Redesigning the Eterna100 for the Vienna 2 folding engine

Rohan V. Koodli[1,*], Boris Rudolfs[2,*], Hannah K. Wayment-Steele[3], Eterna Structure Designers[γ], Rhiju Das[4,5,^]

[1]Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, Berkeley, CA 94720
[2]Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, California, 92110
[3]Department of Chemistry, Stanford University, Stanford, CA 94305
[4]Department of Biochemistry, Stanford University School of Medicine, Stanford, CA 94305
[5]Department of Physics, Stanford University, Stanford, CA 94305

[*]Indicates equal contribution
[γ]Group Author: Membership of Eterna participants and list of player names is provided in Acknowledgments
[^]Corresponding Author: rhiju@stanford.edu

Abstract

The rational design of RNA is becoming important for rapidly developing technologies in medicine and biochemistry. Recent work has led to the development of several RNA secondary structure design algorithms and corresponding benchmarks to evaluate their performance. However, the performance of these algorithms is linked to the nature of the underlying algorithms for predicting secondary structure from sequences. Here, we show that an online community of RNA design experts is capable of modifying an existing RNA secondary structure design benchmark (Eterna100) with minimal alterations to address changes in the folding engine used (Vienna 1.8 updated to Vienna 2.4). We tested this new Eterna100-V2 benchmark with five RNA design algorithms, and found that neural network-based methods exhibited reduced performance in the folding engine they were evaluated on in their respective papers. We investigated this discrepancy, and determined that structural features, previously classified as difficult, may be dependent on parameters inherent to the RNA energy function itself. These findings suggest that for optimal performance, future algorithms should focus on finding strategies capable of solving RNA secondary structure design benchmarks independently of the free energy benchmark used. Eterna100-V1 and Eterna100-V2 benchmarks and example solutions are freely available at https://github.com/eternagame/eterna100-benchmarking.

Keywords: RNA design, secondary structure, benchmark, inverse folding, citizen science

Introduction

Ribonucleic acid (RNA) has significantly expanded past its original proposed role as an intermediate in the genetic code and as a catalytic scaffold for protein synthesis. RNA has been observed to act as a genetic expression regulator [1], perform catalysis [2], be a scaffold for complex formation [3,4], and be used as a guide by several ribonucleoprotein complexes [5–7]. This increased appreciation for the versatile activity of RNA has led to the recent development of several RNA therapies that include the control of pre-mRNA splicing [8], gene editing and expression [6], and aptamers for binding and sequestering target molecules [9]. Furthermore, given the modular nature of RNA motifs [10] and the simplistic pairing rules of nucleic acids, RNA has been used to design novel nanostructures [11–13] and drive the development of methodologies for the design of novel RNA tertiary structures [14,15]. By combining these approaches, it is possible to design RNA molecules with varied function and topology. However, as RNA length and complexity increases, the number of asymmetric and symmetric elements increases, thereby increasing the difficulty of sequence design for these molecules [16].

The RNA secondary structure design problem, also known as the inverse folding problem, involves designing an RNA sequence that folds into a target secondary structure given an energy function [17]. Classic RNA inverse folding algorithms used cost function minimization through adaptive random walk [18], structure decomposition [19], minimization of

the ensemble defect [20], or a genetic algorithm [21]. Performance of these older algorithms was not well characterized as benchmarking occurred internally and was performed on well characterized biological RNA or computationally predicted secondary structures from RNA sequences [19,22,23].

To address the need for a community-wide standard benchmark for RNA design, J Anderson-Lee et al. developed a set of 100 secondary structures published on the Eterna website using Vienna 1.8.5 (We henceforth refer to this original benchmark as "Eterna100-V1"). This Eterna100 benchmark was chosen to showcase secondary structure motifs that were identified as being difficult to design, and the best performing algorithm [21] solved 54/100. Since the benchmark was published, several algorithms have surpassed this mark using convolutional neural networks [24], reinforcement learning [25,26], or a Monte Carlo search optimized for game theory [27]. However, algorithms have been inconsistent in which folding engine (i.e., secondary structure prediction algorithm) they use both in training and in predicting. For instance, EternaBrain [24] used Vienna 1.8, but Meta-LEARNA [26] used the Turner 2004 thermodynamic parameters in Vienna 2.1.8. Despite the fundamental link between the folding engine used to run and evaluate inverse folding algorithms, there has been no systematic evaluation of the effect of folding engines in the training and performance of inverse folding algorithms. This work describes our investigation into the extent of folding engine dependency in the Eterna100. We challenged Eterna participants to determine if all the puzzles in the original Eterna100 could still be solved using an updated set of Vienna parameters (Vienna 2.4, henceforth referred to as "Vienna 2"). Indeed, participants identified 19 of the 100 structures that were deemed to be unsolvable in Vienna 2 (list of puzzles and the structures in Vienna 1 and Vienna 2 are provided in Supplemental File S1). We then challenged the community to adapt these secondary structures to a different parameter set using a minimal number of insertions and deletions, resulting in the Eterna100-V2 benchmark. We discuss key structural motifs that are intractable in one set of thermodynamic parameters, but solvable in the other. We evaluated several state-of-the-art inverse folding algorithms and determined that while their relative performance is unchanged, neural network based methods would benefit from re-training with Vienna 2 parameters. Taken together, this work indicates that consideration of which folding engine is used in operating inverse folding algorithms is critical in their evaluation, even in determining the scope of what structures are fundamentally solvable.

Results

Players' structure modifications

Eterna participants identified 19 secondary structures of the original 100 'unsolvable' in Vienna 2. We computationally verified that the stochastic algorithm NEMO [27], which is currently state-of-the-art in the Eterna100 with the original Vienna 1 folding engine, was also unable to find solutions within 24 hour timeframes for these 19 problems with the Vienna 2 folding engine.

These 19 puzzles vary in both length and relative complexity, but share several structural features that gave rise to this discrepancy (Fig 1a). The change in the relative free energies of internal and stem loops as well as the differences in free energy bonuses lead to several structures with isolated base pairs no longer being solvable in the Vienna 2 parameters (Fig 1a). Furthermore, multi-helix junctions have different free energies of initiation in the default settings of Vienna 2, leading to distinct differences in secondary structure predictions between the two models (Fig 1a). In addition, we found motifs that Vienna 1 penalized more than Vienna 2: internal junctions with 3 branches (Fig 1b).

These results motivated us to develop a new set of puzzles for these 19 problems that would comprise a new Eterna100-V2 benchmark to be used with the Vienna 2 folding engine, which has largely displaced the Vienna 1 folding engine for wide use. We solicited these new problems from Eterna participants.Figure 2 illustrates why we needed redesigns from human participants for the 19 puzzles deemed unsolvable in Vienna 2. We originally considered a different, simpler redesign method, based on taking a set of known puzzle solutions in Vienna 1 and calculating their minimum free energy structures in Vienna 2. However, we noted that these structures did not exhibit similar shapes and difficult features as the original structures posed in the Eterna100 benchmark. To quantify this difference, we used RNAdistance [18], a metric based on string edit distance for measuring differences in RNA secondary structures built into the ViennaRNA suite, to measure the “difference” between the folded sequence in Vienna 1 and the folded sequence in Vienna 2 (Fig 1). In all 19 puzzles, this calculated difference was much larger than the difference in secondary structures in V1 and V2 that players created in parallel.

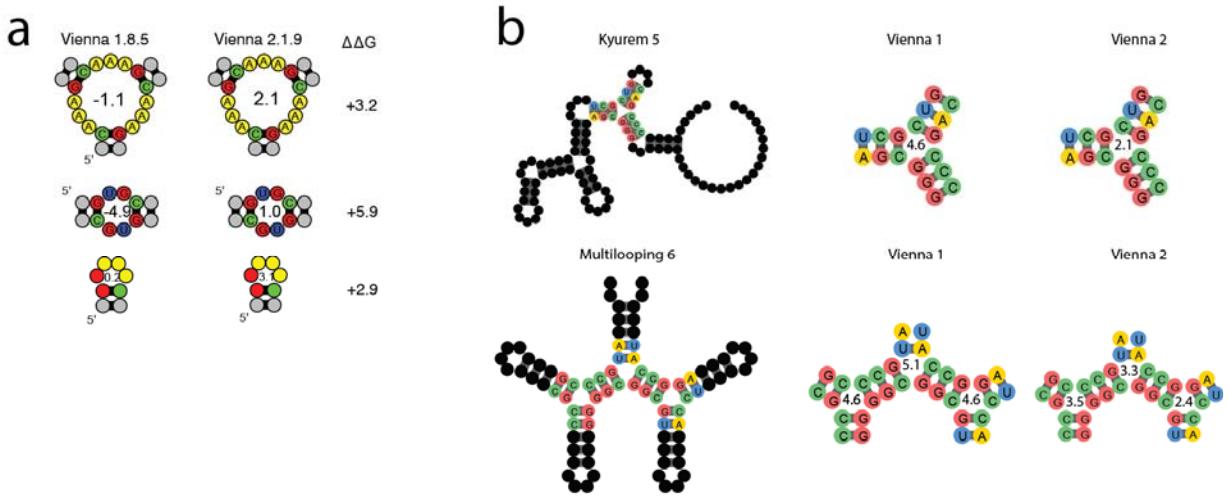


Figure 1. (A) Free energy differences (ΔΔG, kcal/mol) between identical structures in Vienna 1.8 and Vienna 2.4. Topmost: an internal triloop; middle: U-G-U-G superboost with G-C end pairs; bottom: tetraloop hairpin with G-boost and G-C end pairs. (B) Puzzle-specific differences in free energies of structures. The internal tri-junction in Kyurem 5 has a more stable free energy

in Vienna 2. Similar repeated tri-junctions are highlighted in Multilooping 6, with similar changes in free energy from Vienna 1 to 2.

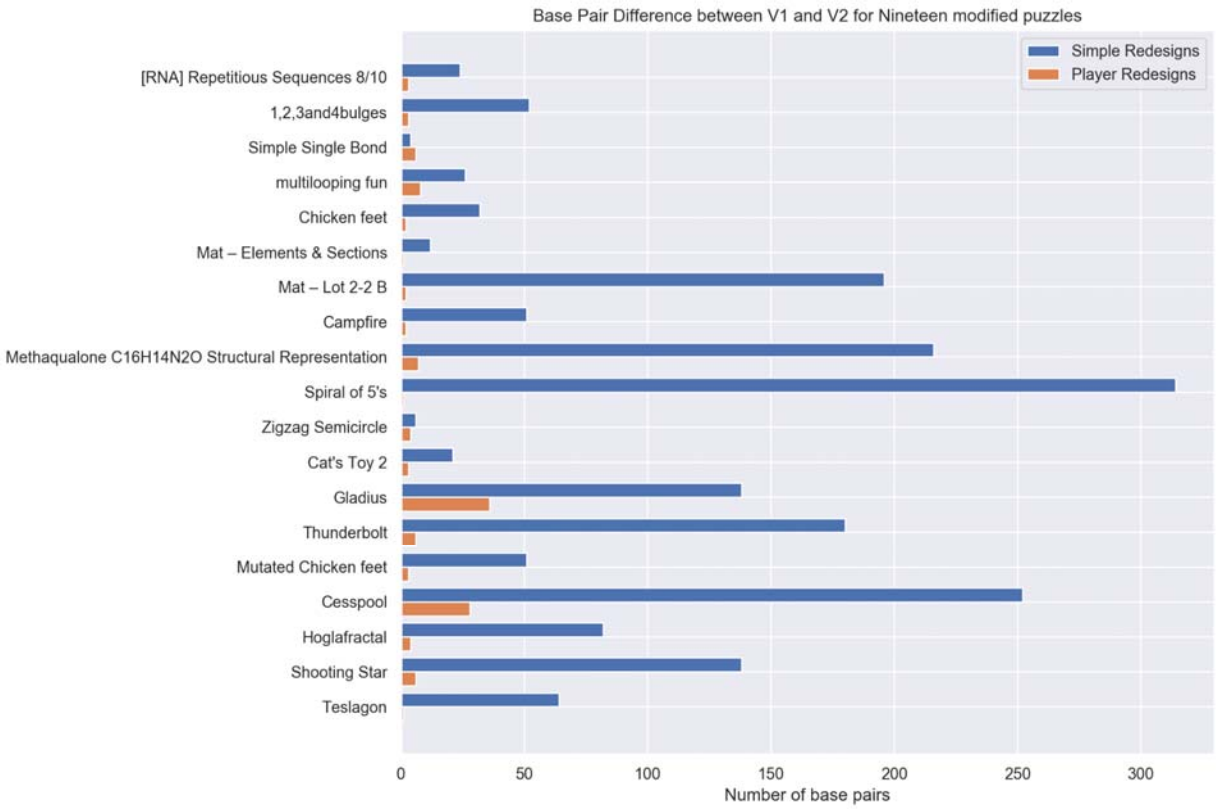


Figure 2. Number of nucleotide modifications for each of the 19 puzzles that were changed to be made solvable in Vienna 2. Blue denotes the RNAdistance calculated distance metric between the simple design algorithm between Vienna 1 and Vienna 2. Orange denotes differences in player designs between Vienna 1 and Vienna 2.

Eterna100-V1 and Eterna100-V2 discrepancies give insight into most difficult motifs

Given that the original Eterna100-V1 benchmark was designed by players, we postulated that players may be able to design modified structures for these puzzles using a minimum number of structure modifications, while maintaining the constraints of the Eterna software platform. Remarkably, many of the secondary structures required minimal modifications to be solvable with the Vienna 2 parameter set. Of particular interest are the puzzles “Teslagon” and “Shooting Star” given they were the most difficult of the original benchmark. “Teslagon” consists of a series of loops around a core 7-way junction and was made solvable through the deletion of a single internal loop base. “Shooting Star” consists of several multi-helix junctions and long helices that contain 29 isolated base pairs and was made solvable through only 3 base pair additions (Fig 3).

In contrast to these two puzzles, the puzzles “Gladius” and “Cesspool” both required a greater number of secondary structure modifications. The number of submissions for the

“Gladius” puzzle was limited. Eterna single-state puzzles enforce a 400 base pair length limit, and many of the submitted structures that minimized the number of modifications were shown to be greater than the 400 bp constraint. “Cesspool” exhibited a structure comprising 38 isolated base pairs and 4 symmetric 6-way junctions; as discussed previously [16], isolated base pairs, repeated motifs, and symmetric junctions make RNA design more difficult. The presence of these previously identified problem features are likely the explanation for why this structure required additional modifications. The total number of base pair changes from Vienna 1 to 2 can be seen in Fig 2.

Kyurem 5 and Multilooping 6 both had junctions without any unpaired bases, which are tougher to solve in Vienna 1 as opposed to Vienna 2. These loops have no unpaired bases, so the chance of a misfold is much higher. In addition, the orientation of the nucleotides in the junction matters and can vary the free energy, as Vienna 1 will penalize these structures more than Vienna 2 (Fig 2b). In both Kyurem 5 and Multilooping 6, the junctions have free energy 4.6 kcal/mol when surrounded by GC pairs in Vienna 1. In Vienna 2, these junctions have energy 3.5 kcal/mol, which allows all 5 algorithms to solve these 2 puzzles in Vienna 2, but only a handful solve these puzzles in Vienna 1.

In addition, large internal loops with several unpaired base pairs were also more difficult in the Vienna 2 folding engine than Vienna 1 folding engine. For example, [RNA] Repetitious Sequences 8/10 (Fig 4a) has 2 such motifs. This puzzle’s structure in Vienna 1 was deemed unsolvable due to the large internal loops being too unstable. This “unsolvability” can be attributed to the increased free energy calculations for these structures. For example, EternaBrain’s solution to [RNA] Repetitious Sequences 8/10’s Eterna100-V1 structure creates two large internal loops both with free energies of -1.0 kcal/mol. However, if the same structure and solution are used in Vienna 2 (Fig 4b), the free energies of both internal loops increases to 2.0 kcal/mol (Fig 2a). The strong G-C bonds in the 2 base pairs (free energy -3.3 kcal/mol total, in both Vienna 1 and Vienna 2) cannot keep the two loops separate, and the structure misfolds. No human or algorithm was able to solve the same Vienna 1 structure in Vienna 2. As a result, in the changed Vienna 2 structure for this puzzle, Eterna players deleted a single base, decreasing the free energy of the open loop enough to stabilize the secondary structure.

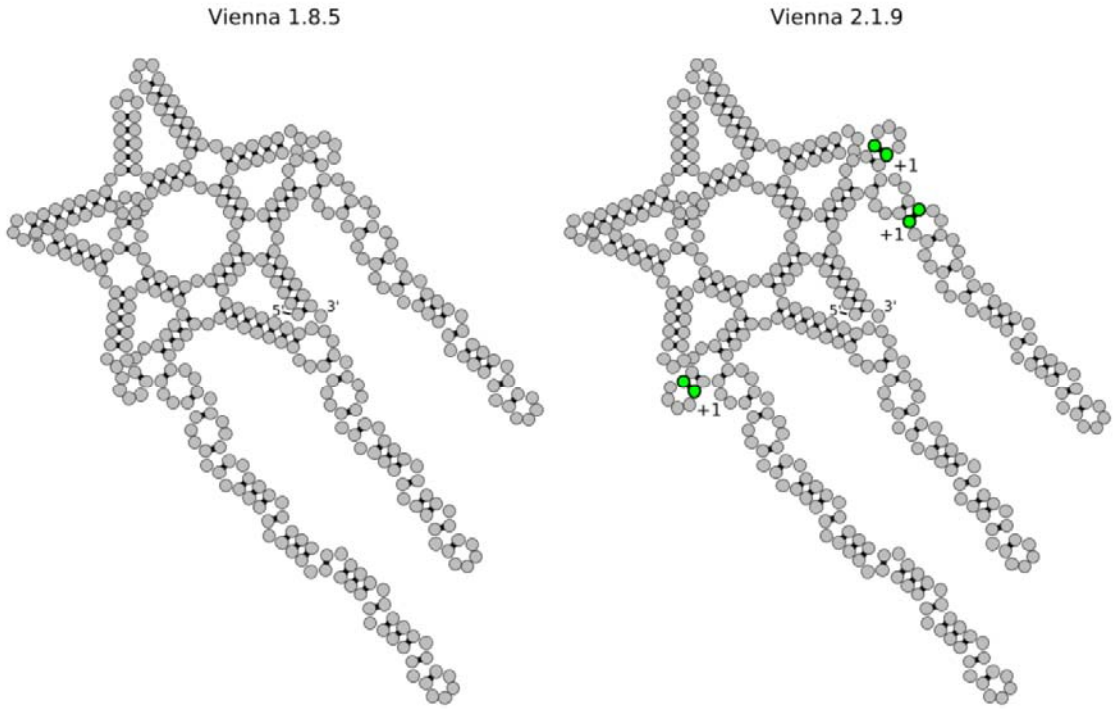


Figure 3. Base pair changes in Shooting Star from Vienna 1 to Vienna 2. The neon green nucleotides indicate the bases that were added to make the structure stable in Vienna 2.

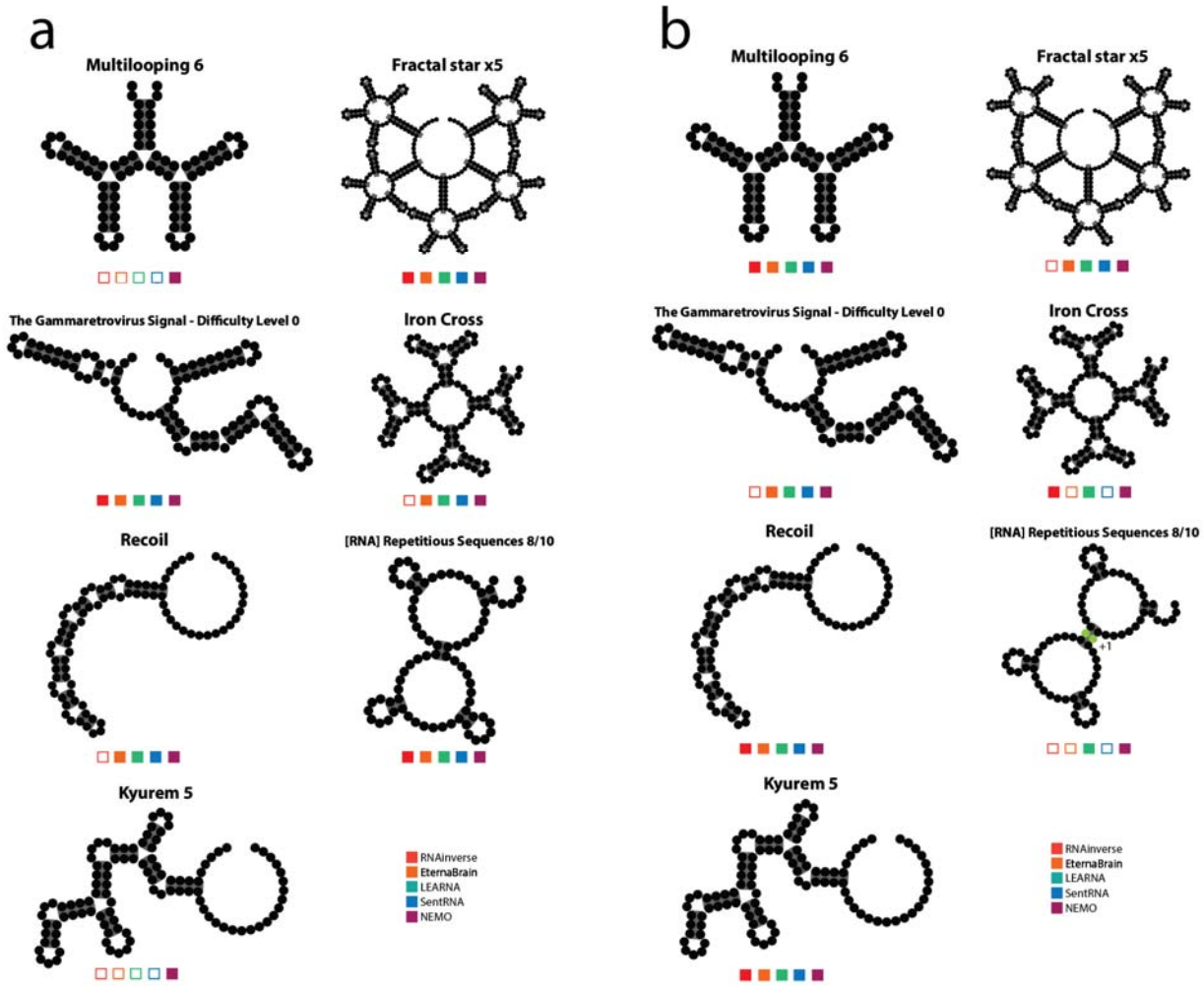


Figure 4. Selected Eterna100-V1 and -V2 puzzles that demonstrate differences in algorithms' puzzle-solving ability. Open squares indicate that a given algorithm was unable to solve that puzzle; filled squares indicate that the algorithm did solve the puzzle. Red: RNAinverse; Orange: EternaBrain; Teal: LEARNA; Blue: SentRNA; Purple: NEMO. (A) Eterna100-V1 puzzles. (B) Eterna100-V2 puzzles.

Inverse folding algorithm performance consistent

After players modified the secondary structures for the 19 Eterna100 puzzles, we assessed the performance of 5 RNA inverse folding algorithms on this updated benchmark (Fig 5) based on performance on the original Eterna100-V1. We chose to include RNAinverse based on historical significance, and we selected EternaBrain, SentRNA, and LEARNA as they are the best-performing neural network-based models. We selected NEMO because it has the highest performance of any algorithm on the Eterna100-V1. The algorithm NEMO had comparable performance against both benchmarks, solving 95 and 94 puzzles in Eterna100-V1 and Eterna100-V2, respectively. We found that the three algorithms based on neural network methods exhibited decreased performance in the parameters they were not trained against. EternaBrain was able to solve 66/100 puzzles on Eterna100-V1, but fewer (59/100) on

Eterna100-V2. SentRNA solved 78/100 on Eterna100-V1 and 69/100 on Eterna100-V2, while LEARNA scored 57/100 on Eterna100-V1 and 68/100 on Eterna100-V2. RNAinverse, a method that does not rely on neural networks, has two versions, Vienna 1 and Vienna 2, which performed similarly on both benchmarks, solving 47 and 49 out of 100 on Eterna100-V1 and Eterna100-V2, respectively.

The 19 structures that appeared unsolvable in Vienna 1 were of particular interest in our benchmark. These structures, in their Vienna 1 form, were some of the hardest secondary structures on the Eterna100-V1. For example, EternaBrain and SentRNA, both trained on Vienna 1 parameters, solved 5 (26%) and 8 (42%) out of 19, respectively, lower than their average across all Eterna100-V1 puzzles (66% and 78%, respectively). The only algorithm to perform well on these structures in the original Eterna100-V1 benchmark was NEMO, which directly uses player strategies within a nested Monte Carlo algorithm, solving 15/19. On the Vienna 2-modified secondary structures of Eterna100-V2, EternaBrain and SentRNA both solved fewer puzzles: EternaBrain solving 1 and SentRNA solving 3. This was expected, as both inferred Vienna 1 solving strategies, either learned via neural networks or via explicitly encoded strategies in the algorithms [24]. Similarly, LEARNA solved four puzzles, two more than its Vienna 1 performance. NEMO was able to solve the same number of these puzzles in Eterna100-V2 as Eterna100-V1, and RNAinverse solved 0 of the 19, the same performance as in the Vienna 1 benchmark.

Eterna100-V1
Puzzle
Simple Hairpin
Arabidopsis Thaliana 6 RNA - Difficulty Level 0
Prion Pseudoknot - Difficulty Level 0
Human Integrated Adenovirus - Difficulty Level 0
The Gammaretrovirus Signal - Difficulty Level 0
Saccharomyces Cerevisiae - Difficulty Level 0
Fractal 2
G-C Placement
The Sun
Frog Foot
InfoRNA test 16
Mat - Martian 2
square
Six legd turtle 2
Small and Easy 6
Fractile
Six legd Turtle
snoRNA SNORD64
Chalk Outline
InfoRNA bulge test 9
Tilted Russian Cross
This is ACTUALLY Small And Easy 6
Shortie 4
Shape Test
The Minitsry
stickshift
U
Still Life (Sunflower In A Vase)
Quasispecies 2-2 Loop Challenge
Corner bulge training
Spiral
InfoRNA bulge test
Worm 1
just down to 1 bulge
Iron Cross
loops and stems
Water Strider
The Turtle(s) Move(s)
Adenine
Tripod5
Shortie 6
Runner
Recoil
[CloudBeta] An Arm and a Leg 1.0
[CloudBeta] 5 Adjacent Stack Multi-Branch Loop
Triple Y
Misfolded Aptamer
Flower power
Kudzu
1,2,3and4bulges
medallion
[RNA] Repetitious Sequences 8/10
Documenting repetitious behavior
7 multiloop
Kyurem 7
JF1
multilooping fun
Multiloop...
hard Y
Mat - Elements & Sections
Chicken feet
Bug 18
Fractal star x5
Crop circle 2
Branching Loop
Bug 38
Simple Single Bond
Taraxacum officinale
Headless Bug on Windshield
Pokeball
Variation of a crop circle
Loop next to a Multiloop
Snowflake 4
Mat - Cuboid
Misfolded Aptamer 6
Snowflake 3
Hard Y and a bit more
Mat - Lot 2-2 B
Shapes and Energy
Spiral of 5s
Campfire
Anemone
Fractal 3
Kyurem 5
Snowflake Necklace (or v2.0)
Methagualone C16H14N2O Structural Representation
Cats Toy 2
Zigzag Semicircle
Short String 4
Gladius
Thunderbolt
Mutated chicken feet
Chicken Tracks
Looking Back Again
Multilooping 6
Cesspool
Hoglafractal
Bullseye
Shooting Star
Teslagon
RNAinverse...41
EternaBrain...66
LEARNA...56
SentRNA...78
NEMO...95

Eterna100-V2
Puzzle
Simple Hairpin
Arabidopsis Thaliana 6 RNA - Difficulty Level 0
Prion Pseudoknot - Difficulty Level 0
Human Integrated Adenovirus - Difficulty Level 0
The Gammaretrovirus Signal - Difficulty Level 0
Saccharomyces Cerevisiae - Difficulty Level 0
Fractal 2
G-C Placement
The Sun
Frog Foot
InfoRNA test 16
Mat - Martian 2
square
Six legd turtle 2
Small and Easy 6
Fractile
Six legd Turtle
snoRNA SNORD64
Chalk Outline
InfoRNA bulge test 9
Tilted Russian Cross
This is ACTUALLY Small And Easy 6
Shortie 4
Shape Test
The Minitsry
stickshift
U
Still Life (Sunflower In A Vase)
Quasispecies 2-2 Loop Challenge
Corner bulge training
Spiral
InfoRNA bulge test
Worm 1
just down to 1 bulge
Iron Cross
loops and stems
Water Strider
The Turtle(s) Move(s)
Adenine
Tripod5
Shortie 6
Runner
Recoil
[CloudBeta] An Arm and a Leg 1.0
[CloudBeta] 5 Adjacent Stack Multi-Branch Loop
Triple Y
Misfolded Aptamer
Flower power
Kudzu
1,2,3and4bulges
medallion
[RNA] Repetitious Sequences 8/10
Documenting repetitious behavior
7 multiloop
Kyurem 7
JF1
multilooping fun
Multiloop...
hard Y
Mat - Elements & Sections
Chicken feet
Bug 18
Fractal star x5
Crop circle 2
Branching Loop
Bug 38
Simple Single Bond
Taraxacum officinale
Headless Bug on Windshield
Pokeball
Variation of a crop circle
Loop next to a Multiloop
Snowflake 4
Mat - Cuboid
Misfolded Aptamer 6
Snowflake 3
Hard Y and a bit more
Mat - Lot 2-2 B
Shapes and Energy
Spiral of 5s
Campfire
Anemone
Fractal 3
Kyurem 5
Snowflake Necklace (or v2.0)
Methagualone C16H14N2O Structural Representation
Cats Toy 2
Zigzag Semicircle
Short String 4
Gladius
Thunderbolt
Mutated chicken feet
Chicken Tracks
Looking Back Again
Multilooping 6
Cesspool
Hoglafractal
Bullseye
Shooting Star
Teslagon
RNAinverse...49
EternaBrain...59
LEARNA...68
SentRNA...69
NEMO...94

Figure 5. Performance of the 5 algorithms mentioned on the Eterna100-V1 and Eterna100-V2. Green: Solved; Orange: Unsolved.

Discussion

In this work, we demonstrated that 19 of the 100 structures in the widely-used Eterna100 benchmark for evaluating RNA inverse design were unsolvable when the thermodynamic parameters of Vienna 1 were substituted with those of Vienna 2. This potentially presents a problem for evaluating inverse folding algorithms, if algorithmic potential is intrinsically limited by the thermodynamic model. To amend this problem, we asked Eterna participants to redesign the 19 puzzles. Participants found strategies to do so that preserved the original challenges of the puzzles in ways that would not have been achievable by simply updating the folding engine. These structure modifications highlight how different energy parameters alter the solvability of RNA secondary structures, with the minimal modifications to the benchmark's secondary structures generating a number of motifs that require more stringent sequence features compared to the original.

We next evaluated state-of-the-art algorithms on this updated benchmark. We found that algorithms based on neural networks (LEARNA, SentRNA, EternaBrain) exhibited worse performance on the benchmark using the folding algorithm they were not trained on. Algorithms with specific Eterna player strategies or strategies that used stochastic iterative folding (NEMO, RNAinverse) exhibited more consistent performance across the two benchmarks. While the neural network models were modified to use Vienna 2 parameters, some of the hard-coded strategies were not modified. This result suggests that for optimal performance, neural network based algorithms will need to be retrained with other parameter sets for use in RNA secondary structure design, or find folding-engine-agnostic methods, such as the stochastic methods used by NEMO and RNAinverse.

Given the number of secondary structures that were unsolvable in the original Eterna100 benchmark with the Vienna2 parameters, it seems likely this benchmark will need to be continuously updated as RNA structure prediction becomes more accurate. Folding engines like Contrafold [28] and Eternafold [29] appear to be more accurate than the Vienna folding engines. However, the newer folding engines do not rely on the same body of experimental results as the Vienna folding engines and are anticipated to deviate even farther from those established in ViennaRNA, leading to even more of these benchmark secondary structures needing to be modified.

Taken together, this work indicates that future RNA inverse folding algorithms should strive to be folding engine independent and we hope that easy availability of Eterna100-V1 and Eterna100-V2 will enable the testing of such algorithms. As folding engines evolve and advance, it will be necessary that inverse folding algorithms keep up with these advancements. Therefore,

it will be necessary in the future for RNA design algorithms to be able to predict designs for multiple folding engines or be easily retrained on different folding engine thermodynamic parameters.

Methods

Design challenges on Eterna platform

Through iterative manual design, we identified 19 secondary structures in the original Eterna100 (Eterna100-V1) benchmark that we hypothesized were not solvable using Vienna 1.8.5. We asked the Eterna community to redesign these 19 puzzles to be compliant with the thermodynamic parameters of the ViennaRNA 2.1.8 software package as implemented in Eterna. Players achieved this through the Eterna "Puzzlemaker" interface, whereby individual base pairs and bases may be deleted or added to a structure (Fig 6). Players submitted 52 total puzzles as modifications of the unsolvable 19. Nineteen of those submissions were chosen by the authors to both maintain the constraints of the Eterna platform and the identity of unique structural elements that existed in the original puzzles.

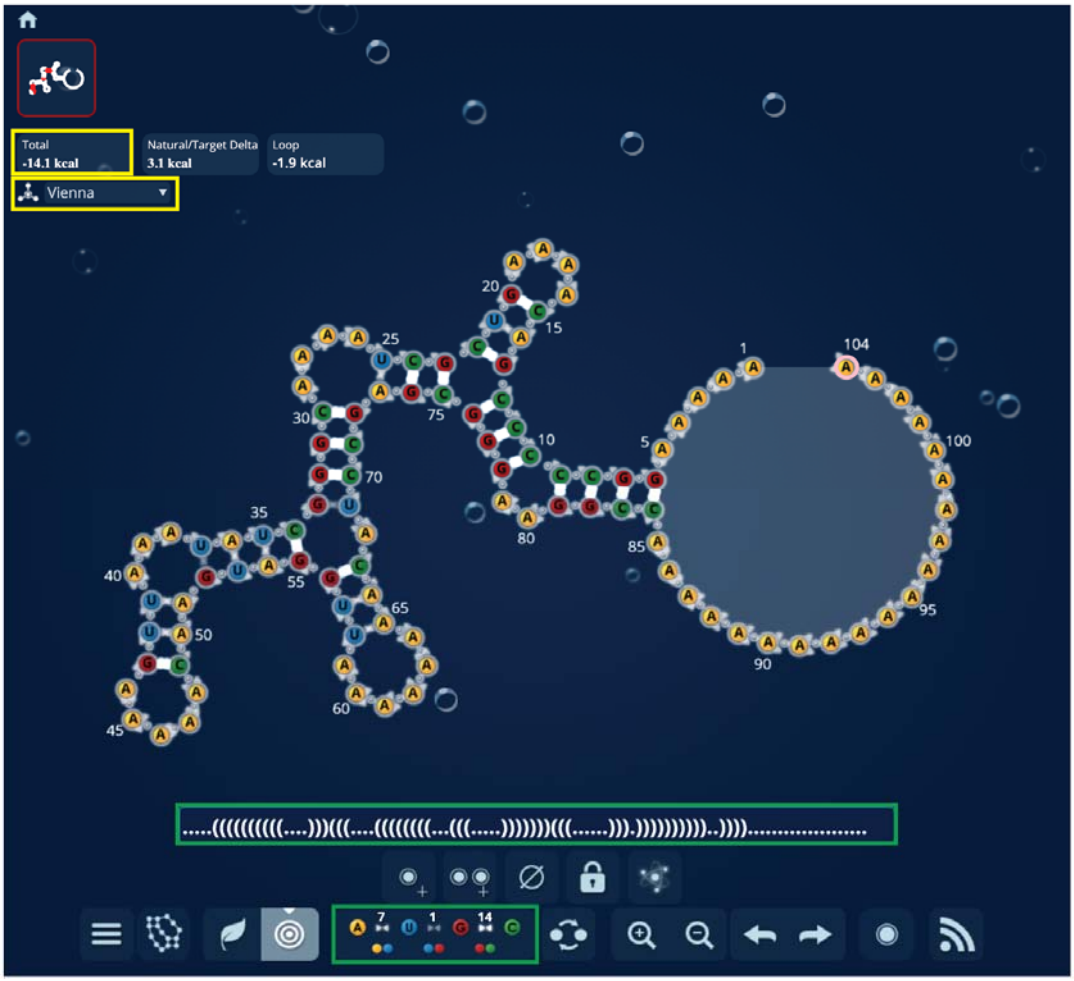


Figure 6. A screenshot of Eterna's "Puzzlemaker" interface. Players can see the free energy and folding engine in the top left corner (yellow rectangles). In the bottom, players can add bases, add base pairs, and interact with the dot-bracket representation of the structure directly (green rectangles).

Automated tests of RNA secondary structure design algorithms

We evaluated the performance of five algorithms (EternaBrain, SentRNA, RNAinverse, LEARNA, and NEMO) on the Eterna100-V1 and on the Eterna100-V2 with the 19 modified puzzles, using Vienna 2.1.8 as the folding engine. EternaBrain, SentRNA, RNAinverse, and LEARNA were run on a Google Cloud instance with 4 CPUs and 10 GB of RAM. Puzzles were benchmarked with a timeout of 2 hours.

EternaBrain uses a combination of a convolutional neural network trained on Eterna player moves and a Single Action Playout (SAP) [24], a depth-1 Monte Carlo Search using Eterna player strategies. To adapt EternaBrain to Vienna 2, the folding engine in the SAP was changed to Vienna 2. EternaBrain's Convolutional Neural Network (CNN) was not retrained, as the player moves used to train the CNN were all using Vienna 1 originally, and insufficient data exist to train a neural network with Vienna 2 player moves.

SentRNA, unlike EternaBrain, does not rely on player moves to train its deep neural networks; instead, the authors define their own set of features in the Methods section of the paper [25]. We retrained SentRNA using the Vienna 2.4.9 energy model, to benchmark on the Eterna100-V2. On both V1 and V2, we trained an ensemble of 20 networks with 300 "solution trajectories" [25]. For LEARNA, no changes were needed to benchmark it against the Eterna100-V2, as the algorithm was already using Vienna 2 as its internal folding engine (via the Anaconda bindings for Vienna 2). To benchmark it on Eterna100-V1, we changed the internal folding engine to Vienna 1 and retrained the reinforcement learning models, keeping the same hyperparameter values.

RNAinverse and NEMO, which are both stochastic methods, were able to be run without modification using either Vienna 1 or Vienna 2 as the internal folding engine.

RNAinverse was benchmarked using standard settings, allowing either Vienna 1.8.5 or Vienna 2.4.8 to be used as the folding algorithm. For each puzzle and Vienna folding engine version (1.8.5 or 2.4.10), NEMO was run for a maximum of 24 hours or 1000 independent trials, using one node per puzzle on the Stanford University Sherlock cluster (16 cores, 64 GB RAM, 100 GB local SSD). Each trial was allowed a maximum of 10^4 Monte Carlo iterations. For the "unsolvable" puzzles, i.e. the 19 original puzzles run in Vienna 2, NEMO was given 10 independent 24-hour periods to attempt to solve. No solutions were found.

Acknowledgements

We thank the Eterna structure designers for creating the updated 19 puzzles in Eterna100-V2 with minimal changes and for solving them (player names: Brourd, Tesla'sDisciple, Astromon,

c-quence, JR, Jieux, jnicol, clollin). Calculations and model training were performed on the Stanford Sherlock cluster and Berkeley CSUA cluster. We acknowledge funding from the National Science Foundation (GRFP to H.K.W.S.) and the National Institutes of Health (R35 GM122579 to R.D.).

Author Contributions Statement

BR and RD formulated the problem. RVK and HKWS ran experiments. RVK and BR carried out analyses and generated figures. Eterna Structure Designers designed the secondary structures. RVK, BR, HKWS, and RD wrote the manuscript.

Figure Legends

Figure 1. (A) Free energy differences (ΔΔG, kcal/mol) between identical structures in Vienna 1.8 and Vienna 2.4. Topmost: an internal triloop; middle: U-G-U-G superboost with G-C end pairs; bottom: tetraloop hairpin with G-boost and G-C end pairs. (B) Puzzle-specific differences in free energies of structures. The internal tri-junction in Kyurem 5 has a more stable free energy in Vienna 2. Similar repeated tri-junctions are highlighted in Multilooping 6, with similar changes in free energy from Vienna 1 to 2.
Figure 2. Number of nucleotide modifications for each of the 19 puzzles that were changed to be made solvable in Vienna 2. Blue denotes RNAdistance calculated distance metric between the simple design algorithm involving calculating structure differences between Vienna 1 and Vienna 2. Orange denotes differences in player designs between Vienna 1 and Vienna 2.
Figure 3. Base pair changes in Shooting Star from Vienna 1 to Vienna 2. The neon yellow nucleotides indicate the bases that were added to make the structure stable in Vienna 2.
Figure 4. Selected Eterna100-V1 and -V2 puzzles that demonstrate differences in algorithms' puzzle-solving ability. Open squares indicate that a given algorithm was unable to solve that puzzle; filled squares indicate that the algorithm did solve the puzzle. Red: RNAinverse; Orange: EternaBrain; Teal: LEARNA; Blue: SentRNA; Purple: NEMO. (A) Eterna100-V1 puzzles.
Figure 5. Performance of the 5 algorithms mentioned in the paper on the Eterna100-V1 and Eterna100-V2. Green: Solved; Orange: Unsolved.
Figure 6. A screenshot of Eterna's "Puzzlemaker" interface. Players can see the free energy and folding engine in the top left corner (yellow rectangles). In the bottom, players can add bases, base pairs and interact with the dot-bracket representation of the structure directly (green rectangles).

References

1. Serganov A, Nudler E. A decade of riboswitches. Cell. 2013;152: 17–24. doi:10.1016/j.cell.2012.12.024
2. Jo B-S, Choi SS. Introns: the functional benefits of introns in genomes. Genomics Inform. 2015;13: 112–118. doi:10.5808/GI.2015.13.4.112
3. Will CL, Lührmann R. Spliceosome structure and function. Cold Spring Harb Perspect Biol. 2011;3. doi:10.1101/cshperspect.a003707
4. Haerty W, Ponting CP. Mutations within lncRNAs are effectively selected against in fruitfly but not in human. Genome Biol. 2013;14: R49. doi:10.1186/gb-2013-14-5-r49
5. Song M-S, Rossi JJ. Molecular mechanisms of Dicer: endonuclease and enzymatic activity. Biochem J. 2017;474: 1603–1618. doi:10.1042/BCJ20160759
6. Doudna JA, Charpentier E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. Science. 2014;346: 1258096. doi:10.1126/science.1258096
7. Chen J-L, Greider CW. Telomerase RNA structure and function: implications for dyskeratosis congenita. Trends Biochem Sci. 2004;29: 183–192. doi:10.1016/j.tibs.2004.02.003
8. Singh RN, Singh NN. Mechanism of splicing regulation of spinal muscular atrophy genes. Adv Neurobiol. 2018;20: 31–61. doi:10.1007/978-3-319-89689-2_2
9. Ng EWM, Shima DT, Calias P, Cunningham ET, Guyer DR, Adamis AP. Pegaptanib, a targeted anti-VEGF aptamer for ocular vascular disease. Nat Rev Drug Discov. 2006;5: 123–132. doi:10.1038/nrd1955
10. Schlick T, Pyle AM. Opportunities and challenges in RNA structural modeling and design. Biophys J. 2017;113: 225–234. doi:10.1016/j.bpj.2016.12.037
11. Leontis NB, Westhof E. Chemistry. Self-assembled RNA nanostructures. Science. 2014;345: 732–733. doi:10.1126/science.1257989
12. Geary C, Chworos A, Verzemnieks E, Voss NR, Jaeger L. Composing RNA Nanostructures from a Syntax of RNA Structural Modules. Nano Lett. 2017;17: 7095–7101. doi:10.1021/acs.nanolett.7b03842
13. Monferrer A, Zhang D, Lushnikov AJ, Hermann T. Versatile kit of robust nanoshapes self-assembling from RNA and DNA modules. Nat Commun. 2019;10: 608. doi:10.1038/s41467-019-08521-6
14. Jain S, Schlick T. F-RAG: Generating Atomic Coordinates from RNA Graphs by Fragment Assembly. J Mol Biol. 2017;429: 3587–3605. doi:10.1016/j.jmb.2017.09.017
15. Yesselman JD, Eiler D, Carlson ED, Gotrik MR, d'Aquino AE, Ooms AN, et al. Computational design of three-dimensional RNA structure and function. Nat Nanotechnol.

2019;14: 866–873. doi:10.1038/s41565-019-0517-8

16. Anderson-Lee J, Fisker E, Kosaraju V, Wu M, Kong J, Lee J, et al. Principles for predicting RNA secondary structure design difficulty. J Mol Biol. 2016;428: 748–757. doi:10.1016/j.jmb.2015.11.013

17. Zuker M. On finding all suboptimal foldings of an RNA molecule. Science. 1989;244: 48–52. doi:10.1126/science.2468181

18. Hofacker IL. Vienna RNA secondary structure server. Nucleic Acids Res. 2003;31: 3429–3431. doi:10.1093/nar/gkg599

19. Andronescu M, Fejes AP, Hutter F, Hoos HH, Condon A. A new algorithm for RNA secondary structure design. J Mol Biol. 2004;336: 607–624. doi:10.1016/j.jmb.2003.12.041

20. Zadeh JN, Steenberg CD, Bois JS, Wolfe BR, Pierce MB, Khan AR, et al. NUPACK: Analysis and design of nucleic acid systems. J Comput Chem. 2011;32: 170–173. doi:10.1002/jcc.21596

21. Taneda A. MODENA: a multi-objective RNA inverse folding. Adv Appl Bioinform Chem. 2011;4: 1–12. doi:10.2147/aabc.s14335

22. Busch A, Backofen R. INFO-RNA--a server for fast inverse RNA folding satisfying sequence constraints. Nucleic Acids Res. 2007;35: W310-3. doi:10.1093/nar/gkm218

23. Griffiths-Jones S, Bateman A, Marshall M, Khanna A, Eddy SR. Rfam: an RNA family database. Nucleic Acids Res. 2003;31: 439–441. doi:10.1093/nar/gkg006

24. Koodli RV, Keep B, Coppess KR, Portela F, Eterna participants, Das R. EternaBrain: Automated RNA design through move sets and strategies from an Internet-scale RNA videogame. PLoS Comput Biol. 2019;15: e1007059. doi:10.1371/journal.pcbi.1007059

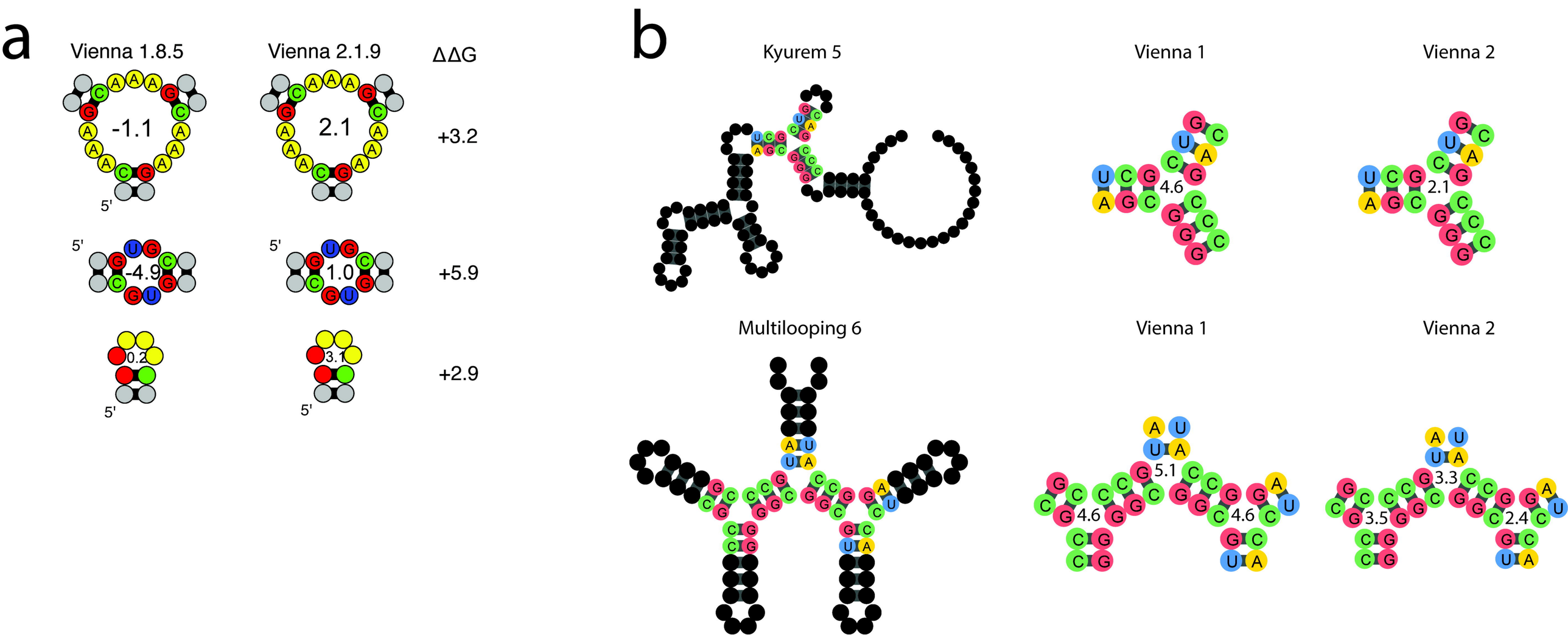
25. Shi J, Das R, Pande VS. SentRNA: Improving computational RNA design by incorporating a prior of human design strategies. arXiv preprint arXiv: 180303146. 2018.

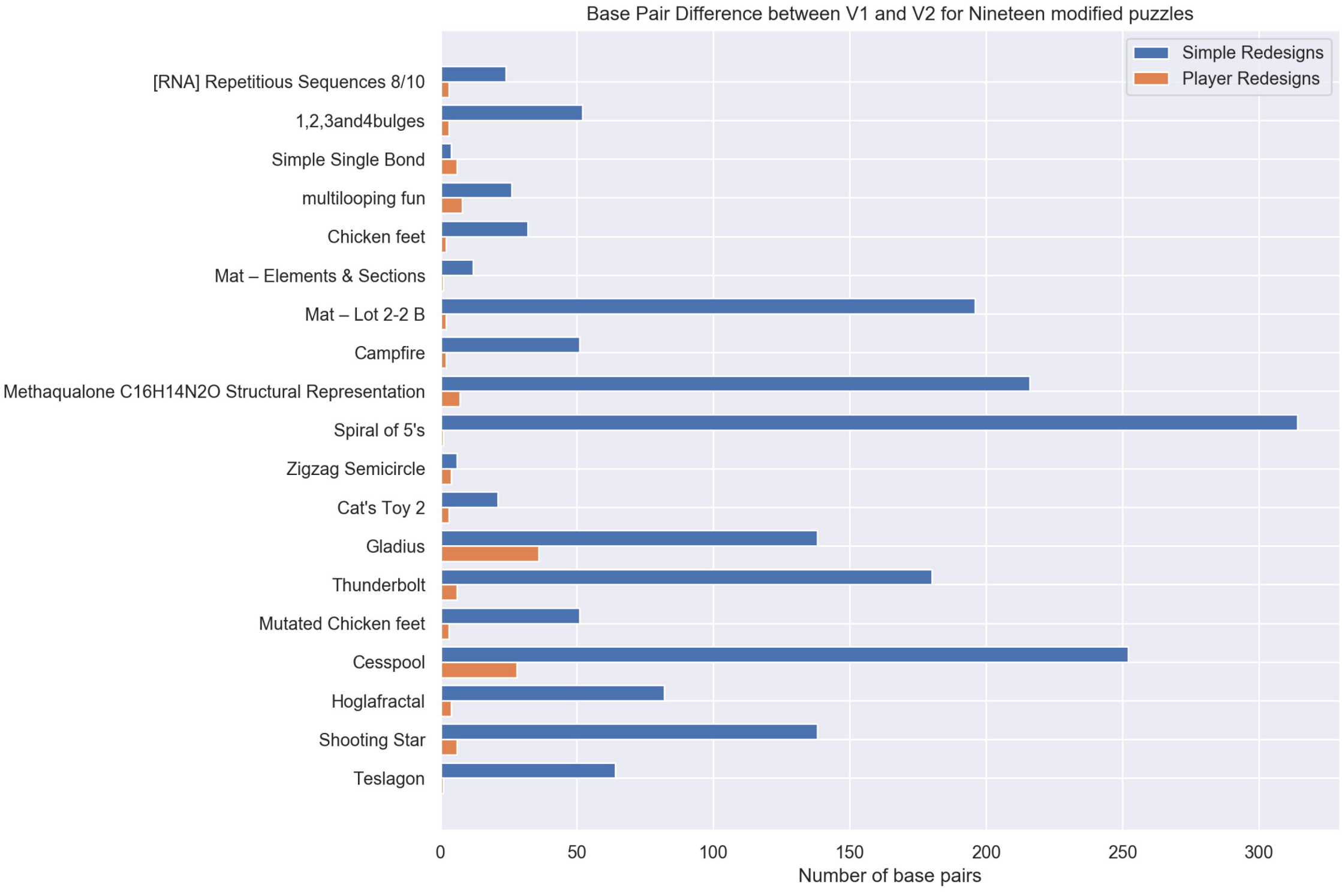
26. Runge F, Stoll D, Falkner S, Hutter F. Learning to Design RNA. ICLR 2019. 2018.

27. Portela F. An unexpectedly effective Monte Carlo technique for the RNA inverse folding problem. BioRxiv. 2018. doi:10.1101/345587

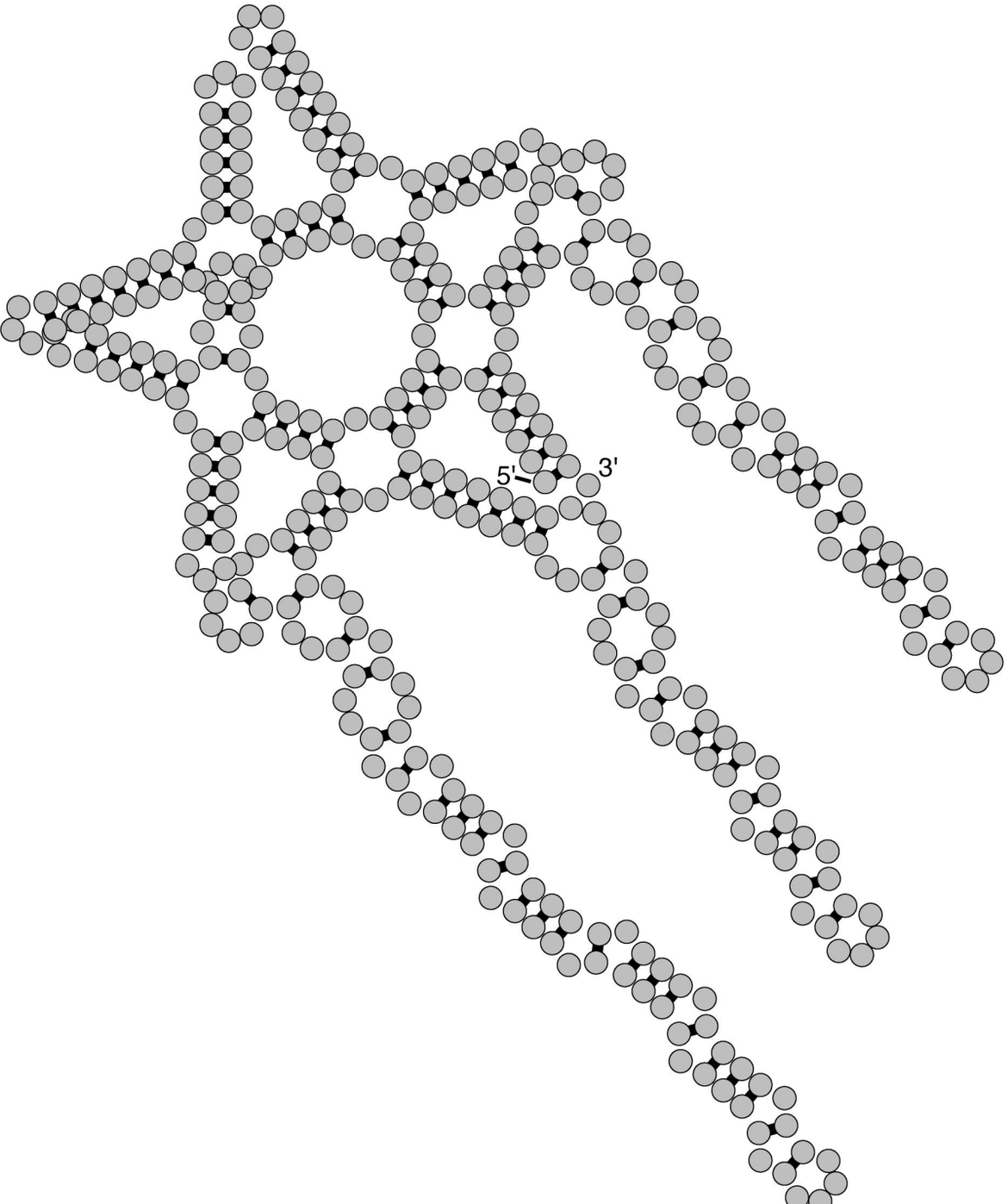
28. Do CB, Woods DA, Batzoglou S. CONTRAfold: RNA secondary structure prediction without physics-based models. Bioinformatics. 2006;22: e90-8. doi:10.1093/bioinformatics/btl246

29. Wayment-Steele HK, Kladwang W, Eterna Participants, Das R. RNA secondary structure packages ranked and improved by high-throughput experiments. BioRxiv. 2020. doi:10.1101/2020.05.29.124511

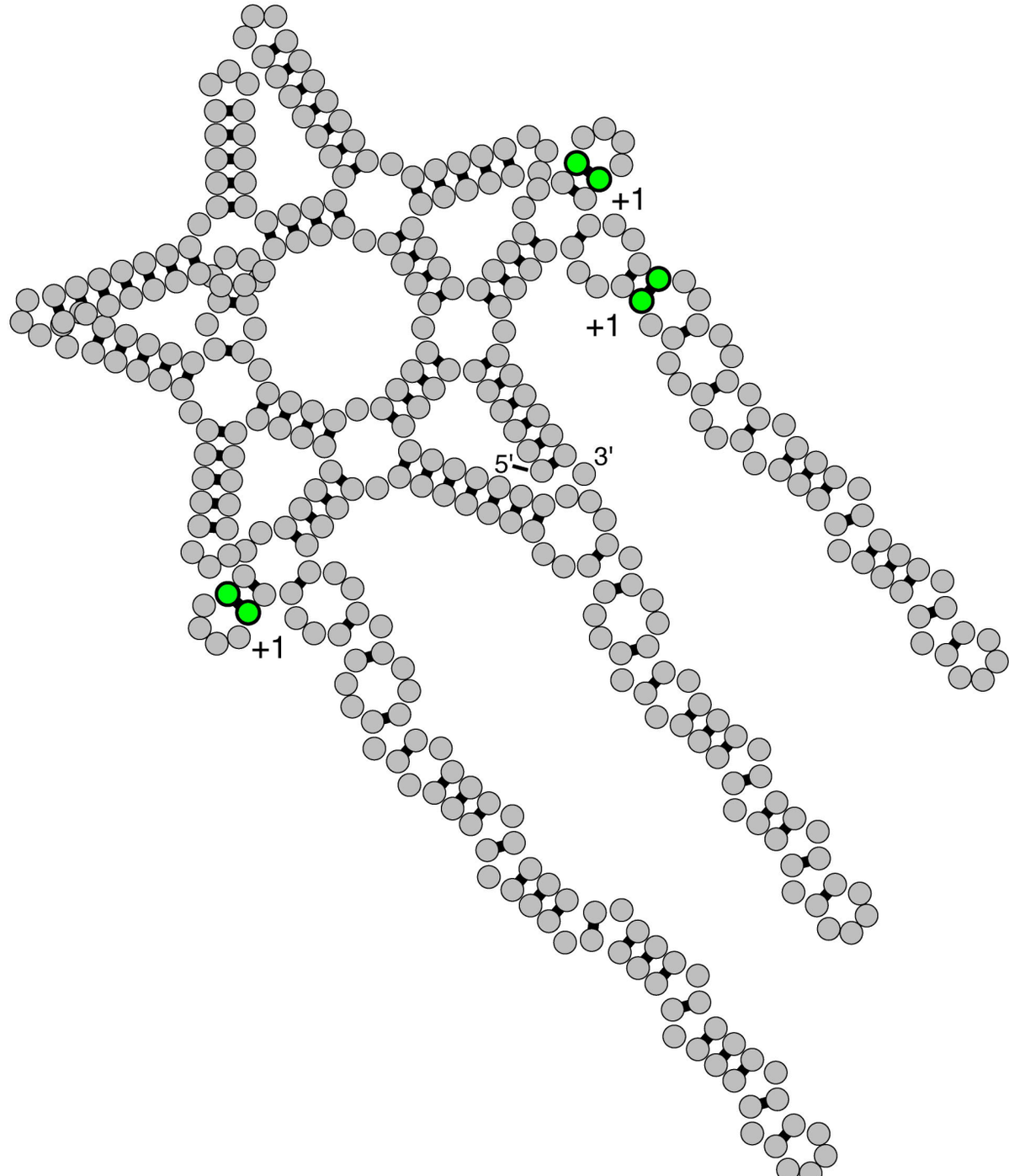

a
Vienna 1.8.5
Vienna 2.1.9
ΔΔG
-1.1
2.1
+3.2
5'
-4.9
1.0
+5.9
0.2
3.1
+2.9
b
Kyurem 5
Vienna 1
Vienna 2
4.6
2.1
Multilooping 6
Vienna 1
Vienna 2
5.1
4.6
4.6
3.3
3.5
2.4

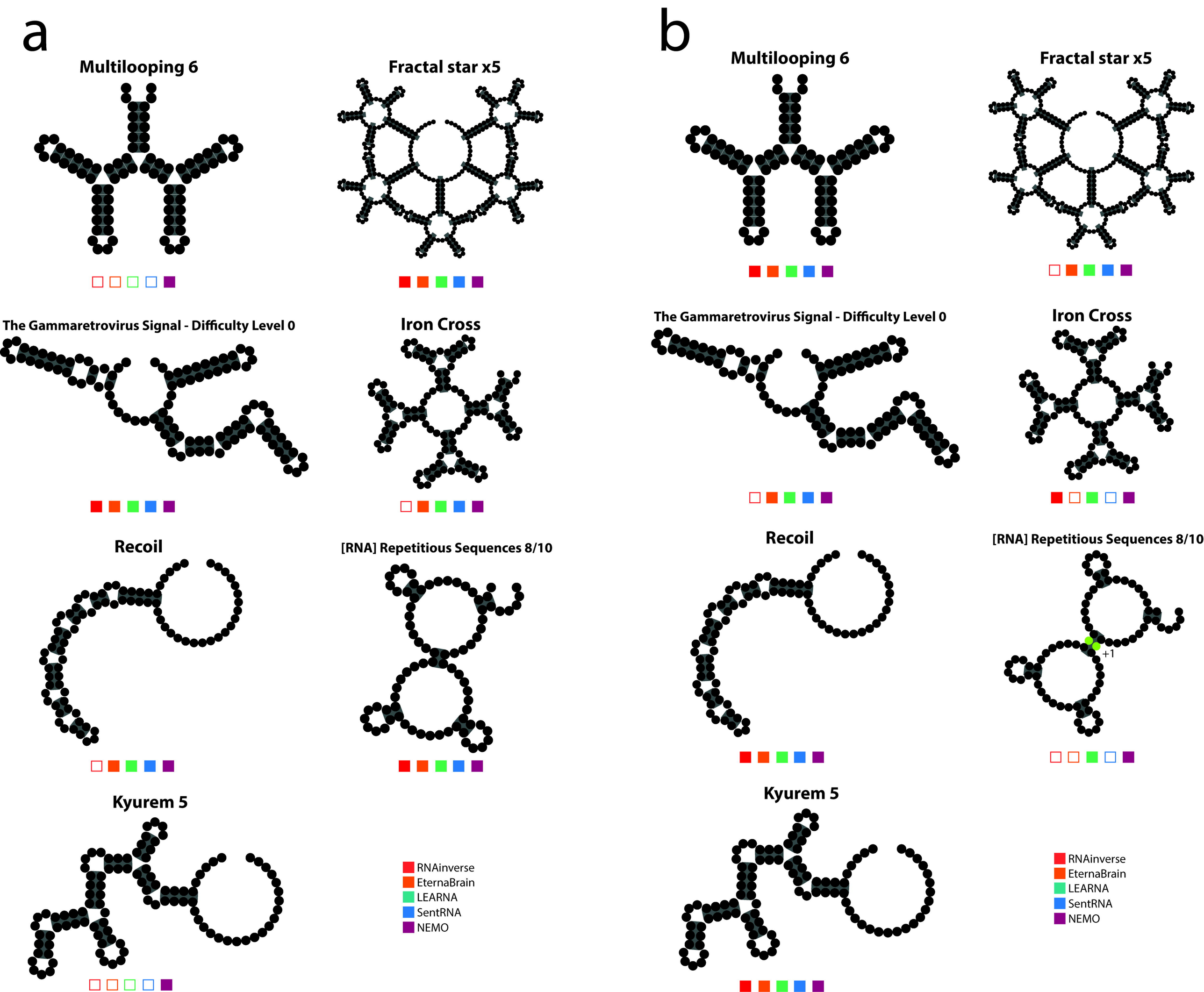

Base Pair Difference between V1 and V2 for Nineteen modified puzzles
Simple Redesigns
Player Redesigns
[RNA] Repetitious Sequences 8/10
1,2,3and4bulges
Simple Single Bond
multilooping fun
Chicken feet
Mat – Elements & Sections
Mat – Lot 2-2 B
Campfire
Methaqualone C16H14N2O Structural Representation
Spiral of 5's
Zigzag Semicircle
Cat's Toy 2
Gladius
Thunderbolt
Mutated Chicken feet
Cesspool
Hoglafractal
Shooting Star
Teslagon
0
50
100
150
200
250
300
Number of base pairs

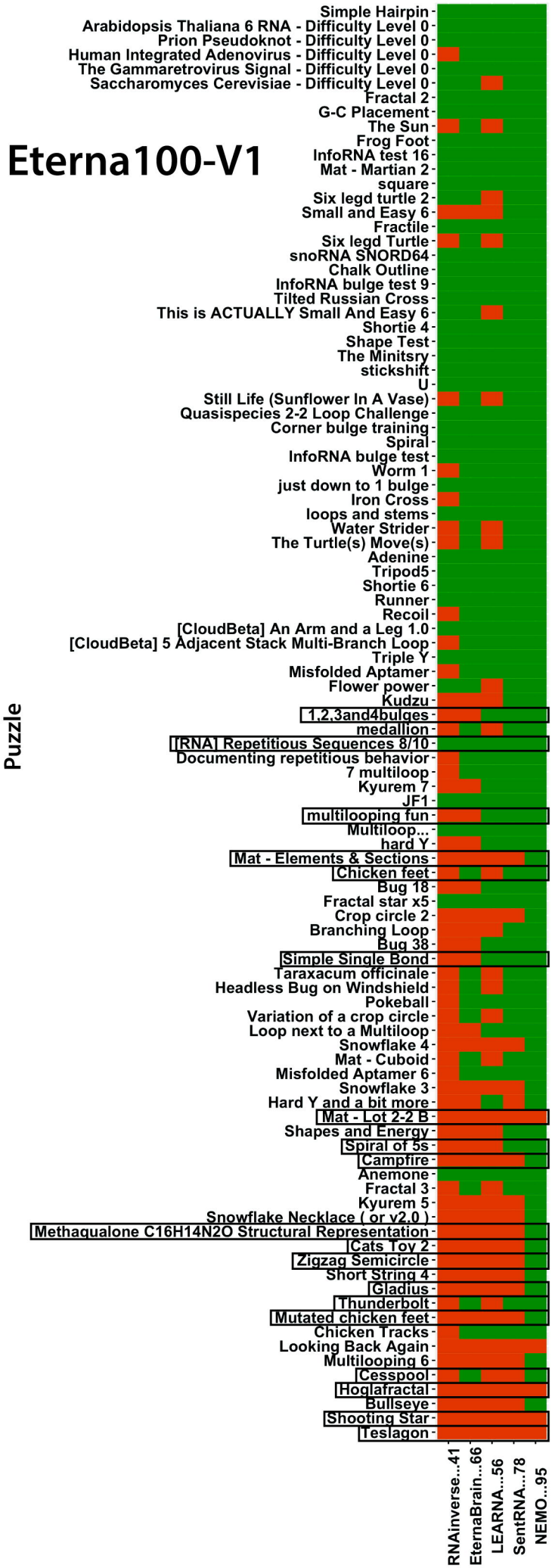
Vienna 1.8.5

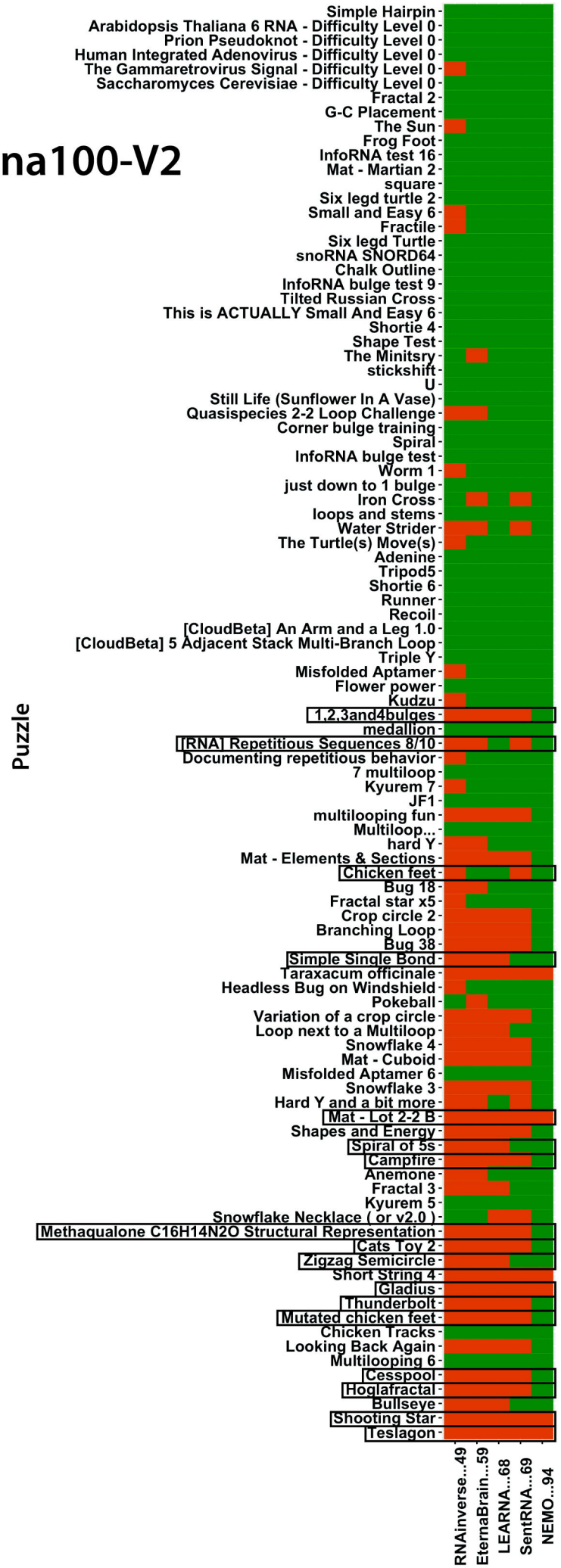


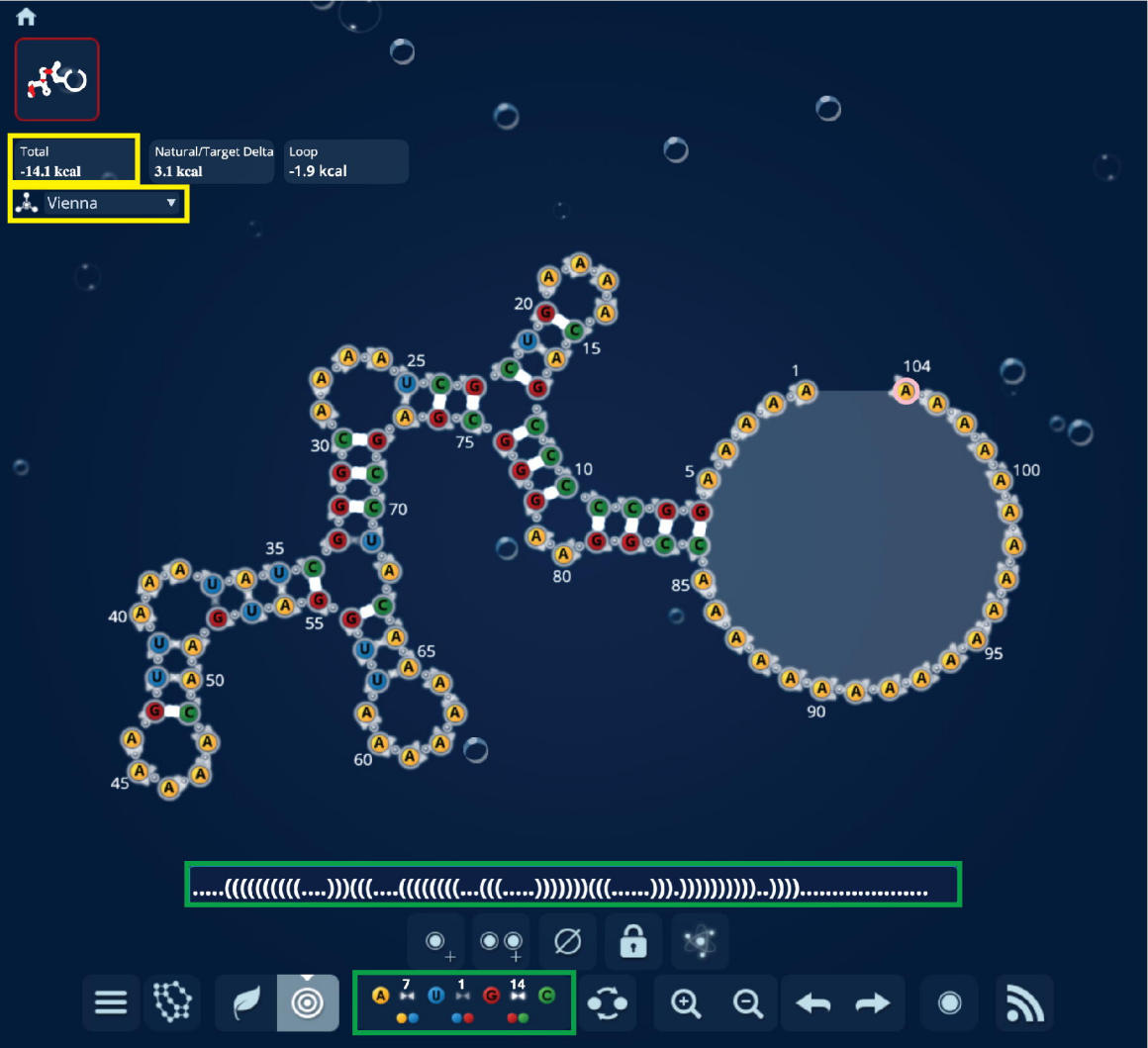
Vienna 2.1.9




a
Multilooping 6
Fractal star x5
The Gammaretrovirus Signal - Difficulty Level 0
Iron Cross
Recoil
[RNA] Repetitious Sequences 8/10
Kyurem 5
RNAinverse
EternaBrain
LEARNA
SentRNA
NEMO
b
Multilooping 6
Fractal star x5
The Gammaretrovirus Signal - Difficulty Level 0
Iron Cross
Recoil
[RNA] Repetitious Sequences 8/10
+1
Kyurem 5
RNAinverse
EternaBrain
LEARNA
SentRNA
NEMO


Eterna100-V1
Puzzle
Simple Hairpin
Arabidopsis Thaliana 6 RNA - Difficulty Level 0
Prion Pseudoknot - Difficulty Level 0
Human Integrated Adenovirus - Difficulty Level 0
The Gammaretrovirus Signal - Difficulty Level 0
Saccharomyces Cerevisiae - Difficulty Level 0
Fractal 2
G-C Placement
The Sun
Frog Foot
InfoRNA test 16
Mat - Martian 2
square
Six legd turtle 2
Small and Easy 6
Fractile
Six legd Turtle
snoRNA SNORD64
Chalk Outline
InfoRNA bulge test 9
Tilted Russian Cross
This is ACTUALLY Small And Easy 6
Shortie 4
Shape Test
The Minitsry
stickshift
U
Still Life (Sunflower In A Vase)
Quasispecies 2-2 Loop Challenge
Corner bulge training
Spiral
InfoRNA bulge test
Worm 1
just down to 1 bulge
Iron Cross
loops and stems
Water Strider
The Turtle(s) Move(s)
Adenine
Tripod5
Shortie 6
Runner
Recoil
[CloudBeta] An Arm and a Leg 1.0
[CloudBeta] 5 Adjacent Stack Multi-Branch Loop
Triple Y
Misfolded Aptamer
Flower power
Kudzu
1,2,3and4bulges
medallion
[RNA] Repetitious Sequences 8/10
Documenting repetitious behavior
7 multiloop
Kyurem 7
JF1
multilooping fun
Multiloop...
hard Y
Mat - Elements & Sections
Chicken feet
Bug 18
Fractal star x5
Crop circle 2
Branching Loop
Bug 38
Simple Single Bond
Taraxacum officinale
Headless Bug on Windshield
Pokeball
Variation of a crop circle
Loop next to a Multiloop
Snowflake 4
Mat - Cuboid
Misfolded Aptamer 6
Snowflake 3
Hard Y and a bit more
Mat - Lot 2-2 B
Shapes and Energy
Spiral of 5s
Campfire
Anemone
Fractal 3
Kyurem 5
Snowflake Necklace (or v2.0)
Methaqualone C16H14N2O Structural Representation
Cats Toy 2
Zigzag Semicircle
Short String 4
Gladius
Thunderbolt
Mutated chicken feet
Chicken Tracks
Looking Back Again
Multilooping 6
Cesspool
Hoglafractal
Bullseye
Shooting Star
Teslagon
RNAinverse...41
EternaBrain...66
LEARNA...56
SentRNA...78
NEMO...95


Eterna100-V2
Puzzle
Simple Hairpin
Arabidopsis Thaliana 6 RNA - Difficulty Level 0
Prion Pseudoknot - Difficulty Level 0
Human Integrated Adenovirus - Difficulty Level 0
The Gammaretrovirus Signal - Difficulty Level 0
Saccharomyces Cerevisiae - Difficulty Level 0
Fractal 2
G-C Placement
The Sun
Frog Foot
InfoRNA test 16
Mat - Martian 2
square
Six legd turtle 2
Small and Easy 6
Fractile
Six legd Turtle
snoRNA SNORD64
Chalk Outline
InfoRNA bulge test 9
Tilted Russian Cross
This is ACTUALLY Small And Easy 6
Shortie 4
Shape Test
The Minitsry
stickshift
U
Still Life (Sunflower In A Vase)
Quasispecies 2-2 Loop Challenge
Corner bulge training
Spiral
InfoRNA bulge test
Worm 1
just down to 1 bulge
Iron Cross
loops and stems
Water Strider
The Turtle(s) Move(s)
Adenine
Tripod5
Shortie 6
Runner
Recoil
[CloudBeta] An Arm and a Leg 1.0
[CloudBeta] 5 Adjacent Stack Multi-Branch Loop
Triple Y
Misfolded Aptamer
Flower power
Kudzu
1,2,3and4bulges
medallion
[RNA] Repetitious Sequences 8/10
Documenting repetitious behavior
7 multiloop
Kyurem 7
JF1
multilooping fun
Multiloop...
hard Y
Mat - Elements & Sections
Chicken feet
Bug 18
Fractal star x5
Crop circle 2
Branching Loop
Bug 38
Simple Single Bond
Taraxacum officinale
Headless Bug on Windshield
Pokeball
Variation of a crop circle
Loop next to a Multiloop
Snowflake 4
Mat - Cuboid
Misfolded Aptamer 6
Snowflake 3
Hard Y and a bit more
Mat - Lot 2-2 B
Shapes and Energy
Spiral of 5s
Campfire
Anemone
Fractal 3
Kyurem 5
Snowflake Necklace (or v2.0)
Methaqualone C16H14N2O Structural Representation
Cats Toy 2
Zigzag Semicircle
Short String 4
Gladius
Thunderbolt
Mutated chicken feet
Chicken Tracks
Looking Back Again
Multilooping 6
Cesspool
Hoglafractal
Bullseye
Shooting Star
Teslagon
RNAinverse...49
EternaBrain...59
LEARNA...68
SentRNA...69
NEMO...94


Total
-14.1 kcal
Natural/Target Delta
3.1 kcal
Loop
-1.9 kcal
Vienna
.....((((((((((....)))(((....((((((((...(((.....)))))))(((.....))).))))))))))..))))).....

Eterna100-V2: Impossible(?) puzzles and call for solutions

Hi Eterna Players,

A few months ago, in the creation of the Eterna100-V2, we asked that players design a number of secondary structures modifications to 19 puzzles in the Eterna100 that made them solvable in the Vienna2 energy model as implemented on Eterna. However, the question as to whether these puzzles were truly unsolvable in these parameter sets was never really put forth towards players. As some of you may have recently seen, there is a new series of puzzles that have been published on the site that try to address this question. The [IMOSSIBLE?] puzzles are those original 19 Eterna100 puzzles that we believe are unsolvable in Vienna2, published in Vienna2. However, determining if a secondary structure is unsolvable is a hard problem, and there is a possibility that we missed potential solutions!

In addition to those puzzles that we believe are unsolvable, there are also the Vienna2 remakes of the Eterna 100 puzzles. Referred to as Eterna100-v2, we are publishing them in Vienna1.8.5 to determine if these structures are backwards compatible in their energy functions. Some of these we have solutions for, but others we are not so sure about!

So we ask that Eterna players try solving these puzzles, and if you can't solve them that's okay. In the comments of the puzzle, you can post which base pairs you were not able to stabilize, and that could provide us with greater insight on what makes these puzzles unsolvable. Below are links to the lists of puzzles that we published on the site.

[IMPOSSIBLE?][Vienna 2][Eterna100-v1]
[Eterna100-v2][Vienna 1]