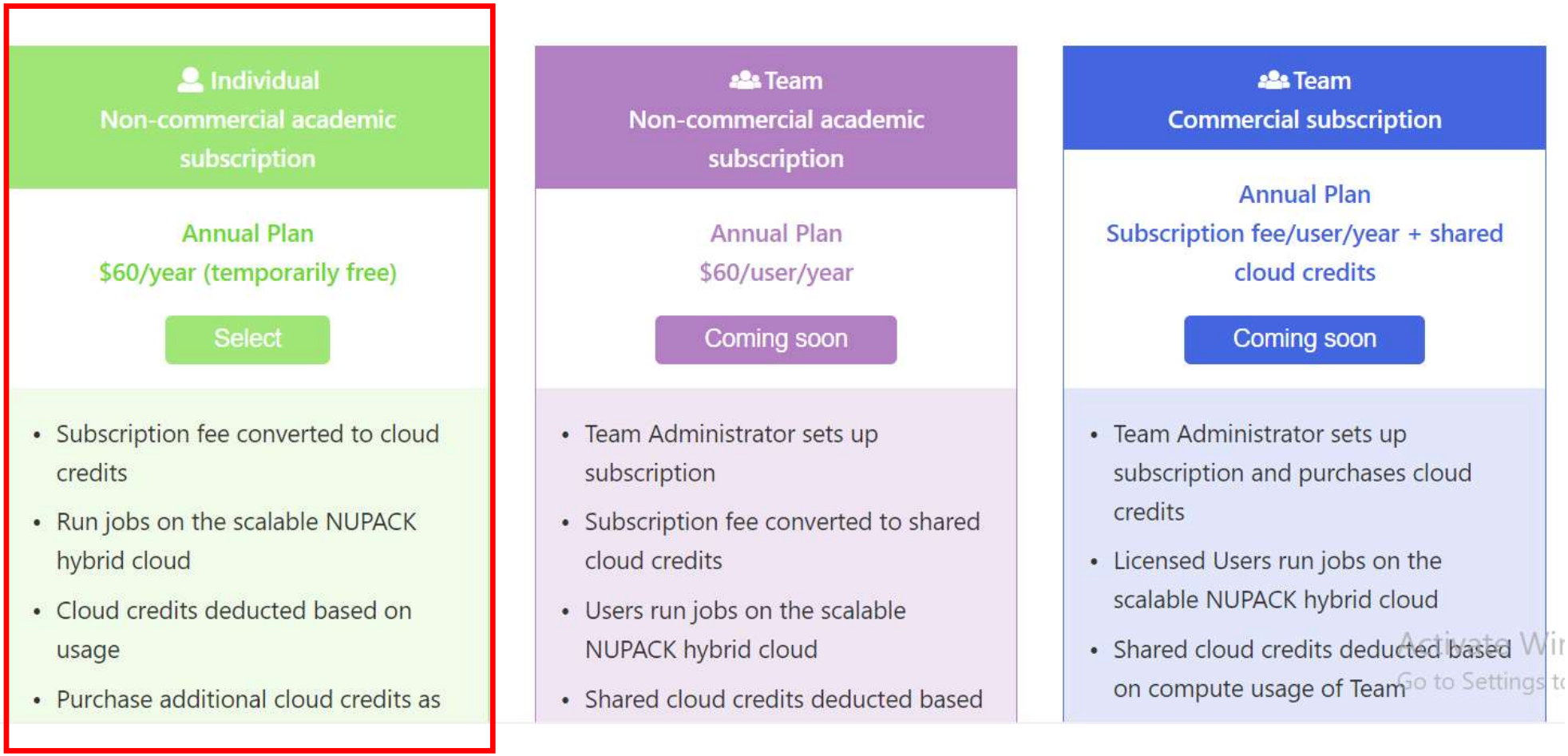
Lab 7 Simulation of DNA hybridizations

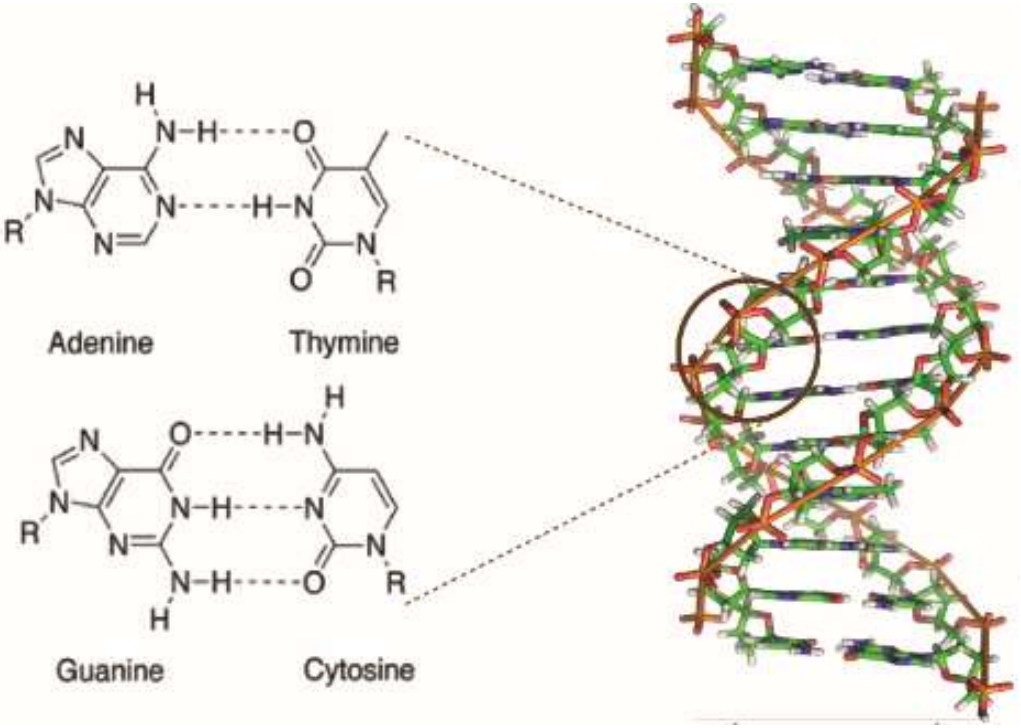
NUPACK Registration (Free now)

- https://www.nupack.org/subscribe



Register your use name on NUPACK Cloud for use it

DNA hybridization energy calculation



ssDNA: single-stranded

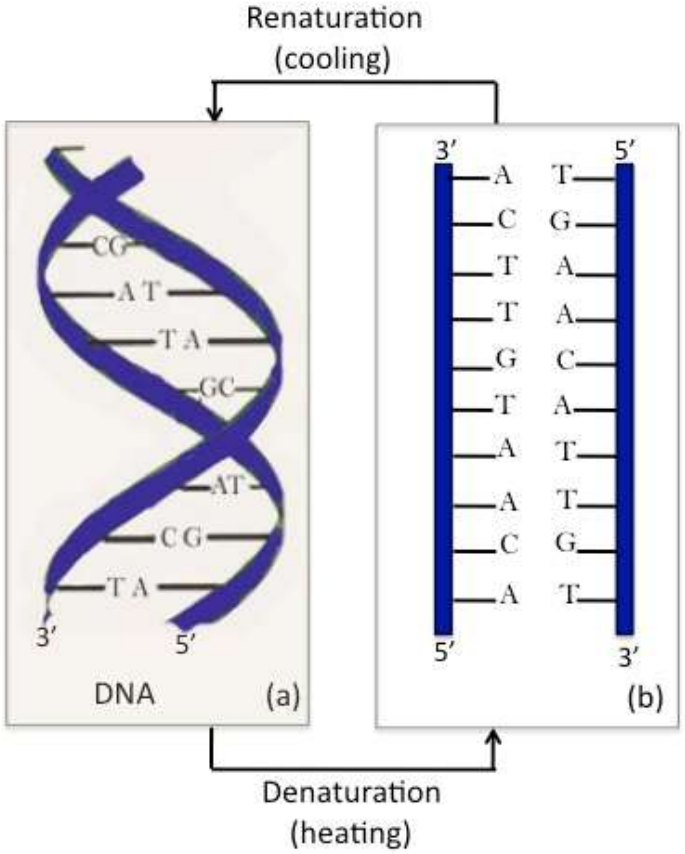
dsDNA: double-stranded

- dsDNA is stabilized by base pairing
- Hybridization energy follows a thermal physical reaction process:

$$\Delta G(total) = \sum \Delta G\ (individual\ base\ pairs) + \sum \pi - \pi\ stacking - \sum charge\ repulsion$$

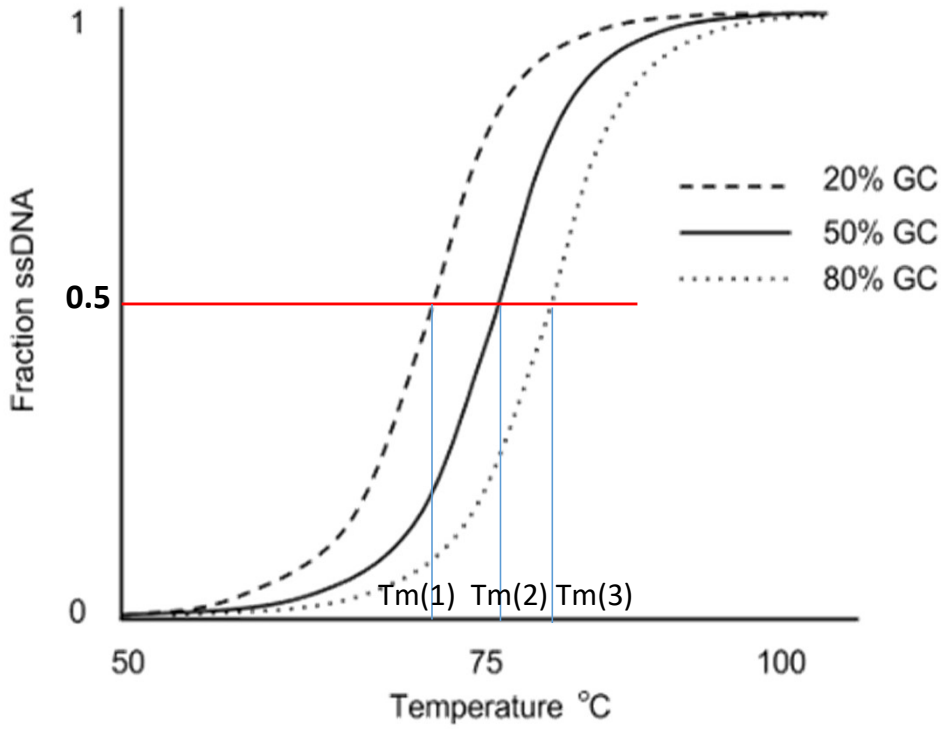
- Melting temperature is an important parameter for DNA hybridization stability

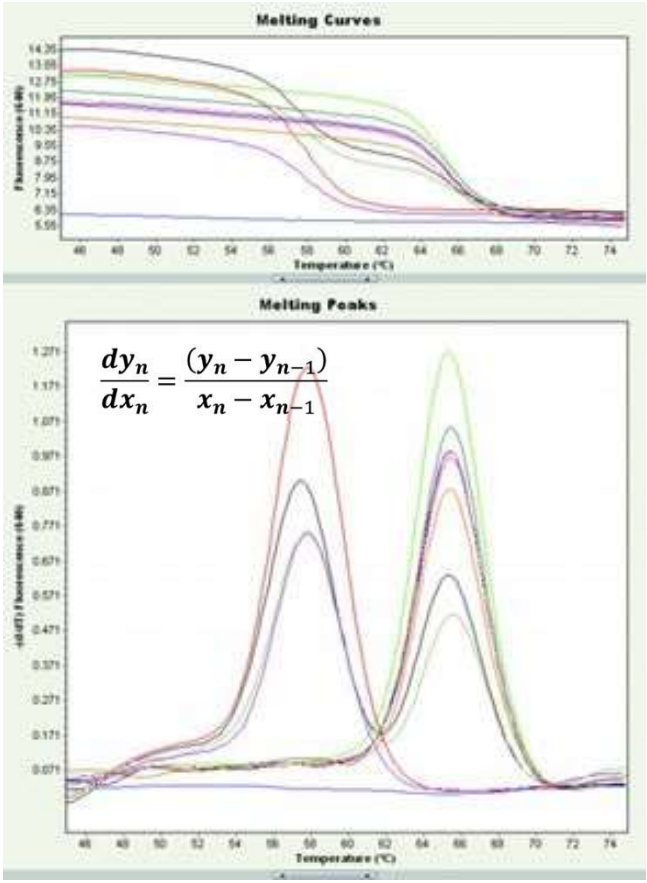
DNA thermal melting (or thermal denaturation): The dsDNA complexes may be dissociated by thermal denaturation, which is referred to as thermal melting.



- Melting temperature (Tm): the melting temperature of DNA refers to the temperature at which 50% of DNA in a sample has denatured from double-stranded DNA (dsDNA) to single-stranded DNA (ssDNA). Sensitive measurement of the melting curve of a sample of DNA can be used to detect single nucleotide differences between two DNA samples.

- Other melting methods: Chemical reagent can also disrupt dsDNA, such as urea, SDS, detergent …
  It is called chemical denaturation





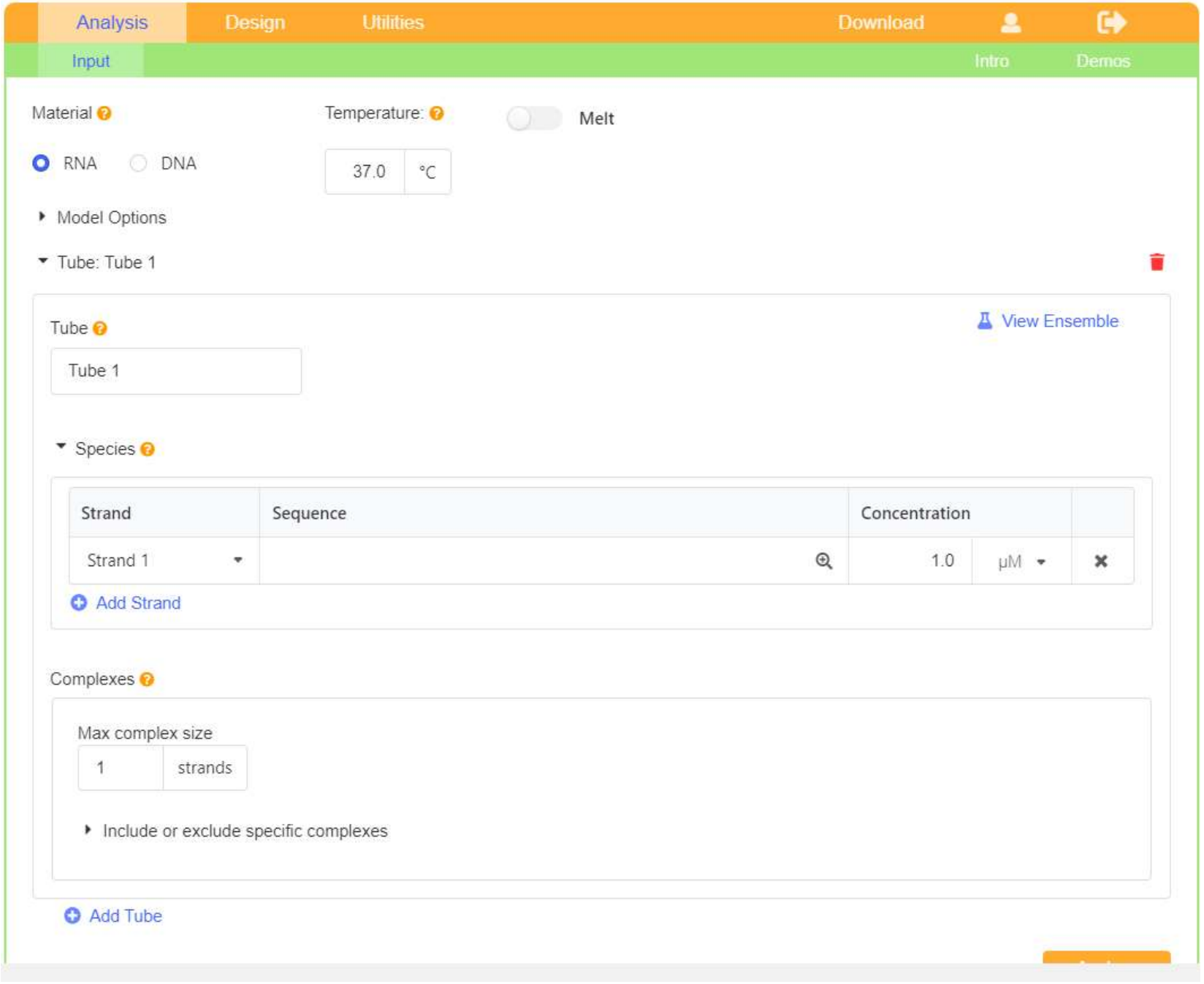
Melting Temperature (Tm) for dsDNA complexes with different GC%. Note, a dsDNA with **high GC%** shows a higher melting temperature due to the increased thermal stability. In addition, **magnesium ions** also play a key role to stabilize dsDNA by shielding the negative charges of phosphate backbone.

Tm, practically, is calculated by slope derivative analysis, that Tm is the temperature at the peak slope value. Generally, the first-order derivative analysis can be used to produce the slope curve and identify Tm value

Lab Goal:

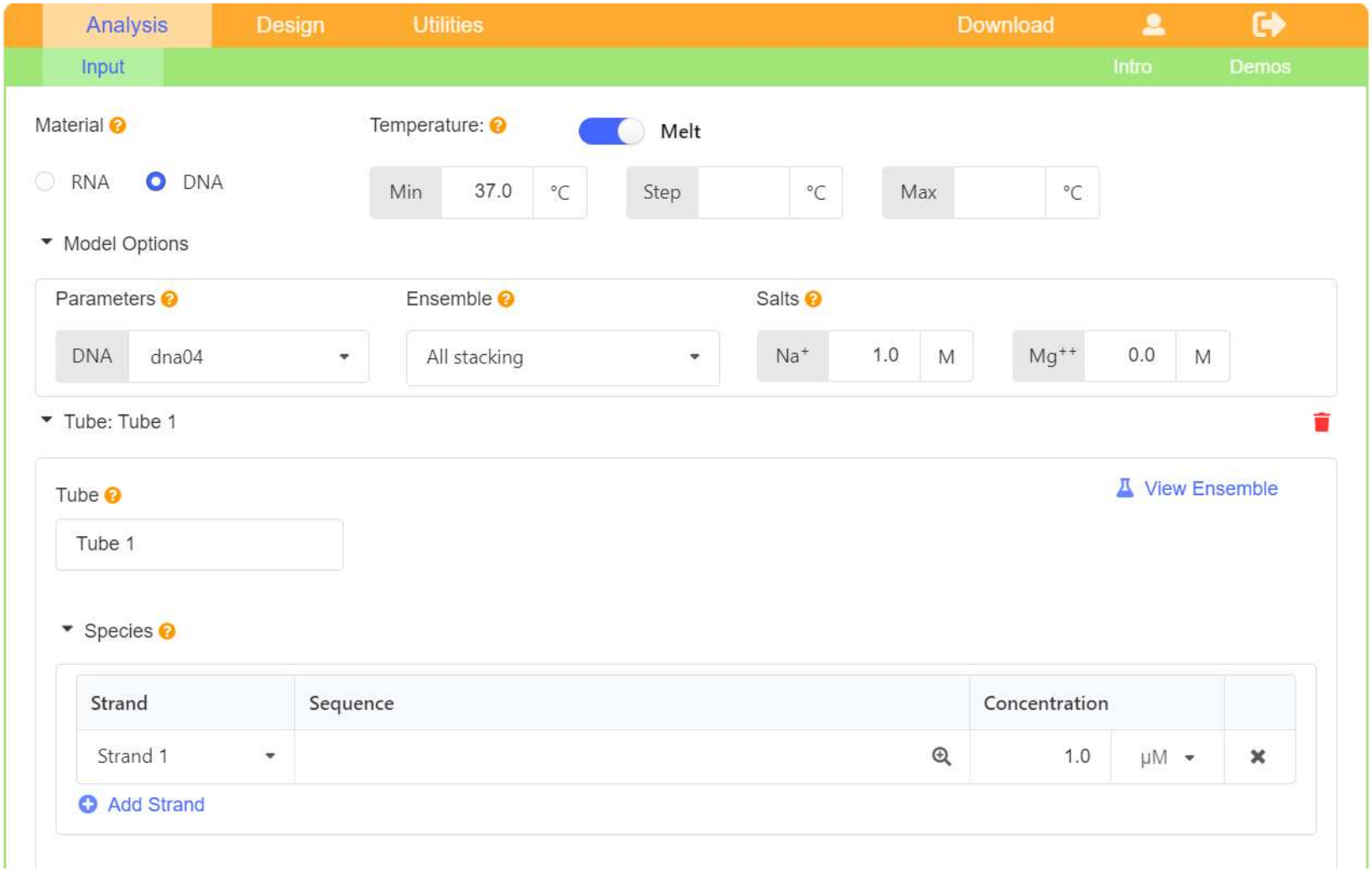
- **Learn NUPACK to simulate DNA hybridizations;**
- **Learn the fitting of DNA melting temperature;**
- **Understand the how DNA melting temperature is impacted by GC%, length, and Mg2+ conc.**

DNA computation and simulation with NUPACK



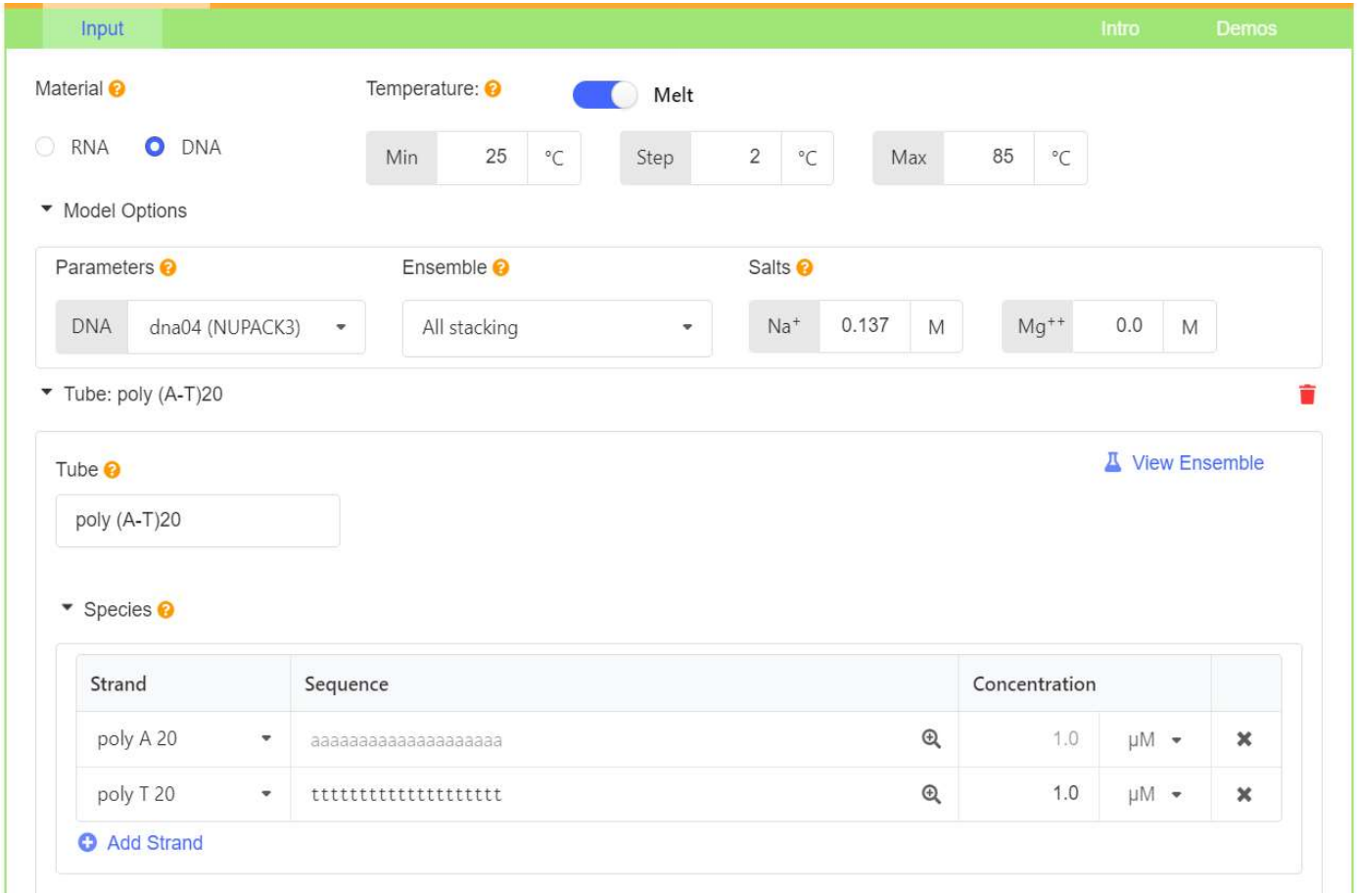
• Set Parameters for NUPACK

- Can analyze DNA and RNA;
- Set temperature and Melting analysis
- Run user-defined sequences and concentration
- In the “More Option”: Set Salt concentrations (Na+ and Mg2+)



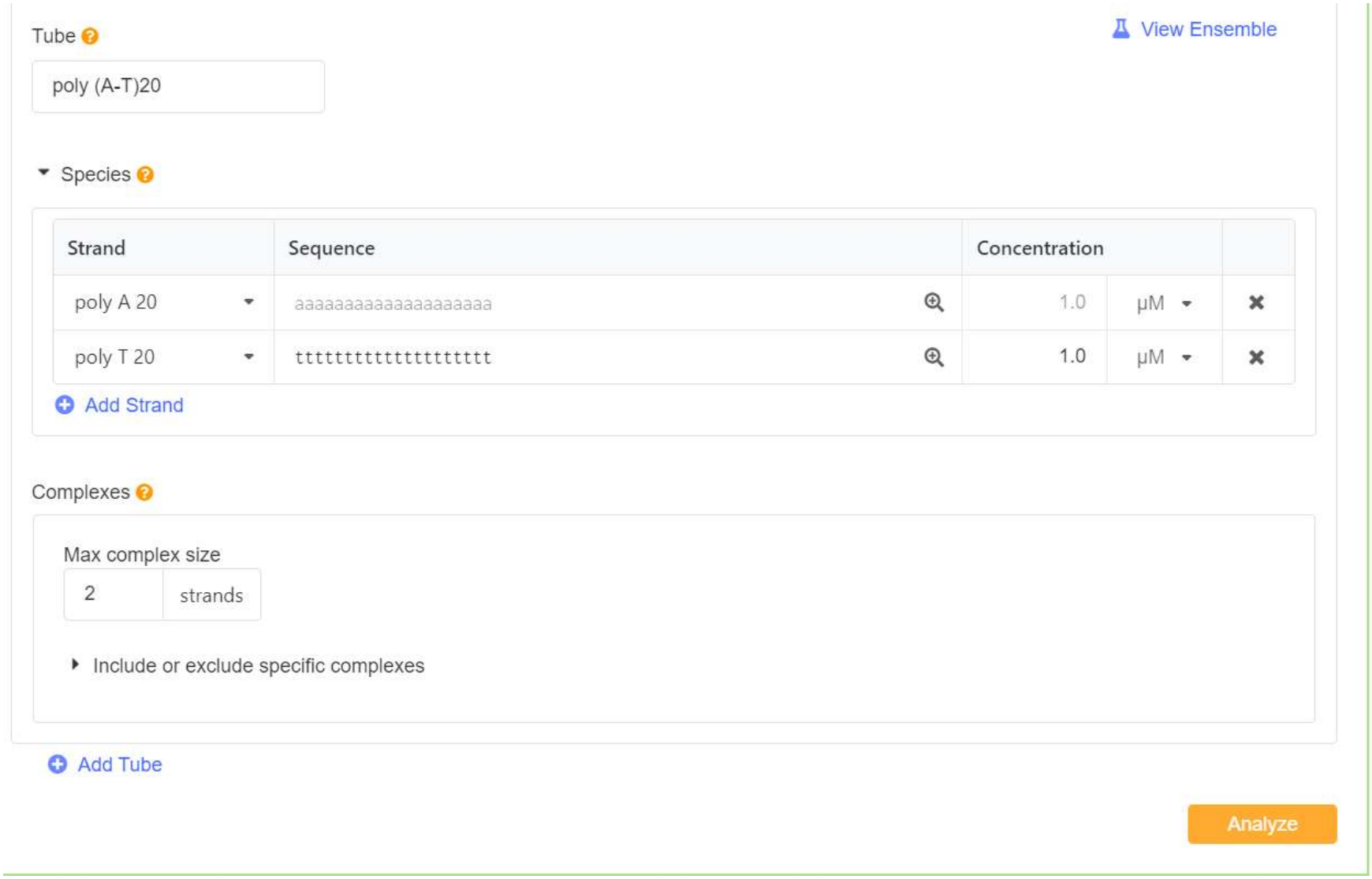
Task 1 Use Nupack to simulate and analyze dsDNA melting graph

- **Set Melting Temp analysis for poly (A-T) 20: Choose “DNA” and “Melt”**
- Set Min 25 C and Max 85 C, step 2 C
- Set “dna04” and “All stacking”; set Na+ to 0.137 M, Mg++ to 0 M
- Set name of “Tube” as: poly (A-T)20
- Set two strands of poly A20 and polyT20. You can use “Add strand” to add the second strand. Conc. = 1 uM



Task 1 Use Nupack to simulate and analyze dsDNA melting graph

- Set “Max Complex Size” to: 2 strands for dsDNA
- Click “Analyze”



Task 1 Use Nupack to simulate and analyze dsDNA melting graph

- Click results from analysis view melting graph.
- You can drag the temperature bar to view the hybridization changes at a specific temperature

Task 1 Use Nupack to simulate and analyze dsDNA melting graph

- Download data from NUPACK simulation
- Click “Download” to save the simulated DNA hybridization data in excel.

A	B	C
Temperature(C)	Fraction of bases unp	
25	0.083582	
27	0.088888	
29	0.095704	
31	0.103196	
33	0.11521	
35	0.124953	
37	0.136696	
39	0.151554	
41	0.171085	
43	0.197727	
45	0.235084	
47	0.287373	
49	0.359036	
51	0.452491	
53	0.564164	
55	0.681804	
57	0.787051	
59	0.865391	
61	0.914699	

• Fitting of “Melt Temp” from downloaded data

- Tm fitting with IC50 online tool: http://www.ic50.tk/index.html
- Copy and paste data into the blank box, click submit. IC50 is the Melt Temp of poly (A-T)20

Paste in X, Y values:

```
25      0.083582479
27      0.088888131
29      0.095704074
31      0.103195933
33      0.115209652
35      0.124953308
37      0.136696167
39      0.151554048
41      0.171084897
43      0.197727253
45      0.235084209
47      0.287372781
49      0.359035691
51      0.452490502
```

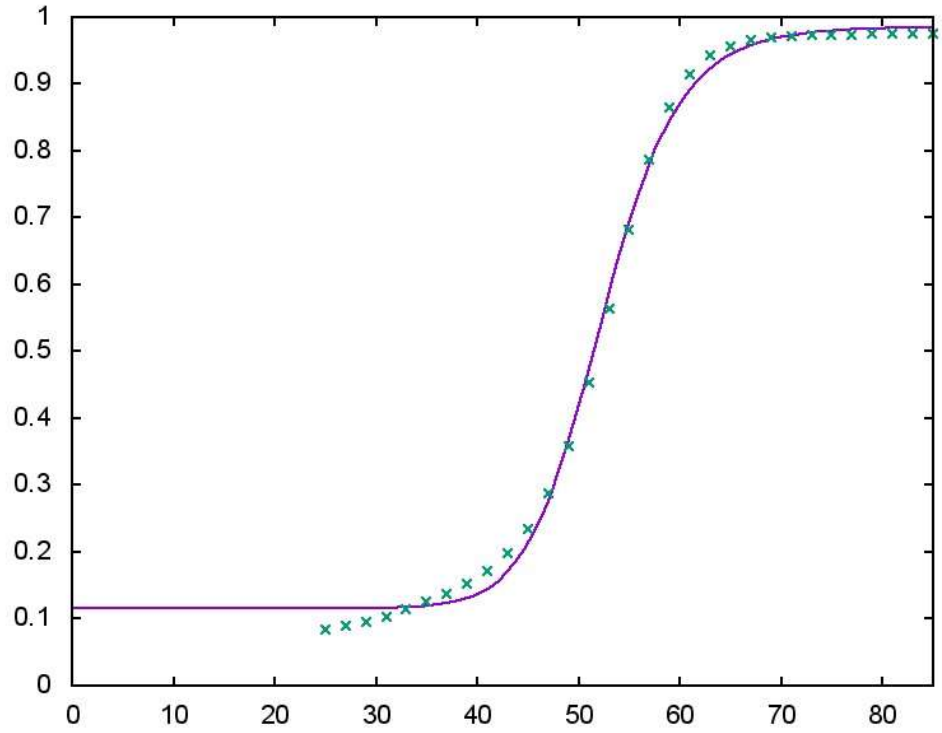
Submit

☐ Tick here and re-submit if curve-fit is bad.

Right-click on image to save.

Results:

Minimum	0.115961 +/- 0.006163 (5.315%)
Maximum	0.986135 +/- 0.006087 (0.6172%)
IC_{50}	52.3513 +/- 0.1722 (0.329%)
Hill coeff.	13.7464 +/- 0.5523 (4.018%)



Task 1 Use Nupack to simulate and analyze dsDNA melting graph

（from 25 –85 C, 3 C per step)

- **GC%:** Thermal melting graph for 1 uM poly(A-T)$_{20}$ and poly (C-G)20; at 0.137M Na+, 0 Mg2+, Fitting for Tm
- **Length:** Thermal melting graph for 1 uM poly(A-T)20 and 1 uM poly(A-T)40; at 0.137 M Na+, 0 Mg2+, Fitting for Tm
- **Mg2+**: Thermal melting graph for 1 uM poly(A-T)$_{20}$ at 0.137 M Na+, 0 M Mg2+, Fitting for Tm.

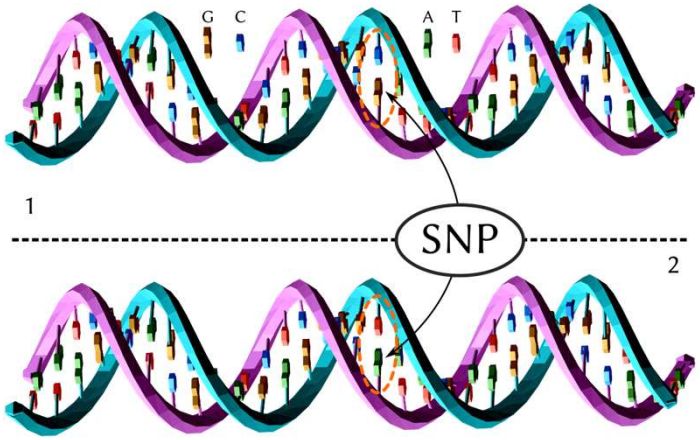
  Do it again for salt conditions at 0.137 M Na+, 0.01 M Mg2+.
- How to conclude the parameters that affect the melt Temp of dsDNA?

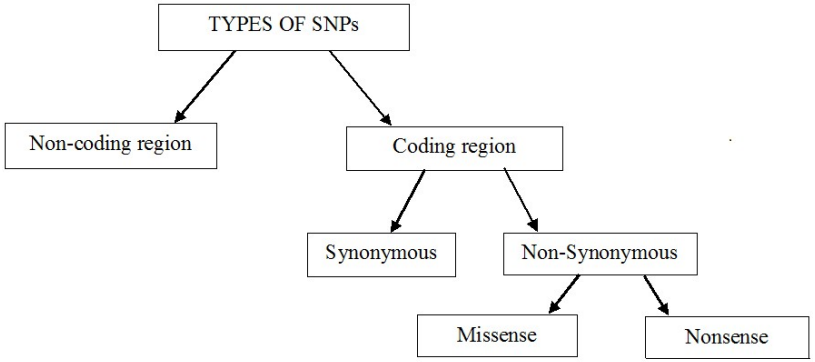
Task 2. DNA hybridization melting temperature to detect SNP/SNV (Single-nucleotide Polymorphism/Variants)

A single nucleotide polymorphism, or SNP (pronounced "snip"), is a variation at a single position in a DNA sequence among individuals. For example, at a specific base position in the human genome, the G nucleotide may appear in most individuals, but in a minority of individuals, the position is occupied by an A. This means that there is a SNP at this specific position, and the two possible nucleotide variations – G or A – are said to be the alleles for this specific position. Although a particular SNP may not cause a disorder, some SNPs are associated with certain diseases.

SNP may or may not cause an amino acid mutation.

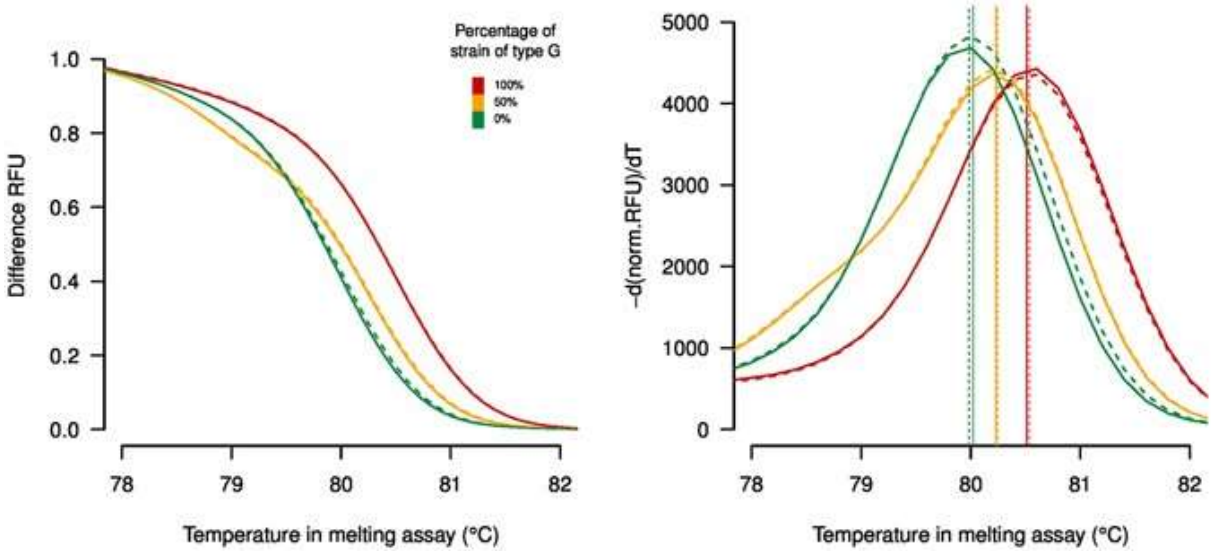
Single-nucleotide variant (SNV) is a variation in a single nucleotide. SNVs differ from SNPs in that a SNV can be somatic[9] and **can be caused by cancer**,[10] but a SNP has to segregate in a species' population of organisms. SNVs also commonly arise in molecular diagnostics such as designing PCR primers to detect viruses, in which the viral RNA or DNA sample may contain SNVs.





https://en.wikipedia.org/wiki/Single-nucleotide_polymorphism#/media/File:Dna-SNP.svg

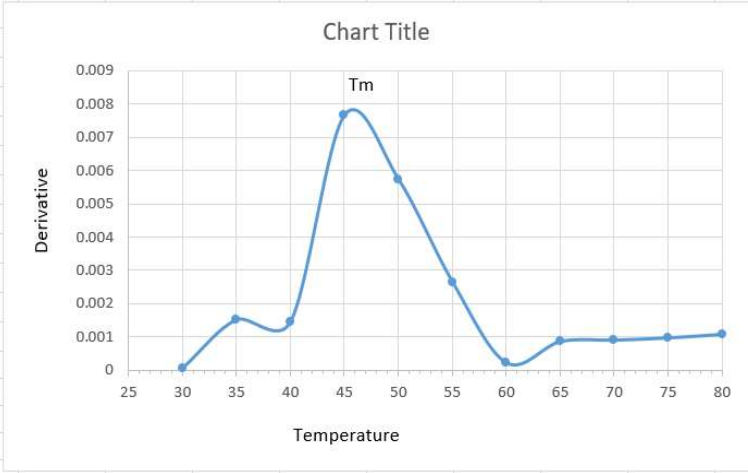
SNP/SNV detection by PCR with melting temperature difference, **first-order derivative analysis**



Derivative Analysis with Excel

$$\frac{dy_n}{dx_n} = \frac{(y_n - y_{n-1})}{x_n - x_{n-1}}$$

Temperature	A260	Derivative
25	1.12084	
30	1.12119	=(B3-B2)/(A3-A2)
35	1.12882	0.001526403
40	1.13616	0.001466918
45	1.17456	0.00768013
50	1.2033	0.00574801
55	1.21659	0.002658391
60	1.21771	0.000224924
65	1.22205	0.00086658
70	1.22659	0.000908089
75	1.23146	0.000974798
80	1.23686	0.001080463
Tm ~ 46 C		



Tm is the temperature with the highest derivative value

Example: CYP1A2 SNP

rs762551, also known as -164A>C or -163C>A, is a SNP encoding the CYP1A2*1F allele of the CYP1A2 gene. The CYP1A2 gene encodes a member of the cytochrome p450 family of proteins, which metabolize nutrients and drugs. One well known substrate of CYP1A2 is caffeine; individuals who carry one or more CYP1A2*1C alleles are "slow" caffeine metabolizers, whereas carriers of the variant CYP1A2*1F are "fast" caffeine metabolizers.

Primer	Sequences
TaqMan probe – TET	5′-CTC TGT GGG C**C**C AGG ACG CAT-3′
TaqMan probe – FAM	5′-TC TGT GGG C**A**C AGG ACG CAT GG-3′
TaqMan Forward primer	5′-TTT CCA GCT CTC AGA TTC TGT GAT-3′
TaqMan Reverse primer	5′-GGA TAC CAG AAA GAC TAA GCT CCA TC-3′
Forward primer*	5′-TTC CCC ATT TTG GAG TGG TC-3′
Reverse primer*	5′-CCG AGA AGG GAA CAG ACT GG-3′

Nordmark et. al. 2002

Home › Search Tool › Search Results › C___8881221_40

SNP ID:	rs762551
Gene	▾CYP1A2
Gene Name	cytochrome P450 family 1 subfamily A member 2
Set Membership:	> HapMap > DME > Validated > Inventoried
Chromosome Location:	-
Polymorphism:	C/A, Transversion Substitution
Context Sequence [VIC/FAM]:	TGCTCAAAGGGTGAGCTCTGTGGGC[C/A]CAGGACGCATGGTA GATGGAGCTTA

Task 2: Use Nupack to predict thermal melting graphs and Tm at 0.137 M Na+ and 0 Mg++

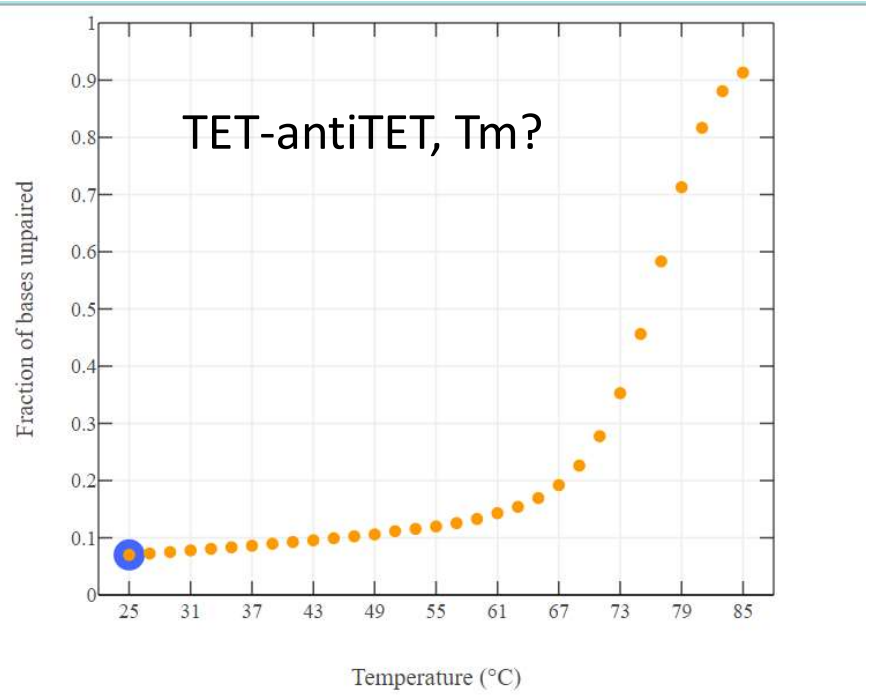
1) Anti-TET + TET seq (1 uM) and anti-TET + FAM seq (1 uM)

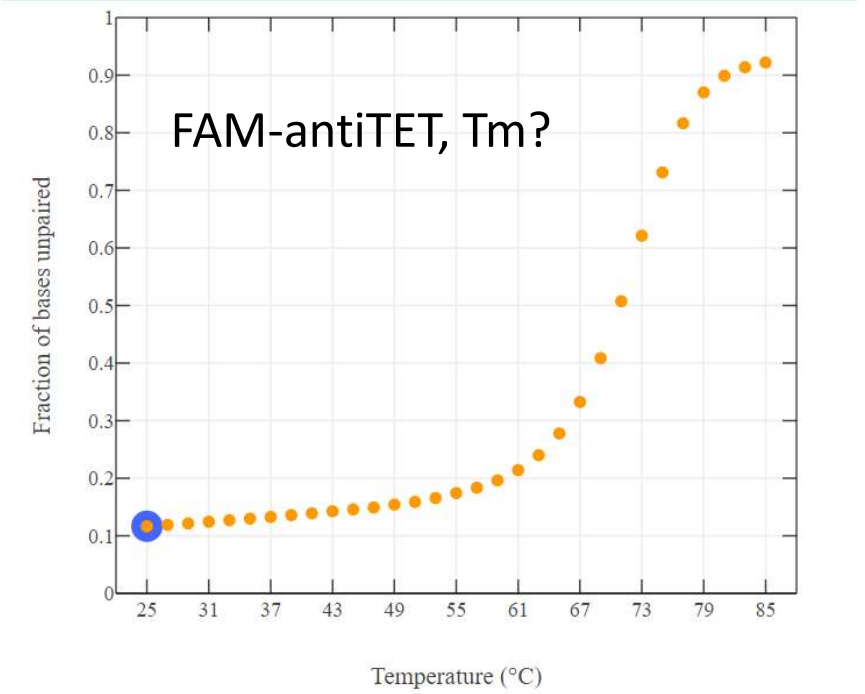
	Sequence	Tm (C)
Target TET	CTCTGTGGGC**C**CAGGACGCAT	
Target FAM	CTCTGTGGGC**A**CAGGACGCAT	
Anti-TET	ATGCGTCCTG**G**GCCCACAGAG	

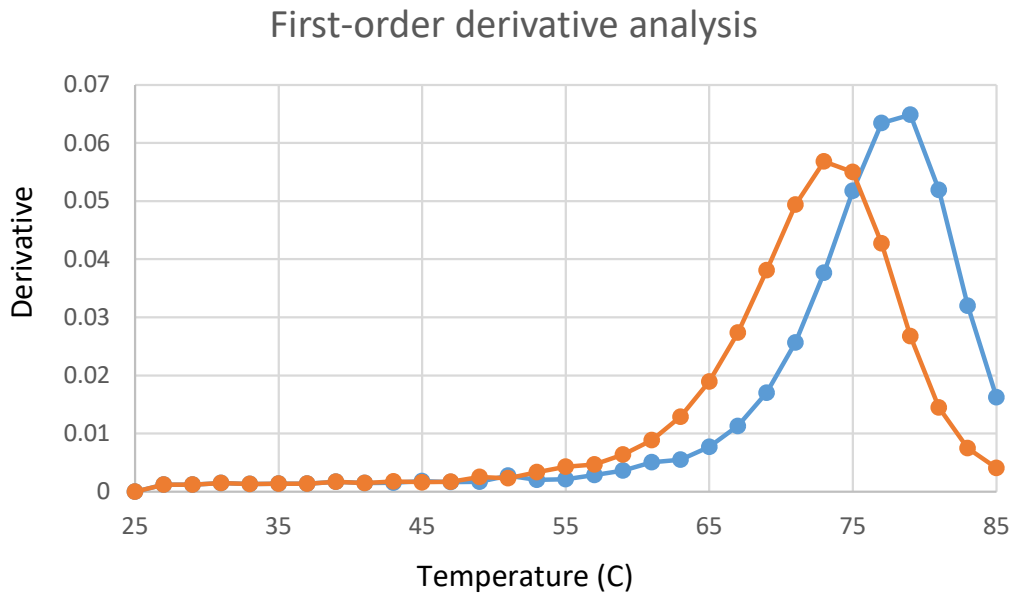
2) Based on Tm difference, suggest **a temperature** that produces the biggest difference for hybridizations between TET/anti-TET and FAM/anti-TET

	Sequence	Test Temperature (C)	Hybridization yield (%)
Target TET	CTCTGTGGGC**C**CAG GACGCAT		
Target FAM	CTCTGTGGGC**A**CA GGACGCAT		

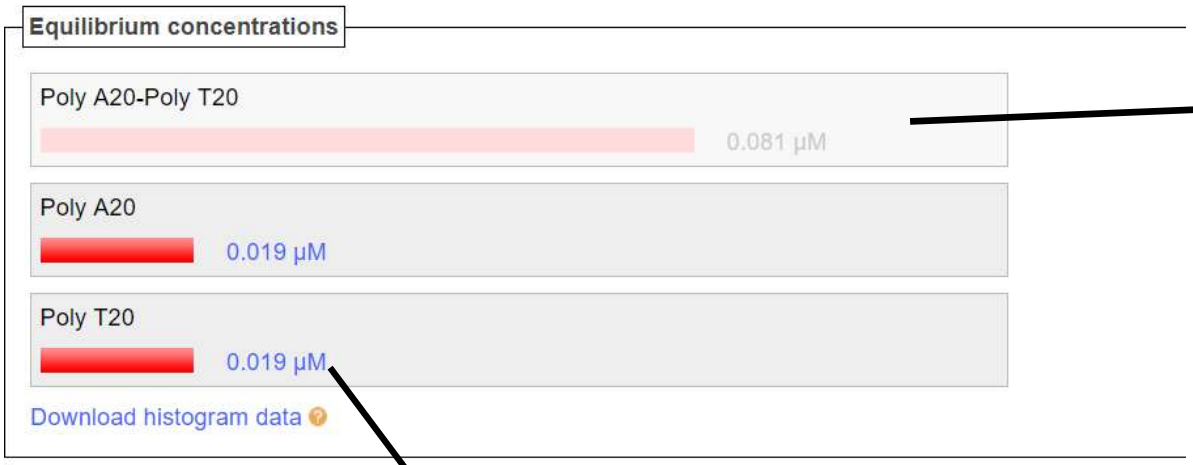
- **Melt Temp by first-order derivative analysis**

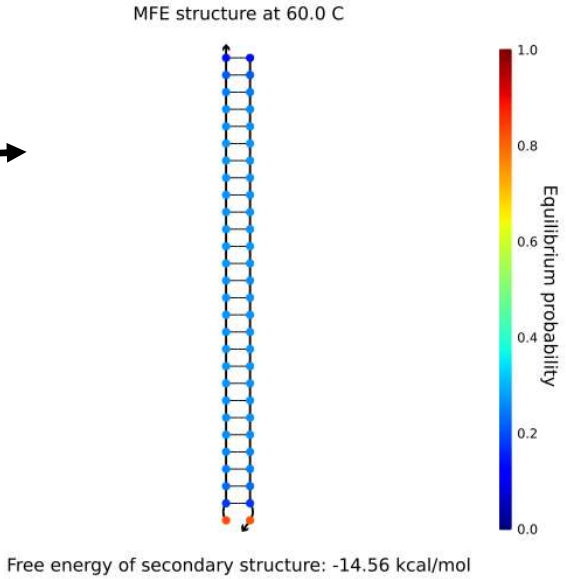


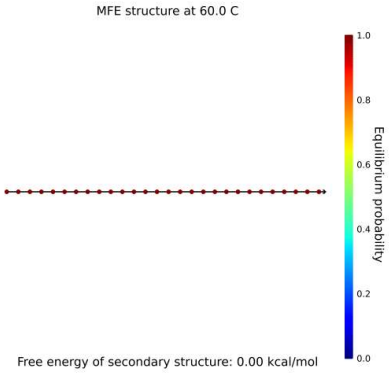




Hybridization Yield%







The color means the probability of the base in the predicted structure: Red, high probability; blue, low probability

Nupack can predict the equilibrium conc. of hybridizations. You can calculate the product yield:

e.g.

[Target dsDNA]/[total ssDNA input] = 0.081/0.1 = 0.81=81%

Task 3. Predict the secondary structure of RNA (RNA mode at 37 C) : the folding structure of SARS-COV-2 (cause COVID-19) amplicons for PCR diagnosis

N1 :

GATAATGGACCCCAAAATCAGCGAAATGCACCCCGCATTACGTTTGGTGGACCCTCAGATTCAACTGG
CAGTAACCAGAAT

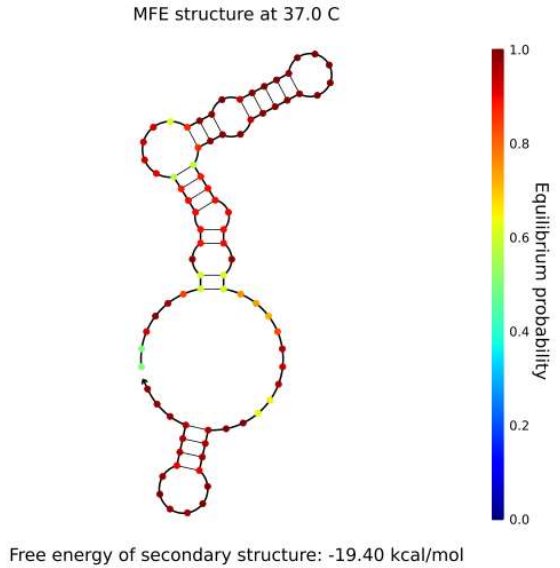
N2:

ACTAATCAGACAAGGAACTGATTACAAACATTGGCCGCAAATTGCACAATTTGCCCCCAGCGCTTCAG
CGTTCTTCGGAATGTCGCGC

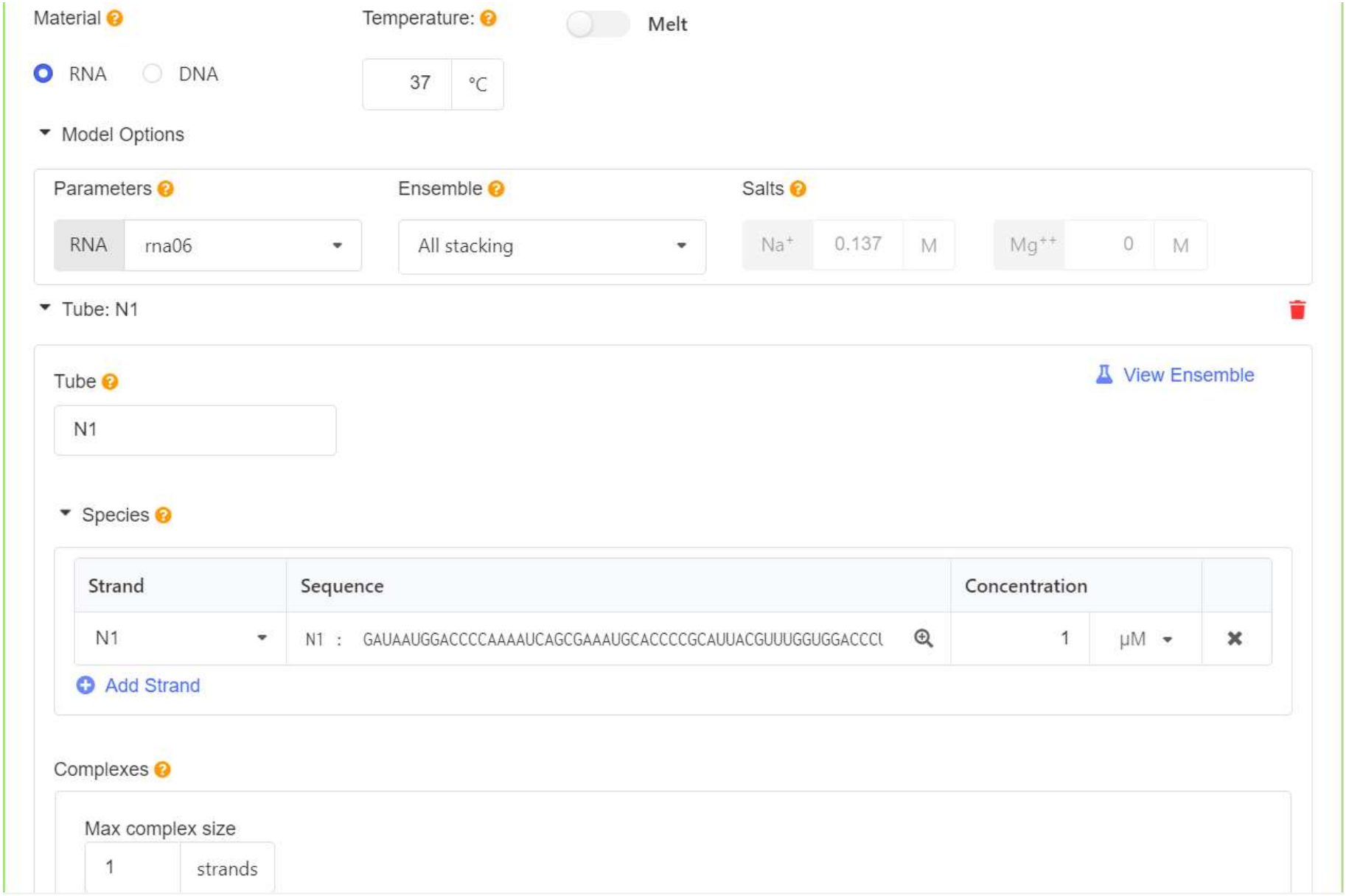
N3:

AGACGGCATCATATGGGTTGCAACTGAGGGAGCCTTGAATACACCAAAAGATCACATTGGCACCCGC
AATCCTGCTAACAATGCTGCAATCGTGCTACAACTTCCTCAAGGAACAACA

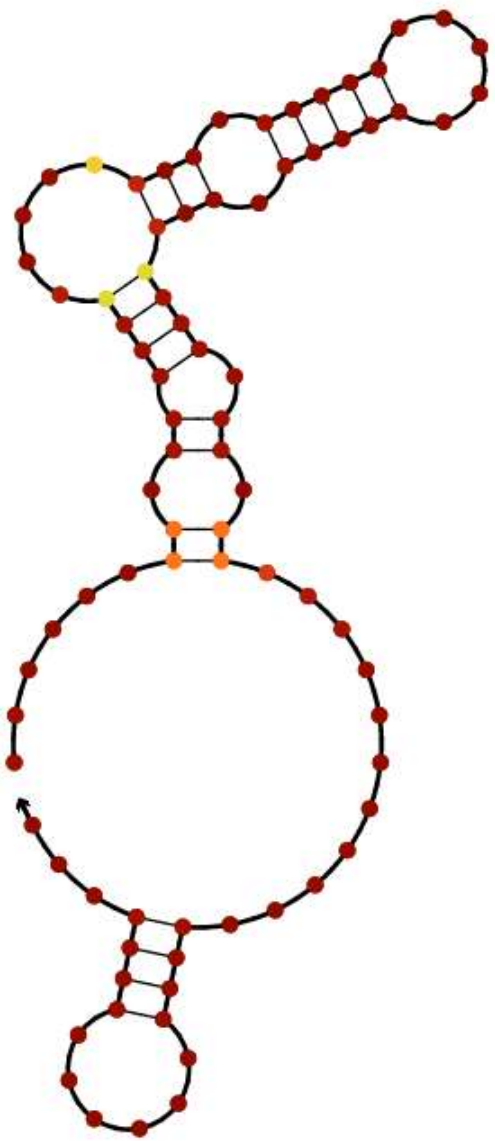
The color means the probability of the base in the predicted structure: Red, high probability; blue, low probability

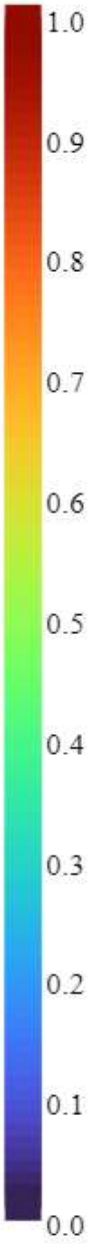


- **Calculation mode for RNA**
- **Under RNA mode, you cannot choose salt concentration**
- **Set Temp to 37 C, 1 strand**



MFE proxy structure at 37°C





Equilibrium Probability

Free energy of the secondary structure: -15.42 kcal/mol

Data sheet

Task 1 (at least 2 C interval, from 20 – 80 C) Na+ fixed at 0.137 M

- Figure 1 Thermal melting curves for poly(A-T)20 vs poly (C-G)20
- Figure 2 Thermal melting curves for poly(A-T)20 vs poly (A-T)40
- Figure 3 Thermal melting curves for poly (A-T)20 at 0 Mg2+ vs 0.01 M Mg2+
- Tm value fitting

Task 2

Use salt condition: 0.137 M Na+, 0 Mg

Thermal meting curves for TET and FAM, Tm value fitting using **1st order of derivative analysis**

Pick up a temperature based on derivative analysis , compare hybridization yield difference between TET and FAM at this fixed temerpature.

Task 3

- Folding structure for N1, N2, N3 amplicons