Norwegian University of
Science and Technology,
Department of Physics

EXAM I COURSE
TFY4310 MOLECULAR BIOPHYSICS

Suggested solutions

Friday, 18 December 2020
Time: kl. 9.00 - 12.00

All questions have the same weight. None of the questions require lengthy answers so answer as precisely and concisely as possible. Good luck!

Exercise 1.

Justify the following (correct) sentences:

1. Molecular modeling using implicit solvent models is not suitable for studying the formation of lipid membranes.

   Answer: In molecular modeling it is common to use implicit solvent models to decrease the computational time. With implicit solvent model, the water molecules are not explicitely taken into account, but the solvent is described as a continuum with a particular dielectric constant.

   The driving force for the formation of lipid membranes is the hydrophobic effect between the apolar parts (hydrophobic tails) of the lipid molecules. This effect arises from the fact that water molecules can not form hydrogen bonds with apolar molecules. Due to the sp^3 hybridization of the oxygen, and the tetrahedral shape of the molecule, each water has the potential of establishing 4 hydrogen bonds. In liquid water, due to the thermal motions, one water molecule forms about 3.4 bonds with other waters. When a polar molecule is placed in a water solution, the water molecules are not able to establish H-bonds with it and will arrange at its surface so not to loose bonds with other waters. This leads to an increase in the number of H-bonds between water molecules that are organized around the apolar molecule. While this leads to a more negative potential energy, it also contributes to a loss in entropy of the water molecules. By 'pushing' apolar molecules together, the apolar area that is exposed to the water is reduced, which leads to the 'release' of some of the organized water molecules, leading to an increase in entropy.

   Since the water molecules are not explicitely taken into account in implicit solvent models, the hydrophobic effect can not be probed using these models. Hydrophobic interaction in these models are usually described using attractive potential energies instead, such as the Lennard-Jones potential.

2. The complexation of lysozyme proteins with DNA is favorable from an entropic point of view.

Answer: At a first glance it may sounds strange that the formation of complexes between lysozyme proteins with DNA is entropically favorable, since the mixing entropy of the proteins and the DNA decreases.

Both lysozyme and DNA molecules possess groups that are dissociated when dissolved in aqueous solution. At neutral pH the lysozyme proteins are overall positively charged and the DNA is negatively charged. However, not all groups are dissociated and many of the counterions will remain associated to the surface of the macromolecules. This phenomenon is called ion condensation, and the associated counterions form the so-called double layer.

When one lysozyme protein associates with one DNA molecule, some of counterions from the lysozyme and some from the DNA are released. Thus, the entropy of mixing from the counterions is much increased upon lysozyme-DNA complexation. In addition, some of the hydration water from the macromolecules may be released too.

3. During the second lab you have seen (in principle) that the fluorescence intensity of the sample decreases when the concentration of protein increases. However, the autofluorescence of the proteins increases instead.

   Answer: During the lab you followed the fluorescence intensity from a fluorophore called GelStar. The assay is based on the fact that the fluorescence intensity of the dye is much higher when it is bound to DNA molecules than when it is in aqueous solution. When lysozyme is added to the DNA + Gelstar solution, the lysozyme binds to the DNA molecules and the dye molecules are released from the DNA into the solution. Differences in the hydrophobicity of the environment leads to the so-called quenching of the dye molecule and corresponding decrease of the fluorescence intensity.

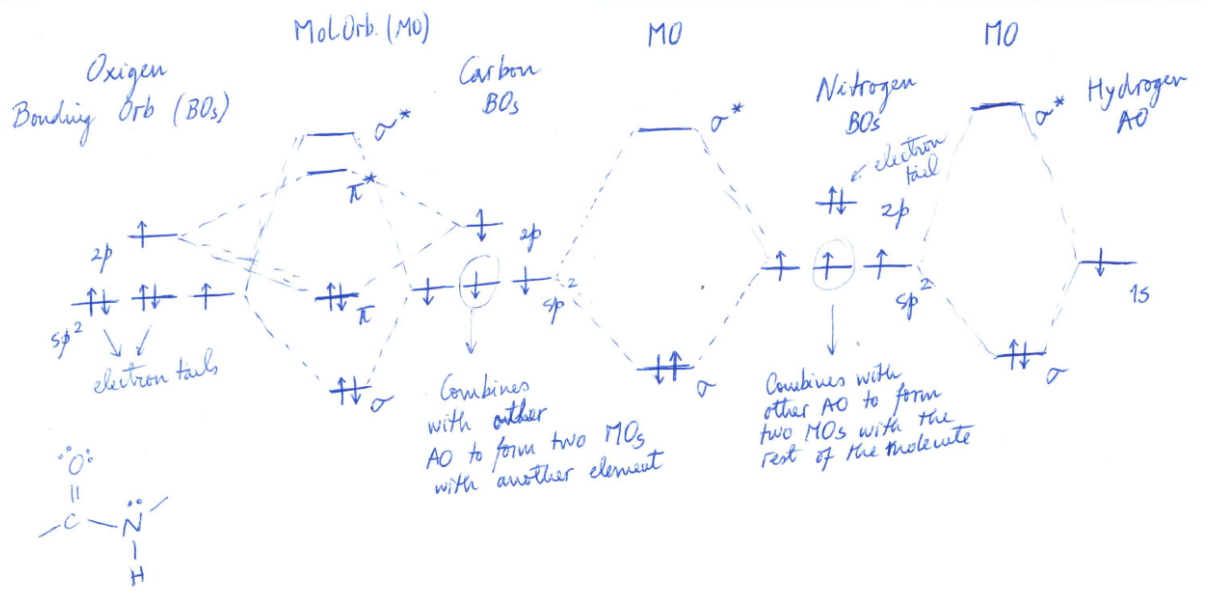
   Due to the tryptophan and tyrosine aminoacids, containing aromatic groups, the lysozyme also shows autofluorescence. This can be followed by exciting the samples with UV light at 295 nm and measuring the fluorescence emitted at 350 nm. The interaction with the DNA does not change significantly the emission intensity from these groups, but the increase in protein concentration leads to the an increase in the autofluorescence.

Exercise 2 and 3. The peptide bond is a recurrent group in biology. What is the hybridization of the carbon, nitrogen and oxygen? Describe (or draw and upload) the bonding orbitals for each of the represented atoms an discuss (or draw) how they combine to form molecular orbitals. Note that this molecule is planar.

Answer: The hybridization of the oxygen and the carbon are sp^2. Because the molecule is planar, the hybridization of the nitrogen must also be sp^2.

The energy-level describing the hybrid and molecular orbitals (MOs) is depicted below. Starting from C=O, one of sp^2 orbitals of the carbon linearly combines with one of the sp^2 orbitals of the oxygen to form a σ and a $\sigma^\star$ molecular orbitals. The lower level one is occupied by the each of the electrons from the bonding orbitals. The other two sp^2 orbitals from oxygen are fully occupied and do not contribute to molecular orbitals. The $2p$ atomic orbitals of the carbon and oxygen linearly combine to form a π and an anti-bonding π molecular orbitals, giving the double bond. The second sp^2 orbital of the carbon combines with another AO to give two MOs with the rest of the molecule, and the third combines with a sp^2 from the nitrogen to form two more MOs (σ and a $\sigma^\star$). The second sp^2 bonding

orbital combines with the $1s$ of the hydrogen and the third one to an AO of the element bound to the nitrogen to form two more MOs. We are left with the $2p$ orbital of the nitrogen, which is fully occupied.



Exercise 4. From the perspective of protecting yourself against infection by corona viruses, it is known that it is most efficient to wash your hand with soap than hand sanitizer. Why? If you were to prepare the soap formulation which characteristics would you chose for its components? Justify.

Answer: Corona viruses are coated with a lipid membrane. The spike proteins that are responsible for attaching to the receptors of the cells and facilitate the entrance of the virus in the cells are membrane proteins associated to this lipid membrane.

The most important component of soaps are surfactants. These are amphiphilic molecules, characterized by the hydrophophilic headgroup and hydrophobic tail. When the viruses are exposed to soap, the surfactant in the formulation penetrate the lipid membrane forming mixed lipid-surfactant aggregates. Depending on the quantity and the characteristic of the surfactant, the lamellar-like structure of the lipid membrane will change, leading for example to the formation of more spherical-like aggregates. This will dissolve the lipid membrane around the virus, contributing to, at least, the detachment of the spike proteins, potentially also to the destruction of the virus.

The surfactant should be chosen so that it changes the packing parameter (CPP) of the lipids as much as possible, that is away from a PCC of 1 and a lamellar-like structure. Considering that lipids are usually rather long, the easiest is to reduce the CPP and induce the formation of normal (as opposed to inverted) surfactant micelles. This could be achieved using relatively short surfactants with large headgroups, such as surfactants with a charged headgroup.

Some students focused their answer on the CMC of the surfactants. A long as the CMC was discussed in terms of the surfactant properties, some points were awarded in the question.

Exercises 5. Calculate the interaction between two oppositely charged proteins (A and B) in an aqueous solution with 100 mM Na_3PO_4. Assume the proteins are in contact.

$Z_A = 15e$, $R_A = 2.0$ nm, $Z_B = -10e$, $R_B = 1.5$ nm, $T = 20°$C.

Answer: Using the screened Coulomb potential:

$$V(r) = \frac{z_1 z_2\, e^2}{4\pi\varepsilon_0\varepsilon r} \exp\left(-\kappa r\right) \ ,$$

where κ is the inverse of the Debye screening length, defined as:

$$\lambda_D = \frac{1}{\kappa} = \left(\frac{\varepsilon_0\, \varepsilon k_{\mathrm{B}} T}{2e^2 I N_{\mathrm{Av}} \cdot 1000}\right)^{1/2} \ ,$$

with $I = \frac{1}{2}\sum_{i=1}^{N} z_i^2 c_{\infty i}$ and $c_{\infty i}$ is the bulk concentration of ions in solution.

First we begin by calculating the ionic strength, which gives $I = \frac{1}{2}(1^2 \cdot 300 + (-3)^2 \cdot 100) = 600$ mM $= 0.6$ M.

The Debye screening length is calculated according to

$$\lambda_D = \frac{1}{\kappa} = \left(\frac{8.854 \times 10^{-12} \cdot 78.4 \cdot 1.38 \times 10^{-23} \cdot 293.15}{2(1.602 \times 10^{-19})^2 \cdot 0.60 \cdot 6.022 \times 10^{23} \cdot 1000}\right)^{1/2} = 3.90 \times 10^{-10} \text{ m} \ .$$

Since the proteins are assumed to be in contact, r the center-to-center distance is taken as the sum of the protein radii, $r = R_A + R_B$.

$$V(r = 2.5\,\text{nm}) = \frac{15 \cdot (-10) \cdot (1.602 \times 10^{-29})^2}{4\pi \cdot 8.854 \times 10^{-12} \cdot 3.5 \times 10^{-9}} \exp\left(-\frac{3.50 \times 10^{-9}}{3.90 \times 10^{-10}}\right) = 1.60 \times 10^{-23} \text{ J} \ .$$

Exercises 6. Justify, shortly, the answer you gave in the previous question.

In the previous question it is assumed that the proteins are in contact. Is this a good assumption?

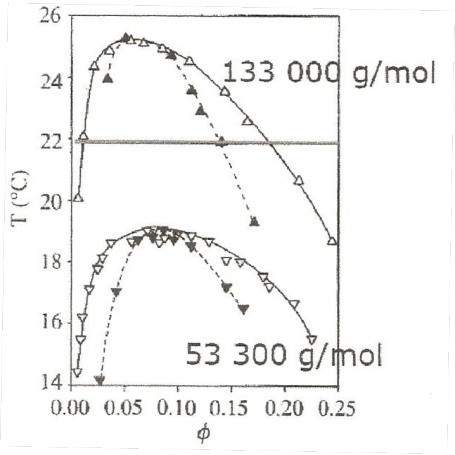
Answer: The proteins are oppositely charged so Coulomb interactions are the most predominant in this system. The proteins are in a salt solution so we need to take the screening effect of the salt ions. The ionic strength was calculated to be 0.6 M, and Deby screening length 3.90 Å and the Coulomb potential energy 1.60×10^{-23} J.

We can see from the result that the potential energy between the proteins in the salt solution is lower than the thermal energy ($1\ k_{\mathrm{B}}T = 1.38 \times 10^{-23} \times 293.15 = 4.045 \times 10^{-21}$ J). Since the interaction is not strong enough to prevent the proteins from drifting away due to collisions with the solvent, it is not a good assumption that they are in contact in this situation.

Exercise 7. The polyisoprene/dioxane system has a strongly asymmetric phase diagram. The figure below shows the phase diagram for two polymers with different molecular weights (M = 133,000 g/mol and M =53,300 g/mol).

- What do the different data points (open and filled triangles) correspond to?

  Answer: The open triangles correspond to the binodal points, that is the polymer composition at which the free energy of mixing of the polymer in the solvent is a minimum, for a particular temperature. It defines the range of polymer concentration at which the polymer is miscible in the solvent. The filled triangles correspond to the spinodal points, that is, the inflexion points in the free energy of mixing of the polymer, calculated by equalizing the second derivative of $\Delta G_{mix}/N$ as a function composition to zero.



- For M=133,000 g/mol at 24 °C, indicate the composition range of miscibility.

  Answer: The miscibility gap of the polymer with M=133,000 g/mol at 24 °C is roughly $\phi = 0.02$ to 0.15, that is, the area between the binodal points. The miscibility regions correspond to polyisoprene/dioxane compositions below $\phi = 0.02$ and above 0.15.

  For the same polymer at T = 17 °C, for example, the polymer shows no miscibility throughout the given range of composition (0 - 0.25).

- Considering the polymer with M = 53,300 g/mol. Assume that the theta temperature is 22 °C. Discuss, justifying, the conformation of the polymer under the following conditions:

  (i) T = 16°C, $\phi = 0.01$;

  Answer: At T = 16°C, the polymer solution is below the theta temperature, which indicates that at that T the solvent is a bad solvent. We can see in the phase diagram that at $\phi = 0.01$ we have a diluted one phase solution, so the polymers will be surrounded by solvent. Since the solvent is a bad solvent, the polymer will be in a compact conformation, with a scaling parameter of 1/3.

  (ii) T = 16 °C, $\phi = 0.10$;

  Answer: In this case, the T is still below the theta temperature, but the system is precipitated, that is, the majority of the polymer chains are in the precipitate in contact with other polymers, not the solvent. According to the definition of an ideal polymer chain, the monomer - solvent interactions are the same as the monomer - monomer interactions. In this case, the monomers and also the solvent, and so, the polymer in a precipitate (or polymer melt) has an ideal behavior, with a scaling parameter of 1/2.

  (iii) T = 22 °C, $\phi = 0.01$;

  Answer: At T = 22 °C, we are at the theta temperature and so the polymer behaves as an ideal chain ($\alpha = 1/2$), independently of the polymer concentration.

  (iv) T = 24 °C, $\phi = 0.20$.

  Answer: Above the theta temperature, the solvent is a good solvent for the polymer, meaning that the interactions between the polymer and solvent and more favorable than the monomer-monomer ones. The chain will thus have a

more expanded conformation and the end-to-end radius of the polymer scales with the monomer length with a scaling parameter of 3/5, independently of the polymer concentration.

Exercises 8 and 9. The molecule CH_2A-CHA-O-CHX-CH_3, with A and X unknown elements, gives rise to a ^{1}H-NMR spectra with 4 peaks:

Multiplicity	δ (ppm)	Relative area
doublet	1.2	3
doublet	2.5	2
quartet	3.8	1
triplet	4.2	1

a. Identify the protons that correspond to each of the peaks.

Answer: Based on the relative areas we can say that the first doublet, with a relative area of 3, corresponds to the CH_3 group in the end of the molecule. The second doublet with a relative area of 2, corresponds to the protons in the CH_2A group. We are left with two peaks with relative areas 1 and we can use the multiplicity to identify the protons, based on the neighboring groups. The quartet indicates a proton that are J-coupled with a group having 3 protons and thus correspond to the CHX group. Finally the triplet, corresponds to the proton that is coupled to a group with 2 protons, that is CH_2A.

b. Which of the unknown elements (A or X) is more electronegative? Justify.

Answer: When comparing the position of the peaks arising from the groups CHX (δ = 3.8 ppm) and CHA (δ = 4.2 ppm), both bound to an oxygen, we see that the proton in group CHA has a larger chemical shift than the one in group CHX. This indicates that the proton is less shielded, that is more exposed to the magnetic field, and thus suggests that A is more electronegative than X.

c. Describe or draw the COSY ^{1}H-NMR spectrum of the molecule.

Answer: A COSY ^{1}H-NMR spectrum is a 2D ^{1}H-NMR spectrum where the two axis refer to the chemical shift (or frequency). Contours are used to indicate the intensity of the peaks. COSY, shortening to COrrelation SpectroscopY, probes, as the common 1D ^{1}H-NMR the J-coupling or through bond correlations between neighboring protons. When plotting a COSY spectra, the diagonal of the plot, will show the peaks present in the 1D ^{1}H-NMR spectrum. Does we will have four diagonal peaks in the positions (1.2, 1.2), (2.5, 2.5), (3.8, 3.8), and (4.2).

In addition to these we expect to see off-diagonal peaks that identify protons that are at a different chemical shift but that are J-coupled. For this particular molecules these peaks would be in the positions: (1.2, 3.8), (3.8, 1.2), (2.5, 4.2), and (4.2, 2.5).

Exercise 10. The partial specific volume of a protein in water at 20 °C is 0.75 cm^3/g and the molecular weight is 50,000 g/mol. The translational diffusion coefficient of the protein is 2.9×10^{-13} m^2/s. The hydration fraction is 0.5.

- Calculate the radius of the sphere that has the same volume as the protein.

Answer: To calculate the radius of the sphere that has the same volume as the protein we use: From the definition of partial specific volume and, assuming the protein to be spherical we can write:

$$\overline{V}_1^{(S)} = v_1 \left(\frac{N_{\rm Av}}{M_1}\right) = \left(\frac{4\pi}{3}R^3\right)\left(\frac{N_{\rm Av}}{M_1}\right)$$

$$\Rightarrow\ R = \left(\frac{3\overline{V}_1^{(S)} M_1}{4\pi N_{\rm Av}}\right)^{1/3} = \left(\frac{3 \cdot 0.75\,[\mathrm{cm^3/g}] \cdot 50,000\,[\mathrm{gmol^{-1}}]}{4\pi \cdot 6.022 \times 10^{23}\,[\mathrm{mol^{-1}}]}\right)^{1/3} = 246\,\mathrm{nm}$$

- Calculate the sedimentation coefficient of the protein.

  Answer: To calculate the sedimentation coefficient one can simply use:

$$s = \left(1 - \overline{V}_1^{(S)}\rho\right)\frac{M_1}{N_{\rm Av} f} = \left(1 - \overline{V}_1^{(S)}\rho\right)\frac{M_1 D_T}{N_{\rm Av} k_{\rm B} T} = 1.40 \times 10^{-12}\ \mathrm{s} = 14\ \mathrm{S}\ .$$

Exercise 11. The following data were obtained from dynamic scattering of a protein as a function of the delay time:

$\tau(\mu s)$	0	20	40	60	80	100	120
$g^{(2)}(\tau)$	1.55	1.45	1.36	1.30	1.24	1.19	1.16

What is the friction coefficient if the protein?

$\theta = 90°$; $n_0 = 1.33$; $\lambda_0 = 456$ nm; $T = 20°$C.

Answer: One can relate the time correlation function with the translational diffusion coefficient using:

$$g^{(2)}(q,\tau) = 1 + [g^{(1)}(q,\tau)]^2 \qquad \text{and} \qquad g^{(1)}(q,\tau) = \exp(-q^2 D_T \tau)\ .$$

where $g^{(2)}(q,\tau)$ corresponds to the normalized time correlation function of the scattered intensity (what is measured) and $g^{(1)}(q,\tau)$ to the electric correlation function. One gets:

$$\ln\left[g^{(2)}(\tau) - 1\right] = -2q^2 D_T \tau.$$

Thus, by plotting the $\ln[g^{(2)}(q,\tau) - 1]$ as a function of τ, one can calculate the translational diffusion from the slope and the relation:

$$\text{slope} = -2q^2 D_T\ .$$

The slope is found to be -0.0104×10^6 s^{-1}.

The scattering vector is calculated using $q = \frac{4\pi}{\lambda}\sin(\frac{\theta}{2})$, where λ is the wavelength of light in the solution and θ is the angle between the incident light and the measured scattered light (detector):

$$q = \frac{4\pi n_0}{\lambda_0}\sin\left(\frac{\theta}{2}\right) = \frac{4\pi 1.33}{456 \times 10^{-9}}\sin\left(\frac{90}{2}\right) = 2.592 \times 10^7\ \mathrm{m^{-1}}$$

Relating the slope with the friction coefficient:

$$D_T = -\frac{\text{slope}}{2q^2}\ .$$

Since $f D_T = k_{\rm B} T$:

$$f = -\frac{2q^2 k_{\rm B} T}{\text{slope}} = -\frac{2 \cdot (2.592 \times 10^7)^2 1.38 \times 10^{-23} \cdot 293.15}{-0.0104 \times 10^6} = 5.23 \times 10^{-10}\ \mathrm{kg/s}\ .$$