Biology 5357

Chemistry & Physics of Biomolecules

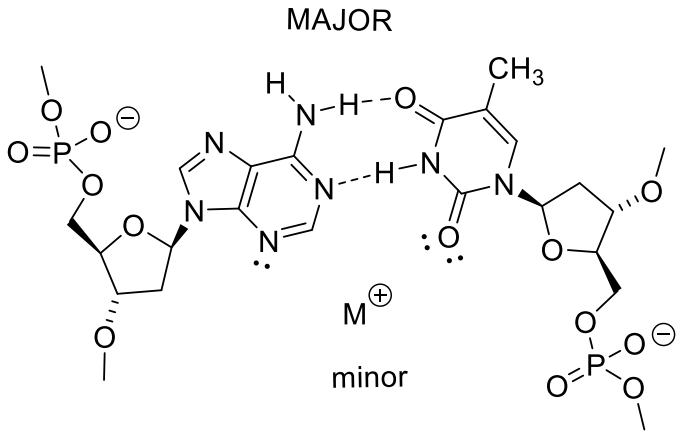
Examination #2

Nucleic Acids Module

November 3, 2017

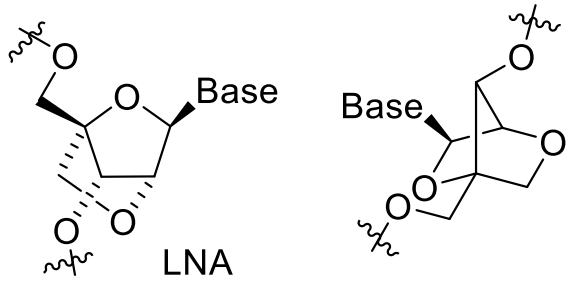
Answer Key

Question 1. (8 points) Draw a Watson-Crick A-T base pair as well as the sugar phosphate backbone with the glycosyl bonds in the conformation found in B DNA ***(1 pt for each nucleotide, 2 pts for H-bonds)***. Also indicate what would be the major ***(1 pt)*** and minor ***(1 pt)*** groove sides of the base pairs in a B DNA duplex. Show where metal ions might interact ***(1 pt)*** strongly with this base pair in B DNA and explain why ***(1 pt, see below).***



The minor groove at AT base pairs has a large negative electrostatic potential due to lone pairs that can stabilize metal cations.

Question 2. (10 points; A-E, 2 points each) In an effort to better understand and control the conformation and thermodynamics of nucleic acid duplex formation, chemists have developed "locked nucleic acids" or LNAs shown below as a structural drawing, and as a 3D perspective drawing.



(A) Why is the term "locked" being used?

The bridging group effectively "locks" the sugar into a single conformation. ***(2 pts)***

(B) What conformation is the sugar ring portion of the molecule in? Be specific using $C_{x'}$-endo/exo nomenclature for twist or envelope conformations.

The sugar ring conformation is C3'-endo. ***(2 pts)***

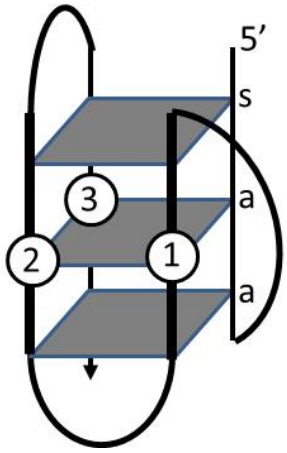
(C) What nucleic acid duplex conformation adopts this type of sugar conformation, if any?

Both A-form DNA and RNA adopt the C3'-endo conformation. ***(2 pts)***

(D) It has been found that the thermodynamic stability of DNA duplexes increases when nucleotides are replaced with LNAs. How might you explain this based on what you know about what contributes to the thermodynamic stability of DNA duplexes.

Because the sugar ring is no longer free to pseudorotate, the number of conformational states of the nucleotide are reduced, and there is less of an entropic penalty going from the single strand to duplex form. ***(2 pts)***

(E) Into which of the positions shown in the following G-quadruplex would you introduce LNAs to stabilize the structure? Why?



Positions 1 & 3, because these sugars should be in the anti glycosyl conformation as they are in strands that are parallel to the first strand which has the G in that quartet in an anti conformation. Anti glycosyl conformations are preferred by the C3'-endo sugar conformation which the LNA enforces. ***(2 pts)***

Question 3. (8 points; A-B, 4 points each) Consider the folding of the strand complementary to the GGGTTA strand of human telomeric DNA which forms the intercalated structural motif shown, where each parallelogram represents a C-C base pair and each circle represents the sugar phosphate unit of the C.

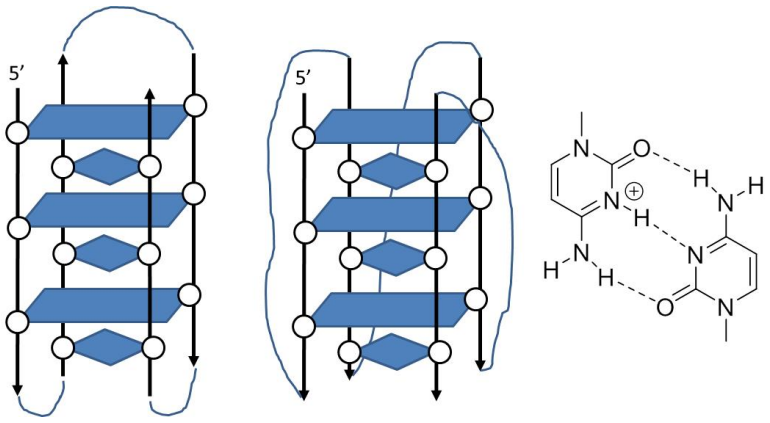
(A) Based on the geometry of the base pair shown propose a folded structure for four repeats of TAACCC. I.e., connect the individual strands with TAA loops

in a way that would accommodate the C-C pairs in which both C's are in *anti* glycosyl conformations.

Basically, the two strands containing the C-C base pair have to be parallel ***(2 pts)****, so either the two pairs of parallel strands have to be parallel or antiparallel to each other, using either lateral loops or propeller loops. Two such structures are drawn, and there are two others using the same connections, but to different strands.* ***(2 pts)***

(B) This motif structure is facilitated at lower pH. Explain. [Hint: What will happen to the base pair shown at low pH, and how might this affect the stability of the base pair ?]

The C-C base pair becomes protonated ***(2 pts)*** *to reduce lone pair-lone pair repulsion and add a third H-bond as shown.* ***(2 pts)***



Question 4. (6 points) The equation below describes the polymer behavior of DNA under tension. k_B is the Boltzmann constant. What are the other six parameters in the equation (***x, L₀, T, F, P*** *and* K_0) ?

$$x = L_0 \left(1 - \frac{1}{2}\left(\frac{k_B T}{FP}\right)^{1/2} + \frac{F}{K_0}\right)$$

x = end-to-end distance ***(1 pt)***
L_0 = contour length ***(1 pt)***
F = force ***(1 pt)***
P = persistence length ***(1 pt)***
K_0 = stretch modulus (enthalpic spring constant) ***(1 pt)***
T = temperature ***(1 pt)***

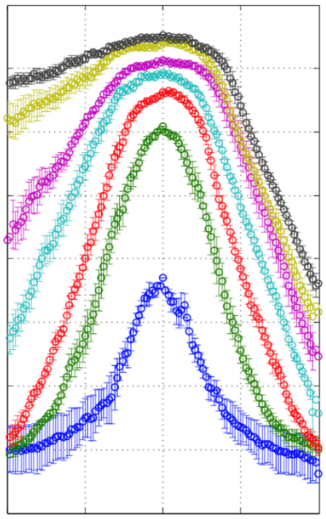
Question 5. (4 points) Describe, using the concepts of entropy and enthalpy, what happens as you use increasing amounts of force (starting from zero force) to stretch a dsDNA molecule.

Entropy dictates that the end-to-end distance of a flexible polymer will be shorter than its contour length. At low forces, the work performed during DNA extension is counteracting the reduction in entropy of the chain. ***(2 pts)*** *As the force continues to increase and the end-to-end distance becomes greater than the contour length, the work performed is counteracting enthalpic penalties caused by deforming the molecular structure of the DNA molecule.* ***(2 pts)***

Question 6. (4 points) Is DNA perfectly straight over its persistence length? Why or why not?

No. ***(2 pts)*** *Persistence length describes the correlation of the tangent vector over the path length of a polymer.* ***(1 pt)*** *It is length over which the correlation drops to 1/e.* ***(1 pt)***

Question 7. (6 points) The below shows rotation-extension curves for dsDNA obtained with the magnetic tweezers under different conditions (different colors). Label the axis of the plot. Provide an explanation for what parameter or parameters may be being varied in the experiment and include a physical description of why changes in this parameter leads to differently shaped curves using the concepts of linking number, twist, and writhe.



X-axis: Super-helical density or number of turns ***(1 pt)***
Y-axis: End-to-end distance ***(1 pt)***

DNA can absorb turns or changes in linking number (Lk) via either twist (Tw) or writhe (Wr). Only writhe leads to a change in DNA extension. Thus, the parameter can be force, temperature, salt concentration, or anything that changes the equilibrium between ssDNA bubbles and dsDNA. ***(2 pts)***

Under conditions where dsDNA is stable, the rotation-extension curve is symmetric for both negatively and positively supercoiled DNA as all turns are absorbed as writhe (i.e. the blue, green, and red curves). However, when ssDNA becomes energetically favored by high force or high temperature or low salt, an asymmetry develops in the curves (i.e. purple. yellow, and black curves). This is because, here, the negative links (turns) can be absorbed by negative twist which does not lead to an appreciable change in the end-to-end distance of the DNA. ***(2 pts)***

Question 8. (10 points; A = 6 pts, B = 4 pts) Two useful quantities to describe the conformations of proteins and DNA are the end-to-end distance and the radius of gyration.

(A) Provide a concise description of these two important quantities using words and/or equations.

End-to-End Distance: *The end-to-end distance vector describes the vector connecting the first and the last monomer of the chain (this is sufficient). If we define R_i as the bond vector connecting the monomer i to the monomer i+1, the end-to-end vector is given by $\overrightarrow{R_{ee}} = \sum_{i=1}^{N} R_i$ with N number of bonds (this is an example of more mathematically detailed answer).*

Radius of Gyration: *The square radius of gyration describes the average square distance between monomers and the center of mass of the polymer. Alternatively, the squared radius of gyration can be linked to the average square distance between each pair of monomers. $R_g^2 = \frac{1}{N}\sum_{i=1}^{N}(r_i - r_{cm})^2$ with r_i and r_{cm} the position vector of each bead and of the center of mass, respectively.*

(B) For an ideal chain, an important relationship connects the mean square end-to-end distance with the mean square radius of gyration. Which one is larger? What is the conversion factor?

The mean square end-to-end distance is larger than the mean square radius of gyration. The conversion factor between the two is 6.

$$\left\langle R_{ee}^2 \right\rangle = 6\left\langle R_g^2 \right\rangle$$

Question 9. (12 points; A-C = 4 pts each)

(A) Scaling exponents, commonly designated by the Greek letter ν, are an important concept in polymer physics. Indicate three different scaling exponents that describe the dependence of the root-mean-squared radius of gyration versus the polymer length (expressed in numbers of monomers N). For each scaling exponent assign the corresponding quality of the solvent and describe the average shape adopted by the polymer when all the configurations are aligned in the same direction.

ν	**Solvent Quality**	**Shape**
3/5	*Good solvent/Athermal solvent*	*Prolate*
1/2	*Theta solvent/Ideal solvent*	*Prolate*
1/3	*Poor solvent*	*Sphere*

(B) Write the scaling equation for the radius of gyration versus the number of monomers N:

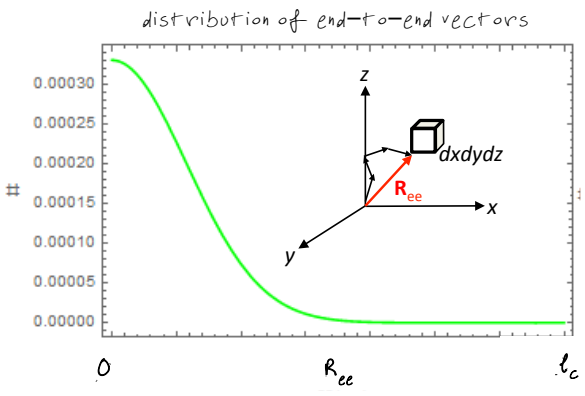
$$\left\langle R_g^2 \right\rangle^{1/2} \propto N^{\nu}$$

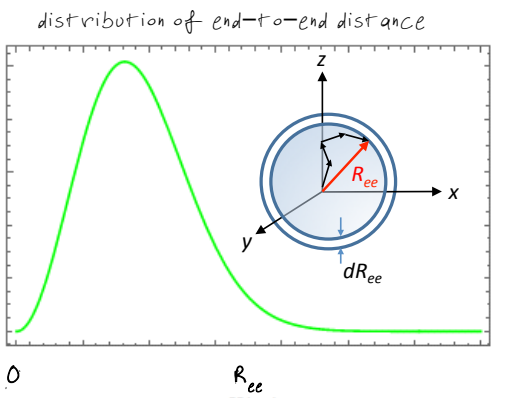
(C) The disordered tails of two proteins have identical repeated sequences. One tail is 60 amino acids long, whereas the other is 120 long. The measured radius of gyration for the tails are indicated as R_g^{60} and R_g^{120}. Using the measured numbers can you estimate the scaling exponent and the solvent quality.

	R_g^{60} (nm)	**R_g^{120} (nm)**	ν	**Solvent Quality**
6M Urea	4.30	6.5	*3/5*	*Good solvent*
50 mM NaP	1.40	1.76	*1/3*	*Poor solvent*

Question 10. (6 points) Describe the probability distribution for the end-to-end vector and for the end-to-end scalar distance of an ideal chain. Which one of the two is centered at zero? Why? Discuss the difference between the two distributions.

The end-to-end vector distribution is zero because the end-to-end vector is a random walk and the mean value of the end-to-end vector is zero. Alternatively, the distribution of the end-to-end vector is a Gaussian distribution centered in zero that represents a random walk.





$$P_{GC}\left(\mathbf{R}_{ee}\right)=\left(\frac{3}{2\pi\left\langle \mathbf{R}_{ee}^{2}\right\rangle }\right)^{3/2}\exp\left(-\frac{3}{2}\frac{\mathbf{R}_{ee}^{2}}{\left\langle \mathbf{R}_{ee}^{2}\right\rangle }\right)^{*}$$

The probability density that the chain end is in a particular location between (x,y,z) and (x+dx,y+dy,z +dz). The most probable termination point is the origin.

$$P_{GC}\left(R_{ee}\right)=4\pi R_{ee}^{2}\left(\frac{3}{2\pi\left\langle \mathbf{R}_{ee}^{2}\right\rangle }\right)^{3/2}\exp\left(-\frac{3}{2}\frac{R_{ee}^{2}}{\left\langle \mathbf{R}_{ee}^{2}\right\rangle }\right)^{*}$$

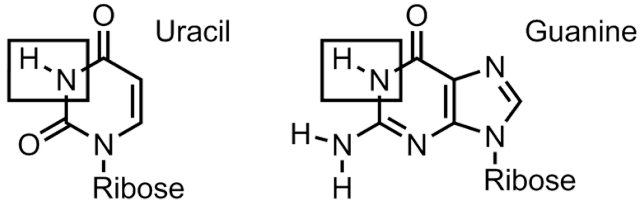
The probability density that the chain end is anywhere in a radial shell between r and r+dr.

Question 11. (8 points) What do imino protons in NMR experiments tell you about the structure and stability of an RNA duplex? Explain, and provide structural examples to illustrate your answer.

G, U (and T) have an imino proton that is hydrogen bonded in a Watson-Crick base pair. The hydrogen bonding protects the imino protons from rapid exchange with protons on water molecules, allowing them to be observed in NMR experiments. In addition to base paiting, an imino proton can be sequestered from exchange with water via a tertiary interaction, either through hydrogen bonding of a structure that excludes water.

A:U and G:C Watson-Crick pairs have one imino proton, while G:U pairs have two. The chemical shifts of these protons depend on the base pair, so to a first approximation, the imino proton NMR spectrum will reveal the number and type of base pairs present.

For a stable structure is stable, the imino proton spectrum has sharp peaks of equal intensity. When the structure is not stable, at the termini or in a flexible loop, the imino protons exchange with H_2O and lose intensity. When a duplex melts, its imino protons disappear from the spectrum. Melting can be monitored by progressive loss of specific peaks as a function of temperature.



Question 12. (16 points; A-B = 8 pts each) Shown below is the purine riboswitch aptomer, in this case with bound Guanosine.

(A) Identify what differences you might expect between this RNA aptamer and a DNA molecule of the same sequence.

(1) DNA duplexes are B-form, which will change the orientation of the loops with respect to each other.

(2) DNA duplexes are less stable than RNA A-form duplexes, so the entire structure will be thermodymamically less stable.

(3) The deoxyribose has a different pucker preference than ribose, which could affect the loop structures and the junction stability.

(4) Loss of the 2′OH could destabilize the junction when G is bound.

(5) Thymine with its –CH3 group could sterically interfere with binding the guanosine/purine.

(6) U's are more "flexible" than T's, since they lack a propensity for stacking. Replacing U with T could make the junction and loops more rigid.

(7) The combination of B-form geometry and the U-to-T replacement could alter the binding site so that G is not bound.

(8) If Mg^{+2} is required for folding, and if the 2′OH is involved, then the Mg^{+2} dependence of stability will be different.

(B) How would you experimentally demonstrate those differences? Illustrate any experimental results if needed for clarity.

(1) Compare the remerature dependence of denaturation of ribo- vs. deoxyribo- switch, in the absence and presence of G (or A). Monitor by measuring A_{260nm} and A_{280nm} as a function of temperature.

(2) Repeat the above in the presence and absence of Mg^{+2} ion.

(3) Use 2-aminopurine (purine) as a probe of binding. Look for fluorescence changes in the presence and absence of Mg^{+2} ion.

(4) Run a native polyacrylamide gel to look for changes in mobility in the presence/absence of purine and presence/absence of Mg^{+2}.