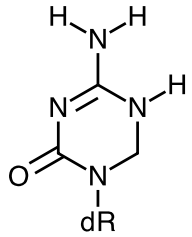
Problem Set #5 – Chemistry 391 **Name**_____

Due in class October 4, 2018

1. KP1212 is an antiviral drug candidate that interferes with replication. It adopts many different tautomers [see Li et al. (2014) paper on web page].



a. Produce tautomers of KP1212 that are capable of Watson-Crick base pairing with each of the four naturally occurring DNA bases, and show the hydrogen-bonding scheme for each (actually, C is pretty tricky – that's an extra credit effort).

b. In a study of mutation rates in DNA replication, it was found that KP1212 introduces guanine to adenine mutations 10% of the time, but no measurable mutation to cytosine or thymine. Why not?

2. van't Hoff analysis of DNA melting is a little different than protein unfolding due to the stoichiometry of one duplex melting into two single strands. To find thermodynamic parameters in DNA melting, one plots $1/T_m$ vs. $\ln[C_T]$, where C_T is the <u>total</u> concentration of ssDNA. The equilibrium may be written as:

dsDNA $\Leftrightarrow$ 2 ssDNA.

At T_m, 50% of the duplex has melted.

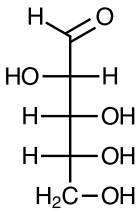
a. Express [dsDNA] and [ssDNA] at the T_m as functions of C_T. Show that $K = C_T$ at T_m

b. Given the above, note that $\Delta G^\circ = \Delta H^\circ - T_m\Delta S^\circ = -RT_m\ln C_T$. Show that a plot of $1/T_m$ vs. $\ln C_T$ will have a slope of $-R/\Delta H^\circ$ and an intercept of $\Delta S^\circ/\Delta H^\circ$.

3. Draw each of the following assuming 3', 5' linkages unless otherwise indicated.

a. The 2-deoxyribodinucleotide with the sequence dAdG.

b. An arabinodinucleotide, aGaI. Note that β-D-arabinonucleotides uses the furanose form of D-arabinose instead of D-ribose – see right.



c. A **2', 5'** dinucleotide of RNA, whose sequence is rCrU.

d. Discuss whether the structures you have drawn in a-c are less susceptible to base hydrolysis than normal RNA.

4. The Kool lab has also worked on an unnatural base pair that forms between methylbenzimidazole (Z) and difluorotoluene (F). No hydrogen bonding takes place between the two, yet stable DNA duplexes form. You can see the complex in the file ZFbp.pse (on web-page).

duplex	ΔH°_{25} (kcal/mol)	ΔS° (eu)	ΔG°_{25} (kcal/mol)
5'-CTTTTC**A**TTCTT 3'-GAAAAG**T**AAGAA	99.0	291	12.4
5'-CTTTTC**Z**TTCTT 3'-GAAAAG**F**AAGAA	47.4	129	8.9

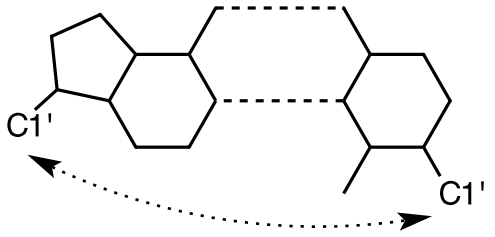
a. Based on the melting data above, which duplex is more stable at 25°C? Which duplex has a higher melting temperature (show calculations to support your conclusion. T_m is where $\Delta G = 0$.).

b. Explain why ΔS° of melting is likely less positive for the Z-F pair than for the A-T base pair.

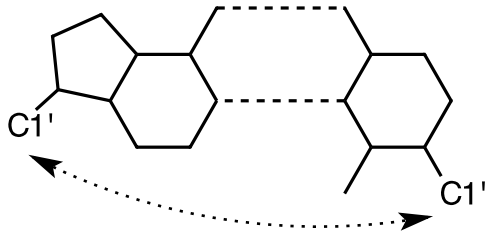
Clearly the AT base pair creates a more enthalpically stable duplex. One argument might be H-bonding stabilizes the AT base pair (not likely). Another argument might be strain destabilizes the Z•F base pair. A crystal structure exists for the Z•F base pair in a DNA duplex (download **ZFbp.pse.** I have colored major groove magenta and minor groove cyan so you can see these first hand. You may want to recolor by element. Objects are available for an AT and ZF base pair).

c. On the templates below, **draw in** the A•T and Z•F base pairs. Identify major groove and minor groove atoms in the ZF base pair and and indicate the distances between atoms indicated with dotted lines, including between C1' atoms.

A•T



Z•F



d. Aside from distances, indicate if there are any other differences in the geometry of the base pairs, either on the diagrams or in a brief description below.

e. Ignoring differences in hydrogen bonding, is there evidence that the A•T base pair duplex have a higher ΔH for melting than the Z•F base pair due to strain?

6. RNA structure reveals that nucleic acids are not limited to simple double-stranded helices. The structure of a 158-base, folded RNA structure can be found in **1HR2.pse**. The initial view will show the entire structure with Mg^{2+} ions as spheres. Of interest is an tertiary interaction that leads to the overall structure of the RNA.

Zoom in on the object **GAAA** and hide the cartoon (and anything else that confuses your view). This is a "tetraloop", a four-base turn that comes at the end of a stem containing regular base pairs. The structure is "GNRA" type loop, where the first and last bases are always a G and A, and the third base is always a purine (R). The second base can be anything (N).

a. Sketch the interaction that clearly shows why the loop must be begin with an G and end with an A.

b. Sketch the interaction that clearly shows why the third base (GN<u>R</u>A) must be a purine.

c. Looking at the loop, explain why the second base can be A,C,G or U, the "N" in GNRA.

d. In this structure, the tetraloop (GAAA) makes interactions with a "receptor region" to stabilize the tertiary fold. Sketch the interaction between the second base of the loop (G<u>A</u>AA), base 151, and adenine 248 (light green) that helps enable the loop-receptor interaction.