1. DNA Structure and Stability (30 pts):

There are a lot of deviations from Watson Crick base pairing and the WC helix, both from nature and from synthesis. Eric Kool of Stanford has been particularly active in the latter field.

(a; 10 pts) Draw two G:A base pairs: (1) a structure in which the Watson-Crick H-bonding faces of G and A interact with each other. (2) Freeform H-bonding of your choice, with at least two H-bonds formed, different from (1). An example might be the G:A pair seen in the tetraloop. You do not need to draw out sugars.

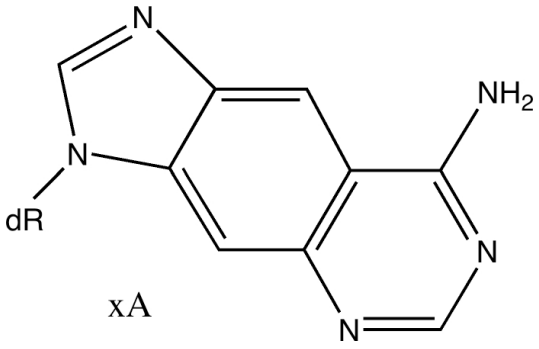
Score for the page_____

(b; 3 pts) We have emphasized over and over what it is that makes the WC base pairs special. What is it?

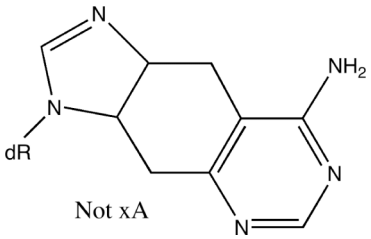
(c; 3 pts) G:A base pair # 1 from part (a) destabilizes B-form DNA. Why? It stacks and H-bonds well.

Kool's group has synthesized the extended adenine at the right (xA) and incorporated the xA:T base pair into DNA oligonucleotides.

(d; 3 pts) Why did they choose to make the xA shown instead of "Not xA" shown below xA?



xA



Not xA

(e; 3 pts) It is observed that the structure and backbone of the extended DNA (with an extendabase in each pair) is essentially a thick B-form, as shown below. This was somewhat surprising, at least to me. Why is it surprising? (Consider the limit of including more benzene "extenders.")

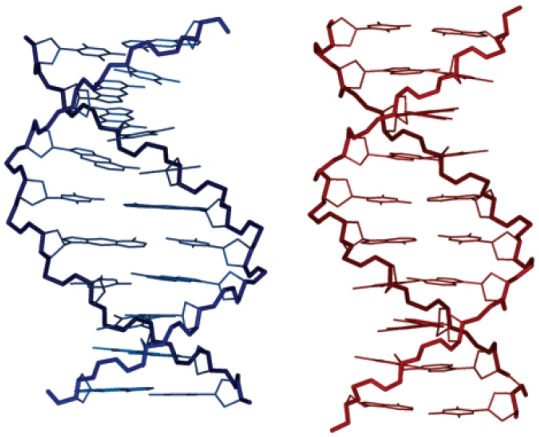


Figure 4. Mean structure of 24 randomly selected final structures of xDNA (at left) and a model structure of the control B-DNA (right).

Score for the page_____

(f; 8 pts) When xA:T is placed in an otherwise normal DNA it is observed to destabilize the helix (T_m goes down). However, the duplex in (e) that includes an extendabase in each base pair is observed to be more stable than normal DNA. For complicated reasons, the change in stability has not been parsed into $\Delta H°$ and $\Delta S°$ contributions yet. Sit on the fence: give a rationale for each possibility, i.e. explain how the extendabase could give a larger negative $\Delta H°$ for hybridization or else a smaller negative $\Delta S°$.

2. Hybridization (25 pts):

(a; 10 pts) Briefly describe the DNA microarray and how it is used to study mRNA expression patterns.

Score for the page_____

Microarrays can also be applied in other ways. The "DNA tiling array" can be used to sequence variants of known genomes or look for things like polymorphisms or splice site changes. The idea is that 25mer probes that span a genome are synthesized on the array, and then labeled DNA is hybridized to the array. The picture below (Bertone *et al.*, *Genome Res.* 2006 16: 271-281) should give you the idea, with some extraneous detail.

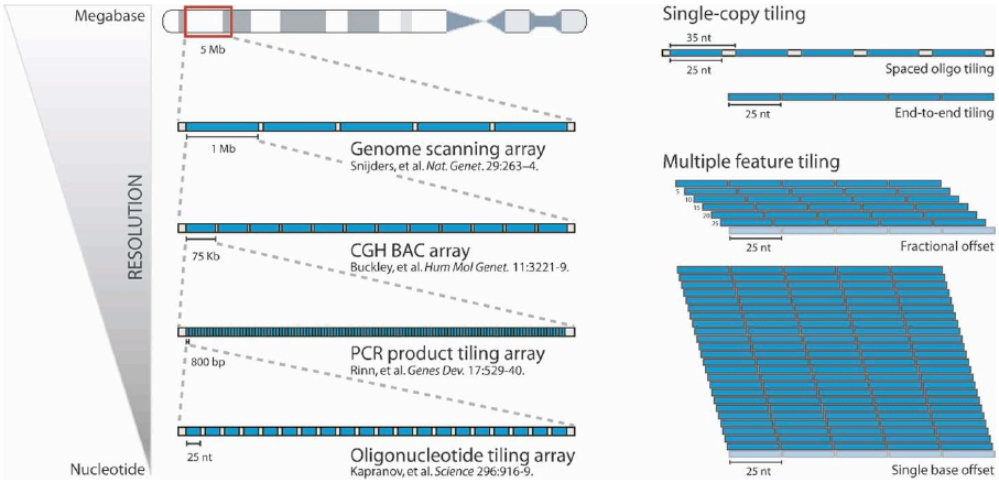


Figure 1. (*Left*) Evolution of genomic tiling arrays. Representing large spans of genomic DNA with bacterial artificial chromosome (BAC) clones facilitates global experimentation using relatively few array features, at the expense of low-tiling resolution. Higher-resolution designs using PCR products or oligonucleotides allow precise mapping of transcripts and regulatory elements, but require labor-intensive or technologically sophisticated approaches to implement. (*Upper right*) Linear feature tiling with gapped and end-to-end oligonucleotide placement. (*Lower right*) Overlapping tiles using fractional offset (e.g., one 25-mer probe placed every 5 nt) and single-base offset placement. The latter strategy provides a finer-resolution tiling of the genomic sequence, and can give a more precise indication of where hybridizing sequences are located on the chromosome.

(b; 6 pts) Experimentally, how could you identify a single nucleotide polymorphism or other slight change in a genome using a tiling array like the one at the bottom right above? In other words, what would be the signal that e.g. a new isolate of the bacterium had a sequence difference?

Score for the page_____

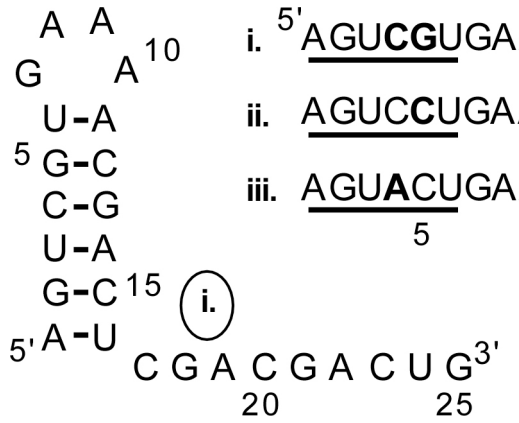
(c; 9 pts) Why 25mers? Why do much longer probes not work well for identifying SNPs? On the other hand, why not make the probes much shorter, so that synthesis would be easier? In general, what two somewhat conflicting goals do we have in any hybridization experiment?

3. RNA Structure (20 pts):

(a; 10 pts) Sketch base-catalyzed hydrolysis of the RNA phosphodiester backbone. Why are there no large RNA genomes?

Score for the page_____

The essential RNA (i) below was proposed to form the structure shown based on computer modeling. Then homologous sequences (ii) and (iii) were discovered. The bases that differ from RNA (i) are indicated in bold. The underlines are hints.



i. 5' AGU**CG**UGAAAA**C**G**A**CUCGA**CG**ACUG 3'

ii. AGUC**C**UGAAAA**C**A**A**CUCGA**G**GACUG

iii. AGU**A**CUGAAAA**C**C**A**CUCGAG**U**ACUG

5 10 15 20 25

(b; 4 pts) Explain the notion of correlated invariants in phylogenetic studies of RNA structure.

(c; 6 pts) The sequences (ii) and (iii) support a structure different from the one shown. Draw the structure, for sequence (i). Why is this structure unexpected?

Score for the page_____

4. DNA Topology (20 pts):

(a; 15 pts) Sketch the three main ways in which negative supercoiling is manifested (one has $\Delta Tw < 0$ and the other two are writhed). Identify the thermodynamically stable form for typical supercoiled B-DNA in the absence of proteins. Give one biological function for each of the two writhed forms.

Score for the page_____

(b; 5 pts) The equation for determining linking number from a flat sketch of single strands winding around each other is Lk = (sum of nodes)/2. We determine writhe from a projected image of a double strand DNA similarly. What is the equation for that (and why is it different)? We know that Lk is always an integer, but writhe is not. How can this be given that the equations are so similar?

Page	Score
1	
2	
3	
4	
5	
6	
7	
8	
Total	

Score for the page_____