Biology 5357

Chemistry & Physics of Biomolecules

Examination #2

Nucleic Acids Module

November 5, 2021

Answer Key

Question 1. (8 points)

This sequence should be intrinsically bent according to a junction model. The A/T tract should adopt a B* conformation with a narrow minor groove bending towards the minor groove, while the G/C tract should be intrinsically bent towards the major groove. Since the minor groove is 180 degrees out of phase with the major groove, it will become in phase with the major groove after ½ a turn, and so the bends induced by the G/C and A/T tracks will be in phase. Since the sequence is 10 bp in total length, oligomers will have the bends in phase.

Question 2.

(A) (4 points)

At the most simplistic level both conformations have two identical based-paired sections consisting of 2 base pair stacks. The double hairpin has in addition two 3-nucleotide hairpin loops, while the long hairpin consists of one 6-nucleotide interior loop. Therefore, at this level of analysis the difference in free energy would be ascribed to the difference in free energy between a 3-nucleotide hairpin loop and a 6-nucleotide interior loop.

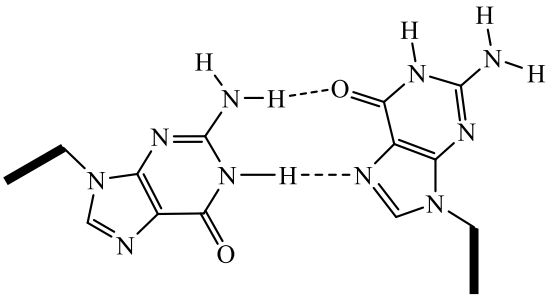
$$\Delta G^{o}_{long} = 2\Delta G^{o}_{stack} + \Delta G^{o}_{hairpin\ loop} + \Delta G^{o}_{interior\ loop}$$

$$\Delta G^{o}_{double} = 2\Delta G^{o}_{stack} + 2\Delta G^{o}_{hairpin\ loop}$$

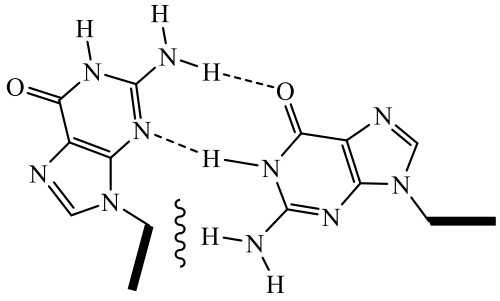
$$\Delta G^{o}_{long\text{-}double} = \Delta G^{o}_{hairpin} - \Delta G^{o}_{hairpin\ loop}$$

(B) (4 points)

Either one of the following would be acceptable answer.



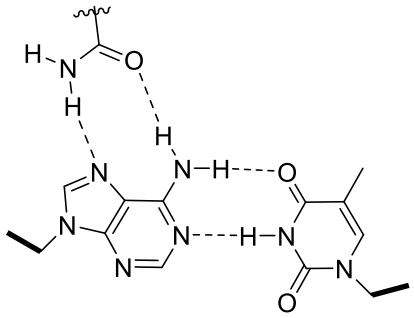
reverse Hoogsteen



may not be feasible
because of steric clash

Question 3. (8 points)

The structure below shows the hydrogen bonding in the A-T base pair, and in the upper left indicates the interaction of the side chain of Asn or Gln with the A.



Question 4. (8 points)

a = unfolded, as it is present before addition of K^+

b = chair, since b has the highest efficiency and the chair would appear to have the shortest distance between donor and acceptor

c = could in principle be either the two tetrad basket structure or the hybrid structures but is likely to be the two tetrad basket since it is expected to be less stable due to having one tetrad base stack and one potassium ion and would have a shorter lifetime

d = most likely the hybrid-1 or hybrid-2, as it is the most long lived and the hybrid structures are expected to be the most stable having two postassiums and two tetrad stacks

The trace always shows the unfolded species appearing between the same or different conformations indicating that the DNA unfolds completely and then refolds to the same or a different conformation, and not directly changing from one conformation to the next.

Question 5.

(A) (4 points)

Continuous 1000, Dashed 10, Dotted 1000, Dot-Dashed 100.

(B) (4 points)

The theta temperature for the dashed, dotted and dot-dashed curves is approximately 300 K, while the theta temperature for the continuous curve is approximately 400 K.

(C) (4 points)

Poor solvent below the theta temperature, good solvent above theta temperature.
Exponent in poor solvent: $R^2 \sim N^{2/3}$ or $R \sim N^{1/3}$ (same if they used R_g)
Exponent in good solvent: $R^2 \sim N^{6/5}$ or $R \sim N^{3/5}$ (same if they used R_g)
Alternative answer in good solvent is with exponent 2*0.588 or 0.588.

(D) (4 points)

Poor solvent: 100^(2/3)/1000^(2/3) = 0.2154
Theta solvent: 100^(1)/1000^(1) = 0.1
Good solvent: 100^(6/5)/1000^(6/5) = 0.063

It is fine if the answer computed the inverse:

Poor solvent: 1000^(2/3)/100^(2/3) = 4.6415
Theta solvent: 1000^(1)/100^(1) = 10
Good solvent: 1000^(6/5)/100^(6/5) = 15.84

Rounding up the digits is not a problem. Award half-credit if the answer computed the ratio for the root mean squared values instead of the mean square.

Question 6.

(A) (4 points)

$N = 120$
$B = 0.38$ nm
$l_c = b*N = 45.6$ nm

(B) (4 points)

$R^2 = 2\, l_p\, l_c = N_k\, b_k^2$
$l_c = N_k b_k$
$b_k = l_c / N_k = 2\, l_p = 1.0$ nm
$N_k = l_c / b_k = 45.6$

(C) (4 points)

$R^2 = 2\ l_p\ l_c = N_k\ b_k^2 = 45.6\ nm^2$
$R_g^2 = R^2/6 = 7.6\ nm^2$

$P(R) = Z^{-1}\ R^2\ Exp[-(3/2)\ (R^2/45.6)]$
$P(R_g) = Z^{-1}\ R_g^6\ Exp[-\ (7/2)\ (R_g^2/7.6)]$

where Z are normalization factors.

Z^{-1} for the Gaussian Chain is equal to: $4\ \pi\ (3/(2\ \pi\ 45.6))^{3/2}$

Question 7.

(A) (6 points)

Starting 2AP35 fluorescence intensity in the secondary structure is high, so it probably isn't stacking with another nucleobase. When Mg^{+2} is added, 2AP fluorescence intensity decreases. Therefore, it has changed its environment to become at least partially (geometrically or temporally) stacked. This new environment could be local, if Mg^{+2} has caused the loop has become structured, or there could be a more global rearrangement where the two loops are interacting (and the position of Mg^{+2} is not specified).

When adenine is added, there is an additional quench of the 2AP, indicating that 2AP35 has become more stacked with another nucleobase, presumably as shown in the structure where the nucleobases from the two loops are stacking with each other. Thus, ligand binding is necessary for stable tertiary structure of the aptamer. Mg^{+2} alone is not enough to induce stable folding.

Note that the response to Mg^{+2} addition is at the limit of resolution of the experiment, while binding by adenine is slower, suggesting that the conformational rearrangement of the riboswitch is also slower.

(B) (4 points)

The binding site for adenine is an unstructured junction, so it's possible that an adenine could bind there and lock down the flexibility of the stems. If so, then P2 and P3 might be juxtaposed to allow the loops to interact. If that happens, then 2AP35 fluorescence could decrease if the nucleobases can stack. Subsequent addition of Mg^{+2} could stabilize the bound adenine and thus stabilize the tertiary

structure, so a larger fluorescence intensity decrease would be observed. This would be a two-step folding process.

However, if the junction is too flexible to allow productive adenine binding, then addition of adenine alone would not affect 2AP35 fluorescence. Now addition of Mg^{+2} in the presence of adenine would lead to loss of fluorescence intensity as both ligands bind. Perhaps there will be a single step loss of fluorescence, or the two events could be separable in time.

Question 8.

(A) (6 points)

The UUCG tetraloop on the left is more stable than the GAAA tetraloop, so the hairpin on the left is more stable than the one on the right.

The UUCG tetraloop (with its CG loop-closing base pair) has an intricate network of hydrogen bonds involving U6 and G9 as well as the ribose of C8. In addition, base stacking stabilizes the structure. Its structure in solution is extremely stable, and its thermodynamic stability is anomalously high (large negative ΔG).

The GAAA tetraloop has a flexible structure and its thermodynamic stability is not unusual.

(B) (4 points)

The UUCG tetraloop acts as a staple in the folding of an RNA, since it is a local structure that forms very quickly (microseconds) and even with a two base pair stem, it is stable. No proteins are known that bind to it, neither is it known to interact with any RNA sequence or structure.

The GAAA tetraloop is one half of a tertiary interaction – the adenines flip out of the loop to make hydrogen bonding interactions within the tetraloop receptor. It is perhaps the most common tertiary interaction in RNAs. The loops are more generally GNRA, where N is anything and R is a purine, but GAAA predominates.

Question 9.

(A) (6 points)

As we discussed in class, ARM peptides need a deformed A-form helix in order to bind. The duplex above is deformed in the middle by the C:C and A:G noncanonical

base pairs and the adjacent G:U. Therefore, the peptide will use this site to wiggle into the duplex.

Once bound, the peptide can slide within the major groove to maximize contacts with the bases and riboses. It will use its R, W, and Q amino acids to make different contacts with the RNA.

(B) (4 points)

The arginine sidechain is uniquely able to offer an electrostatic and potential hydrogen bonding moiety to bases, phosphates, and riboses. Its positive charge can also help to neutralize the phosphates in a duplex, since its long chain allows its guanidinium group (pKa of 13.8, so it's always protonated) to reach out of the major groove, where it can potentially displace ions that are proximal to the nonbonded phosphate oxygens.

Question 10.

(A) (6 points)

The following lists some advantages of cryo-EM over X-ray crystallography:

- crystals are not required for cryo-EM, which allows it to better deal with disordered systems; growing crystals for X-ray analysis can be difficult or impossible
- multiple differentconformational states can be captured in a single experiment
- cryo-EM reduces radiation damage to the sample, and maintains the native activity and functional state of samples, including posttranslational modifications
- only very small amounts of material are needed, while optimization of crystallization conditions to get X-ray quality crystals can use much larger amounts of material
- cryo-EM can be effective on very large biological structure and complexes, while X-ray crystallography is typically more suited (with some exceptions!) to small molecules through individual protein chains
- since we have a "lens" for electrons, unlike X-rays, it is possible to more directly image samples in cryo-EM; in X-ray crystallography, a diffraction pattern of

intensities is collected, and then other methods are needed to get “phase” information

- cryo-EM is more easily applicable to membrane proteins and their complexes

(B) (2 points)

The Fourier transform is a mathematical process that performs the same function as a physical lens in image processing and in the analysis of diffraction data. The Fourier transformation performs a convolution operation that converts diffraction data in “reciprocal space” into an object in normal 3D “real space”. (Or analogously, an inverse Fourier transform does the reverse operation, and goes from real space to reciprocal space.)

(C) (2 points)

The contrast transfer function (CTF) mathematically describes how aberrations in a electron microscope modify the image of a sample. It provides a quantitative method to translate the exit wavefunction from the microscope to a final image. Application of the CTF allows correction of a raw cryo-EM image to obtain high-resolution structures.