

# What makes an in silico puzzle hard?

Vineet asked if I also had a guide to what makes an in silico puzzle hard too. So here I compile a list of the things we have found that regularly kills the bots. This is meant as a help for the bot competition. Inspiration have been drawn from work with Mat and JandersonLee on [JandersonLee's bot](#). Much of the main content is in here, but some of the sections are still messy and needs to get swapped around. Also I will try add in some more links.

## Stems

- Equal length
- Short stems
- Non-hairpin stems

## Multiloops

- Adjacent multiloops
- Elements close to multiloops
- Elements directly attached to multiloops

## Bulges

- 1+ bulges
- Zigzags

## Internal loops

- 2-2
- Many loops close together

## End loops

- Trilops

## Symmetry

Repeat of structure

Bent shapes

Thinking out of the box

## Sum up of things that makes in silico puzzles hard

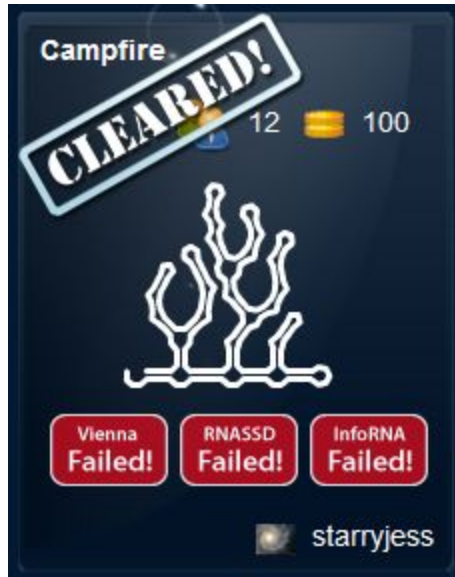
Short sum up of things that in general make puzzles hard. Usually several of these factors are in play at once. And the more of them, the worse.

- Many short stems, 3 nts and below. (4-5 can also be a problem if there are many adjacent multiloops or spots)
- Many equal length stems - if short
- Adjacent multiloops
- Short non hairpin stems
- Elements very close to or directly attached to multiloop (Few exceptions 1-1 loop)
- Many bulges - in particular if between short stems.
- Many internal loops
- Many loops and bulges with short distance between them. 2 base pair between or less, the less the worse
- Zigzags
- Big end loops on very short stems
- Multiloops very close to each other.
- Internal loops with single bonds
- Repeat of similar structure, the more the worse
- Symmetry, the more the worse
- The more combined, the worse
- Triloops make stems behave as if they were 1-2 base pair shorter. They make a puzzle harder.

# Stems

## Short stems and equal length stems

(okay, this one also got a lot of adjacent multiloops - find better picture)

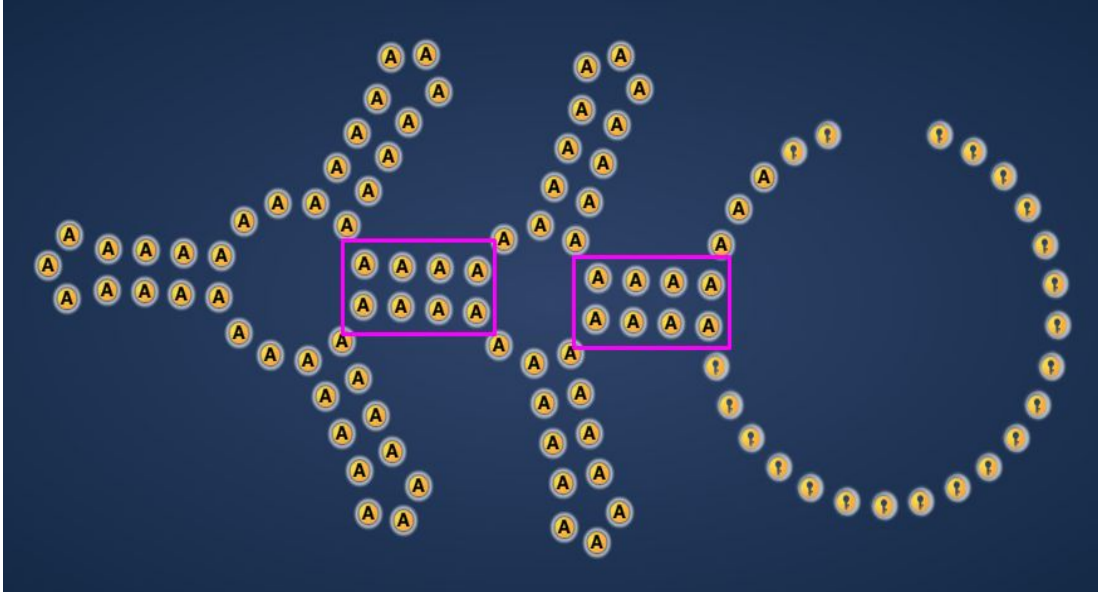


Short stems makes it hard for the bots too. Imagine you have 20 short stems only 3 base pair long. Adjacency makes things worse. If there were single basepairs between them in multiloops things would be loosened a bit more. But even in that case, you need to imagine that you have 20 stems that just as much like to fold up with one another in a different structure from what you want forming. If you just have 5 stems that is 3 base pair long, they have fewer options to mispair. So it is the sole number of participants in a puzzle that makes it harder. If you have stems of varied length, there is a much higher likelihood that you can get the structure folding properly. For short stems it will mostly take GC-pairs for solving, whereas you with longer stems, can vary your stem sequence much more. (same in lab) So for a puzzle with 20 equal length short stems, you will have very few all GC-variations to choose from for each stem, and thus the probability of mispairing gets much higher. I think there can be made statistics on this. It simply becomes much more unlikely that you hit exactly the right sequence that will not misfold into a shape that has a more advantageous energy.

## Non-hairpin stems

Short non hairpin stems - the more the worse. In particular if combined with adjacent multiloops.

Non-hairpin stem marked in picture. Term coined by JandersonLee.



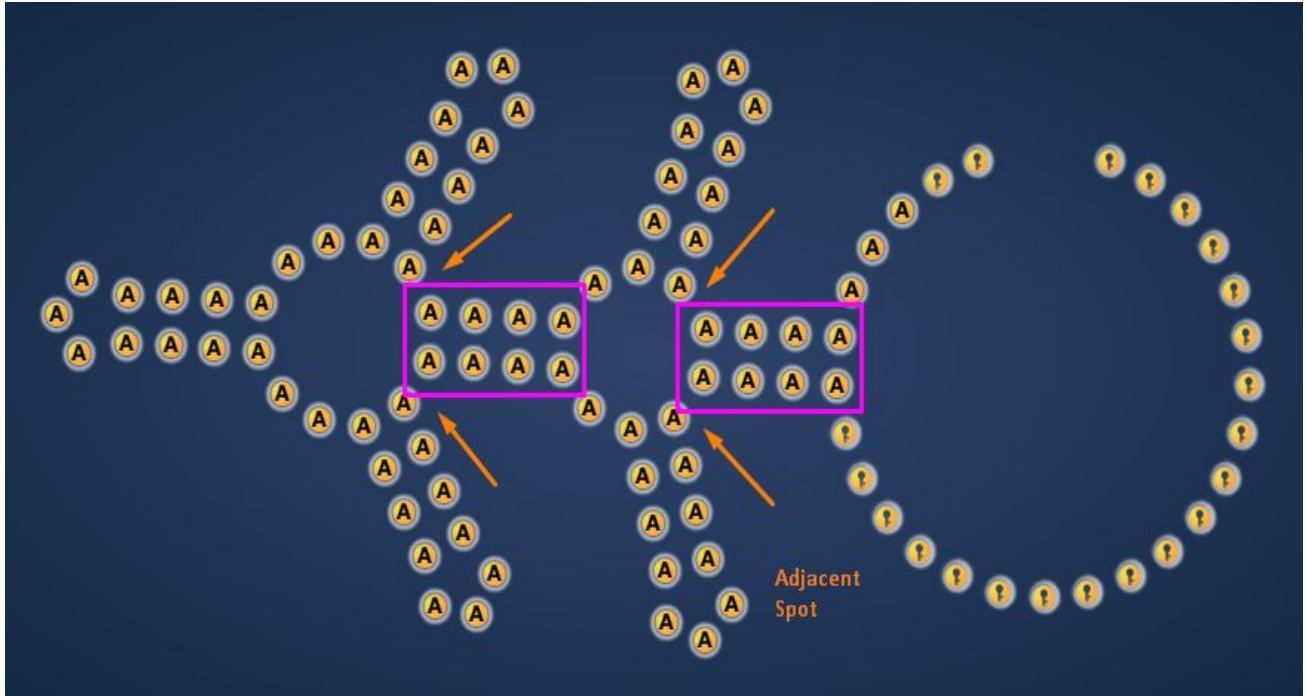
# Multiloops

## Adjacent multiloops

The image displays six panels, each representing a different Kyurem species (Kyurem 8 through Kyurem 3). Each panel shows the result of an RNA secondary structure prediction using three different algorithms: Vienna, RNASSD, and InfoRNA. The results are consistently 'Failed!' for all three algorithms across all species. The panels also show a 'CLEARED!' status and a score (e.g., 17/100) for each species. The RNA structures are visualized as white lines on a dark blue background, showing the complex folding patterns of the RNA molecules. The structures become increasingly complex and feature more adjacent multiloops and shorter stems from Kyurem 8 to Kyurem 3.


Kyurem	Score	Vienna	RNASSD	InfoRNA
Kyurem 8	17 / 100	Failed!	Failed!	Failed!
Kyurem 7	17 / 100	Failed!	Failed!	Failed!
Kyurem 6	17 / 100	Failed!	Failed!	Failed!
Kyurem 5	8 / 100	Failed!	Failed!	Failed!
Kyurem 4	13 / 100	Failed!	Failed!	Failed!
Kyurem 3	12 / 100	Failed!	Failed!	Failed!

Generally the more adjacent multiloops and adjacent spots, the worse. In particular if main part of the stems are short.



- Two multiloops close to each other, separated by a short non-hairpin stem (1 basepair - check)

**[RNA] Repetitious Sequences 8/10**



37 100 10 Jul 2013  
 paramodic

Vienna Failed!	RNASSD Solved!	InfoRNA Failed!
-------------------	-------------------	--------------------

## Bulges

Bulges (the bigger the worse - check) if between short stems and close to other elements.  
(Bulges included) As Paramodic [describes it](#):

- Bulges: A new observation: Infobot seems to have a great deal of difficulty with puzzles featuring 1 NT bulges. The details require a lot more fleshing out at the moment, but it seems to follow this pattern:

- The bulges are on the same side of the fold as each other (ex: Left side)
- The bulges are close to an internal loop
- The bulges appear on sequential or adjacent strings.

Or as Freywa puts it: I propose an InfoRNA-foiling structure: corners (bulges) and multiloops placed close to each other. (Add link)

Edward Lane also did this test. It looks like the more bulges the worse. I know that adding bulges also means more short stacks. (Add link)

<p>123 bulges 2&amp;3 reflected</p> <p>👤 54 💰 100</p>  <p>Edward Lane</p> <p>Vienna Failed! RNASSD Failed! InfoRNA Failed!</p>	<p>2 bulge turned other way</p> <p>👤 69 💰 100</p>  <p>Edward Lane</p> <p>Vienna Failed! RNASSD Solved! InfoRNA Failed!</p>	<p>just down to 1 bulge</p> <p>👤 74 💰 100</p>  <p>Edward Lane</p> <p>Vienna Failed! RNASSD Solved! InfoRNA Solved!</p>
<p>1 and 2 bulges</p> <p>👤 68 💰 100</p>  <p>Edward Lane</p> <p>Vienna Failed! RNASSD Failed! InfoRNA Failed!</p>	<p>123 bulges (easier for bots)</p> <p>👤 56 💰 100</p>  <p>Edward Lane</p> <p>Vienna Failed! RNASSD Failed! InfoRNA Failed!</p>	<p>1,2,3and4bulges</p> <p>👤 48 💰 100</p>  <p>Edward Lane</p> <p>Vienna Failed! RNASSD Failed! InfoRNA Failed!</p>

Quote Jieux on multiple bulge solutions (Add link)

## End loops

Triloops if on short stems 3-4 nt's long and shorter

## General adjacency

- General adjacency, not only in multiloop, but also in hooks. 0-0-0, 0-0-0-0 and bigger (or total adjacency is easier solving in puzzles than I suspect they will be in lab although there were some slack for 4 ways junctions in lab)

## Symmetry

Symmetry - the more the worse, and in particular if there are many short stems.

- Many equal length stems attached to same multi loop/hook - if short stems.

## Zigzags

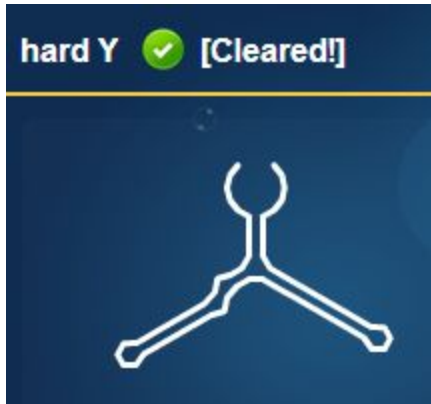
Zigzags as discovered by Joshua Weizman. (Add link) The bots can't solve zigzags.



The image shows a screenshot of a puzzle interface. At the top, the puzzle title is "Hard Y and a bit more" with a green checkmark and "[Cleared!]" next to it. Below the title, there is a white RNA secondary structure diagram on a dark blue background. To the right of the diagram, the puzzle statistics are displayed: a person icon followed by "22", a coin icon followed by "100", and the date "10 Jun 2011". Below the statistics, the name "Joshua Weizman" is shown with a small profile picture. At the bottom, there are three red buttons with white text: "Vienna Failed!", "RNASSD Failed!", and "InfoRNA Failed!".

Though I saw RNASSD get one by chance.





## Internal loops

- 2-2 loops if closely placed - with one base pair between them. InfoRNA usually has no problems, it just goes all GC on them.



## Many loops close together

- Lots of closely spaced loops (Quasispecies)

**Crop circle 4** ✓ [Cleared!]

25 100 22 Jun 2011

Eli Fisker

Vienna Failed! RNASSD Failed! InfoRNA Failed!

Many different loops followed close after each other with 2 or less base pairs between them. This puzzle is a long row of 9-9 loops with 2 basepairs between them:

**Documenting repetitious behavior** ✓ [Cleared!]

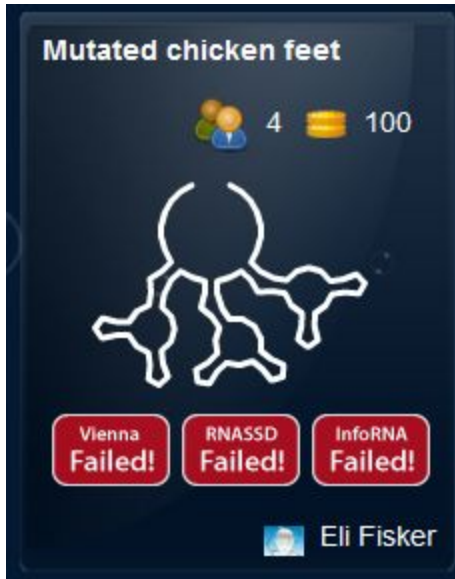
33 100 27 Jul 2013

paramodic

InfoRNA Failed!

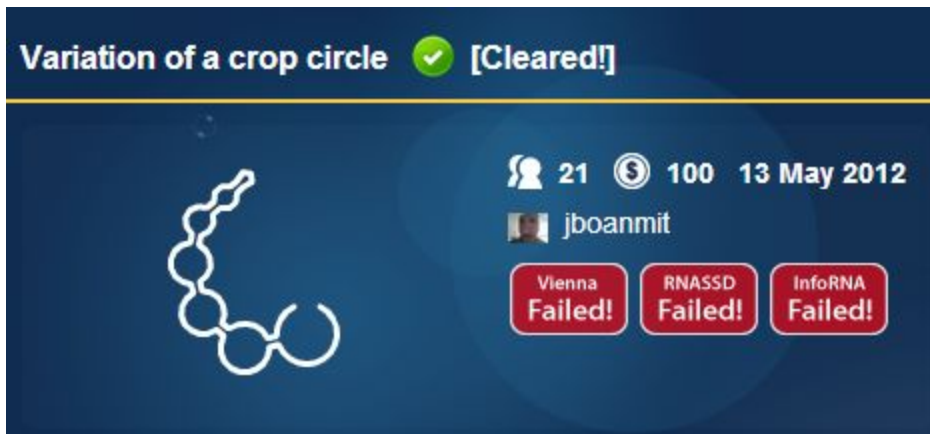
### Identical structures repeated multiple times

Multiple repeats of the same structure makes a puzzle harder. Both because it means a higher amount of similar stems, but also because it is harder to vary the sequence if the structures are alike. Gets worse if the stems are short.




## Bent shapes


Bent structures with lots of bigger internal loops close together (crop circles)





## Catalyst points

- Single bonds closing of a big internal loop or [Catalyst points](#) - takes knowledge of how to lower the energy inside an internal loop with uracil boost. RNA SSD can take this puzzle type down.

Simple Single Bond  [Cleared!]



60  100 24 Jul 2011


 starryjess

Vienna Failed! RNASSD Solved! InforNA Failed!

### Mixed elements

- Any element x number of base pair close to a multiloop
- Lots of closely spaced elements, 2 base pair between and below. Bending.

InfoRNA bulge test ✓ [Cleared!]




74 \$ 100 06 Jan 2012  
 paramodic

Vienna Failed! RNASSD Solved! InfoRNA Failed!

### Structure Notation

.....((((.....((((.....((((.....((((.....((((.....))))))..))))))..))))))..

- +



[Paramodic](#): Now that I think about it, having bulges in sequence with internal loops (particularly of the uneven variety) seems to also mess with infobot's success level.

- bulges directly attached to multiloops
- internal loops directly attached to multiloops. (Describe which types causes problems, not all of them)
- There is something something with orientation of bulges according to Paramodic and Edward Lane. Bots should get an easier time solving one way compared to the other. (Add link)

## Neck

Short blunt end necks like this one can also make a puzzle hard. Generally just a few dangling bases on a neck makes it much easier to stabilize without having to “boost” to optimize stability. See Drake’s [Advanced Puzzle Solving Guide](#). Especially the last section.



## Creativity - Thinking out of the box



The robots have trouble with some of the easiest of puzzles as they requires thinking a bit out of the box.

Do you have something to add on what makes a puzzle hard to solve for a robot? Feel free to add your thoughts in the forum post [Nature of the bots](#).

*Eli Fisker*

For more about what makes in silico puzzles and lab puzzles hard read:

[Nature of the bots](#)

[Pressured designs](#)

[What makes an RNA design hard?](#)

## [What makes a RNA design hard? - Part two](#)

*Started 9 Aug 2013, 11 Aug last update*

### *Working notes*

Many short strings and strings of equal length. Equal length strings opens up to repetitive sequence, which leads to more mispairing.

For example, they (bots) seem to also have issue with structures formed by repetitive sequences, so not just symmetrical structures, but repetitive ones.

Nothing opens more up to repetition than many strings of the equal length. Be they really short or even mid length.