

Using multiplayer online videogames for structure prediction and design

Meera Srinivasan
Jesse Min
Seth Hildick-Smith

Table of Contents

1. Protein Structure Prediction

2. RNA Design

3. Multiplayer online collaborative platforms

Predicting protein structures with a multiplayer online game

Seth Cooper, Firas Khatib, Adrien Treuille, Janos Barbero, Jeehyung Lee,
Michael Beenen, Andrew Leaver-Fay, David Baker, Zoran Popovic &
Foldit players

Presented by Meera Srinivasan

Protein Structure Prediction: How do you predict the structure of a protein from its amino acid sequence?

1. **Homology or template-based modeling:** given a particular protein sequence, search for a homolog
2. ***Ab initio* structure prediction:** when no homologs, combine deterministic & stochastic approaches to predict structure

Introduction: Rosetta for Protein Structure Prediction

Searches conformational space using fragment assembly

- Deterministically finds various protein fragments
- Applies Monte Carlo methods to position fragments
- Evaluates empirically-derived knowledge-based Rosetta energy function to do so



Refines structure

- Using Monte Carlo methods and Rosetta all atom energy function

This cycle is repeated many times.

Introduction: Rosetta's Components

- Deterministic: fragment finding and assembly
- Stochastic: energy minimization; structure perturbation & refinement

Challenges with Rosetta

- Free energy landscape is large
- Suboptimal conformations pursued in effort to minimize energy

Experimental Hypothesis:

What if we replaced stochastic components with human decision making and retained deterministic components?

FoldIt: Game Development for Protein Structure Prediction

The screenshot displays the FoldIt game interface. At the top, it shows 'Rank: 317' and 'Score: 2534' for a 'Soloist' in 'Beginner Puzzle 8 (<150): Fruit Fly'. The main area features a 3D protein structure with various colored segments and arrows (1-12) pointing to specific parts. On the right, there are two leaderboards: 'Group Competition' and 'Soloist Competition'. The bottom of the screen has a control panel with various actions like 'Shake Sidechains', 'Wiggle All', 'Freeze Protein', etc., and a chat window on the far right.

Group Competition

#	Group Name	Score
1	Rice Biochemistry	9174
2	Team Commonwealth	9168
3	Ukraine	9088
4	Team Canada	9085
5	Firebird BioChem	9073
6	SETI.Germany	9030
7	Boinc.be	9001

Soloist Competition

#	Player Name	Current	Best
1	Mike Crunching for Physics	-	9242
2	weltzen	-	9235
3	ys719	-	9222
4	jmarfc	-	9211
5	issin_karpus	-	9186
6	JINXter	-	9185
7	ob.eric	-	9183

Control Panel

- Shake Sidechains
- Wiggle All
- Wiggle Backbone
- Wiggle Sidechains
- Freeze Protein
- Remove Bands
- Disable Bands
- Align Guide
- Reset Structures
- Reset Puzzle
- Help
- Glossary

Chat Window

- Chat - Group (auto show)
- Chat - Puzzle (auto show)
- Chat - Global (auto show)
- Notifications (auto show)

Image: (Seth et. al.)

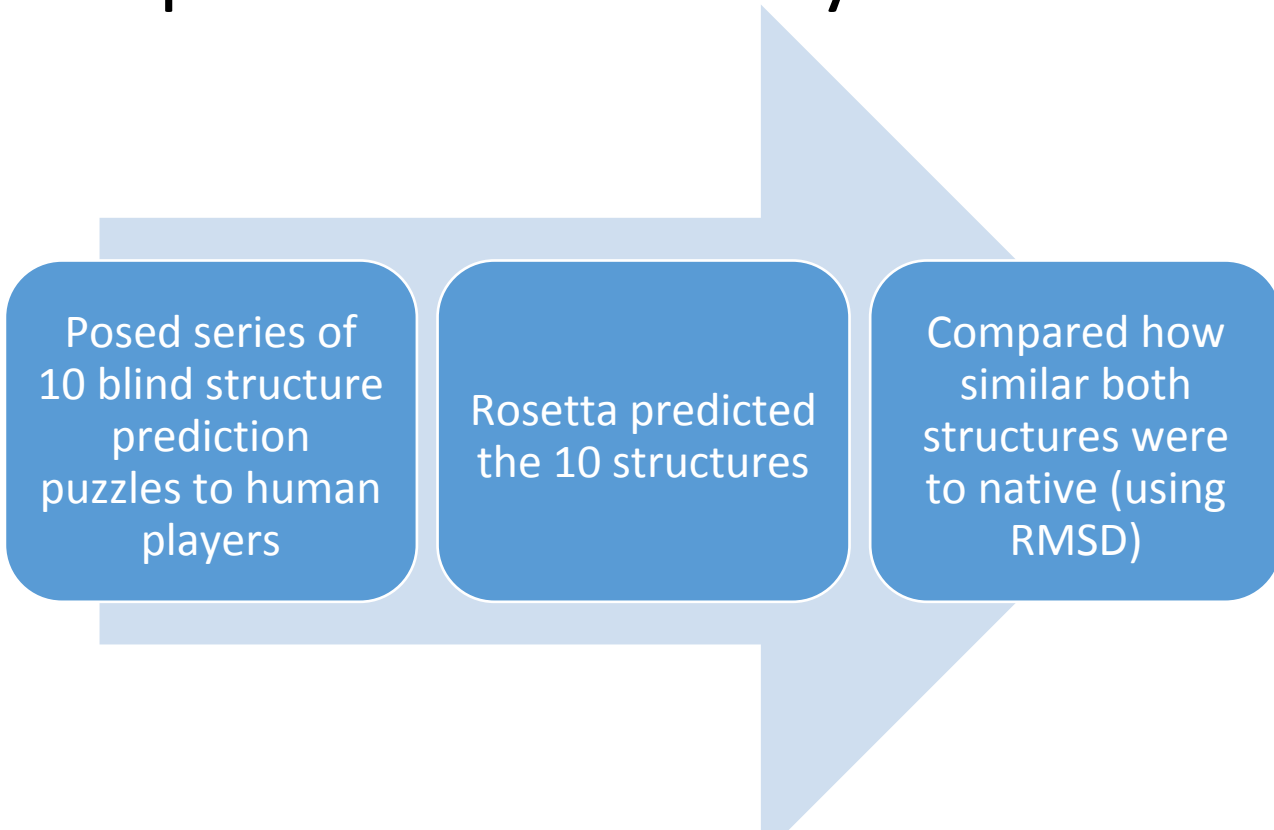
FoldIt: Game Development for Protein Structure Prediction

- User tools reflect deterministic Rosetta algorithms
 - Meant to minimize local energy
 - Scoring: computes energy same way Rosetta does, but multiplies by 100; goal is to get higher score.
- Player intentions parallel stochastic processes
 - Searches conformational space instead of computer doing so
- Meant to be approachable for players from all backgrounds

FoldIt: Game Development for Protein Structure Prediction

- <https://www.youtube.com/watch?v=IGYJyur4FUA>
- 1.40-1.55 important

Experimental Methods: Comparing human structure prediction accuracy to Rosetta's



Experimental Methods: RMSD (Root mean squared deviation)

- Measure of similarity between two structures for a given protein
- Must align the two structures being compared before measuring RMSD

Results: Human intuition is very effective

Table 1 | Blind data set

Puzzle ID	Foldit C α r.m.s.d.	Rebuild and refine C α r.m.s.d.	Native	Method	Number of residues	Figure(s)
986875	1.4	4.5	2kpo	NMR	99	3a–c, Supplementary 4
986698	1.8	3.7	2kky	NMR	102	3d, e
986836	5.7	6.6	3epu	X-ray	136	2c, Supplementary 6d
987088	3.5	4.3	2kpt	NMR	116	2a, b, Supplementary 6a, b
987162	4.5	5.2	3lur	X-ray	158	Supplementary 6c
987076	3.3	3.5	2kpm	NMR	81	2e, Supplementary 5c
986629	3.5	3.3	2kk1	NMR	135	Supplementary 5b
987145	2.6	2.3	3nuf	X-ray	105	2d, Supplementary 5a
986844	6.9	5.8	2ki0	NMR	36	Supplementary 10a
986961	10.6	5.7	2knr	NMR	118	Supplementary 10b

A listing of all the Foldit puzzles run in the blind data set. A C α r.m.s.d. comparison to the native structure is given between the best-scoring model produced by Foldit players and the best-scoring model produced by the Rosetta rebuild and refine protocol, given the same starting model(s). Solutions considerably better with one method than the other are indicated in bold. The solved structures (which were released after each puzzle ended) are represented by their Protein Data Bank (PDB) codes. Results from these Foldit puzzles can be accessed on the Foldit website by replacing ID with the corresponding Foldit puzzle ID in <http://fold.it/portal/node/ID>. 2kky, 2kpt, 2kpm, 2kk1 and 2knr were taken from the CASD-NMR experiment¹⁰. 2kpo was provided by N. Koga and R. Koga. 2ki0 and 3epu were found by searching for unreleased structures on the PDB website (<http://www.rcsb.org/pdb/search/searchStatus.do>). 3lur and 3nuf were provided by the Joint Center for Structural Genomics (JCSG). The location of figures containing results for each puzzle are provided in the last column.

*Image: (Seth et. al.
2010)*

Results: Scenarios in which humans performed better

- Cases with substantial backbone remodeling so exposed hydrophobic residues faced inwards

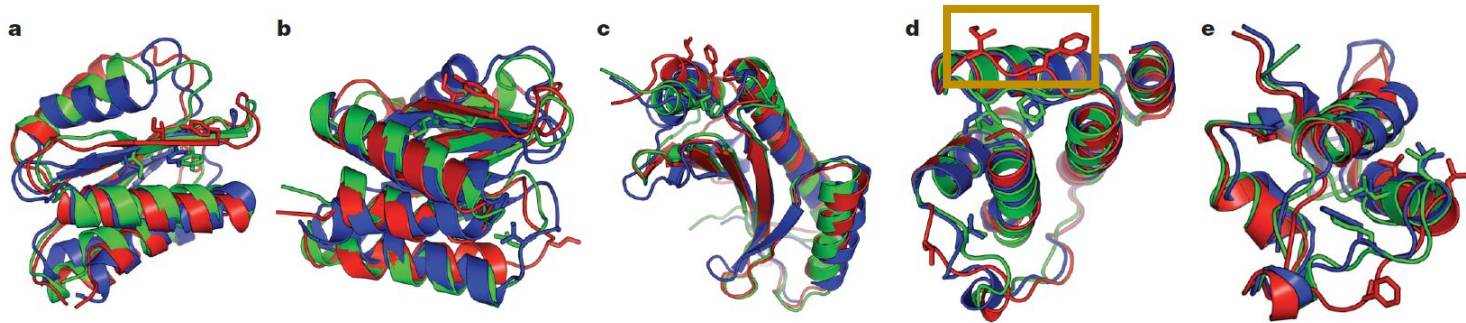


Figure 2 | Structure prediction problems solved by Foldit players.

Examples of blind structure prediction problems in which players were successfully able to improve structures. Native structures are shown in blue, starting puzzles in red, and top-scoring Foldit predictions in green. **a**, The red starting puzzle had a register shift and the top-scoring green Foldit prediction correctly flips and slides the β -strand. **b**, On the same structure as above, Foldit players correctly buried an exposed isoleucine residue in the loop on the bottom right by remodelling the loop backbone. **c**, The

top-scoring Foldit prediction correctly rotated an entire helix that was misplaced in the starting puzzle. **d**, The starting puzzle had an exposed isoleucine and phenylalanine on the top, as well as an exposed valine on the bottom left. The top-scoring Foldit prediction was able to correctly bury these exposed hydrophobic residues. **e**, Another successful Foldit helix rotation along with a remodelled loop that correctly buries an exposed phenylalanine. Images were produced using PyMOL software¹¹.

Results: Humans more likely to “take risks” sampling than Rosetta

- Rosetta’s stochastic Monte Carlo methods aim to **minimize energy** while humans are more likely to **create high energy intermediates**, knowing they will lead to **lower energy structures**.
- Humans can distinguish which starting point is most useful.
- Human players restructure significantly to improve hydrophobic burial and hydrogen bond quality.

Foldit's social & psychological dimensions

- Complexity, creativity, & collaboration in search process
- Inter-group competition
- Retaining engagement & “thrill” factor
- Motivation & rewards structure

Foldit's social & psychological dimensions

GROUP	PUZZLE	SCORE
Anthropic Dreams	1329: Unsolved ... 96	9,240
BIOL 4030 Winter 2017	Beginner Puzzle...ign	8,877
Beta Folders	1328: 85 Residu...int	10,621
Contenders	1327: Revisitin...one	9,327
Anthropic Dreams	Beginner Puzzle...nis	9,254
Kotocycle	Beginner Puzzle...yle	8,388
Gargleblasters	Beginner Puzzle...ign	10,942
Kotocycle	Beginner Puzzle...ity	14,439
Kotocycle	Beginner Puzzle...zle	8,582
		FULL

SOLOISTS	EVOLVERS	GROUPS	TOPICS
PLAYER		PUZZLE	SCORE
Bruno Kestemont	11 12	1329: Unsolved ... 96	9,239
AeonFluff	100 316	Beginner Puzzle...ign	8,879
bertro	23 7	1328: 85 Residu...int	10,620
gitwut	10 4	1327: Revisitin...one	9,327
jfryk	100 474	Beginner Puzzle...nis	9,254
d_chasy	100 3100	Beginner Puzzle...yle	8,598
cinnamonkitty	100 251	Beginner Puzzle...ign	11,068
cinnamonkitty	100 251	Beginner Puzzle...ity	14,530
florashaman	100 114	Beginner Puzzle...zle	8,680
			FULL

<https://fold.it/portal/info/fa>
q

Limitations of Foldit

- Players are reliant on Foldit's user tools (e.g. visual cues, available moves)
 - Had trouble folding an extended protein chain
 - Tools may limit breadth of intuition
- Maintaining user engagement & sustaining interest

Possible Future Directions

- Psychological studies of player motives and behavior to improve existing PSP algorithms
- Applications to drug discovery-related research areas

Introduction to RNA Design and EteRNA

Table of Contents

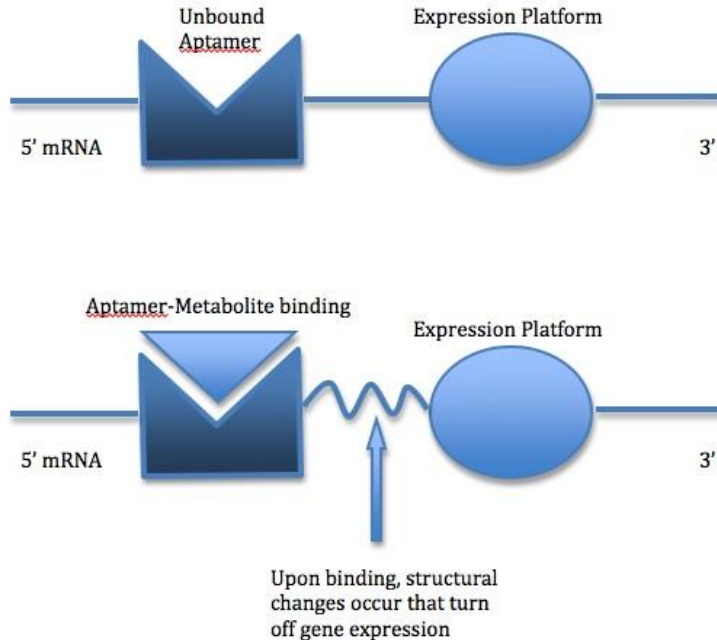
1. RNA Design

- a. Importance of RNA Structure
- b. RNA Design: Computational Determination Methods

2. EteRNA

Importance of RNA Structure (1)

Riboswitch & “miRNA alterations are related to cancer” (Calin et al., 2006)



MicroRNA signatures in human cancers

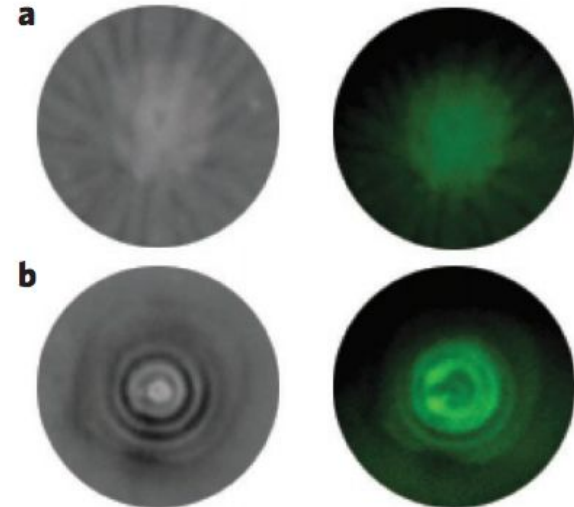
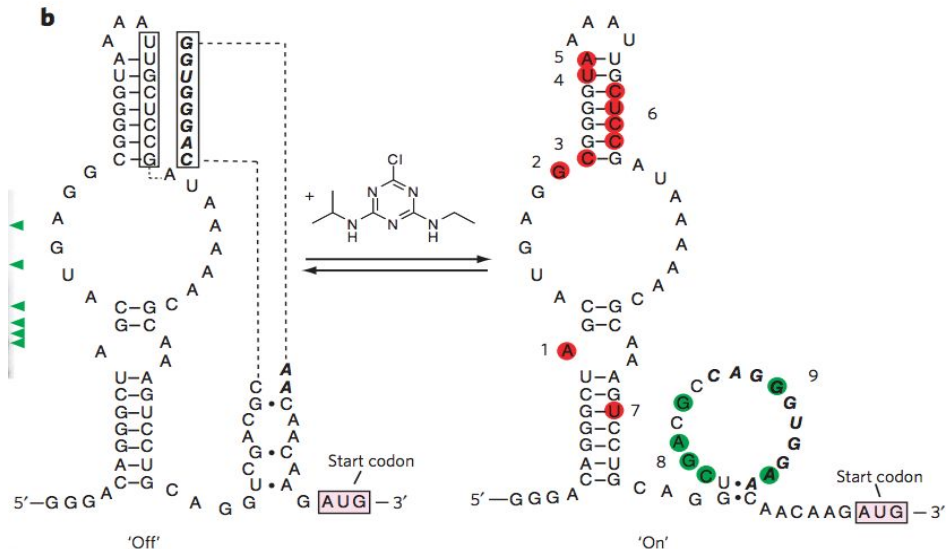
George A. Calin and Carlo M. Croce

Abstract | MicroRNA (miRNA) alterations are involved in the initiation and progression of human cancer. The causes of the widespread differential expression of miRNA genes in malignant compared with normal cells can be explained by the location of these genes in cancer-associated genomic regions, by epigenetic mechanisms and by alterations in the miRNA processing machinery. MiRNA-expression profiling of human tumours has identified signatures associated with diagnosis, staging, progression, prognosis and response to treatment. In addition, profiling has been exploited to identify miRNA genes that might represent downstream targets of activated oncogenic pathways, or that target protein-coding genes involved in cancer.

Importance of RNA Structure (2)

Sinha et al., (2010) Reprogramming bacteria to seek and destroy an herbicide

- Manipulate E. Coli's RNA to follow herbicide



RNA Design: Computational Determination

SIAM J. APPL. MATH.
Vol. 35, No. 1, July 1978

© Society for Industrial and Applied Mathematics
0036-1399/78/3501-0006 \$01.00/0

ALGORITHMS FOR LOOP MATCHINGS*

RUTH NUSSINOV,† GEORGE PIECZENIK,‡ JERROLD R. GRIGGS¶
AND DANIEL J. KLEITMAN§

Abstract. A simplified (two-base) version of the problem of planar folding of long chains (e.g., RNA and DNA biomolecules) is formulated as a matching problem. The chain is prescribed as a loop or circular sequence of letters A and B , n units long. A matching here means a set of A - B base pairings or matches obeying a planarity condition: no two matches may cross each other if drawn on the interior of the loop. Also, no two adjacent letters may be matched. We present a dynamic programming algorithm requiring $O(n^3)$ steps and $O(n^2)$ storage which computes the size of the maximum for the given A - B base sequence and which also allows reconstructing a particular folded form of the original string which realizes the maximum matching size. The algorithm can be adapted to deal with sequences with larger alphabets and with weighted matchings.

An algorithm is also presented for a modified problem closer to the biochemical problem of interest: We demand that every match must be adjacent to another match, forcing groups of two or more parallel matches.

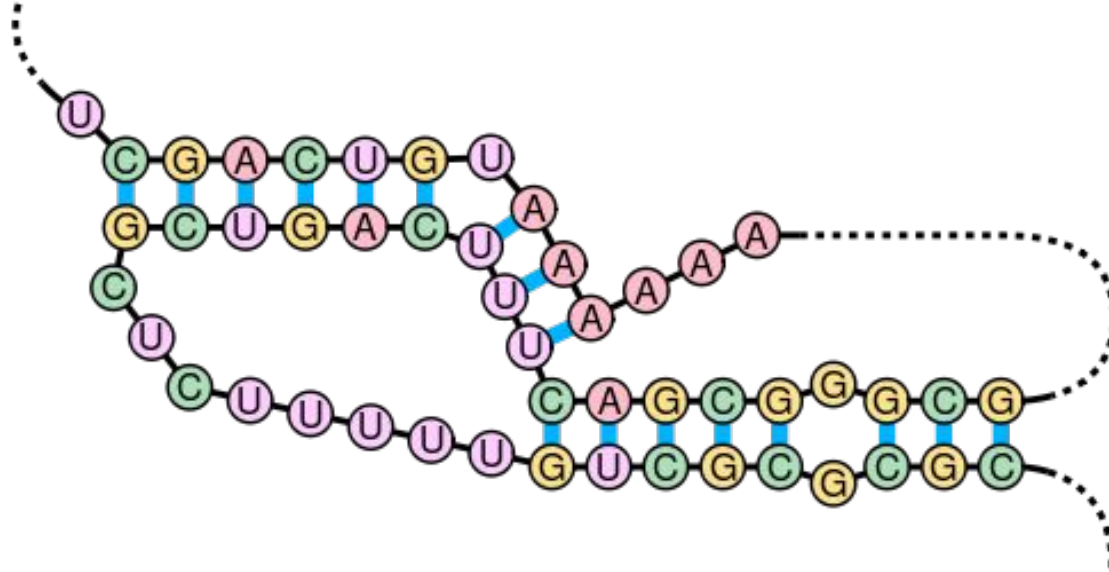
Some results on the expected maximum matching size are presented. As $n \rightarrow \infty$, at least 80% of the vertices can be matched on the average on an A - B string of size n .

We briefly discuss the practical application of the algorithm by using contracted versions of very long molecules with a preliminary block construction. A maximum matching is presented for the J-gene of the ϕ X174 DNA virus. We conclude by stating some problems requiring further study.

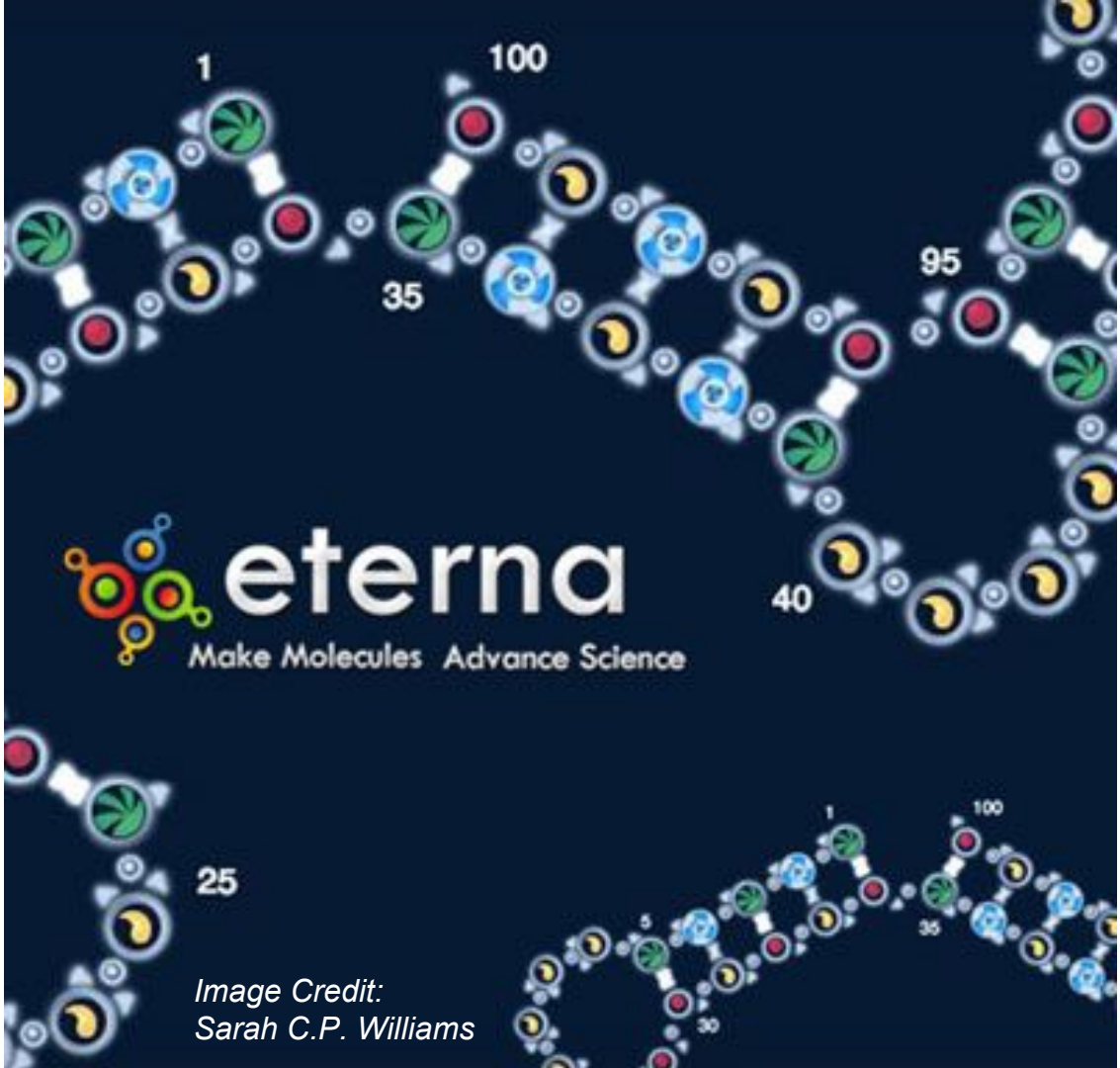
Table 4. Nearest neighbor parameters for G·U base-pairs

Nearest-neighbor	ΔG_{37}° (kcal/mol)	Error	ΔH° (kcal/mol)	Error	ΔS° (eu) ^c	Error ^d
5' AG 3' 3' UU 5'	-0.55	0.32	-3.21	2.76	-8.6	8.45
5' AU 3' 3' UG 5'	-1.36	0.24	-8.81	2.10	-24.0	6.44
5' CG 3' 3' GU 5'	-1.41	0.24	-5.61	2.13	-13.5	6.53
5' CU 3' 3' GG 5'	-2.11	0.25	-12.11	2.22	-32.2	6.81
5' GG 3' 3' CU 5'	-1.53	0.27	-8.33	2.33	-21.9	7.14
5' GU 3' 3' CG 5'	-2.51	0.25	-12.59	2.18	-32.5	6.67
5' GA 3' 3' UU 5'	-1.27	0.28	-12.83	2.44	-37.3	7.47
5' GG 3' 3' UU 5'	+0.47 (-0.5) ^b	0.96	-13.47	8.37	-44.9	25.65
5' GU 3' ^a 3' UG 5' ^a	+1.29	0.56	-14.59	4.92	-51.2	15.08
5' GGUC 3' ^a 3' CUGG 5' ^a	-4.12	0.54	-30.80	8.87	-86.0	23.70
5' UG 3' 3' AU 5'	-1.00	0.30	-6.99	2.64	-19.3	8.09
5' UG 3' 3' CU 5'	+0.30	0.48	-9.26	4.19	-30.8	12.86
Each Terminal G·U ^d	+0.45	-	+3.72	-	+10.5	-

RNA Design: Pseudoknots



Pseudoknot found in telomerase
(Wikipedia)

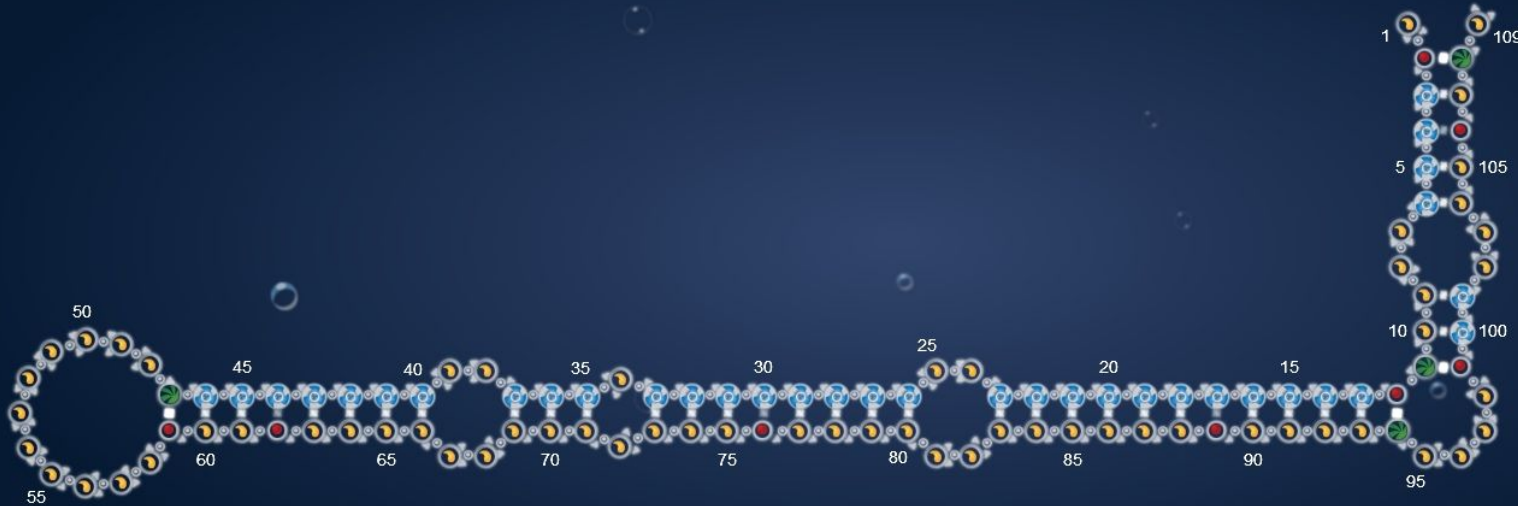


eterna
Make Molecules Advance Science

*Image Credit:
Sarah C.P. Williams*

Total:
-18.6 kcal

 /25
  /15
 



Laranja 

Chat Players Online (40)

labs [7:14 PM]
dw.thewilliams33: thats just what i need [7:14 PM]
dw.thewilliams33: oh [7:14 PM]
Zanna: it was all very confusing [7:14 PM]
Jieux: insanity has its privileges. [7:14 PM]
dw.thewilliams33: see I just started this game like 5 days ago [7:14 PM]
dw.thewilliams33: so im kinda new [7:14 PM]

Image Credit: Ivy Lily Crafts
<http://bit.ly/2jip8m>











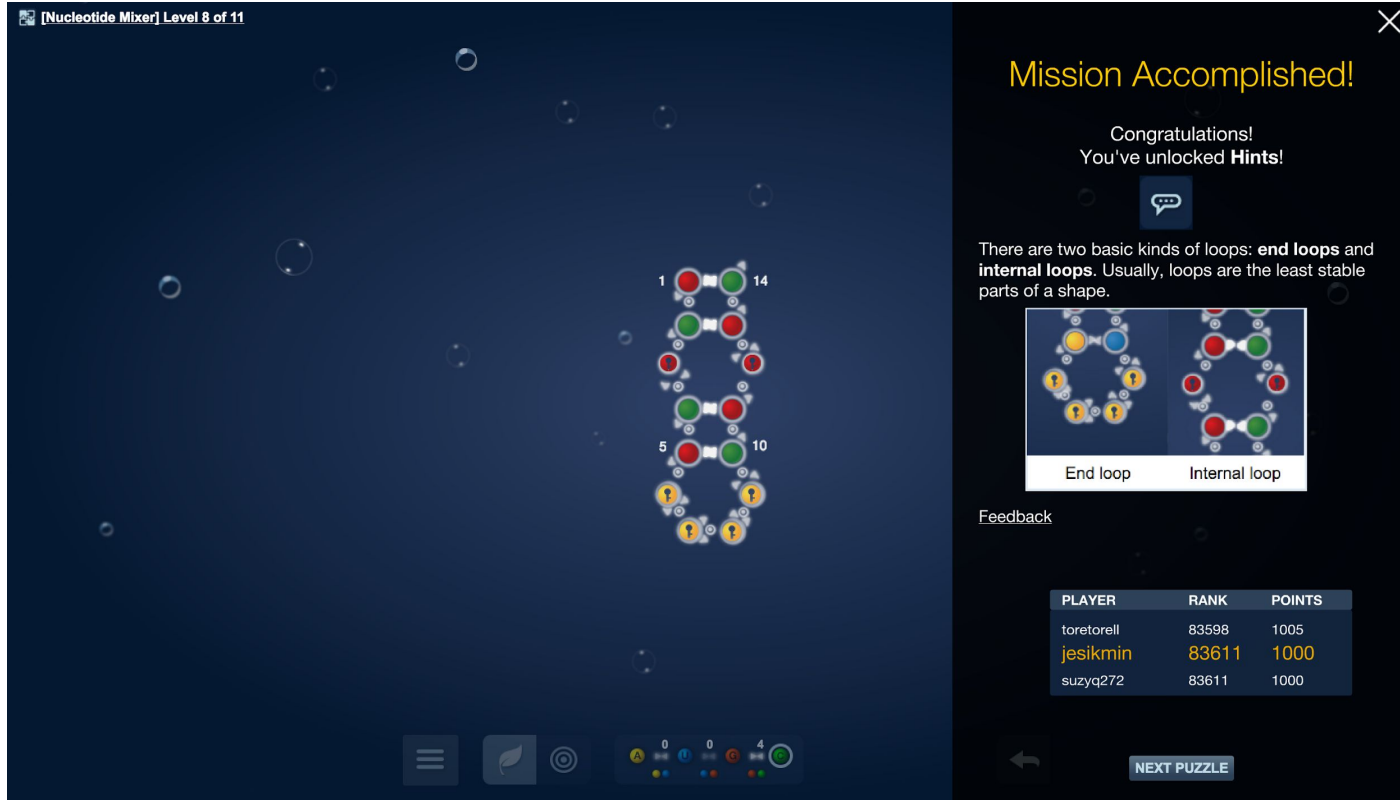
 31
  4
  4
 

Design Interface: Tutorial / Badge Collections



Design Interface: Lesson

[Nucleotide Mixer] Level 8 of 11



The screenshot shows the 'Nucleotide Mixer' game interface. On the left, a puzzle is displayed on a dark blue background with floating bubbles. The puzzle consists of a vertical chain of nucleotides (represented by colored circles: red, green, blue, yellow) with numbers 1, 5, 10, and 14 indicating positions. On the right, a 'Mission Accomplished!' overlay is shown, featuring a congratulatory message, a hint icon, and a lesson about end loops and internal loops. Below the lesson is a table showing player rankings and points.

Mission Accomplished!

Congratulations!
You've unlocked Hints!

There are two basic kinds of loops: **end loops** and **internal loops**. Usually, loops are the least stable parts of a shape.

End loop Internal loop

Feedback

PLAYER	RANK	POINTS
toretorell	83598	1005
jesikmin	83611	1000
suzyq272	83611	1000

NEXT PUZZLE

Design Interface: Challenge Puzzles

The screenshot displays the Eterna game interface. At the top, the logo for 'eterna' is on the left, and 'OpenTB HELP CREATE A NEW WAY TO TARGET TUBERCULOSIS' is on the right. A user profile for 'jessikmin' with rank #52268 and 2300 coins is visible. The main content area is titled 'Challenges' and includes a 'Go To Player Puzzles' button. Below this, there are sorting options: 'Post date', 'Number of people cleared', 'Reward', and 'Length'. Filter buttons for 'Single only', 'Switch only', and 'Uncleared' are present, along with a search bar. The challenges are arranged in a grid:

- Simple Hairpin**: 63163 people, 25 coins. Diagram shows a simple hairpin structure.
- Hairpin Series**: 54823 people, 35 coins. Diagram shows a hairpin structure with a shaded area.
- Hairpin RNA**: 53005 people, 50 coins. Diagram shows a hairpin structure with a shaded area.
- Loop Strategy - Learn about the 1-1 Loop**: 37244 people, 55 coins. Diagram shows a loop structure.
- Y RNA - Difficulty Level 0**: 42922 people, 100 coins. Diagram shows a Y-shaped RNA structure.
- Synthetic RNA - Difficulty Level 0**: 42446 people, 100 coins. Diagram shows a complex RNA structure.
- Oryza sativa 4 - Difficulty Level 0**: 31628 people, 100 coins. Diagram shows a complex RNA structure.
- Oryza sativa 5 - Difficulty Level 0**: 28244 people, 100 coins. Diagram shows a complex RNA structure.
- Chlamydomonas reinhardtii - Difficulty Level 0**: 22706 people, 100 coins. Diagram shows a complex RNA structure.

On the right side, there is a navigation menu with 'HOME', 'NEWSFEED' (with a red notification badge '8'), 'COMMUNITY', 'RESOURCES', and 'ABOUT'. Below the menu is a 'LIVE CHAT' section with 'ONLINE (27)' users. The chat log shows messages from 'cynwulf28', 'Warren09', 'LinkBot', 'caro9923', 'MasterStormer', 'xx_kemi_xx12', and 'Astromon', including a URL to a Google Docs document.

Demo!

**RNA Design Rules
from a
Massive Open Laboratory**

Jesse Min

Table of Contents

1. Emergence of RNA Folding Models: Vienna RNA Package
2. Citizen Science: Massive Open Laboratory
3. Interfaces
4. Results
5. Significances
6. Limitations

Emergence of RNA Folding Models: ViennaRNA

Hofacker et al., (2004) RNA secondary structure analysis using the Vienna RNA package

- 1) **Classical minimum free energy algorithm** (Zuker and Stiegler., 1981)
- 2) **Partition function algorithm** (McCaskill., 1990)
- 3) **Suboptimal folding algorithm** (Wuchty et al., 1999)

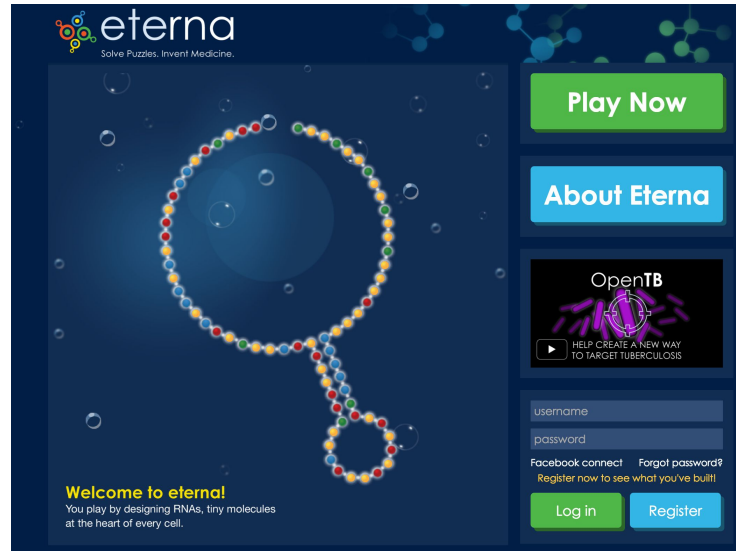
Citizen Science: Massive Open Laboratory

Limitations of current approach

- Small group of professional scientists must interpret empirical data
- How about with ML and visualization tools?

Massive Open Laboratory

- 1) Simulated molecules
- 2) Remote experimental pipeline
- 3) Design RNA structures



The screenshot shows the Eterna website interface. At the top left is the Eterna logo with the tagline "Solve Puzzles. Invent Medicine." Below the logo is a large, colorful RNA structure visualization. On the right side, there are several buttons: a green "Play Now" button, a blue "About Eterna" button, and a video player for "OpenTB" with the text "HELP CREATE A NEW WAY TO TARGET TUBERCULOSIS". Below the video player is a login form with fields for "username" and "password", and buttons for "Log in" and "Register". There are also links for "Facebook connect" and "Forgot password?". At the bottom left, there is a "Welcome to eterna!" message: "You play by designing RNAs, tiny molecules at the heart of every cell."

Design Interface: Challenge Puzzles

The screenshot displays the Eterna game interface. At the top, the logo for 'eterna' is on the left, with the tagline 'Solve Puzzles. Invent Medicine.' Below it. To the right is the 'OpenTB' logo with the text 'HELP CREATE A NEW WAY TO TARGET TUBERCULOSIS'. Further right, a user profile for 'jessikmin' is shown with a rank of #52268 and 2300 coins, along with a 'Logout' button.

The main content area is titled 'Challenges' and includes a 'Go To Player Puzzles' button. It features a sorting menu with options for 'Post date', 'Number of people cleared', 'Reward', and 'Length'. Below the sorting menu are three filter buttons: 'Single only', 'Switch only', and 'Uncleared', followed by a search input field.

The challenge grid consists of nine items, each with a title, a player count icon, a reward icon, and a puzzle diagram:

- Simple Hairpin**: 63163 players, 25 reward. Diagram shows a simple hairpin structure.
- Hairpin Series**: 54823 players, 35 reward. Diagram shows a hairpin structure with a shaded area.
- Hairpin RNA**: 53005 players, 50 reward. Diagram shows a hairpin structure with a shaded area.
- Loop Strategy - Learn about the 1-1 Loop**: 37244 players, 55 reward. Diagram shows a loop structure.
- Y RNA - Difficulty Level 0**: 42922 players, 100 reward. Diagram shows a Y-shaped RNA structure.
- Synthetic RNA - Difficulty Level 0**: 42446 players, 100 reward. Diagram shows a complex RNA structure.
- Oryza sativa 4 - Difficulty Level 0**: 31628 players, 100 reward. Diagram shows a complex RNA structure.
- Oryza sativa 5 - Difficulty Level 0**: 28244 players, 100 reward. Diagram shows a complex RNA structure.
- Chlamydomonas reinhardtii - Difficulty Level 0**: 22706 players, 100 reward. Diagram shows a complex RNA structure.

On the right side, there is a vertical navigation menu with buttons for 'HOME', 'NEWSFEED' (with a red notification badge showing '8'), 'COMMUNITY', 'RESOURCES', and 'ABOUT'. Below this menu is a 'LIVE CHAT' section with 'ONLINE (27)' users. The chat log shows several messages, including a 'no worries' message, a 'boost' message, a 'LinkBot' message with a Google Docs link, and several other user interactions.

Lab Interface (1)

The screenshot displays the Eterna OpenTB web interface. At the top, the Eterna logo is on the left, and the OpenTB logo with the tagline "HELP CREATE A NEW WAY TO TARGET TUBERCULOSIS" is on the right. A user profile for "jessikmin" with rank #52269 and 2300 coins is visible. Below the header, a "Paper Labs" section features a puzzle titled "Single-input switches, revisited" by Elnando888, featuring by Stanford University. The puzzle description discusses the challenge of designing efficient switches for various inputs like tryptophan and arginine. A "Participate" button is located at the bottom of the puzzle description. Below this, another puzzle titled "OpenTB - Flush Stack Experiment" by calebgeniesse is partially visible. On the right side, a vertical navigation menu includes links for HOME, NEWSFEED (with a red notification badge), COMMUNITY, RESOURCES, and ABOUT. At the bottom right, a "LIVE CHAT" window shows a list of online users (27) and a scrollable chat log with messages from users like cymwulF28, Warren09, LinkBot, caro9923, MasterStormer, xx_kemi_xx12, Astromon, and alecpkachu.

eterna Solve Puzzles. Invent Medicine.

OpenTB HELP CREATE A NEW WAY TO TARGET TUBERCULOSIS

Rank #52269 2300 Logout

Paper Labs

Participate in amazing projects created by both Eterna players and professional scientists.
Total 11000 designs will be selected for synthesis on **February 28, 2017 9:00PM PST**.

sort by [Puzzle post date](#) [Synthesis slot](#)

Single-input switches, revisited

featured by [Stanford University](#)  Elnando888



As we are waiting for experiments to complete on the most recent round of player's OpenTB designs, we are taking a refresher here and going back to some simpler, single-input puzzles. The Eterna community previously succeeded in creating 'perfect' switches that respond to a small drug-like molecule called flavin mononucleotide.

Now, by testing if we can get efficient switches with a larger variety of inputs -- tryptophan, arginine, theophylline -- we can prepare for designing eventual diagnostics that could sense small molecules and drugs in the bloodstream. And we can compare our results to prior work from expert labs on these similar problems!

Good luck! These should be much easier than OpenTB, so we are looking for designs from **as many players as possible**. We may have a few more puzzles out in this series in early 2017.

[Participate](#)

OpenTB - Flush Stack Experiment

 calebgeniesse



HOME

NEWSFEED 8

COMMUNITY

RESOURCES

ABOUT

LIVE CHAT ONLINE (27)

cymwulF28: no worries (4:06 PM)
Warren09: .boost (3:02 AM)
LinkBot:
https://docs.google.com/document/d/1rL_M72Sj8_j9Zo44h71YXQkYP_wurfXcqxL18NXbi-8 (3:02 AM)
caro9923:
http://www.eternagame.org/sites/default/files/chat_screens/239236_1484658965.png (6:16 AM)
caro9923: Hello, can you please help me? I need to remove a G-C (5:16 AM)
MasterStormer: hi (8:03 AM)
MasterStormer: are you still here? (8:04 AM)
xx_kemi_xx12: HI (11:14 AM)
Astromon: hi (1:34 PM)
alecpkachu:
http://www.eternagame.org/sites/default/files/chat_screens/238564_1484694237.png (3:09 PM)

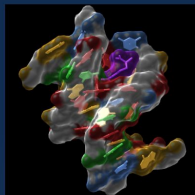
Lab Interface (2)



[jesikmin](#) rank #52269 2300 [Logout](#)

Single-Input switches, revisited

[Follow](#)



As we are waiting for experiments to complete on the most recent round of player's OpenTB designs, we are taking a refresher here and going back to some simpler, single-input puzzles. The Eterna community previously succeeded in creating 'perfect' switches that respond to a small drug-like molecule called flavin mononucleotide.

Now, by testing if we can get efficient switches with a larger variety of inputs -- tryptophan, arginine, theophylline -- we can prepare for designing eventual diagnostics that could sense small molecules and drugs in the bloodstream. And we can compare our results to prior work from expert labs on these similar problems!

Good luck! These should be much easier than openTB, so we are looking for designs from **as many players as possible**. We may have a few more puzzles out in this series in early 2017.

HOME

NEWSFEED 8

COMMUNITY

RESOURCES

ABOUT

Lab info

Admin : [EINando888](#)

Research affiliation : [Stanford University](#)

Synthesis slots : **3600**

Selection method : **User voting**

Number of your votes : **0**

Designs you submitted : **0**

Designs submitted : **1168**

News from Lab Admins

Current Round

Same State - Tryptophan A 300



Exclusion - Tryptophan A 300



Same State - Tryptophan B 300



Exclusion - Tryptophan B 300



Same State - Theophylline A 300



Exclusion - Theophylline A 300



LIVE CHAT ONLINE (27)

cynwulf28: no worries (4:06 PM)
Warren09: .boost (3:02 AM)
LinkBot:
https://docs.google.com/document/d/1rLM72Sj8_j9Zo44h71tYXQkYPwufXoqx0L18NXbI-8 (3:02 AM)
caro9923:
http://www.etergame.org/sites/default/files/chat_screens/239236_1484658965.png (3:16 AM)
caro9923: Hello, can you please help me? I need to remove a G-C (3:18 AM)
MasterStormer: hi (8:03 AM)
MasterStormer: are you still here? (8:04 AM)
xx_kemi_xx12: Hi (11:14 AM)
Astromon: hi (1:04 PM)
alecpikachu:
http://www.etergame.org/sites/default/files/chat_screens/238564_1484694237.png (3:09 PM)

Lab Tracker



Accepting Submissions



Ordering DNA Template



Synthesizing RNA



Getting Data



Results Posted

Synthesis Results

This lab does not have synthesized designs yet.

Voting Interface

Shape Library 102 : The Cross

rank #305 tartan 550 logout

Me Puzzle RNA Lab Community Why RNA?

Title	Designer	Votes		My Votes		Description
		min	max	min	max	
search	search					search
The cross R3-1 Solution 2	Chesterfield	0		0		Blue flu-go green
The Cross	deivad	1		0		No comment
David attempt 1	Gene	1		0		My first lab result.
mat late one	deivad	0		0		My first try. What do you...
mat - last one	mat747	0		0		No comment
new solution-1	mat747	0		0		No comment
Third try	Chesterfield	0		0		times square
My quick solution	anneromaine	0		0		No comment
Cross -45.5 kJ	alimpaecher	0		0		I put this together quickly
Cross 2.222	Ipso	0		0		No comment
Extra GUs + Mod...	anneromaine	0		0		Second Try
Extra GUs	singinst	-		-		Added extra GUs and modif...
crossV3	singinst	0		0		67 degrees
CrossV2	Lloyd	0		0		No comment
Try Number 1	Lloyd	0		0		No comment
BMG Cross 2.0	TheGummer	0		0		No comment
CrossV1	BootsMcGraw	-		-		Started with all A-U, the...
Mat cross design v2	Lloyd	0		0		No comment
finnish?	mat747	-		-		Similar strategy to my/ot...
Jewelad Jest	bestcreator	0		0		i think my mission is com...
IT WORKS	Johny.Manic	-		-		An alternating usage of G...
GlobusCruciger-A	Devon	0		0		I fixed it....
T-Bone	Vexelius	0		0		A new design, created fro...
Pick me	Stefan	0		0		Yeah. First try, nothing ...
	mjm128	0		0		No comment

chat: players online (12)

can debug it

Chesterfield: Hope you can. that sounds like it is related to my location?

jeehyung: That could be possible, but I can't be sure. This is very weird that submission successfully goes through but you don't get a response..

Jee Jee: oh hey one more thing. When you are in the game does, your points/rank show up correctly in the upper right corner?

Chesterfield: yes,

You have 8 votes left.

You have 3 solution slots left.

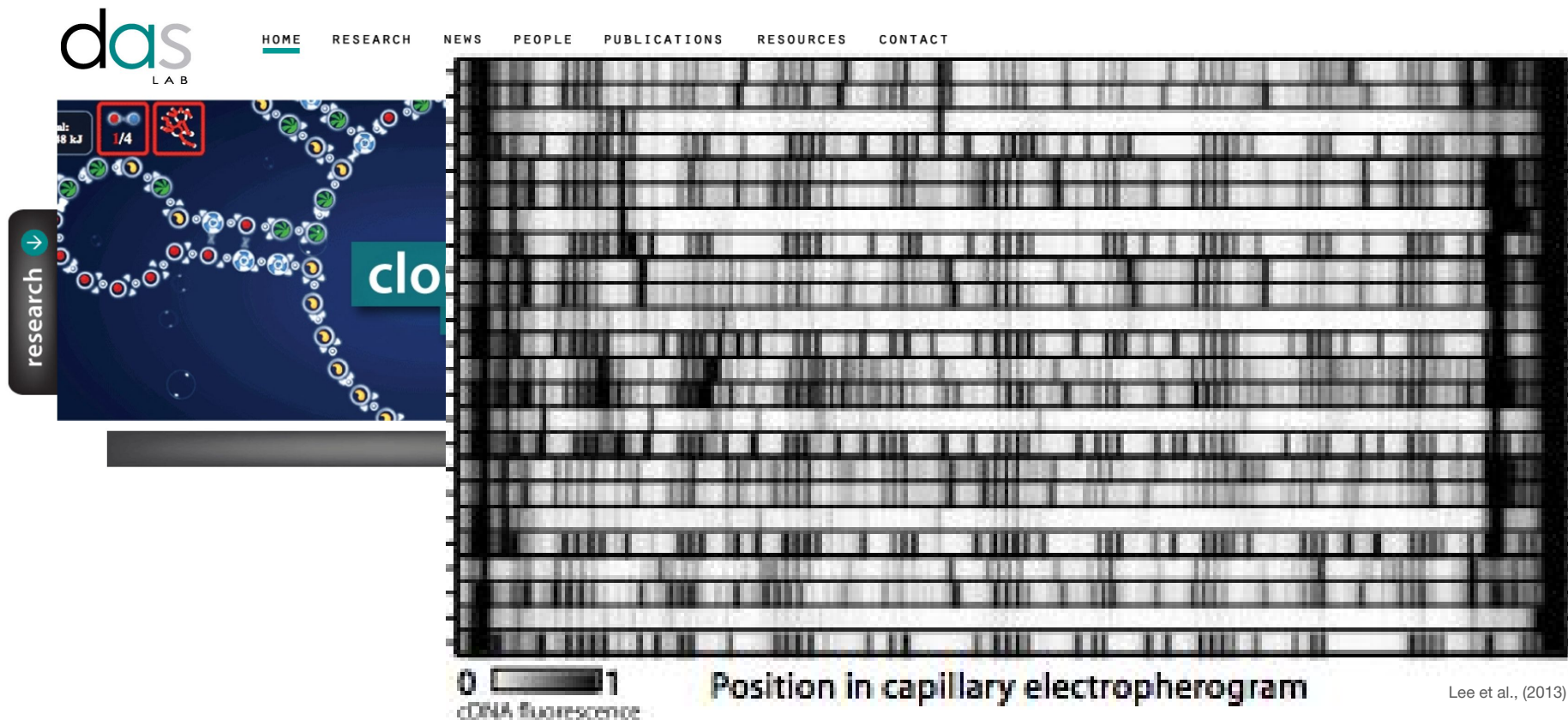
How does the lab work?

How do I see experimental results?

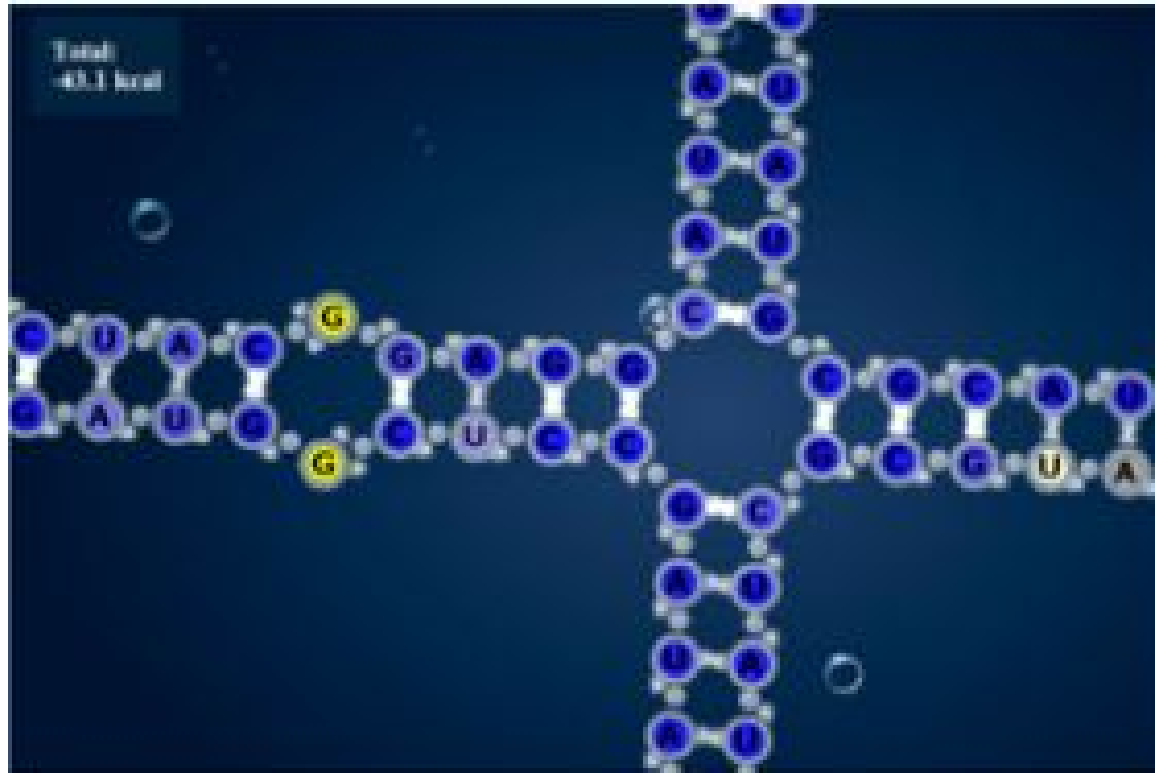
Discuss the puzzle and designs with others.

Click on a design to vote.

Remote Lab













Result Viewer



Results & Rule Collections

Winners, (22 out of 34)

User	Name	Puzzle Title	Solution Title	Score
	Eli Fisker	[A]/[C] - DEC	Rocketdog 42 mod 3	100
	jandersonlee	[A]/[C] - INC	jl-acinc 7061024 G53+G85 Eli Fisker mod #MUTATE2	100
	jandersonlee	[A]/[C] - INC	jl-acinc 7061024 G52+U53 Eli Fisker mod #MUTATE2	100
	jandersonlee	[A]/[C] - INC	jl-acinc 7061024 U52+U53 Eli Fisker mod #MUTATE2	100
	jandersonlee	[A]/[C] - INC	jl-acinc 7061024 G53+U56 Eli Fisker mod #MUTATE2	100
	jandersonlee	[A]/[C] - INC	jl-acinc 7061024 G52+G85 Eli Fisker mod #MUTATE2	100
	jandersonlee	[A]/[C] - INC	jl-acinc 7061024 U52+G53 Eli Fisker mod #MUTATE2	100
	joy45	[B]/[C] - INC	B/C INC J1	100
	mat747	[A]/[C] - INC	Single #MUTATE 7061024 G85 Eli Fisker mod	100
	mat747	[A]/[C] - INC	Single #MUTATE 7061024 U53 Eli	100

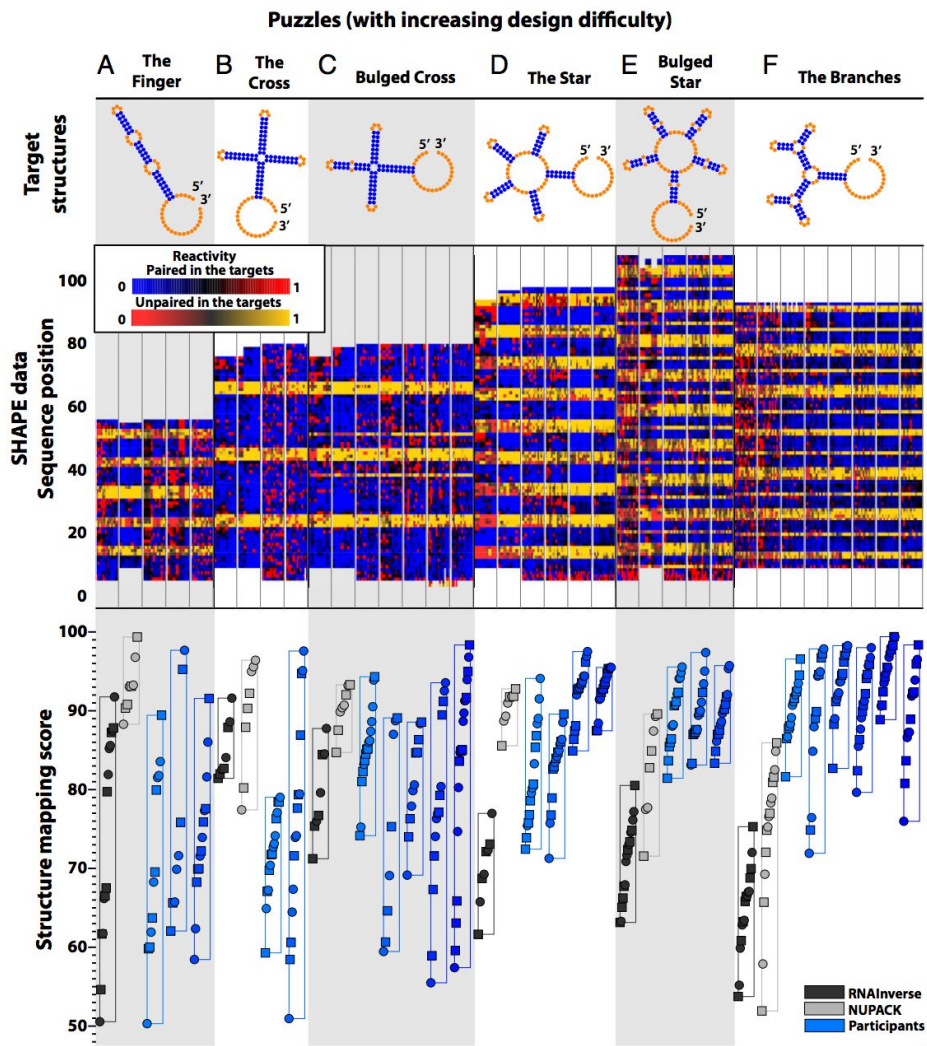
[Review results](#)

[Lab Info](#)

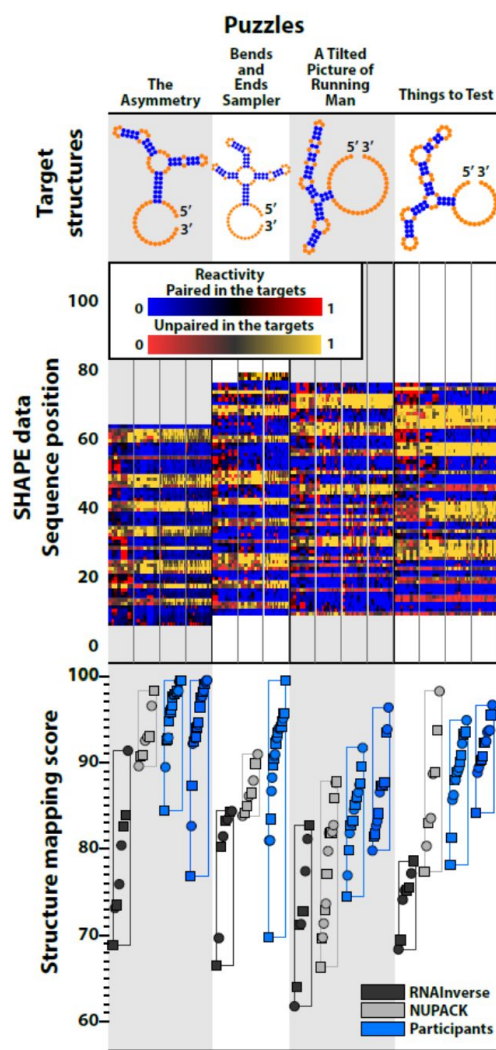


Lee et al., (2013)

Result: Phase(1)



Result: Phase(1) to Phase(2)



EteRNABot

A Basic Test

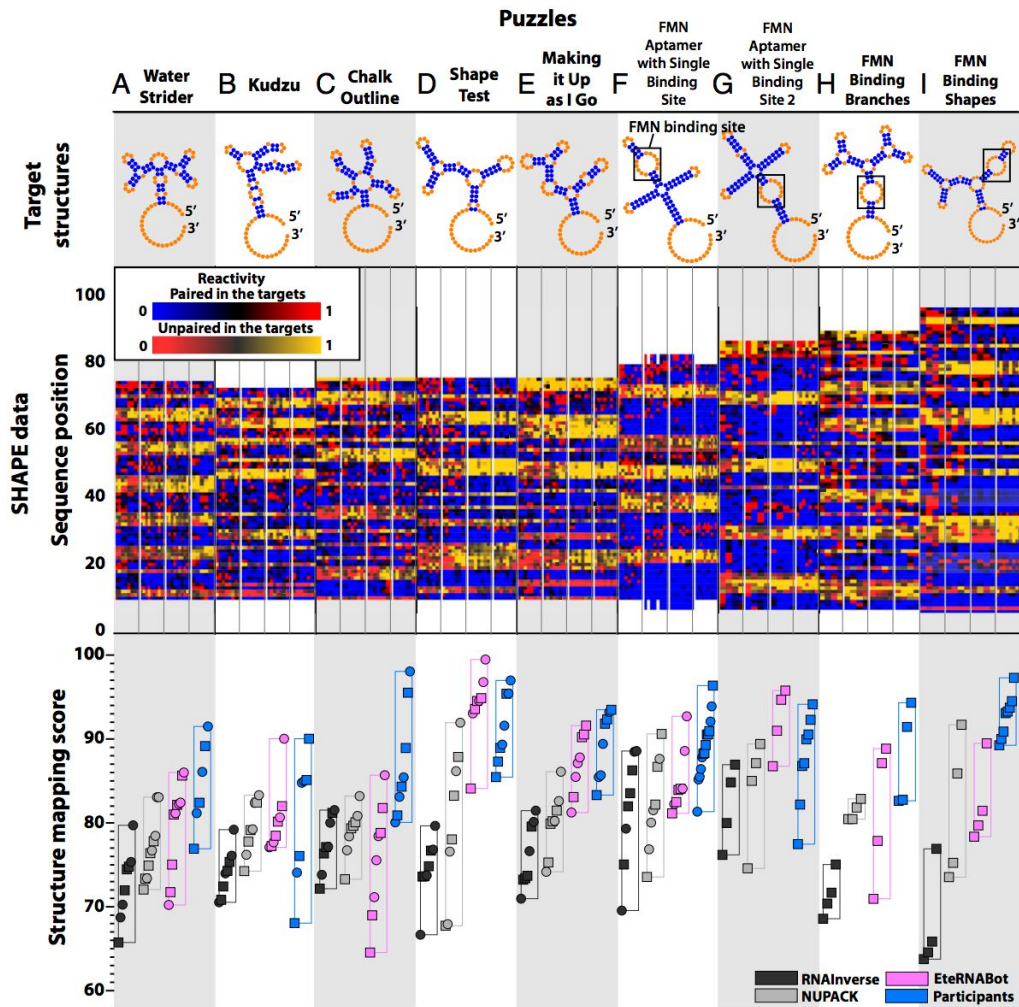
Clean Plot, stack caps, and safe GC

Direction of GC-pairs in multiloops + neckarea

Berex Test

Numbers of Yellow Nucleotides per Length of String

Result: Phase(2)



Significance: User-created Design Rules

User-created design rules

- [Player-created Guides](#)
- [Multi-loop Guides](#)
- **Negative Design Rules**
 - Penalties on special structures (repetitions, tetraloop similarity, and twisted base pairs)

Significance: Feedback from Real Experiments

Previous Efforts

Year	Paper	Authors
2001	Can distributed volunteers accomplish massive data analysis tasks?	Kanefsky et al.
2008	Galaxy Zoo: The large-scale spin statistics of spiral galaxies in the Sloan Digital Sky Survey	Land K et al.
2011	Algorithm discovery by protein folding game players	Khatib F et al.
2012	Increased Diels-Alderase activity through backbone remodeling guided by Foldit players	Eiben CB et al.

Lee et al., (2013)

- Collectively **generate and test hypotheses through actual experiments**

Limitations: What is the next step?

Engineering Cost

Theoretical Explanation

More Incentive

Social Prejudice

Limitations: What is the next step?

Treuille et al., (2014) Scientific rigor through videogames

- Budget
- Career risk
- Generalization
 - Schindel et al., (2005) DNA barcoding a useful tool for taxonomists
 - Helmstaedter et al., (2013) Cellular-resolution connectomics: Challenges of dense neural circuit reconstruction

Principles for Predicting RNA Secondary Structure Design Difficulty

...

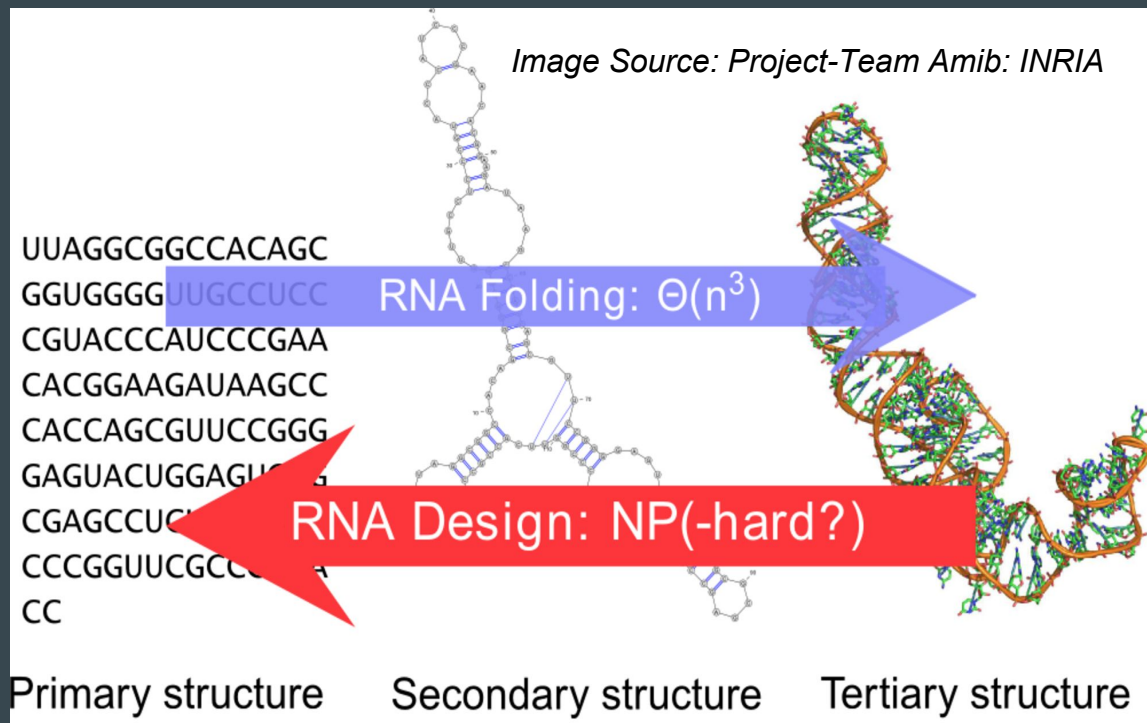
Jeff Anderson-Lee, Eli Fisker, Vineet Kosaraju, Michelle Wu, Justin Kong, Jeehyung Lee, Minjae Lee, Mathew Zada, Adrien Treuille, Rhiju Das, Eterna Players

Presentation Overview

- Paper Focus/Main Idea
- Difficult Structures
 - Short Stems
 - Symmetry
 - Structural Motifs
- Impact
- Weaknesses

Inverse Folding Problem

- Know to be hard
- Unknown which structures are tractable
- Paper set forth measures for difficulty



Research Process

- Although “RNA design rules from a massive open laboratory” was revolutionary with 37,000 authors, the lead authors were scientists
- In “Principles for Predicting RNA Secondary Structure Design Difficulty” the lead authors are Eterna players: the revolution continues
- Jeffrey Anderson-Lee: computer systems manager
 - “It’s really a bit amusing, I guess it’s stretching [the journal’s] boundaries a bit.”
- Eli Fisker: Librarian
 - “When the scientist is away, the lab rat will play”

Overview

- Paper Structure/Main Idea
- Difficult Structures
 - Short Stems
 - Symmetry
 - Structural Motifs
- Impact
- Weaknesses

Difficult Structures

- The paper gathers knowledge from the EteRNA player community
- Compares player input to automated design algorithm performance

RNAinverse

INFO.RNA

RNA.SSD


NUPACK

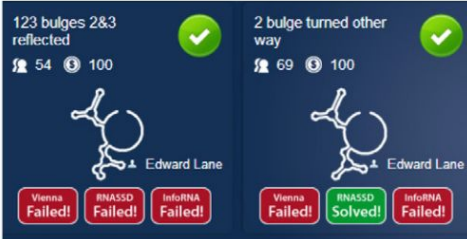
DSS.Opt

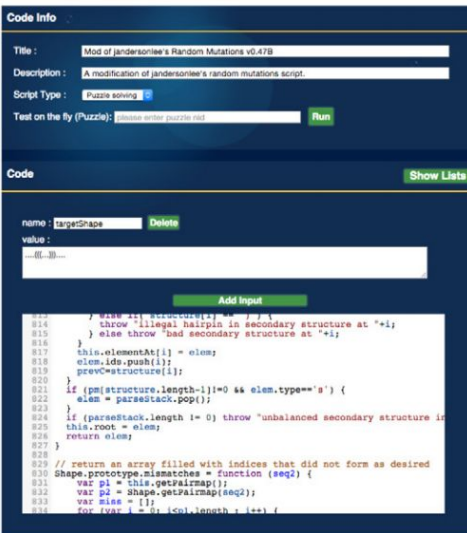
MODENA

Identifying Difficult Problems: Tools

(a) 

(b) 

(c) 

(d) 

Code Info

Title:

Description:

Script Type:

Test on the Fly (Puzzle):

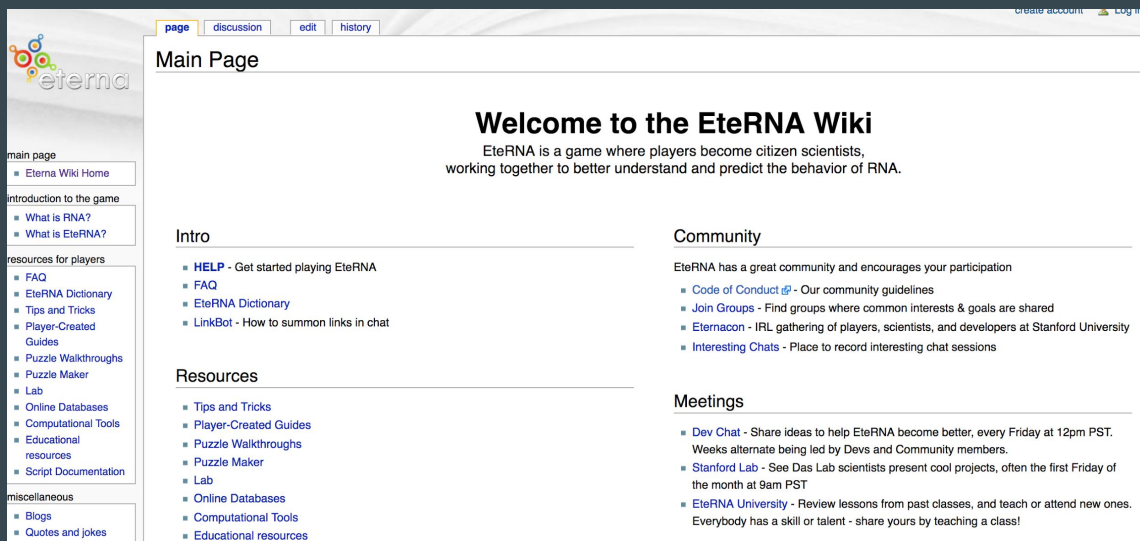
Code

```
name: targetShape 
value:
---[[]]---


112 / with {} structure[] = [] {
113   throw "illegal hairpin in secondary structure at "+i;
114 } else throw "bad secondary structure at "+i;
115 }
116 }
117 this.elementAt[i] = elem;
118 elem.ids.push(i);
119 prevC=structure[i];
120 }
121 if (pn[structure.length-1]===0 && elem.type==='a') {
122   elem = parseStack.pop();
123 }
124 if (parseStack.length != 0) throw "unbalanced secondary structure in
125 this.root = elem;
126 return elem;
127 }
128 }
129 // return an array filled with indices that did not form as desired
130 Shape.prototype.mismatches = function (seq2) {
131   var p1 = this.getFairMap();
132   var p2 = shape.getFairMap(seq2);
133   var mism = [];
134   for (var i=0; i<seq1.length; i++) {
```

Identifying Difficult Problems

- Player's use problem solving, problem creation, algorithm integration and scripting tools to test hypotheses about difficulty
- Player insight is synthesized online in community wiki: crucial to research



The screenshot shows the main page of the EteRNA Wiki. At the top left is the EteRNA logo. The page title is "Main Page". The main heading is "Welcome to the EteRNA Wiki", followed by a sub-heading: "EteRNA is a game where players become citizen scientists, working together to better understand and predict the behavior of RNA." The page is organized into several sections: "Intro" (with a "HELP" link), "Community" (with a "Code of Conduct" link), "Meetings" (with a "Dev Chat" link), and "Resources" (with links to "Tips and Tricks", "Puzzle Walkthroughs", and "Puzzle Maker"). A left sidebar contains a navigation menu with categories like "main page", "introduction to the game", "resources for players", and "miscellaneous".

page discussion edit history

create account Log in

Main Page

Welcome to the EteRNA Wiki

EteRNA is a game where players become citizen scientists, working together to better understand and predict the behavior of RNA.

Intro

- [HELP](#) - Get started playing EteRNA
- [FAQ](#)
- [EteRNA Dictionary](#)
- [LinkBot](#) - How to summon links in chat

Community

EteRNA has a great community and encourages your participation

- [Code of Conduct](#) - Our community guidelines
- [Join Groups](#) - Find groups where common interests & goals are shared
- [Eternacon](#) - IRL gathering of players, scientists, and developers at Stanford University
- [Interesting Chats](#) - Place to record interesting chat sessions

Meetings

- [Dev Chat](#) - Share ideas to help EteRNA become better, every Friday at 12pm PST. Weeks alternate being led by Devs and Community members.
- [Stanford Lab](#) - See Das Lab scientists present cool projects, often the first Friday of the month at 9am PST
- [EteRNA University](#) - Review lessons from past classes, and teach or attend new ones. Everybody has a skill or talent - share yours by teaching a class!

Resources

- [Tips and Tricks](#)
- [Player-Created Guides](#)
- [Puzzle Walkthroughs](#)
- [Puzzle Maker](#)
- [Lab](#)
- [Online Databases](#)
- [Computational Tools](#)
- [Educational resources](#)

main page

- [Eterna Wiki Home](#)

introduction to the game

- [What is RNA?](#)
- [What is EteRNA?](#)

resources for players

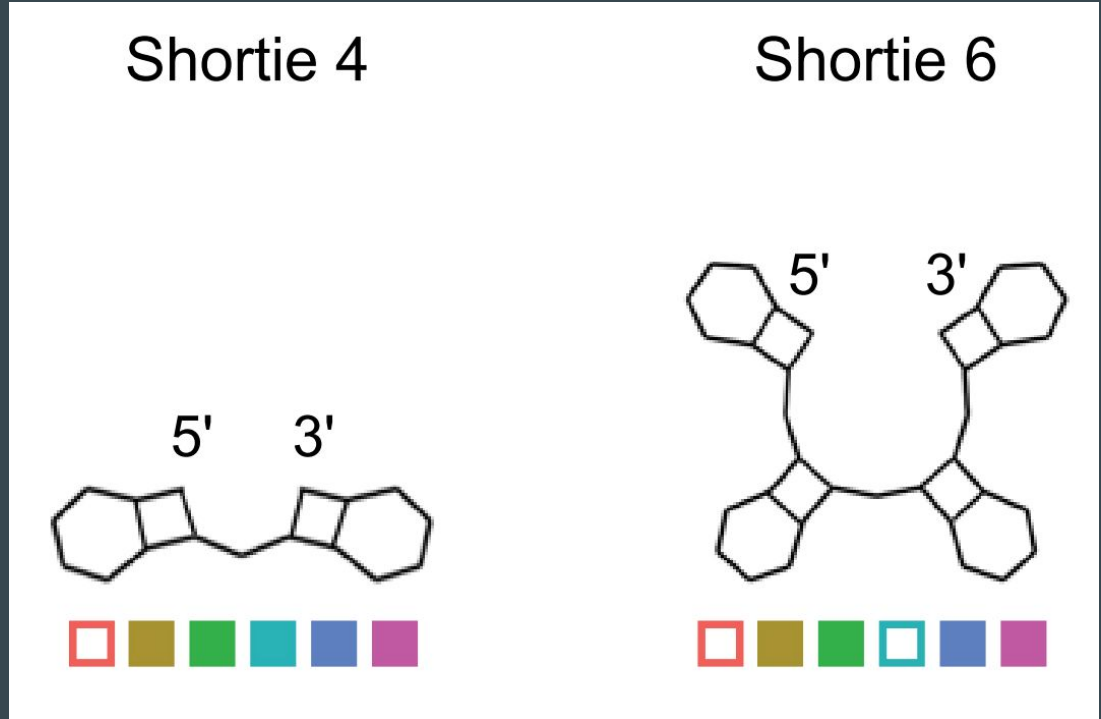
- [FAQ](#)
- [EteRNA Dictionary](#)
- [Tips and Tricks](#)
- [Player-Created Guides](#)
- [Puzzle Walkthroughs](#)
- [Puzzle Maker](#)
- [Lab](#)
- [Online Databases](#)
- [Computational Tools](#)
- [Educational resources](#)
- [Script Documentation](#)

miscellaneous

- [Blogs](#)
- [Quotes and jokes](#)

Short Stems

- Structures joined by short stems are difficult
- Multiplicity of short stems affects free energy



Symmetry

- True symmetry is rare in naturally occurring RNA structures
- Authors postulate this is due to difficulty of design

Water Strider

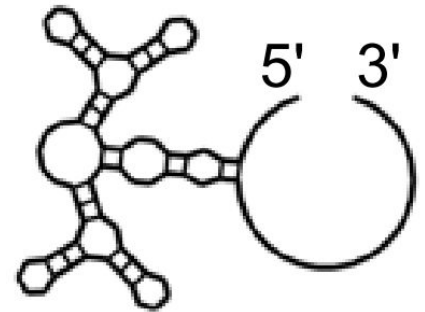
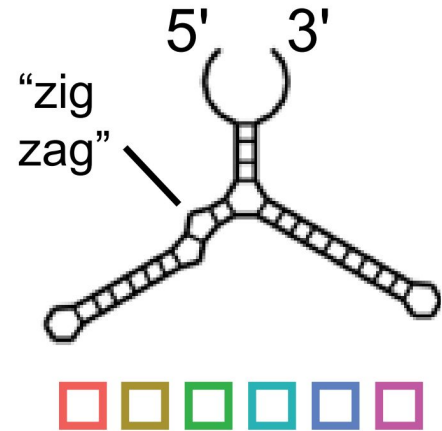


Image Source: Principles for Predicting RNA Secondary Structure Design Difficulty

Structural Motifs

- Player designated “zig-zags” are an example of player discovered difficult structural motifs
- “Zig-zags” close to multi-loops cause current design algorithms to fail
- Again, rare in nature

Image Source: Principles for Predicting RNA Secondary Structure Design Difficulty



Mutated Chicken Feet

Overview

- Paper Structure/Main Idea
- Difficult Structures
 - Short Stems
 - Symmetry
 - Structural Motifs
- Impact
- Weaknesses

Impact

- Paper set forward the Eterna100 benchmark of secondary structure design challenges
- Allows for common benchmark of *in silico* design algorithms



Image Source: Principles for Predicting RNA Secondary Structure Design Difficulty

Overview

- Paper Structure/Main Idea
- Difficult Structures
 - Short Stems
 - Symmetry
 - Structural Motifs
- Impact
- Weaknesses

Weaknesses

- Focus on in-silico test
- No *in vitro* or *in vivo* testing
 - The authors do not include benchmark evaluation on EteRNABot, an algorithm designed for in vitro success

Questions?

Additional Slides