Norwegian University of
Science and Technology,
Department of Physics

Contact during the exam:
Rita de Sousa Dias
Phone 47155399

EXAM I COURSE
TFY4310 MOLECULAR BIOPHYSICS

Suggested solutions

Monday, 2 December 2019
Time: kl. 15.00 - 19.00

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 **six** (6) of the following correct sentences:

a. The interaction between two (spherical) dipole molecules, described as two electric charges separated by 0.1 nm, with a radius of 0.14 nm, at room temperature (T=20 °C) is strong enough to orient the dipoles. Due to the proximity of the two dipoles assume that $\varepsilon = 1$.

Answer: The charge-dipole potential energy can be calculated according to:

$$V(r, \theta_1, \theta_2, \phi) = -\frac{u_1 u_2}{4\pi\varepsilon_0 \varepsilon r^3}[2\cos\theta_1 \cos\theta_2 - \sin\theta_1 \sin\theta_2 \cos\phi] \ .$$

where $u_1 = u_2 = ql = 1.602 \times 10^{-19} \cdot 0.1 \times 10^{-9} = 1.602 \times 10^{-29}$ Cm.

In order to maximize the interaction we consider that the dipoles are as close as they can be without overlapping ($r = 2 \cdot 0.14 = 0.28$ nm) and that the dipoles are oriented so that $\theta_1 = \theta_2 = 0$ and thus $\cos\theta_1 = \cos\theta_2 = 1$ (all other orientations will decrease the strength of the interaction). The value of ϕ is irrelevant. Taking the data from the formula in the end of the exam:

$$V(0.28, 0, 0, \phi) = -\frac{2 \cdot (1.602 \times 10^{-29})^2}{4\pi(8.854 \times 10^{-12})(0.28 \times 10^{-9})^3} = -2.10 \times 10^{-19} \quad \text{J} \ .$$

To assess if this interaction is strong enough to orient the dipole molecules, we need to calculate the thermal energy: $1\ k_B T = 1.38 \times 10^{-23} \times 293.15 = 4.045 \times 10^{-21}$ J. The interaction energy between the dipoles is strong enough to oppose the thermal motion and orient the dipoles.

b. The overlap concentration of a polyethelene -[-CH_2-]-$_n$ with 10,000 monomers is 1.1 M. Assume that the Kuhn length is about 1 nm, corresponding to seven CH_2 units and that the chain is under θ conditions (behaves like an ideal chain). Treat the chain as a freely-jointed model.

Answer: The overlap concentration is defined as the polymer concentration at which the polymer coils begin to overlap. The volume of the polymer coil can thus be simply deduced from the volume of the solution divided by the number of polymer molecules, N, in solution, which gives the equation (also present in formula list):

$$C^\star = \frac{3N_p}{4\pi N_{\mathrm{Av}}} \frac{1}{R_G^3} \,,$$

where N_p is the number of monomers and R_G the radius of gyration. Since we have an ideal chain, we can use the given relation between R_G and R_{ee}. Starting with the end-to-end distance, the mean square is given by

$$\left\langle R_{\mathrm{ee}}^2 \right\rangle = C_n Q^2 n^{2\alpha} \,,$$

where $C_n = 1$ for the freely jointed model and α, the scaling factor is 1/2 for ideal chains. n the number of Kuhn segments is 10,000/7=1,428.6 and the Q the length of the Kuhn segment, 1 nm:

$$\left\langle R_{\mathrm{ee}}^2 \right\rangle = 1 \cdot 1^2 \cdot 1,428.6 = 1,428.6\,\mathrm{nm}^2 \,.$$

Since the chain is ideal we can use the relation

$$\left\langle R_{\mathrm{ee}}^2 \right\rangle = 6 \left\langle R_G^2 \right\rangle \,,$$

and obtain

$$\left\langle R_G^2 \right\rangle = \frac{1,428.6}{6} = 238.1\,\mathrm{nm}^2 \,.$$

and

$$\left\langle R_G^2 \right\rangle^{1/2} = 15.43\,\mathrm{nm} \,.$$

Finally, the overlap concentration is calculated from

$$C^\star = \frac{3N_p}{4\pi N_{\mathrm{Av}}} \frac{1}{R_G^3} = \frac{3 \cdot 10,000}{4\pi \cdot 6.022 \times 10^{23}\,[\mathrm{mol}^{-1}]} \frac{1}{(15.43 \times 10^{-9}\,[\mathrm{m}])^3}$$

$$= 1079\,\mathrm{mol} \cdot \mathrm{m}^{-3} = 1.08\,\mathrm{mol} \cdot \mathrm{dm}^{-3} \,.$$

c. Increasing the charge density of the polyelectrolyte network (charged gel) leads to an increase in the swelling equilibrium of the network.

Answer: Two competing phenomena occur when a (neutral) polymer network is mixed with a solvent. The mixing free energy generally leads to the expansion (swelling) of the network and the elastic contribution limits such swelling. The equilibrium swelling is reached when these two contribution balance each other. When the network is charged, counterions to the network are drawn into the network, which in turn leads to the movement of water into the network and consequent swelling of the network. Increasing the charge density of the network leads to an increase in concentration of counterions and water inside the gel, leading to a larger swelling.

d. In Raman IR spectroscopy, bands corresponding to C=O stretching vibrations appear at longer wavenumbers (shorter wavelengths) than those of C–O.

Answer: The strength of the covalent bond is larger in the C=O group than in the C–O group, and so a larger energy is required to excite the vibrational mode of the C=O bond. Since the energy is inversely proportional to the wavelength, the band corresponding to C=O stretching vibration will appear at shorter wavelengths or at longer wavenumbers.

e. The fluorescence emission spectrum of a molecule is (to some extent) independent of the wavelength of excitation.

Answer: In fluorescence, a molecule is first excited, by absorbing a photon, from its ground electronic state to one of the various vibrational states in the excited electronic state. In the majority of the cases, the excitation undergoes internal conversion to the ground vibrational level of the excited electronic state. The molecule then drops down to one of the various vibrational levels of the ground electronic state again, emitting a photon in the process. In other words, it does not matter to which vibrational state the molecule is excited to (*i. e.*, independence of the wavelength of excitation) since the emission from the excited state will initiate from the ground vibrational state.

f. Lamm's equation, which describes the sedimentation and diffusion of macromolecules, is valid only for dilute solutions.

Answer: Lamm's equation is based on Fick's second law, which assumes that the translational diffusion coefficient of the molecules under study is independent of the concentration. This is only valid if the sample are diluted. We can also see, in Lamm's equation, that both diffusion coefficient and sedimentation coefficient are independent of time and the spatial distribution of the molecules in the cuvette, which would not be the case if the samples were concentrated.

g. In moving boundary experiments (sedimentation centrifugation) the rotor speed is kept high.

Answer: In centrifugation experiments there are two competing mass transport phenomena. There is the sedimentation of molecules towards the bottom of the cell, which creates a concentration gradient. The formation of the concentration gradient, leads to the diffusion of macromolecules in the opposite direction, and consequently, the 'blurring' of the boundary (between macromolecule solution and solvent). In moving boundary one wants to create sharp boundaries for the evaluation of concentration either at the boundary or the plateau (depending on the measuring technique), thus by keeping the rotor speed high one minimized the diffusion at the boundaries.

h. Light scattering sometimes interfere seriously with absorption measurements on large proteins or nucleic acid particles. One solution is to add sucrose to the solution.

Answer: Addition of sucrose raises the index of refraction of the solvent and thus lowers then contrast between the macromolecule and the solvent. This substantially reduces scattering.

Exercise 2. The translational diffusion coefficient of hemoglobin in water at 20 °C is 6.37×10^{-11} m^2/s. The specific partial volume is 0.75 cm^{3}g^{-1} and the molecular mass 64,500 g mol^{-1}.

a. Assuming hemoglobin to be a hydrated sphere, calculate the translational friction coefficient of the hemoglobin and its hydrodynamic radius.

Answer: The translational friction coefficient can be calculated from:

$$f = \frac{k_B T}{D_T} = \frac{1.38 \times 10^{-23}\,[\mathrm{J/K}] \times 293\,[\mathrm{K}]}{6.37 \times 10^{-11}\,[\mathrm{m^2 s^{-1}}]} = 6.351 \times 10^{-11} \quad \mathrm{Kg/s}\ .$$

The hydrodynamic radius, R_h is determined from:

$$f = 6\pi\,\eta\,R_h \quad \Rightarrow \quad R_h = \frac{6.343 \times 10^{-11}\,[\mathrm{Kg/s}]}{6\pi \cdot 1.0 \times 10^{-3}\,[\mathrm{kgm^{-1}s^{-1}}]} = 3.369 \times 10^{-9}\,\mathrm{m} = 3.37\,\mathrm{nm}\ .$$

b. Calculate the radius of the non-hydrated hemoglobin, assuming this form is also spherical.

Answer: From the definition of partial specific volume and, assuming the protein to be spherical we can write:

$$\overline{V}_1^{(S)} = v_1 \left(\frac{N_{\mathrm{Av}}}{M_1}\right) = \left(\frac{4\pi}{3}R^3\right)\left(\frac{N_{\mathrm{Av}}}{M_1}\right) \Rightarrow R = \left(\frac{3\overline{V}_1^{(S)} M_1}{4\pi N_{\mathrm{Av}}}\right)^{1/3} = 2.68\,\mathrm{nm}$$

As expected the non-hydrated radius is smaller than the hydrodynamic radius.

c. Using the answers obtained for questions 2a. and 2c., calculate the number of water molecules that hydrate a single hemoglobin molecule.

Answer: First we need to calculate the hydration fraction, using:

$$v_{\mathrm{h},1} = \left(\overline{V}_1^{(S)} + \delta\overline{V}_0^{(S)}\right)\frac{M_1}{N_{\mathrm{Av}}} \Rightarrow \left(\frac{4\pi}{3}R_h^3\right) = \left(\overline{V}_1^{(S)} + \delta\overline{V}_0^{(S)}\right)\frac{M_1}{N_{\mathrm{Av}}}$$

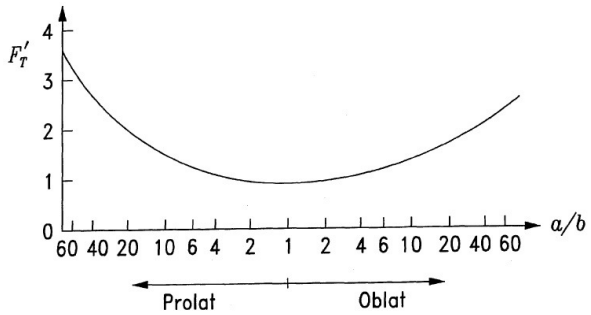
$$\left(\frac{4\pi}{3}(3.37\times10^{-7}\,[\mathrm{cm}])^3\right) = \left(0.75\,[\mathrm{cm^3 g}] + \delta\cdot 1.0\,[\mathrm{cm^3 g}]\right)\frac{64,500[\mathrm{g\cdot mol^{-1}}]}{6.022\times10^{23}[\mathrm{mol^{-1}}]}$$

$$1.497\,[\mathrm{cm^3 g}] = 0.75\,[\mathrm{cm^3 g}] + \delta\cdot 1.0\,[\mathrm{cm^3 g}] \Rightarrow \delta = 0.747\,\frac{\mathrm{g\ water}}{\mathrm{g\ protein}}$$

Each protein has $m_1 = M_1/N_{Av} = 1.071\times10^{-19}$ g. The amount of water per protein is $0.747\cdot 1.071\times10^{-19} = 8.00\times10^{-20}$ g. The number of water molecules is $8.00\times10^{-20}/m_{\mathrm{w}} = 8.00\times10^{-20}\cdot N_{Av}/M_{\mathrm{w}} = 2676$ water molecules.

Another way of solving the exercise would be to calculate the volume of the hydrated protein and subtract the volume of the dry one to calculate the volume occupied by the water and, using the density and molar mass of water, calculate the number of water molecules. The final result will be slightly different, due to rounding numbers.

d. Assume now that the protein is not hydrated but its oblate-like shape is responsible for the observed differences in the translation coefficient. Taking into account the figure below, what is (roughly) the axial ratio of the protein?



Answer: To assess the axial ratio from the figure we need to calculate F_r', which is f/f_0, where f is the friction coefficient obtained from the diffusion coefficient of the protein and f_0 is the friction coefficient of the protein assuming it is spherical and non-hydrated:

$$\frac{f}{f_0} = \frac{R_h}{R} = \frac{3.37}{2.68} = 1.26$$

From the graph we can see that an oblate with such F_r' has an axial ratio of approximately 6.

e. How can one obtain the translational diffusion coefficient of hemoglobin, using a scattering technique?

Answer: One can obtain the translational diffusion coefficient using dynamic light scattering. We are given the following equations

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

$g^{(2)}(q,\tau)$ corresponds to the normalized time correlation function of the scattered intensity (what is measured) and $g^{(1)}(q,\tau)$ to the electric correlation function. To obtain the translational diffusion coefficient one has to find a relation between D_T and $g^{(2)}(q,\tau)$:

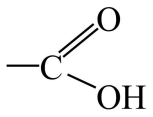
$$\ln\left[g^{(2)}(\tau) - 1\right] = -2q^2 D_T \tau.$$

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:

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

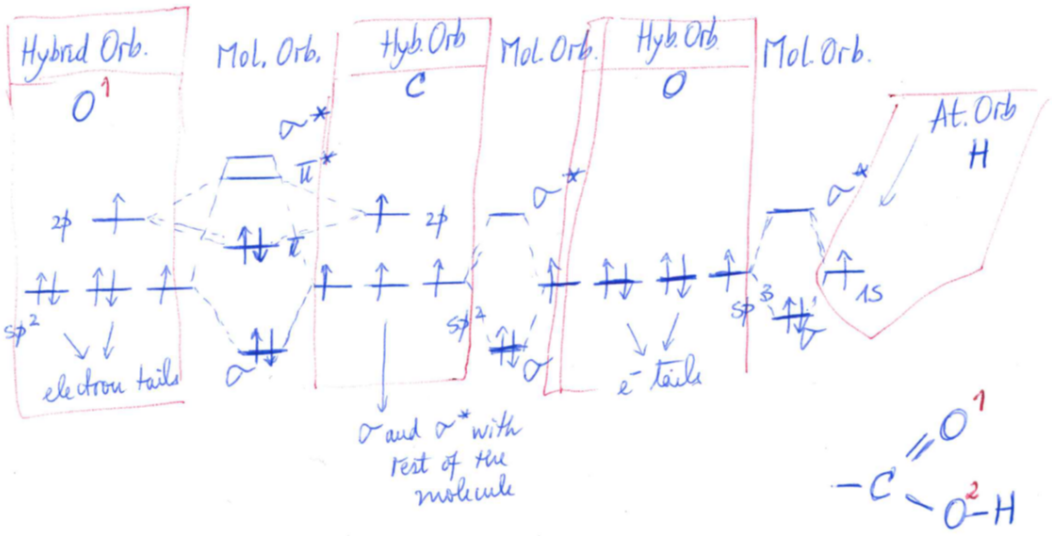
The scattering vector is calculated using $q = \frac{4\pi}{\lambda}\sin(\frac{\theta}{2})$, where λ is the wavelength of light in the solution and θ is the angle between the incident light and the measured scattered light (detector).

f. The carboxylic group (-COOH, see the representation below) is recurrent in biology. What is the hybridization of the carbon and of each oxygen atom? Draw an energy-level describing the hybrid orbitals for each of these three atoms and discuss (or draw) how they combine to form molecular orbitals.



Answer: The hybridization of the carbon is sp^2. The oxygen bound to the carbon and the hydrogen has a sp^3 hybridization, and the other, bound only to the carbon, has a sp^2 hybridization.

The energy-level describing the hybrid and molecular orbitals is depicted below. Starting from the left-hand side of the molecular group shown in the question, one of sp^2 orbitals of the carbon linearly combines with another atomic or hybrid orbital (of an unknown atom) to form a sigma and an anti-sigma molecular orbital, one of the sp^2 orbitals linearly combines with one of the sp^2 orbitals of the oxygen (no. 1) and the third sp^2 orbital combines with one sp^3 orbital of oxygen no. 2. We have then two more sigma and two more anti-bonding sigma molecular orbitals. Two of the sp^3 hybrid orbitals oxygen 1 are fully occupied with electrons and do not contribute to molecular orbitals. The forth sp^3 orbital linearly combines with the $1s$ atomic orbital of the hydrogen to make a σ and a $\sigma^\star$ molecular orbitals. The two sp^2 orbitals from oxygen 1 are also fully occupied and do not contribute to molecular orbitals. Finally the $2p$ atomic orbitals the carbon and oxygen 1 linearly combine to form a π and an anti-bonding π molecular orbitals.



Exercise 3.

In the lab exercises you have studied the interaction of DNA (a negatively charged polyelectrolyte) with a cationic surfactant, CTAB. We have discussed that the DNA condenses due to the formation of surfactant micelles in its vicinity.

a. What is the driving force for the formation of micelles?

Answer: The formation of micelles is driven by hydrophobic forces. This is an entropically driven force that occurs to minimize the contact between the hydrocarbon tails of the surfactants and water. Hydrocarbon tails are apolar and form no hydrogen bonds with water, that is, they are hydrophobic. To prevent turning any of the four H-bonding groups to the hydrocarbon, water will organize itself around it. Above a certain concentration the surfactant molecules are pushed towards each other to decrease the hydrophobic area exposed to water, and thus release some of the water molecules that are organized around the individual surfactant tails, which leads to a large increase in entropy. The larger the hydrophobic hydrocarbon groups the stronger this effect will be and thus, surfactants with longer chains lengths have a lower critical micellar concentration.

b. When performing dye exclusion assays, one of the control experiments you have performed was measuring the fluorescence of a solution consisting of GelStar (fluorophore) + 50 μM CTAB (the largest studied concentration). This control sample showed a very low fluorescence.

(i) What was the purpose of this control?

(ii) Taking into account the self-assembly properties of the CTAB, was this a good control experiment? How could you make it better? Justify. (*Hint:* The critical micellar concentration of CTAB is 0.9 mM).

Answer:

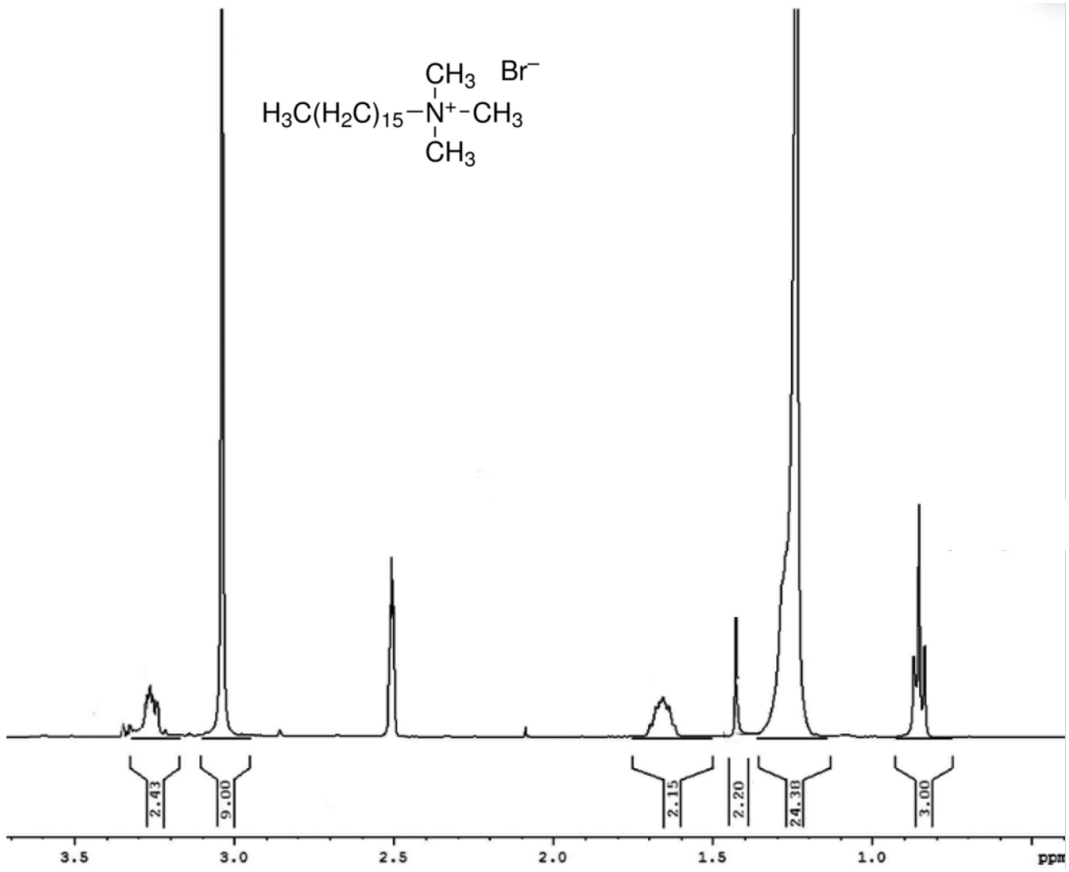
(i) The purpose of the control was to guarantee that the variations in fluorescence intensity observed when the concentration of CTAB was increased, were due to

the weaker access of the dye to the DNA and concomitant quenching of the fluorescence, due to the polar environment of the buffer solution, and not to the direct interaction of the dye with the CTAB molecules. If there was some variations in the fluorescence due to the presence of CTAB, it would be very difficult to separate the two effects: dye exclusion from the DNA and dye interaction with the CTAB.

(ii) The so-called quenching of the fluorophore occurs when it is placed in an aqueous solution. This is why the fluorescence of the fluorophore increases dramatically when bound to DNA, and thus in a non-polar environment. As discussed, the surfactant forms micelles in the vicinity of the DNA. Thus, it is possible that the fluorophore, when excluded from the DNA, goes into the hydrophobic environment of the core of the micelles, where it would presumable suffer less quenching than in water. Now, in the absence of the DNA the CTAB does not form micelles at a concentration of 50 μM. Thus, in order to check if the fluorescence intensity is affected by the presence of the CTAB micelles, a control solution should have also been performed with a CTAB above the CMC, that is, above 0.9 mM.

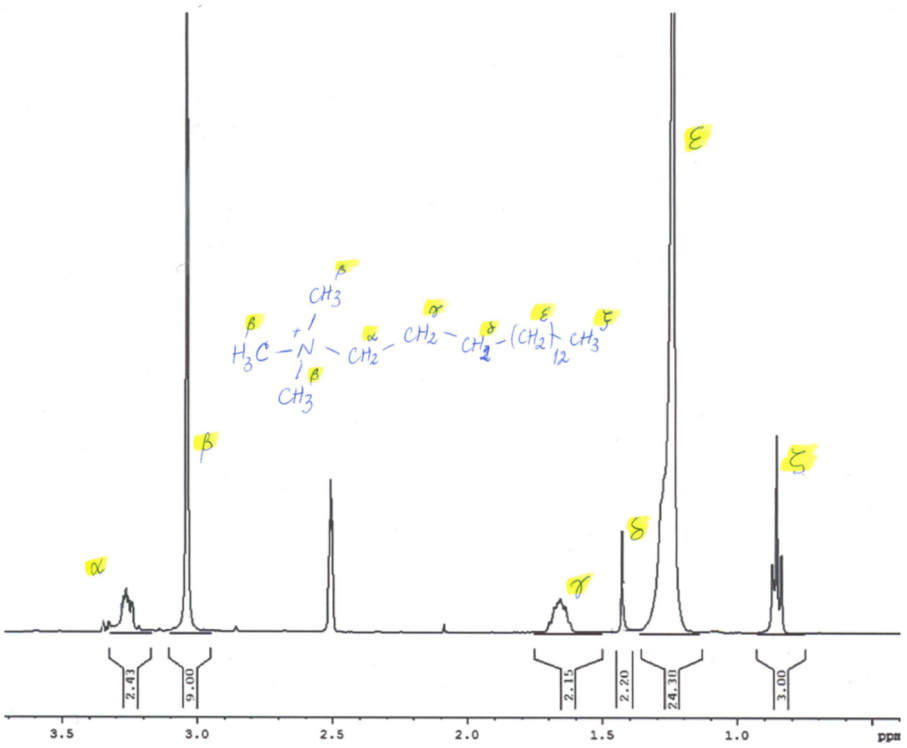
In reality, the GelStar shows nearly no difference in fluorescence intensity in the presence of CTAB monomers or micelles.

c. Below is the ^{1}H-NMR spectrum of cetyltrimethylammonium bromide, as well as the structure of the molecule. Assign each peak in the spectrum to the protons in CTA^+. (Hint: The numbers below each peak refer to the area of the peaks. The peak at 2.5 ppm corresponds to the solvent).



Answer: The peak splitting in the spectrum is, for most of the peaks, not easy to

interpret. The peak around 0.8 ppm correspond to the CH_3 group in the end of the alkyl chain. It (i) is a triplet due to the two protons in the neighboring carbon, (ii) has an area of 3, corresponding to the 3 protons, and (iii) appears at low ppm values, as is usual for these groups. The other peak that is easily identifiable is the one at 3 ppm. It is a singlet with an area of 9. This refers to a 9 protons with the similar chemical environment, and considering the large ppm values, it close to the N, which is an electronegative atom. This draws the electrons to it, leaving the protons more de-shielded and more exposed to the magnetic field, leaving to a larger precession and a large ppm value. Ignoring the solvent peak at 2.5, we are left with 4 peaks, with a total area of 30 protons, the number of protons that are left. There is a very large peak with an area corresponding to 24 protons, appearing close to the peak corresponding to the protons in the CH_3 group. This suggests that this peak corresponds to 12 CH_2 groups closest to end of the hydrocarbon chain. The other 3 peaks, each with an area of 2, must correspond to the CH_2 groups that are left, starting with the one closest to the N atom, for the peak with the largest ppm values towards the end of the hydrocarbon chain. To summarize:



d. Why is the acceptance rate of the different Monte Carlo moves important within the Metropolis algorithm?

Answer: Monte Carlo simulations rely on a making a proper sampling of the system to calculate the average of a property of interest. The Metropolis algorithm was develop to make the sampling more efficient by sampling more often system configurations with higher weights (lower potential energies) and less often configurations that contribute very little or are less relevant to the property of interest (for example, very stretched polymer configurations when calculating the average of the radius of gyration). The Metropolis algorithm relies on generating configurations, not completely randomly, but from a previous accepted (in terms of energy) one. The generation of new configurations

is most commonly done by translating each particle from its position in a random direction up to a maximum distance. Such moves are accepted or not based on energetic arguments. Accepted moves give rise to new configurations while rejected ones leave the system in the same configuration. If many moves are rejected (below 20%), the sampling becomes very limited and we might not be sampling the complete potential energy landscape, that is, one might be 'stuck' in a local energy minimum.