Supporting Information

Evidence of an Unusual Poly(A) RNA Signature Detected by High-throughput Chemical Mapping

Roger Wellington-Oguri, Eli Fisker, Mathew Zada, Michelle Wiley, Jill Townley, Eterna Players*

Eterna Massive Open Laboratory, Stanford University, B400 Beckman Center, Stanford, CA 94305

*Corresponding author, email: omei@stanford.edu

Table of Contents

Data selection details

The strategy for selecting data from RMBD (https://rmdb.stanford.edu/)

1. Identify those puzzles that had the most poly(A) sequences relevant to a specific question/hypothesis.
2. If there was more than one RMDB file (each representing one experimental run) containing data for those puzzles, choose the one with the best average signal-to-noise ratio.
3. Include all the relevant sequences in the analysis, irrespective of the signal-to-noise ratio for each sequence.

Figures 1 and 2: The goal was to identify the puzzles that had the most poly(A) subsequences of length 8 or more in the same base positions. Three puzzles stood out from the rest, but one of those (Anaconda About to Poop) contained long stretches of poly(U) as well as poly(A), resulting in ambiguity in which poly(A) sub sequences were single stranded. The two remaining puzzles were Triangle of Doom and Eli's Big Hairpin, which occurred in the same data set. The RMDB file with the highest average signal-to-noise ratio was RMDB Accession ID ETERNA_R00_0000.

Figure 3: The goal was to find a set of puzzles in one experimental data set with a large number of poly(A) subsequences, with lengths ranging from 5 or 6 to at least 11, regardless of whether they all occurred at the same base positions. Only one experimental data set fit that criteria -- a series of small puzzles that had been measured previously with gel electrophoresis experiments and were used to check the validity of the high throughput experimental process. These small puzzles are identified by their title starting with "Project :", and are included in the data set with RMDB Accession ID ETERNA_R00_0000. A special note on these puzzles: in this initial experiment, all puzzles were padded out to a common length using poly(A) on the 5' end of the segment of the RNA that underwent chemical modification. It was observed, however, that the data for the padding molecules were extremely variable, probably because of design-specific interactions with the trailing CAAAAGAAACAACAACAACAAC sequence. In subsequent experiments, any padding necessitated by the preparation for chemical mapping was removed prior to mapping. For figure 3, the data for the padding sequences were not included.

Figure 4: The goal was to find puzzles with a significant number of subsequences matching the regular expression [^A]A{*n,n*}.A{*n,n*}[^A] with *n* as large as possible. The puzzle Intrinsical – Frequency 8, best satisfied that criteria, with *n* equal to 8, with data in RMDB Accession ID ETERNA_R80_0000. The signal-to-noise ratio for this data set in general was low, and so sequences with signal-to-noise ratios of less than 1.0 were filtered out. In addition, some sequences had poly(A) stretches long enough that the small sequencer reads made it impossible to assign independent reaction values for each base. These sequences were also filtered out.

All data was generated as part of a large Eterna project called the Cloud Lab, which began in 2013 and was the first high throughput lab, running thousands of sequences in one experiment. The Das lab first ran a pilot round with many puzzles but only a few sequences tested. From then on, each round typically contained 20

puzzles. Each player could submit up to three sequences per puzzle, with those to be synthesized and measured selected by player vote.

All experiments were conducted at the Das Lab following the laboratory protocols described in "Standardization of RNA Chemical Mapping Experiments", Biochemistry, 2014, 53 (19), 3063-3065.

Eterna images

Image to the left shows the design sequences, image to the right shows the lab results for the sequence. SHAPE data is displayed in Eterna with the colors blue and yellow. Blue bases indicate pairing and yellow bases are unpaired.

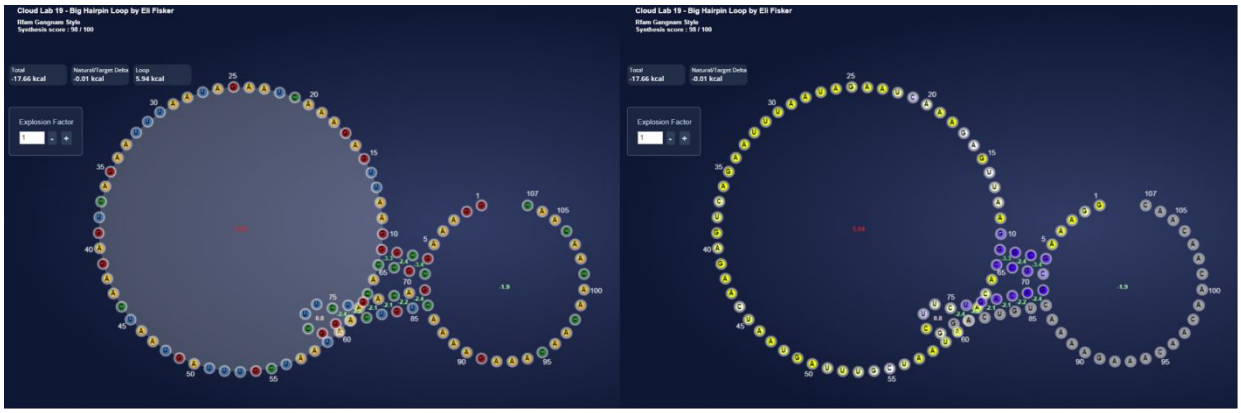

Figure S1. Triangle of Doom with mixed bases in loop, no unusual SHAPE data.

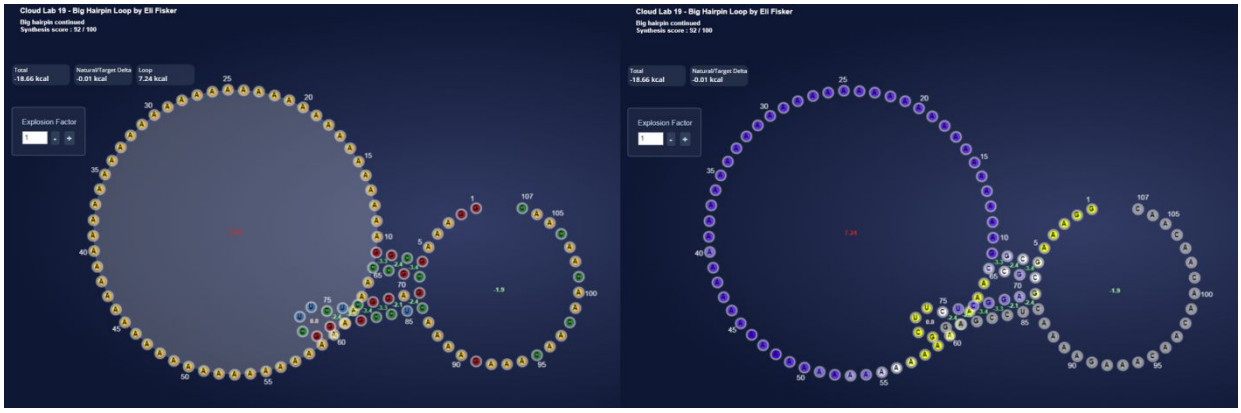

Figure S2. Triangle of Doom with poly(A) loop bases displaying unusual SHAPE data.

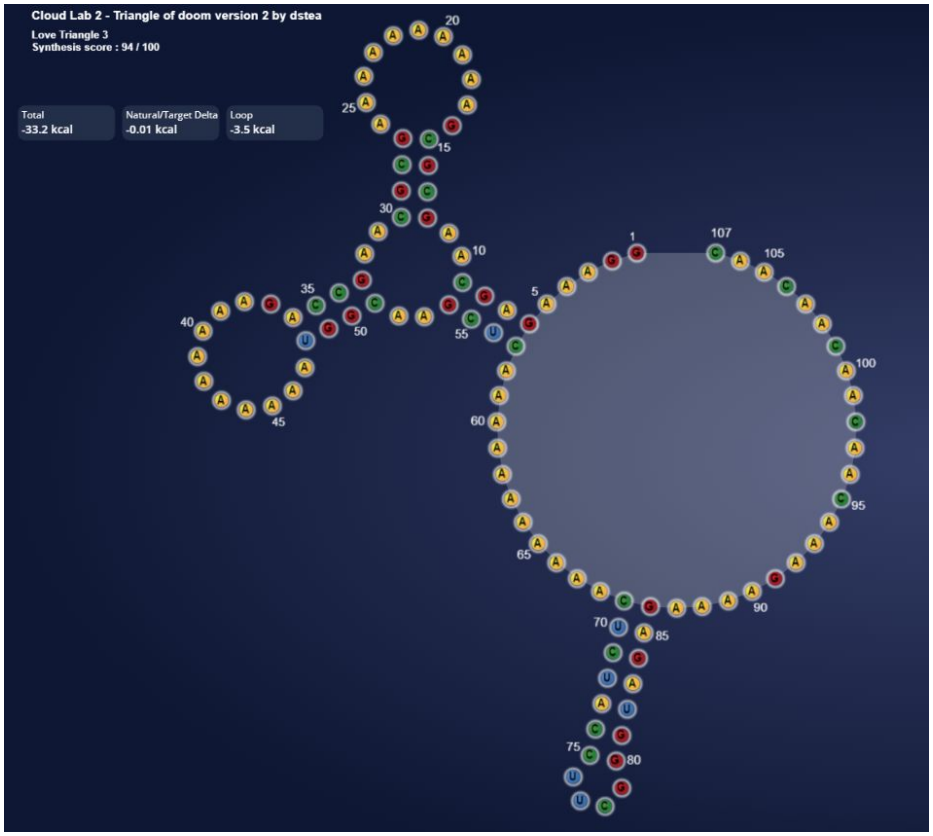

Figure S3. Example player solution for Triangle of Doom puzzle.

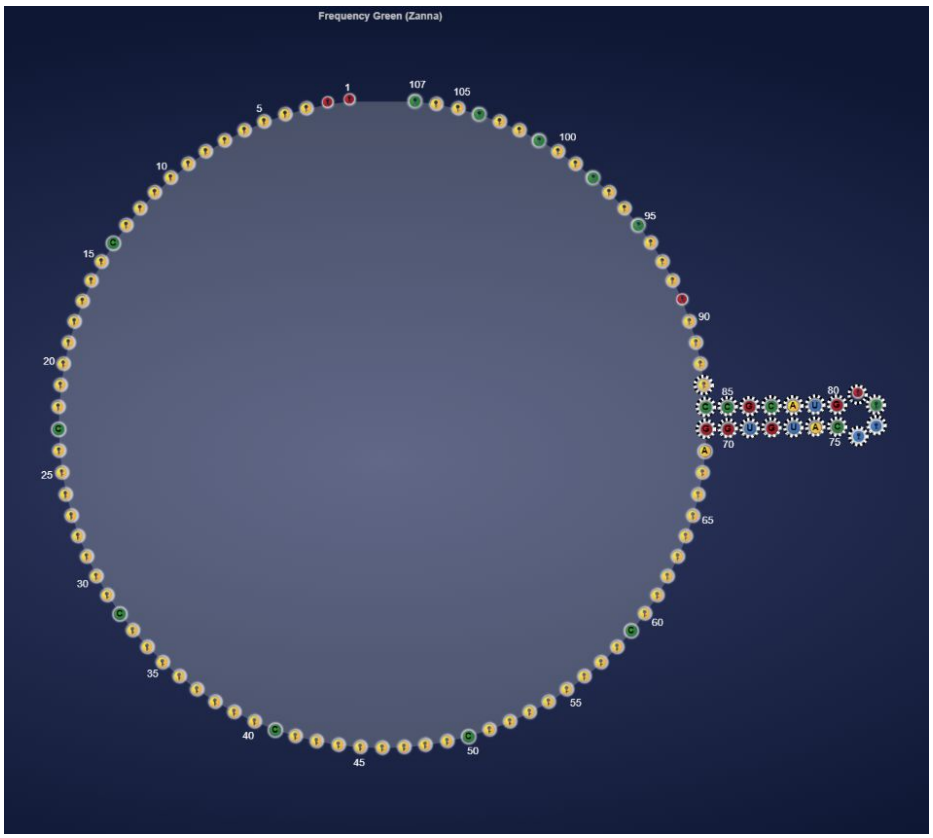

Figure S4. Example player solution for Intrinsical - Frequency 8 puzzle.

Poly(A) models by players

A reviewer suggested we enclose models of what players had speculated a poly(A) structure might look like.

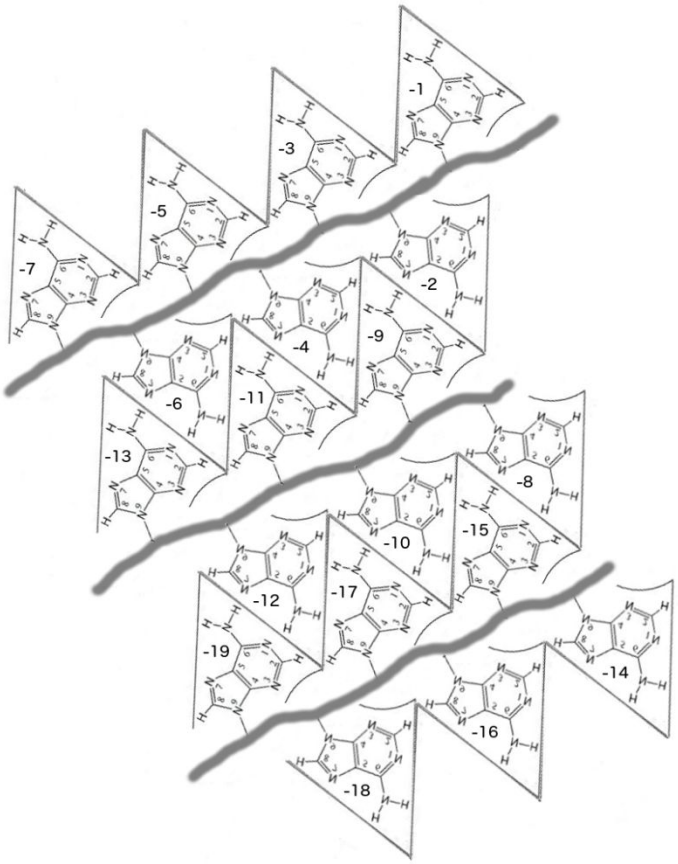

Figure S5. A 2016 poly(A) model created by player Omei, displayed flat.

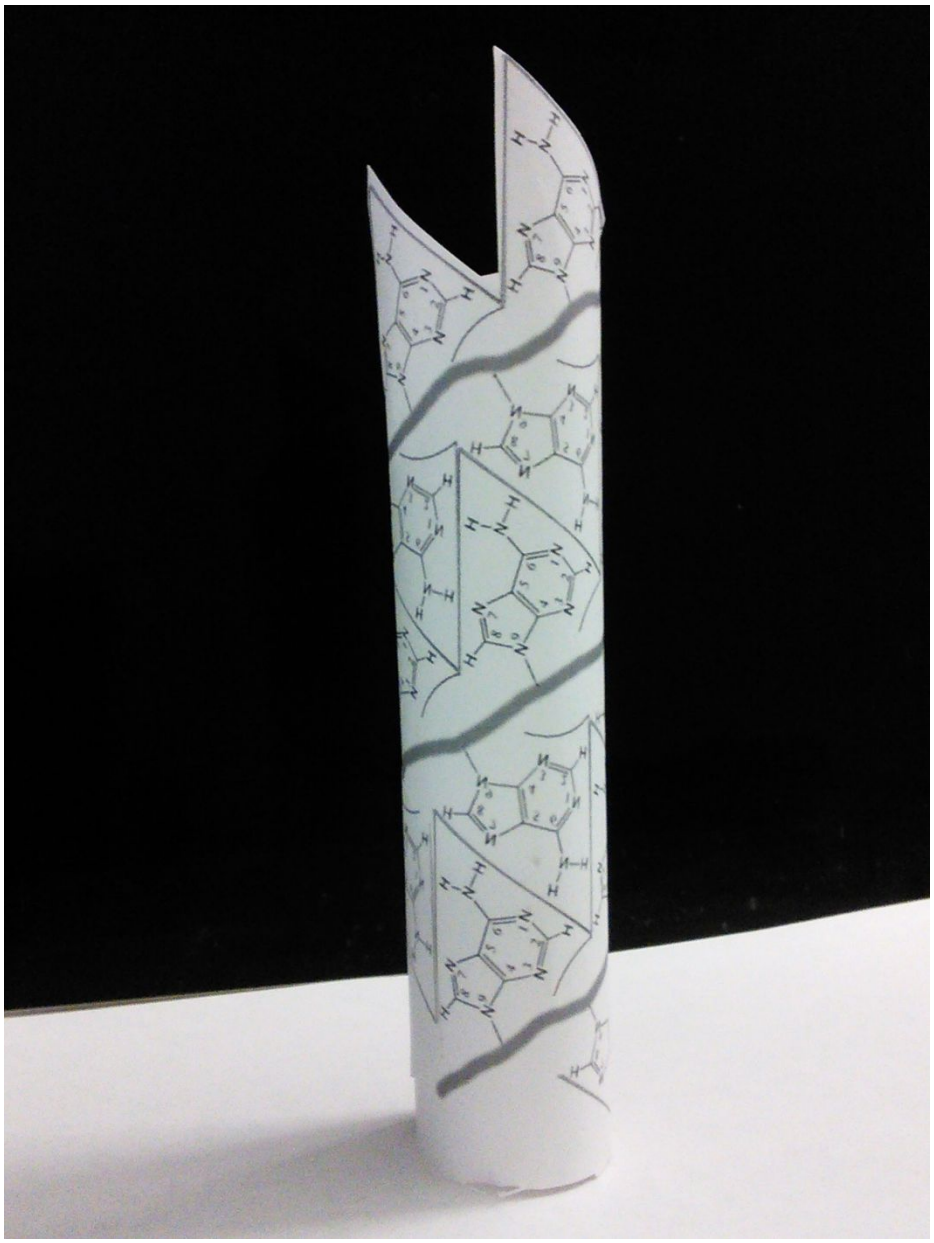

Figure S6. A 2016 poly(A) model created by player Omei, rolled up.

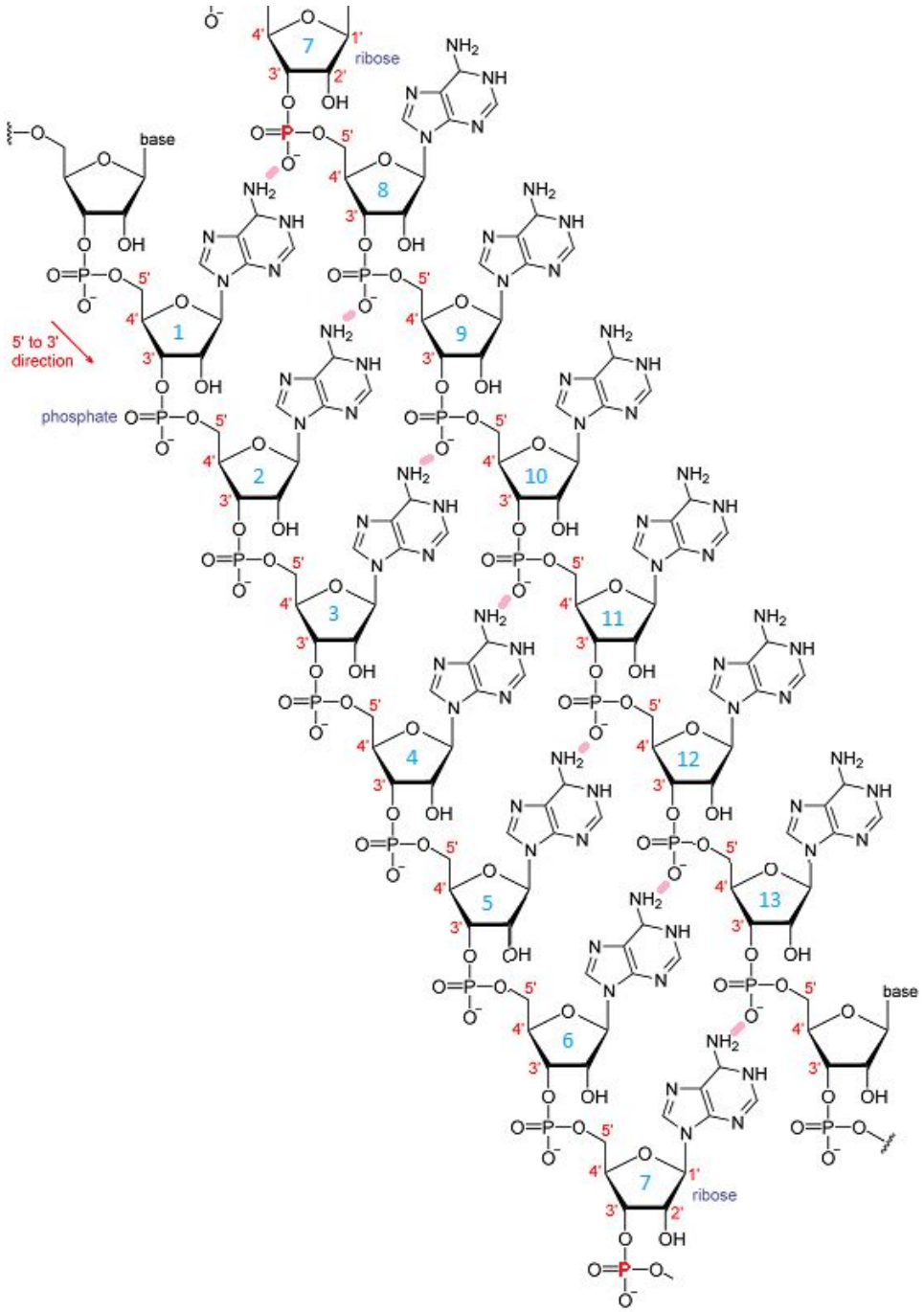

Figure S7. A 2018 poly(A) model created by player jandersonlee, displayed flat.

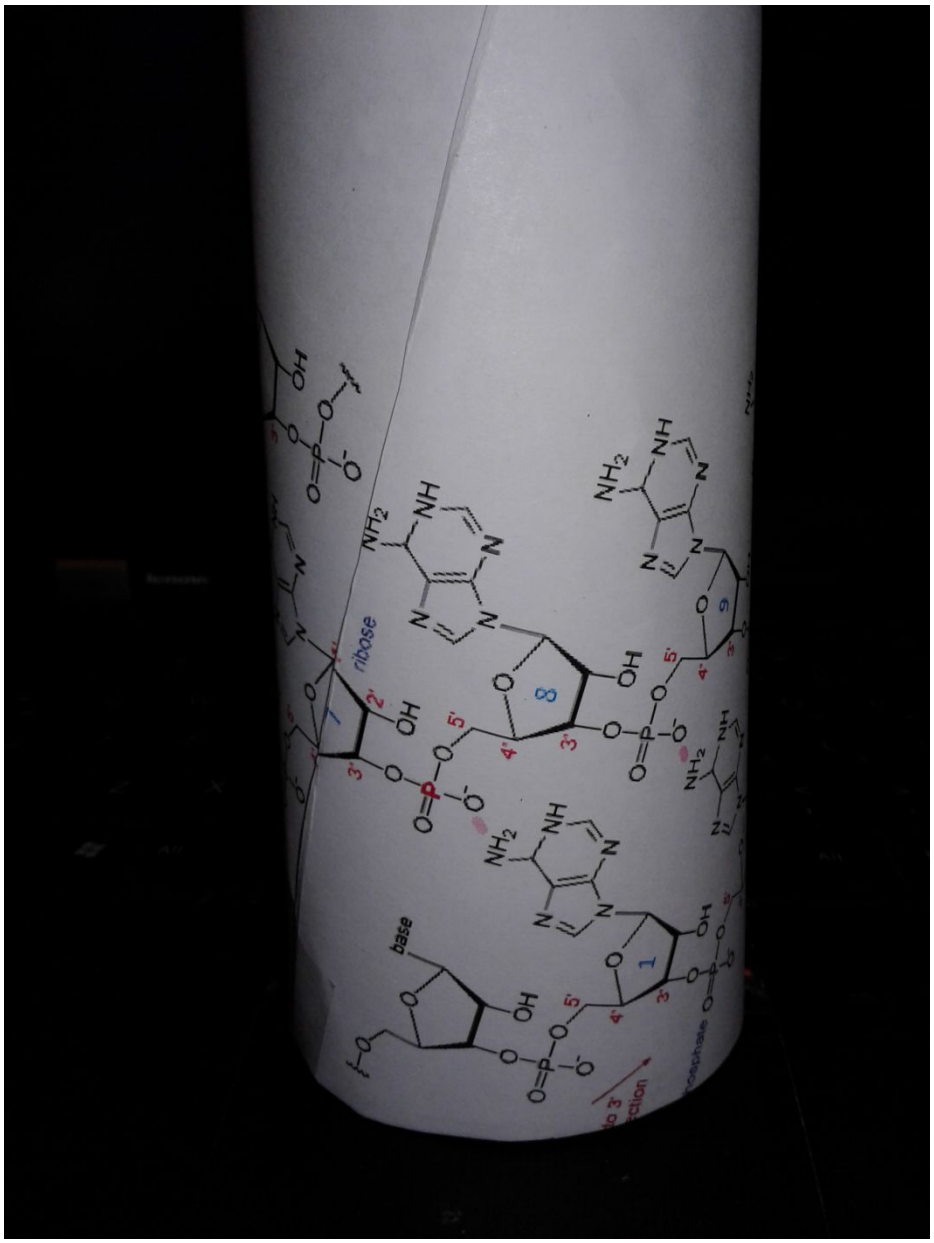

Figure S8. A 2018 poly(A) model created by player jandersonlee, rolled up.

Eterna players contributing to the paper

The following lists the 174 Eterna players (with design sequence counts) who contributed to the paper by submitting sequences for the puzzles used in the analysis.

```
  1 555shaun
  6 77Tennifry
  1 9FireStar
  2 aakspuicomterixx
  1 aberg011
  1 acetom14
  3 Akamu
 88 akhyatt
  1 akros
  1 alice21
  1 aliciaS
  3 AndrewKae
  2 andrewsokolow
  1 anwyn7
  1 arjunadhikari
  4 armin
  3 arxidia
  1 ATPro
  2 awegwert
  3 bigcheese
 25 Billy Reuben
  1 bill zhu
 10 boganis
  1 brambleboy
322 Brourd
  7 cataway
  4 Chesterfield
  1 chrislaser
  1 clollin
  6 c-quence
  1 d1abet1c
  3 daffy
  2 Daniel B.
  3 Darkfire47
  2 Darksite
  1 daskalska
  1 dave2045
  5 davidpat
  3 Deedie
 13 Dennis9600
  6 DHammond
  3 Dogs_Like_Cake
  4 DPope
  4 dw.thewilliams33
  1 Dysprosium
 12 Edward Lane
  3 Edward_Lane
  1 EFER
100 Eli Fisker
  1 EpicShorts
 10 eternacac
```

```
  2 evmasuta
  3 faf13
  2 feldbaum
  1 FleurDA
  2 fluffy3
 43 Frater Wroth
  2 freddydog
 12 Freywa
  2 Fullsail
  1 fwizzybee42
  1 Gagarin-Brat
  7 garydfisher
  1 Genghis Jones
  2 Gilden
  2 goerch
  1 hbovis
 16 hoglahoo
 11 Homebrew
  3 hotcreek
  3 huskerdad3
 15 Hyphema
  6 jal
  1 James Francisco
 62 jandersonlee
  9 janelle
 11 janetmason
  1 JCUAce
  1 jeehyung
  9 JerryP70
 23 Jieux
  1 jkoprivsek
  1 jmason
  2 jmf028
  5 jnicol
  3 joeK2_45
  7 Joshua Weitzman
 12 JR
  5 jruaya
  1 JTBones
  1 katica
  7 kcabral28
  1 kforce214
  3 khar
  1 Kminttech
  2 kneeonlite
  1 Krobar
  1 Kromst
  1 lanceman
  2 LazyBug
  2 LiquidOvar
 15 lroppy
  1 madmax337
  5 Malcolm
226 mat747
  1 matosfran
  2 Matt Indykiewicz
  7 Max Goff
```

```
  1 Mearth
  6 Meechl
 45 merryskies
  1 mistressgieddbrytta
  2 naprat
 15 nascarnut
  1 NE5480
  3 necet
  6 nelulon
  1 Nick Keller
  1 Noah Katcher
  1 nodnerb93
  2 okrazerback
  6 Omei
327 oolong
  1 paloma
 34 paramodic
  1 Pegasus1207
  3 Peter Stampfli
  1 player4596
  1 psychemist
  1 ptrw
  1 qball
126 Quasispecies
  4 Quxwozing
  1 RachelRose
  6 randl
  3 redsoxwy
  5 RedSpah
  1 rhys
 12 ribonucleic
  2 rlmarchal
  6 rnjensen45
 12 robu-san
  1 Sadler
  1 shaheenj
  1 Sherlock
  4 SpaceFolder
  1 sriopelle
 12 starryjess
  3 steven123505
 52 stevetclark
  1 Stormy
  9 Tesla'sDisciple
  1 TheBoogeyMan247
  1 The Dev1
  5 theravin
  1 TheSkedaddling
  1 timl
  1 tjlampo
  2 TL-TBO
  1 tmhsu
  3 tomoedachii
  2 TreyBrown
  2 tryoon93
  2 UNCZack
  1 ValentinPerm
```

```
13 ViennaUCT
 1 Vladimir Semenov
17 waterontheMoon
 8 wilf
17 wisdave
 1 wtoy
 1 xenosapien
 1 xiaofanjin
 3 Zanna
```