Lab 8. Computation-aided design and simulation of DNA sensing circuit

I. Introduction

Nucleic acid detection is an important tool in molecular diagnostics, biosensing, and in-vivo labeling. Although traditional methods rely on enzyme-based signal amplification (e.g. PCR), this lab will discuss some detection methods which rely entirely on conformational (physical) changes in the nucleic acids. To achieve this, we exploit the *specificity* of nucleotide-nucleotide interactions; by changing the primary structure of an oligonucleotide (a short chain of DNA or RNA between ~10 and ~100 nucleotides), we can determine the presence and sometimes even the concentration of specific targets.

One of the earliest and simplest such systems is called "in-situ hybridization." These systems use just two oligonucleotides and rely on a physical phenomenon called FRET (Fluorescence Resonance Energy Transfer). Simply put, two oligonucleotides are covalently linked to two different fluorophores with unique excitation/emission spectra (**Figure 1**). In the absence of the target, the two fluorophores are free in solution; when one fluorophore is excited (e.g. cy3), we only observe that fluorophore's emission. In the present of the target nucleic acid sequence, however, the two strands will hybridize to the target or "scaffold" and hold their fluorophores close together. Now when cy3 is excited, some of the absorbed energy will be transferred to the cy5 molecule which will, in turn, also fluoresce. We will now observe both the emission spectrum of cy3 and that of cy5 (depending on the efficiency of energy transfer).

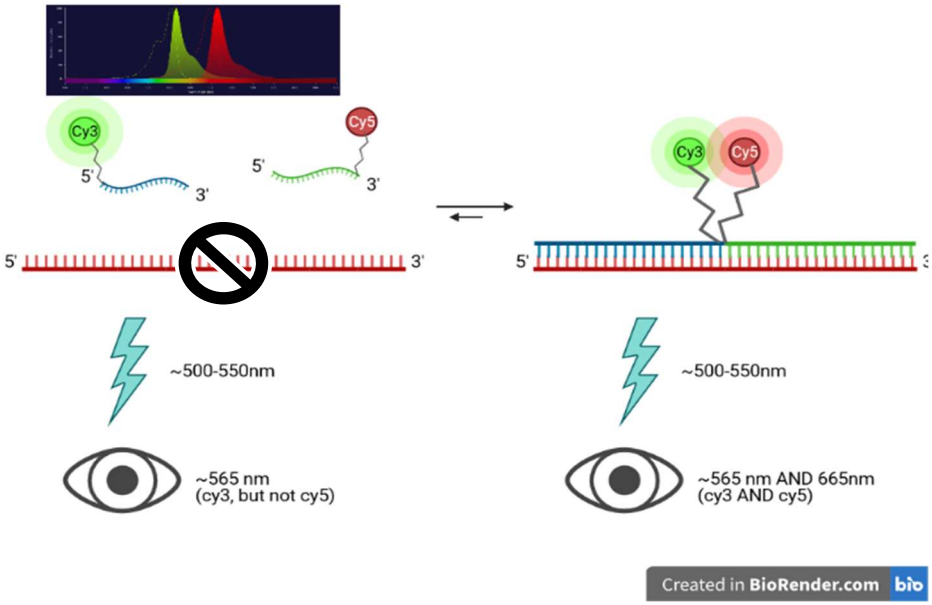


Figure 1. Illustration of in-situ hybridization. In solution, the fluorophores are too far apart for FRET; we only observe the spectrum of the fluorophore which is excited (cy3 here). In contrast, when the two fluorophores are colocalized on the scaffold strand, excitation at 500-550nm will cause both cy3 and cy5 to fluoresce due to FRET. Spectra from thermo fisher, illustration created with bio-render.

Computation-aided workflow for rapid prototyping and evaluating nucleic acids hybridizations. Today, many computational programs have been developed to facilitate the design of DNA hybridizations and to simulate the result. For example, we will use a free and online software NUPACK, which was developed by the Pierce Lab at CalTech.[1] NUPACK can also be downloaded and incorporated into a Python code for customized use. Other complimentary computational tools include Tiamat,[2] IDT biophysics[3] and CanDo.[4] **Figure 2** shows the example of using NUPACK to simulate the folding, hybridization equilibrium of DNA/RNA structures.

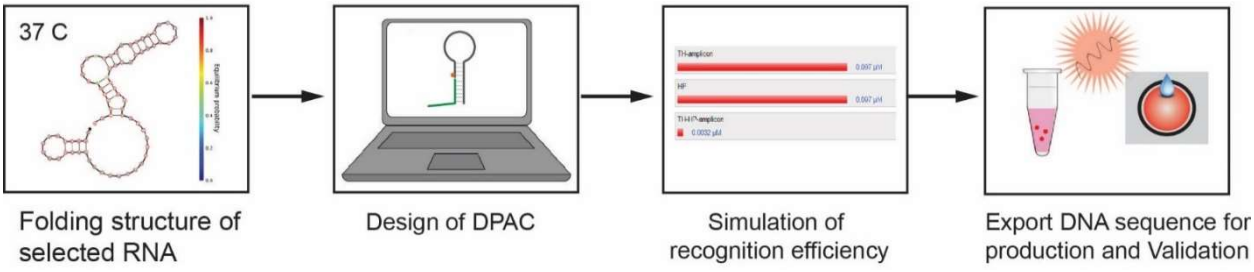


Figure 2. Computation-aided design and rapid prototyping.

II. Lab Goals

a. Learn how to use NUPACK for the basic DNA/RNA calculation and simulation
b. Identify and simulate folding structure of virus amplicons for diagnosis
c. Design anti-amplicon probes
d. Design in-situ hybridization
~~e. Design strand displacement-triggered DPAC sensor~~

III. NUPACK for predicting and simulating nucleic acids structures

a. Nupack can be used online via http://www.nupack.org/. In Biochem-I, we have used NUPACK to simulate the melting temperature of dsDNA, do you still remember that?

b. SARS-COV2-viral amplicon

For molecular diagnostic applications it is first necessary to identify an appropriate target sequence. Due to homologies between virulent and benign species, these sequences must be chosen carefully to avoid false-positives. This can be an involved bioinformatic process. In nucleic acid diagnosis, viral amplicons are the unique and concervative segments (e.g. E and N gene in SARS-COV-2) that are shorter than a few hundred nucleotides.

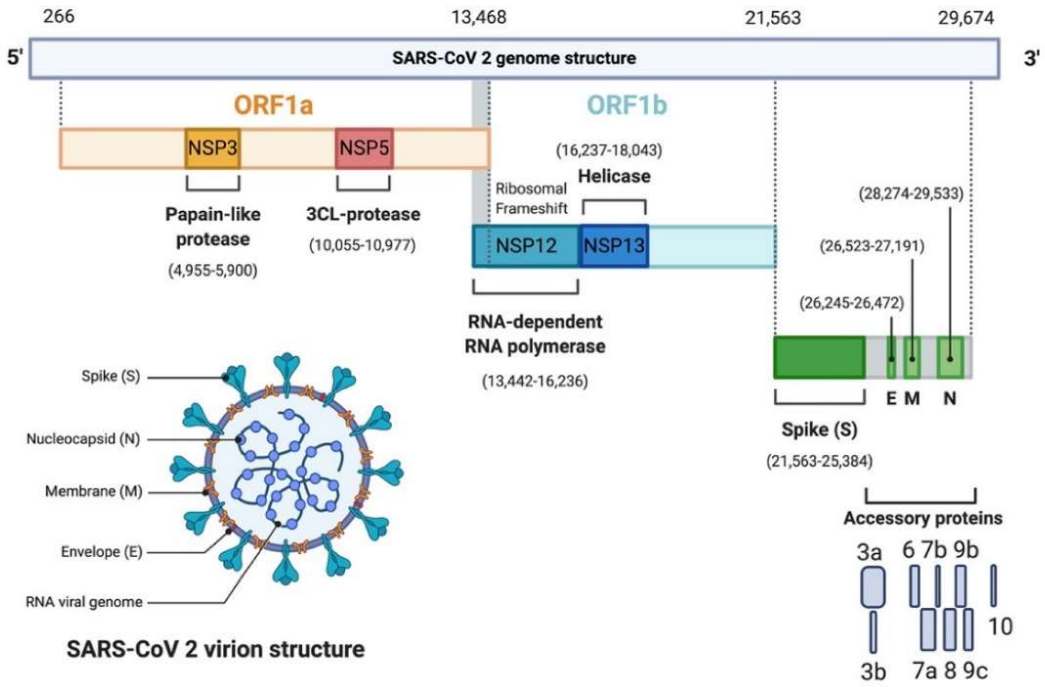


Figure 3. SARS COV-2 genome. Diagnosis amplicons use the unqiue and conservative gene at E and N gene, coating for Envelope and Nucleocapsid. Pathogens 2020, 9(5), 331; https://doi.org/10.3390/pathogens9050331

Fortunately, target sequences have already been identified for many important pathogens. **Table 1** lists the three virral amplicons from N gene for identifying SARS-COV-2 virus that causes COVID-19 pandemic.

Table 1. Three viral amplicones for identifying SARS-COV 2 virus. Red colors are forward primers and anti-reverse primers for the viral genome amplification by PCR. **Green color is the probe sequence** for detection.

ion	Forward Primer	Reverse Primer	Amplicon
N1	GATAATGGACCCCAAAATCAGCGAAATG	ATTCTGGTTACTGCCAGTTGAATCTGAG	GATAATGGACCCCAAAATCAGCGAAATG CACCCCGCATTACGTTTGGTGGACCCT CAGATTCAACTGGCAGTAACCAGAAT
N2	ACTAATCAGACAAGGAACTGATTACAAA	GCGCGACATTCCGAAGAACGCTGAAGCG	ACTAATCAGACAAGGAACTGATTACAAA CATTGGCCGCAAATTGCACAATTTGCC CCCAGCGCTTCAGCGTTCTTCGGAATGT CGCGC
N3	AACTGAGGGAGCCTTGAATACACCAAAAGA	GTTGTAGCACGATTGCAGCATTGTTAGCAG	AGACGGCATCATATGGGTTGCAACTGAG GGAGCCTTGAATACACCAAAAGATCAC ATTGGCACCCGCAATCCTGCTAACAAT GCTGCAATCGTGCTACAACTTCCTCAAG GAACAACA

Task 1. Use NUPACK to simulae the folding structure of RNA amplicons (N1, N2, N3) at 37 C. (*You should have alrady done it in Lab 7***)**

Task 2. Evaluate anti-probe hybridizations for RNA amplicons (N1, N2, N3)

a. *Produce anti-probe sequence by using reverse complement tool at* https://www.bioinformatics.org/sms/rev_comp.html

E.g. N1 probe: 5'- **CACCCCGCATTACGTTTGGTGGACCCT**

Anti-N1 probe (***reverse complement***):5' - AGGGTCCACCAAACGTAATGCGGGGTG

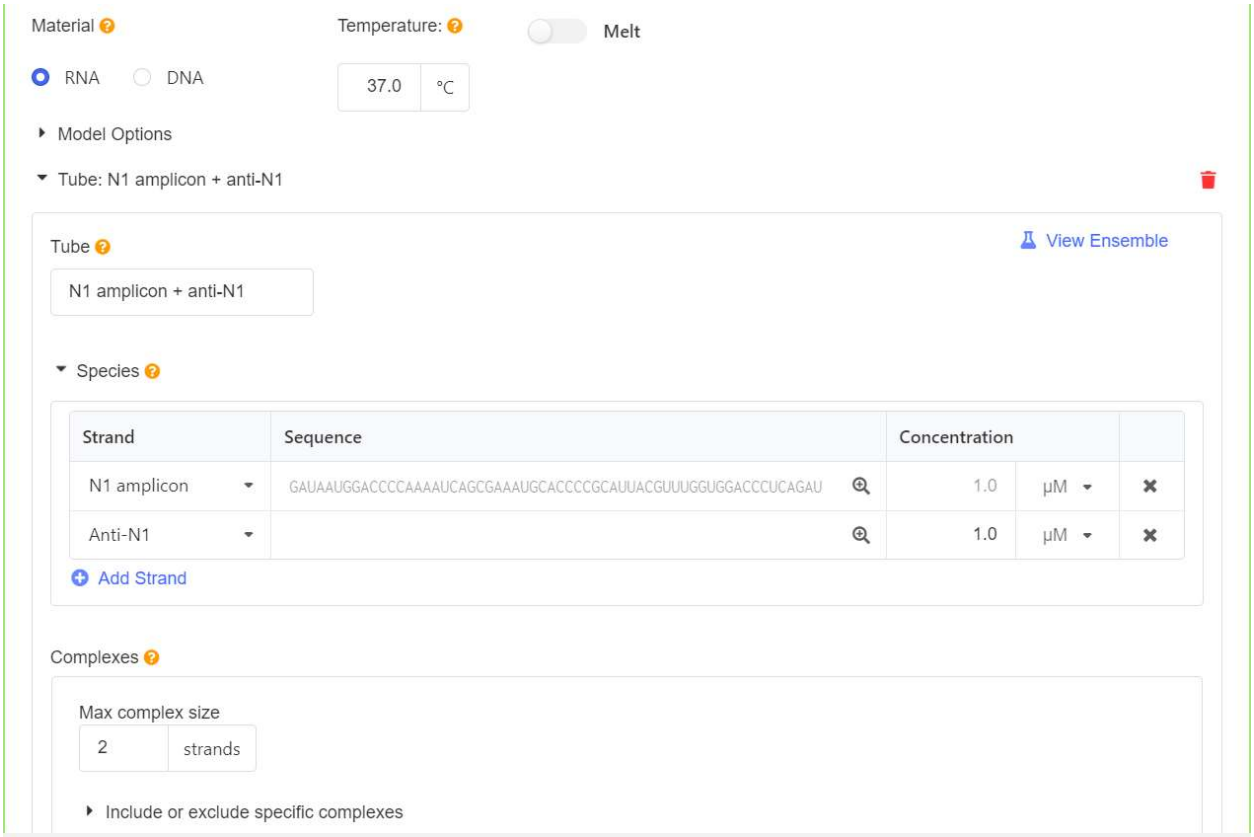
Generate sequences for anti-probes below:

Amplicon	Amplicon Probe	Anti-probe sequences
N1	**CACCCCGCATTACGTTTGGTGGACCCT**	
N2	**CATTGGCCGCAAATTGCACAATTTGCCCCCAG**	
N3	**GCCTTGAATACACCAAAAGATCACATTGGCACCCGCAATCCTGCTAACAATG**	

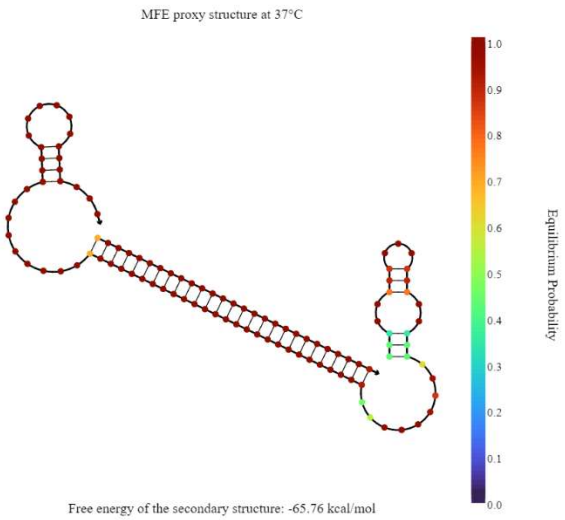
b. Evaluate anti-probe complement for hybridizing with viral amplicons.

(Note: If NUPACK simulation is failed in RNA mode, it can use DNA mode alternatively)

E.g. Anti-N1 probe + N1 amplicon at 37C



Generate hybridization structure to confirm the recognition of viral amplicon.



Task 3. Design in-situ hybridization for targeting probe sequences in N1, N2, N3 amplicons

a. Split anti-probe reverse complement into "half-half"

E.g. N1 probe: 5'- CACCCCGCATTACGTTTGGTGGACCCT

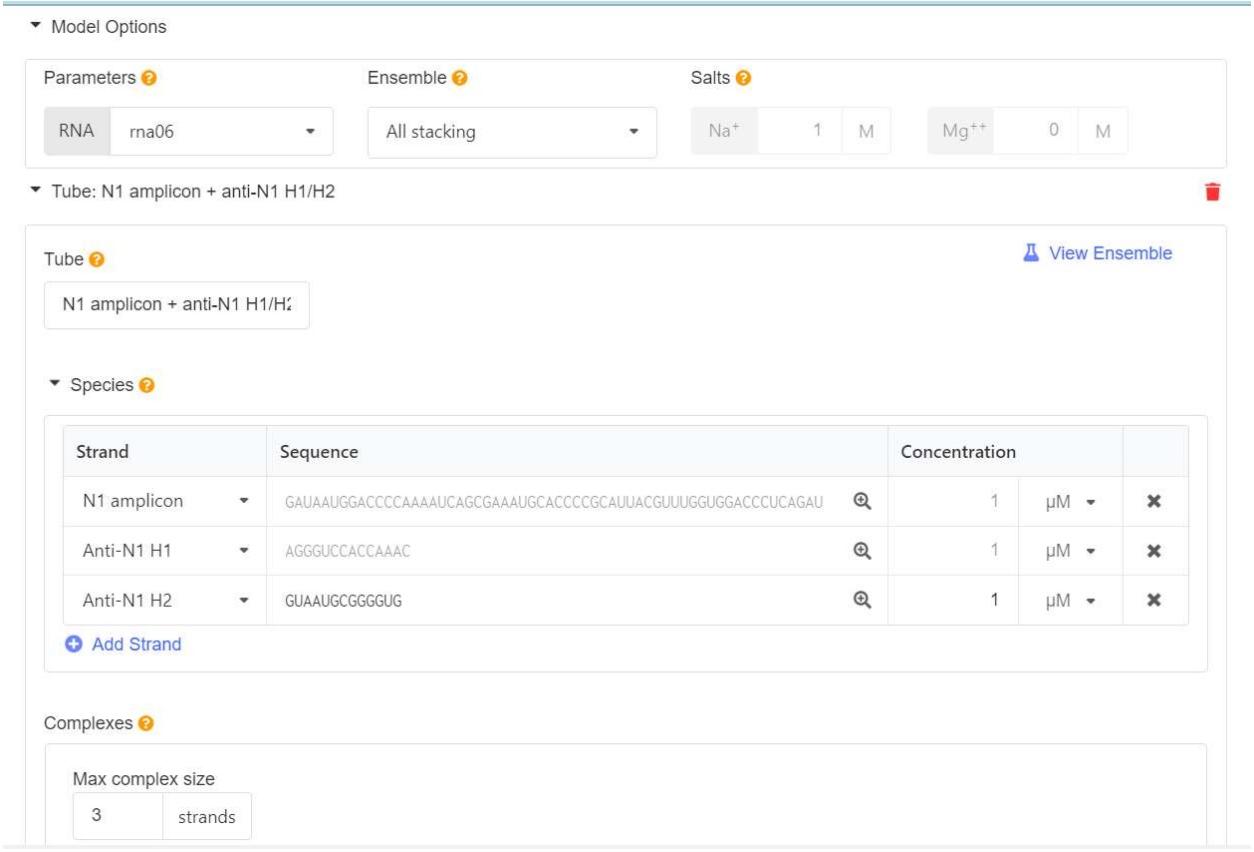
Anti-N1 probe (reverse complement):5' – AGGGTCCACCAAACGTAATGCGGGGTG

Anti-N1-H1: AGGGTCCACCAAAC

Anti-N1-H2: GTAATGCGGGGTG

The two halves can serve as the detection sensors for in-situ hybridization described in **Figure 1.**

b. Use NUPACK to simulate the in situ hybridization of H1, H2 with RNA amplicon (N1).



c. Evaluate the yield and specificity of in situ hybridization

	Hybridization conc. with amplicons (H1-H2-Nx)			Specificity (Target Yield /Nontarget Yield)
Sensor Pair	N1	N2	N3	
Anti-N1 H1/H2				
Anti-N2 H1/H2				
Anti-N3 H1/H2				

E.g. **Anti-N1 H1/H2** target is N1, and if hybridized conc. is 1000 nM; **Anti-N1 H1/H2** to N2 amplicon conc. is 10 nM; then specificity is 1000/10 = 100.

Data report is required for Task 1, 2 , 3.

Reference

(1) Zadeh, J. N.; Steenberg, C. D.; Bois, J. S.; Wolfe, B. R.; Pierce, M. B.; Khan, A. R.; Dirks, R. M.; Pierce, N. A. NUPACK: Analysis and design of nucleic acid systems. *Journal of Computational Chemistry* **2011**, *32*, 170-173.

(2) Williams, S.; Lund, K.; Lin, C.; Wonka, P.; Lindsay, S.; Yan, H.: Tiamat: A Three-Dimensional Editing Tool for Complex DNA Structures. In *DNA Computing*; Goel, A., Simmel, F., Sosík, P., Eds.; Lecture Notes in Computer Science; Springer Berlin Heidelberg, 2009; Vol. 5347; pp 90-101.

(3) Owczarzy, R.; Tataurov, A. V.; Wu, Y.; Manthey, J. A.; McQuisten, K. A.; Almabrazi, H. G.; Pedersen, K. F.; Lin, Y.; Garretson, J.; McEntaggart, N. O.; Sailor, C. A.; Dawson, R. B.; Peek, A. S. IDT SciTools: a suite for analysis and design of nucleic acid oligomers. *Nucleic Acids Research* **2008**, *36*, W163-W169.

(4) Castro, C. E.; Kilchherr, F.; Kim, D.-N.; Shiao, E. L.; Wauer, T.; Wortmann, P.; Bathe, M.; Dietz, H. A primer to scaffolded DNA origami. *Nat Meth* **2011**, *8*, 221-229.