Development of platform for novel molecular method for detection of Coxsackievirus B3 in recycled water

Andrew Townsend[1], Jonathan Auyong[1], William Andreopoulos[1], **Marelyn Negrete[2],** and Leila Khatib[2]

1 Department of Computer Science
2 Department of Biological Science

SJSU SAN JOSÉ STATE UNIVERSITY

California approach- One Water

- EVERY DROP COUNTS
  - Stormwater capture
    - Expanding green areas and controlling flow
  - High quality water
    - Sea level rise- mobilization of legacy mercury
  - Storage capacity- Reliable source
    - Almaden Lake
  - Groundwater wells
    - SF- deeper
  - Recycled water- Advanced water treatment
    - “Fit for purpose”
    - City of San Francisco Recycled Water Ordinance

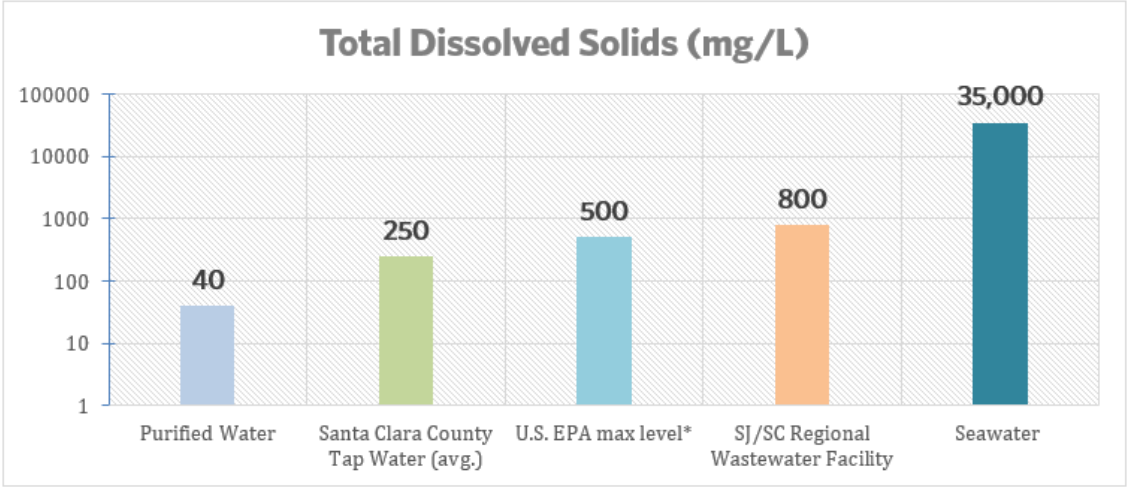
Recycled water supply

- Water and wastewater supplies
  - United States
    - 7-8% water recycled
    - California leads with planned reuse projects
  - Valley Water
    - By 2025, 10% of supply will be recycled water
  - SFPUC
    - By 2025, 2 MGD delivered

Recycled water quality

- Water quality standards for infectious disease
  - Non-potable reuses
    - No fecal coliform detected /ml over 7 day median
      unrestricted use
    - 200 fecal coliform detected/ml over 7 day median
      restricted use

- TDS standards



*U.S. Environmental Protection Agency secondary standard as recommended maximum TDS level for drinking water.

https://beheard.valleywater.org/purifiedwaterproject/widgets/34076/faqs

Water regulations

- Clean Water Act
  - Discharge into navigable waters
  - US EPA ECHO
  - Sewage

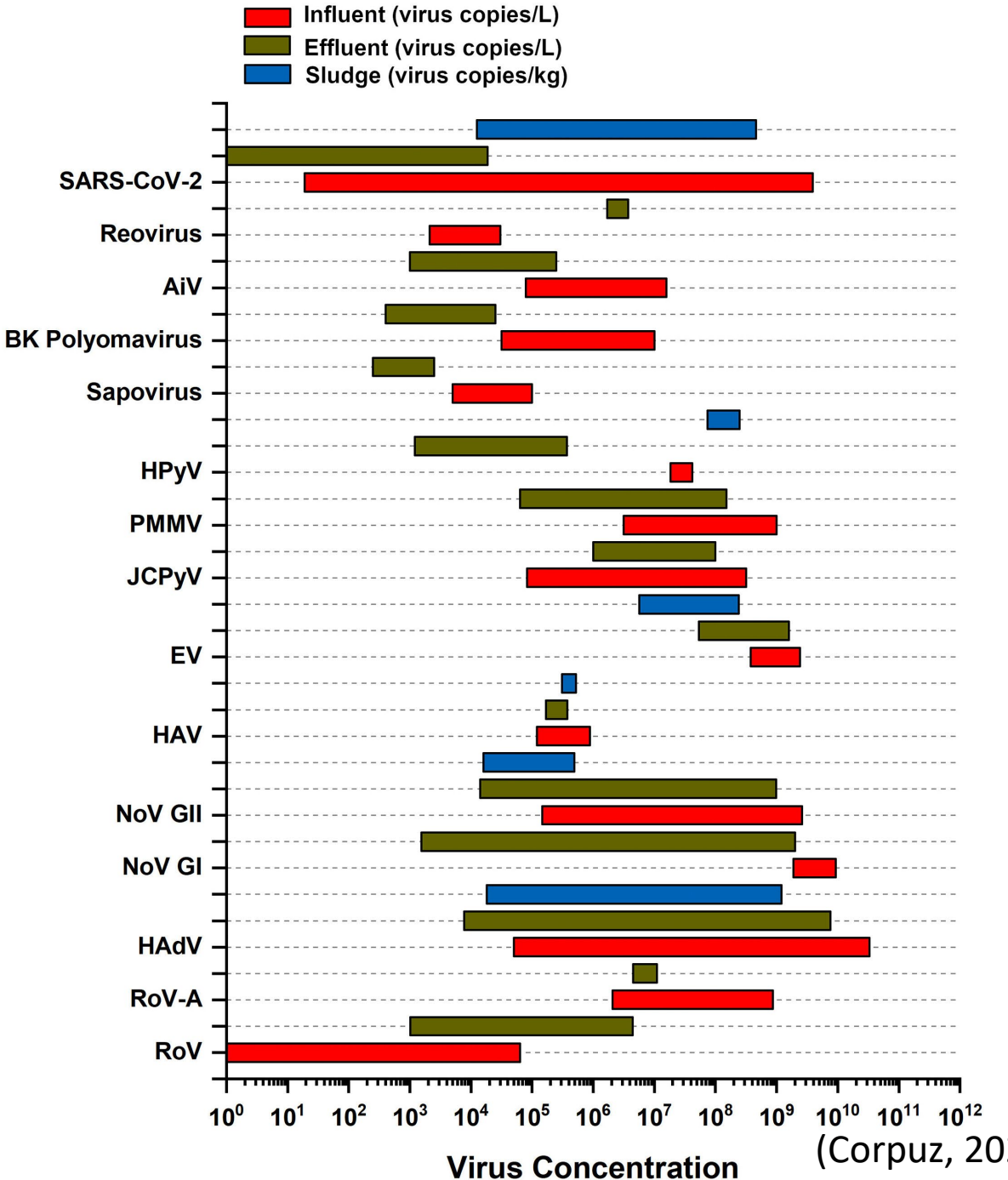
- Safe Drinking Water Act
  - Sets safe drinking water standards
  - MCL
    - Viruses = 99.99% removal/inactivation
  - California: reported via CCR
  - Enhanced Surface Water Treatment Rule- promulgated for states to set standards

- Title 22 California- Recycled water
  - 12-log reduction (99.9999999999%) of viruses for Indirect Potable Reuse
  - 1:10000 risk of viral infection

Viruses in Sewage

- AiV- Aichi virus
- PMMV- Pepper mild mottle virus
- EV – Enterovirus
- HAV- Hepatitis A
- NoV – Norovirus
- RoV- Rotavirus

Virus nucleic acid stable and present in effluent



(Corpuz, 2020)

US EPA direction

Viral surrogates- needed in water reuse as supplies come online

Viral surrogates	Virus/100 ml sewage	Genome
F specific bacteriophage	10^5 -10^7 PFU	+ssRNA or dsDNA
Somatic bacteriophage	10^5-10^6 PFU	dsDNA
CrAssphage	10^6-10^9 gc	dsDNA
Tomato mosaic virus	~10^6 gc	+ssRNA
PMMoV- Pepper mild mottle virus	10^6 gc	+ssRNA

Monitoring methods for viruses in water and wastewater

Method	Time to results	Infrastructure/reagents
US EPA 1615	>2 days	High
US EPA 1645	>2 days	High
Cell culture	>1 day	High
ICC-PCR	>1 day	High
RT-qPCR	>2 hours	Moderate
PCR	>2 hours	Low
Riboswitch	15-60 minutes	Low

Biosecurity preparedness findings

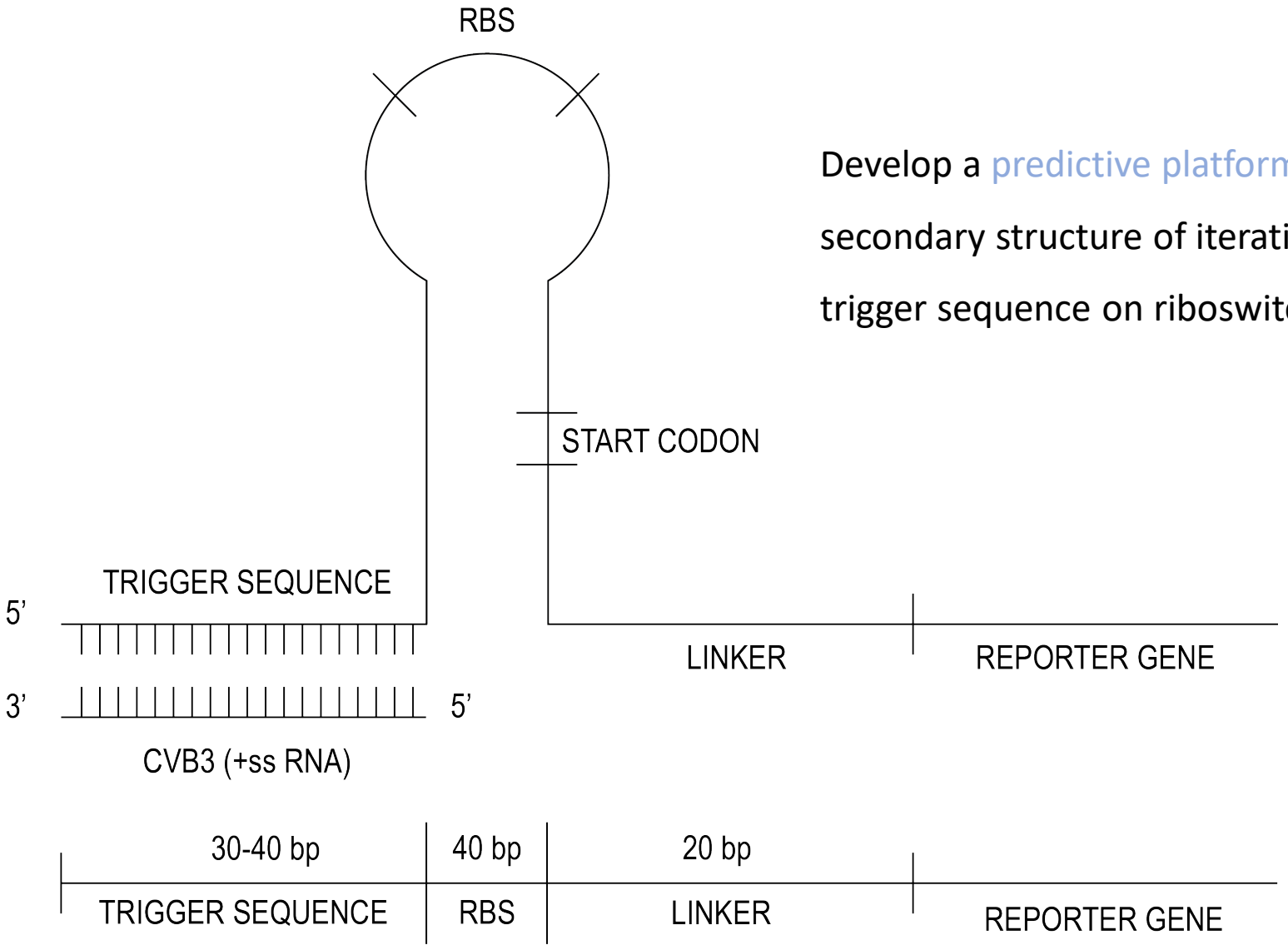
- GHSI findings
  - No county fully prepared
  - ~7% prevention
  - ~19% detection
  - ~5% rapid response

- Factors for preparedness
  - Prevention
  - Detection
  - Rapid Response
  - Health system
  - International norms
  - Risk environment

Alternative Detection

- Need for New Technology without the Technology

Riboswitch detection of +ssRNA viruses



Develop a predictive platform to assess the secondary structure of iterative changes to the trigger sequence on riboswitches

Approach

- Coxsackievirus B3= model organism
  - Trigger sequence development
- Published riboswitch sequences
- Three open access platforms
  - Modified with alternative code specific for nucleic acid/nucleic acid interactions
  - Prediction of Minimum Free energy
- Interface development

Ribologic

- Predicts the correct secondary structure with associated ligand
- Limitation- does not predict secondary structure when a random point mutation occurs
- Limitation- Does not always produce the exact base pairing when point mutations or changes in the sequence are produced.
  - Change one base and the exact base pairing will not be produced.
- Salinity- could possibly be incorporated if base pairing is resolved

Wu et.al., 2019 *ACS Synthetic Biology* **2019** *8* (8), 1838-1846

Vienna RNA

- Predicts secondary structure of nucleic acids
- Code is C++- We are working in Python.
- Does salinity (added 2023)

http://rna.tbi.univie.ac.at/

Nu Pack

- Open source code
- Predicts correct secondary structure with introduction of mutations
- Easy to use but not represent different different levels of salinity- Only 1M NaCl
- Output variables for analysis
  - Single strandedness score
    - How likely the secondary structure available for binding or trigger sequence for annealing
      - 1 all unpaired bp;
      - 0 all paired bp (secondary structure)
  - Normalized ensemble defect– how far secondary structure away from true secondary structure
    - low number for high ensemble (true secondary structure; proper bp matching)
  - Switch minimum free energy- promoter to gene to express
    - <-9.5 free energy for stable structure
- Results thus far
  - 70% predicted accuracy for the secondary structure
    - Develop suite of riboswitch/trigger sequence combination and test in the lab
  - Code developed to reduce time when iterative trigger sequences are introduced
    - 32 hours to 4 seconds

https://www.nupack.org/

Example of output

Riboswitch index	Trigger index	Riboswitch	30bp Trigger switch CVB3	Riboswitch single-strand	Trigger single_strand	Riboswitch ensemble	Trigger ensemble	Riboswitch/trigger /target ensemble	Free Energy score ensemble (MFE)
0	0	GGGGAAGAGUCUAUUG AGCUAGUUGGUAGUCCU AUACAGAAACAGAGGAG AUAUAGGAUGACCAACT AGAACCUGGCGGCAGCG CAAAAGAUGCGUAAA	AGGACTACCAACTAGCTCAATAGACTCTTC	0.75	0.55	0.17	0.29	0.13	-66.9
0	1	GGGGUUUCAUUUUAUU CCUAUACUGGCUGCUUA AUACAGAAACAGAGGAG AUAUTAAAUGGCCAGTAT AAACCUGGCGGCAGCGC AAAAGAUGCGUAAA	TAAGCAGCCAGTATAGGAATAAAATGAAAC	0.48	0.69	0.20	0.12	0.13	-43.5
0	2	GGGGAAGAGUCUAUUG AGCUAGUUGGUAGUCCU AUACAGAAACAGAGGAG AUAUAGGAUGACCAACT AGAACCUGGCGGCAGCG CAAAAGAUGCGUAAA	ATAAGCAGCCAGTATAGGAATAAAATGAAA	0.69	0.67	0.17	0.12	0.14	-43.5
0	3	GGGGAAGAGUCUAUUG AGCUAGUUGGUAGUCCU AUACAGAAACAGAGGAG AUAUAGGAUGACCAACT AGAACCUGGCGGCAGCG CAAAAGAUGCGUAAA	GTATAAACCCAACAAAGGGATATATAATAG	0.69	0.38	0.21	0.40	0.18	-38.3

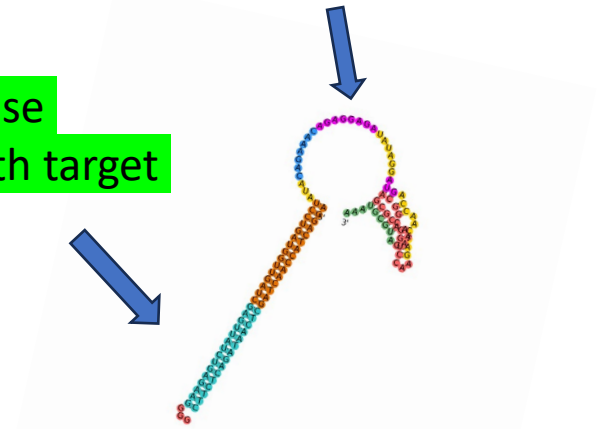
When we change the trigger sequence, the parameters for single-strandedness, ensemble defects and free energy vary

Changing the trigger sequence changes the structure and stability of the Riboswitch/trigger sequence and the Riboswitch/trigger sequence complex.

Riboswitch structure optimization

RBS- available for expression

Correct base pairing with target (CVB3)

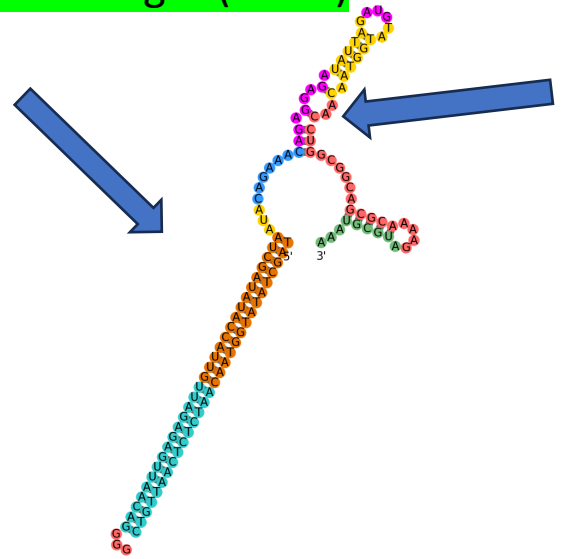


Complex 0

- Complex ensemble (Trigger sequence and riboswitch to target sequence)

Correct base pairing with target (CVB3)

RBS incorporated in secondary structure

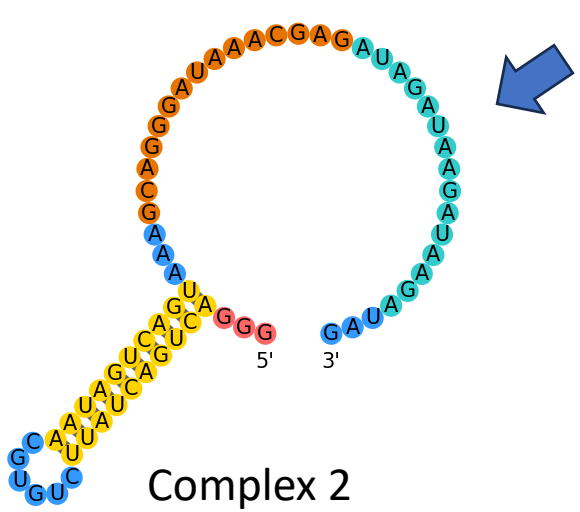


Complex 3-CVB3

Cyan + Orange + Red= trigger sequence
Cyan + Orange + Blue= target sequence (CVB3)
Magenta + Blue = Ribosomal Binding Site
Magenta= start codon
Green= first 10 nt of reporter gene
Yellow= filler sequence

Riboswitch structure optimization platform

- Single strandedness for trigger sequence-
  - Probability of the sequence free to anneal



Complex 2

Free to anneal to target sequence

Cyan + Orange= trigger sequence

Magenta + Blue = RBS

Magenta= start codon

Green= first 9 nt of reporter gene

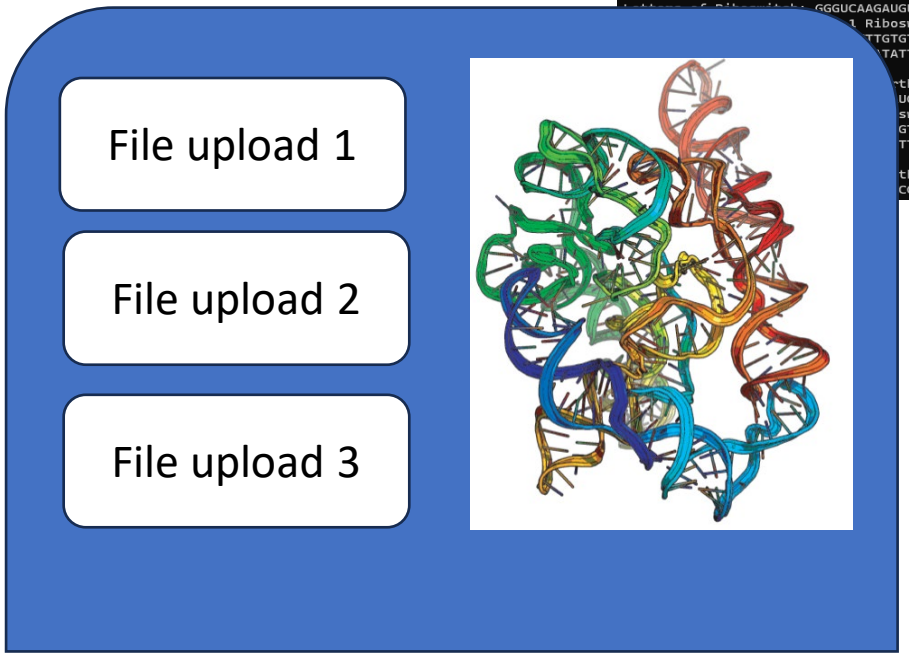
Yellow= filler sequence

Data interface

- Developed script:
  - Quick categorization and storage of data of RNA sequences uploaded through Excel Files

- Next step: Develop script for:
  - Input- Uploading files with target, trigger sequence, linker, and riboswitch sequence
  - Output- free energy calculations

```
Primer row: 82 Linker Row: 1 Riboswitch Row: 71
Target Sequence: AATAATTAATATTTGTGTAGTGAATGATGGAATTG
Trigger Sequence: TTTGAAGAGGCTATATTTTCCAAGTATATAGGAAA
Linker: AACCTGGCGGCAGCGCAAAAG
Description of Riboswitch: 144 orthoganal first gen
Letters of Riboswitch: GGGAUGGAGAUUGAUUAUGAUUGGAUGUGCUUAAACAGAGGAGAUAAGCAAUGCCAAUCAUAAACCUGGCGGCAGCGCAAAAGAUGCGUAAA
Primer row: 82 Linker Row: 1 Riboswitch Row: 72
Target Sequence: AATAATTAATATTTGTGTAGTGAATGATGGAATTG
Trigger Sequence: TTTGAAGAGGCTATATTTTCCAAGTATATAGGAAA
Linker: AACCTGGCGGCAGCGCAAAAG
Description of Riboswitch: 144 orthoganal first gen
Letters of Riboswitch: GGGAGUAAGAAUUGUGAUAAAGUAAUGUGCGUGAACAGAGGAGACACGCAAUGUACUUUAUCAACCUGGCGGCAGCGCAAAAGAUGCGUAAA
Primer row: 82 Linker Row: 1 Riboswitch Row: 73
Target Sequence: AATAATTAATATTTGTGTAGTGAATGATGGAATTG
Trigger Sequence: TTTGAAGAGGCTATATTTTCCAAGTATATAGGAAA
Linker: AACCTGGCGGCAGCGCAAAAG
Description of Riboswitch: 144 orthoganal first gen
Letters of Riboswitch: GGGUAAGAUGAUAAGAGUAUAGAUAUGUUGAUGGACAGAGGAGACAUCAAAUGAUCUAUACUAACCUGGCGGCAGCGCAAAAGAUGCGUAAA
Primer row: 82 Linker Row: 1 Riboswitch Row: 74
Target Sequence: AATAATTAATATTTGTGTAGTGAATGATGGAATTG
Trigger Sequence: TTTGAAGAGGCTATATTTTCCAAGTATATAGGAAA
Linker: AACCTGGCGGCAGCGCAAAAG
Description of Riboswitch: 144 orthoganal first gen
Letters of Riboswitch: GGGAUAUGAUGAUGAACCUUUGAUAUGGAUUUGGACAGAGGAGACAAAUCAUGAUCAAAGGUAACCUGGCGGCAGCGCAAAAGAUGCGUAAA
Primer row: 82 Linker Row: 1 Riboswitch Row: 75
Target Sequence: AATAATTAATATTTGTGTAGTGAATGATGGAATTG
Trigger Sequence: TTTGAAGAGGCTATATTTTCCAAGTATATAGGAAA
Linker: AACCTGGCGGCAGCGCAAAAG
Description of Riboswitch: 144 orthoganal first gen
Letters of Riboswitch: GGGAAUAGUAAAUGUCGAUGUGCUGUGUAGUUAAACAGAGGAGAUAACUAAUGACACAUCGAACCUGGCGGCAGCGCAAAAGAUGCGUAAA
Primer row: 82 Linker Row: 1 Riboswitch Row: 76
Target Sequence: AATAATTAATATTTGTGTAGTGAATGATGGAATTG
Trigger Sequence: TTTGAAGAGGCTATATTTTCCAAGTATATAGGAAA
Linker: AACCTGGCGGCAGCGCAAAAG
Description of Riboswitch: 144 orthoganal first gen
Letters of Riboswitch: GGGUCAAGAUGUAUAGAUUAGUGUAUGUUAGUGAACAGAGGAGACACUAAAUGACACUAAUCAACCUGGCGGCAGCGCAAAAGAUGCGUAAA
                                                    1 Riboswitch Row: 77
                                                    TTGTGTAGTGAATGATGGAATTG
                                                    TATTTTCCAAGTATATAGGAAA
                                                    rthoganal first gen
                                                    UGAGACCGAAUUGAAUGAUGUUAGACAGAGGAGAUAACAUAUGUCAAUUCGGAACCUGGCGGCAGCGCAAAAGAUGCGUAAA
                                                    switch Row: 78
                                                    GTAGTGAATGATGGAATTG
                                                    TTTTCCAAGTATATAGGAAA
                                                    thoganal first gen
                                                    CGAAAGCUAAGUGUAUGUGUCGGGACAGAGGAGACCGACAAUGACACUUAGCAACCUGGCGGCAGCGCAAAAGAUGCGUAAA
```



Courtesy of Jonathan Auyong

Conclusions

- Platform can predict proper base pairing.
- Predictive platform generates free energy prediction based on the riboswitch/trigger sequence.
- Iterative changes initially were at 32 hours
  - New code= predictions made in 4 seconds
- Free energy varies with iterative changes of the trigger sequence.