ChE/BE 163
Problem Set #2
Due: 1 pm Thursday, November 7, 2019

Problem 1 (Test tube design, 20 pts).
Your primary reference is the paper Wolfe and Pierce, *ACS Synth Biol*, **4**, 1086–1100, 2015. Consider DNA sequence design at 23 °C for an on-target complex with target secondary structure:

```
((((((.....((((((((+)))))))).....((((((+)))))).((((((+)))))).))))))
```

and target concentration 1.0 μM.

a) (3 pts) Write down the mathematical expression for the test tube ensemble defect C. Which term would become large if the on-target dimer is not dominated by the target secondary structure? Which term would become large if off-target monomers or tetrameters form at appreciable concentration in the test tube?

b) (7 pts) Use the Design page at nupack.org to design sequences over a test tube ensemble Ψ containing all off-target complexes of up to 4 strands. Provide your final sequences, the final value of the test tube ensemble defect, and your design script. The script for Demo 1 for concentration-based design will help you to get started. (Note that the online designer makes use of v1 of the NUPACK scripting language, as illustrated in the demo.)

c) (5 pts) Use the Analysis page to document the salient properties of the test tube and complex ensembles for your final design.

d) (5 pts) Now use the Design page to design for the target secondary structure without considering off-target complexes (i.e., use "complex design" instead of "test tube design"). How does the new design compare to your previous design when you evaluate its properties over the ensemble Ψ containing all complexes of up to 4 strands?

Problem 2 (Mechanism design for molecular logic, 15 pts).
Design the mechanism for DNA logic gates to evaluate the conditional `(x AND y) OR z`. Here, `x`, `y`, and `z` are all short unstructured DNA strands with *unrelated* sequences. Your primary reference is the paper Seelig et al., *Science*, **314**, 1585–1588, 2006. Making use of sequence translator and logic gate motifs from the paper, sketch the secondary structure of the gates for your problem and annotate the structures to make toehold and domain complementarity clear.

Problem 3 (Sequence design for reaction pathway engineering: orthogonal HCR amplifiers, 30 pts).
Your primary reference is the paper Wolfe et al., *J Am Chem Soc*, **139**, 3134–3144, 2017 and the corresponding Supplementary Information.

a) (5 pts) What is the conceptual advantage of multistate test tube design compared to multistate complex design in terms of design paradigms?

b) (5 pts) Consider sequence design for a library of 2 orthogonal DNA HCR amplifiers intended to detect different initiator sequences and operate independently in the same test tube at 23 °C. The HCR mechanism schematic is shown in Figure S1 (each DNA hairpin has a 12 nt toehold, 24 bp stem, and 12 nt loop) and the corresponding target test tubes are described in Sections S2.2.1 and

S2.2.2 of the Supplementary Information. Sketch the target test tubes for your design (including target secondary structure and target concentration for each depicted on-target complex). How many tubes are there in your design ensemble?

c) (15 pts) Use executable `multitubedesign` of NUPACK3.2 (described in the NUPACK3.2 User Guide) to design 2 orthogonal HCR amplifiers over this multistate test tube design ensemble using a 2% stop condition. Use sequence constraints to ensure GC content in the range 45%-55% and to prevent patterns `AAAA,CCCC,GGGG,TTTT`. Leave all nucleotide, complex, and tube defect weights at the default value of 1 except for the global crosstalk tube which is assigned a weight of 2 (because you are designing $N = 2$ orthogonal systems) to prevent the effect of crosstalk from being diluted in the design objective function as the number of orthogonal systems increases. Include your final designs and your design script. Note that `multitubedesign` uses v2 of the NUPACK scripting language (see Section 3.3.3 of the NUPACK3.2 User Guide). This design script provides a useful starting point (see Example 7 of the User Guide; note that you will need to change the design material and temperature).

d) (5 pts) Plot the residual defects for your design ensemble (see Supplementary Information Section S2.4 and Figure S13 for an example with 4 orthogonal HCR amplifiers).

Problem 4 (Sequence design for reaction pathway engineering: catalytic hairpin assembly of a 4-arm junction, 35 pts)**.**
Your primary reference is the paper Wolfe et al., *J Am Chem Soc*, **139**: 3134–3144, 2017 and the corresponding Supplementary Information. Consider sequence design for calalytic hairpin assembly (CHA) with the elementary steps depicted in Figure 1.

a) (15 pts) Use the definitions and conventions of Section S2.2.1 and S2.2.4 to define the multistate test tube ensemble for CHA with $L_{\max} = 2$ for all tubes. For clarity, use the same set names, domain names, and presentation format as Section S2.2.4 to present your specification. Note that the CHA example in Section S2.2.4 is closely related to your design problem and that there is an error in that example: X·B should be included in the sets $\Psi^{\text{exclude}}_{0_n}$ and $\lambda^{\text{cognate}}_n$.

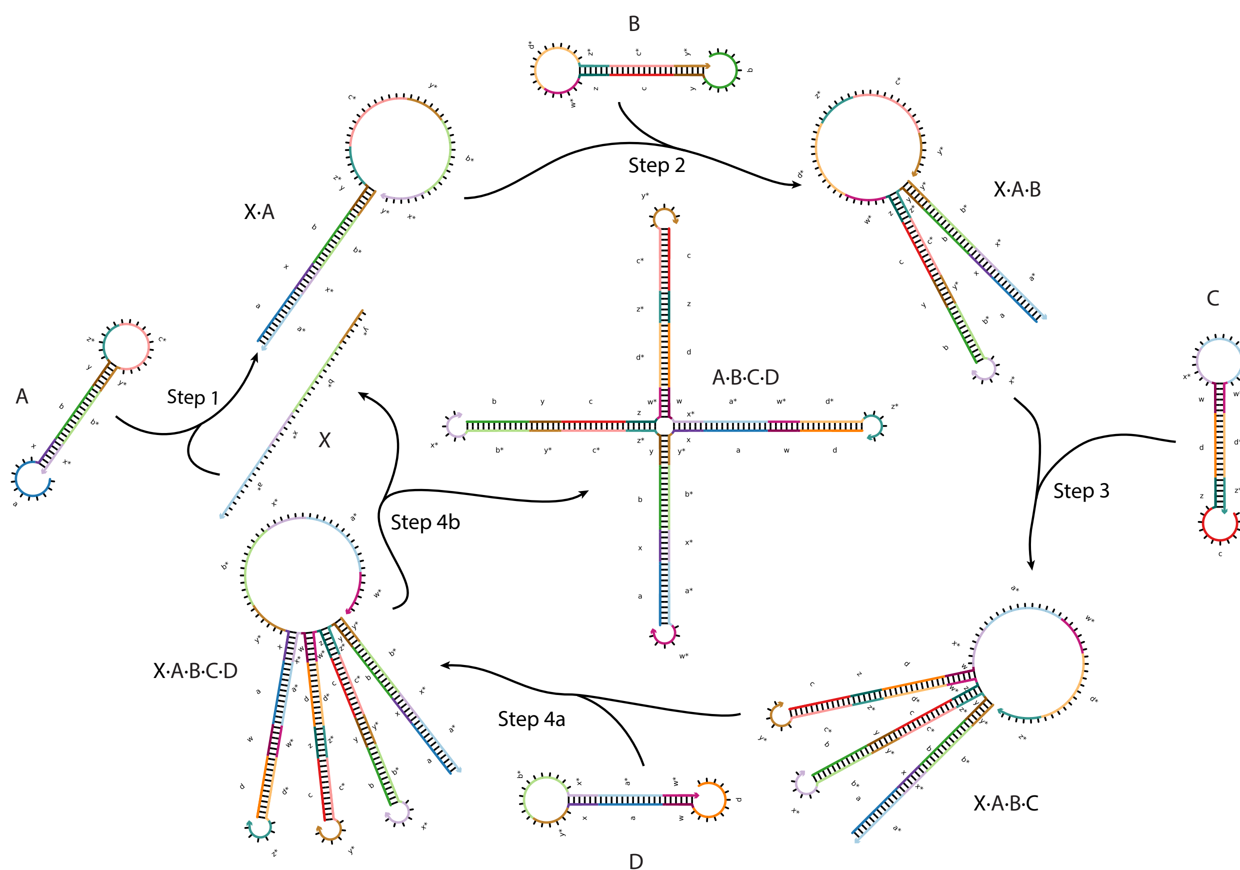
b) (10 pts) Use executable `multitubedesign` of NUPACK3.2 to design sequences over this multistate test tube design ensemble for DNA at 23 °C with a 2% stop condition. Include your final designs and your design script.

c) (5 pts) Update your script to stipulate that X is a subsequence of the *tpm3* mRNA sequence:

```
gaacacuauuagcuauuuguaguacucuaaagaggacugcagaacgcaucgcaguagugg\
ugaaaagccgugcgugcgcgugaaacaucugauccucacguuacuuccacucgcucugcg\
uuugacuuguuggcggggcguuggugccuuggacuuuuuuuuccuccuucucuucuucgc\
ggcucgguccacuacgcugcucgagaggaaucugcuuuauucgaccacacuacuccuaaa\
guaacacauuaaaauggccggaucaaacagcaucgaugcaguuaagagaaaaaucaaagu\
uuuacaacagcaagcagaugaggcagaagaaagagccgagauuuugcagagacaggucga\
ggaggagaagcgugccagggagcaggcugaggcagagguggcuucucugaacaggcguau\
ccagcugguugaggaggaguuggaucgugcucaggagagacuggccacagcccugcaaaa\
gcuggaggaagccgagaaggccgcagaugagagcgagagagggaugaaggugauugagaa\
cagggcucugaaggaugaggagaagauggagcugcaggagauccagcuuaaggaggccaa\
gcacauugcugaggaggcugaccgcaaauaugaagagguggcucguaagcuggugaucgu\
ugagggagaguuggagcguacagaggagagagcagagcuugcagagagccaugucaagca\
```

```
gauggaggaggagcugagagcucuugaccagacacugaagacucuucaggccucagagga\
gaaguauucccagaaggaggacaaguaugaggaagaaaucaagauccucacugauaagcu\
gaaggaggcugagacccgugcagaguuugcugagaggucuguggccaaacuggagaaaac\
cauugaugauuuggaagagaaacugagagaugcuaaagaggagaacaucaagauccaugc\
uacuuuggaccagacccugagcgagcucaauaguuucuaaagaagaccuggagcagaaaa\
aaggccuuuucuucccuucuugacucccucaucucauuuugguuucuuugucucugcaca\
ucugauucucccccuuuuuuuuucuucucuucuucugcuggaggauaagcucaccaagcc\
aaccagcaaaaauguggugccucucaauuuuuccaaacuacuauuccaagugauuugaga\
aaugaucuacuacgauacucccucaagagucaaauguugaccucggggagccuuuuuuggu\
auugcuccaugaucagagcuuuacgagcuaguguuuuuucugcauaucagcccaaacucu\
caaugauaauuuuacuggaggcugauuuuuguaaaauuuugugccauaaaagccuuguug\
gcuugucucuugcuuggcuuuagaucauucucaagccauuuuuuuccugcuguugcucug\
acacagguuguuuuugcuggucuuguuggugccugauccacugcuauccuuuucacaccu\
cuuuuuuuuuuuucuucauccugcacaaguuucugcugccuguuagucggcaucaccggu\
uuugggaccaaaccacaucaugugggucuguaacaguaugcacaaccaugccgugaggac\
caaauuuguuuuauuauuguuauuauuauuaaaagccuuugcuuccauucggaguuuguu\
uuuuugaguaauauauguauucauuguuugggucgaauccccuugcuuuuuuaacacaaa\
uguuuugcaaaccacuauuugaaauggugcacuguuaugggcuuauggugagcagaugag\
gccaagucaugguuucuucauuauaauuuucuuuucauuugcuuuaaagagccauauucu\
acccagggaagaaagguugaaguuguuuuguuuuuuuaccgugaguucaaagcaguggca\
cugccagauuuaaaagguucaaaagccgugcagaucuaaaauauguauuaugaacacagu\
aaugggagcgaauuguaacacuuaauaguauacaaauuuaagaaacaggggugaacacau\
aguuuuaacuggaaaaagcccacaaugauguguaaucacuuuguuacugucuguaucuug\
uguaaugauaccuaaauucuuuuuuuaaauaaaaaccaugauuuuuacugucacugaaaa\
aaaaaaaaaaaaaaa\
```

d) (5 pts) Plot the residual defects for your design ensemble with and without the mRNA sequence constraint (see Supplementary Information Section S2.4 and Figure S15 for an example with 3-arm junction CHA).



Step	Reaction	Function	Mechanism
1	X + A → X·A	assemble with catalyst X	toehold/toehold nucleation, 3-way branch migration
2	X·A + B → X·A·B	assemble	toehold/toehold nucleation, 3-way branch migration
3	X·A·B + C → X·A·B·C	assemble	toehold/toehold nucleation, 3-way branch migration
4a	X·A·B·C + D → X·A·B·C·D	assemble	toehold/toehold nucleation, 3-way branch migration
4b	X·A·B·C·D → X + A·B·C·D	disassemble from catalyst X and assemble 4-arm junction	intracomplex blunt-end strand invasion, 3-way branch migration

Figure 1: Reaction pathway for self-assembly of a 4-arm junction via catalytic hairpin assembly (CHA) (Yin et al., *Nature*, **451**:318-322, 2008 [see Figure 2]). Target X catalyzes self-assembly of metastable hairpins A, B, C, and D into 4-arm junction A·B·C·D. Sequence domain lengths: |a|=|b|=|c|=|d|=12 nt, |x|=|y|=|z|=|w|=6 nt. Top: Reaction pathway schematic. Bottom: Elementary step details.