solvation is hydrophobic (the mutant involved apolar amino-acids): Glu484Ala, Asn501Tyr. Other favourable solvations are probably due to the electrostatic component

Name: Date:

d. Improved sidechain packing

Recognized by a favourable (negative) $\Delta\Delta G_{vdw..}$ Gln493Lys, Gly496Ser, Gln498Arg

e. Decreased hydrophobic effect

Recognized by a positive (unfavourable) $\Delta\Delta G_{solv}$ but only where the major component of the solvation is hydrophobic: Tyr505His.

d) Considering that Omicron Spike is efficient in infection through the binding to ACE2, what can be said about the results obtained?

Considering that Omicron strain contains ALL the mutations shown in the table, and assuming that the effects are additive, the total ΔΔG would be +5.99 Kcal/mol, indicating that the complex ACE2-RBD would ~24,000 fold LESS stable with respect to the WT case. Since the experimental result is that the complex is formed, the only conclusion is that the FOLDX ***energy values are overestimated****, and the theoretical analysis should be re-evaluated.*

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=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}

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

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

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

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

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

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