Methods for improved lab designing

I often mention in the descriptions of my lab designs, how I have done my designing and what designs I have used parts from. But recently I have been thinking, that I haven't demonstrated in pictures what I do. I decided to go in more details with what I do in lab. I will mention several other different and great methods for improving lab work by other players.

Remember that I'm sharing what I know and think I know. Not everything will be correct. However I think we will be far better of making changes in designs and designing with intent, going about it with methods, than making random changes in designs.

Also see Ways to deal with voting in big labs

Written May 2013, latest update Oct 2015

Eli Fisker

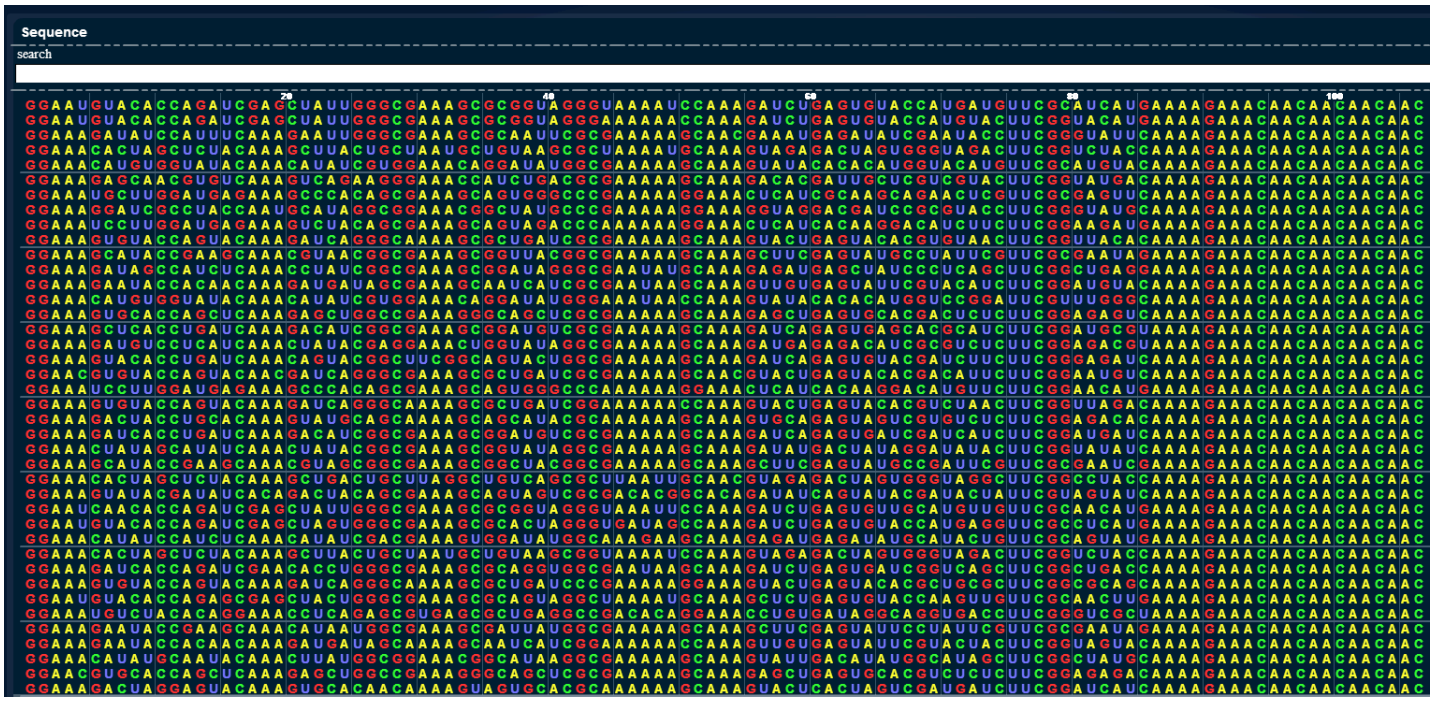
Document shortcuts

Using the misfolds
True Forms

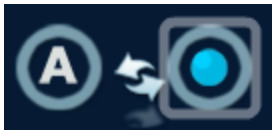
Getting the best out of the Lab Result Interface

Also I have been starting to work in a different way, that I think is more helpful now we have many results to pick from. Earlier I would just look through the lab designs by hand from either the same lab, or from a lab that had a similar element or working spot for what I was interested in. I also have used the CSE tool first, where possible.

Here is what I have been doing lately to isolate potentially good fixes. I have used the SHAPE data part of the Lab Result interface to hunt down good local solves.



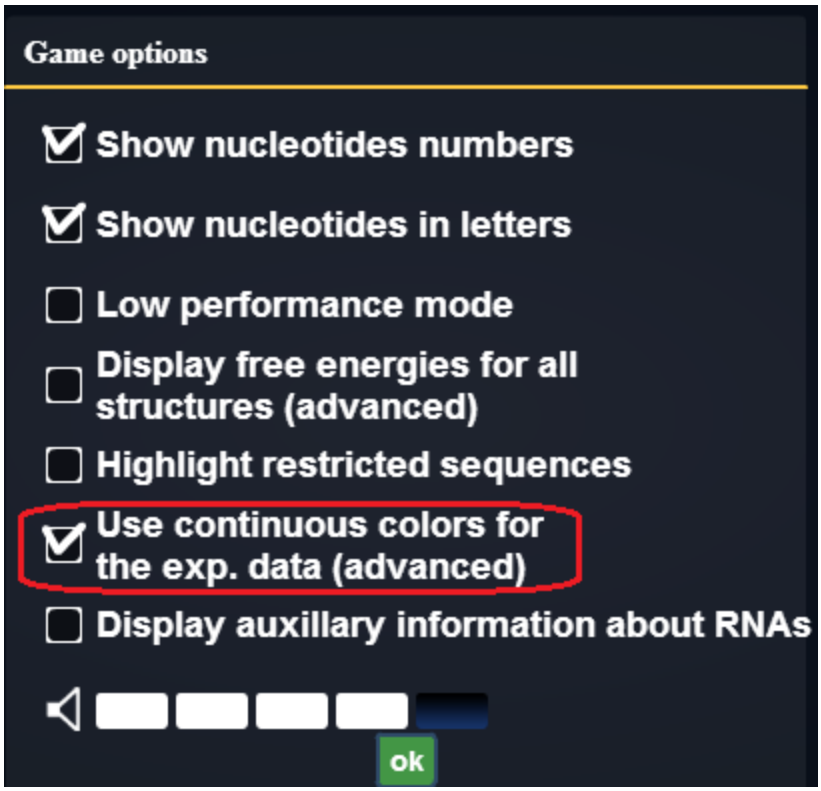
If you switch the above setting to the below, you will get a very different look of the lab results.

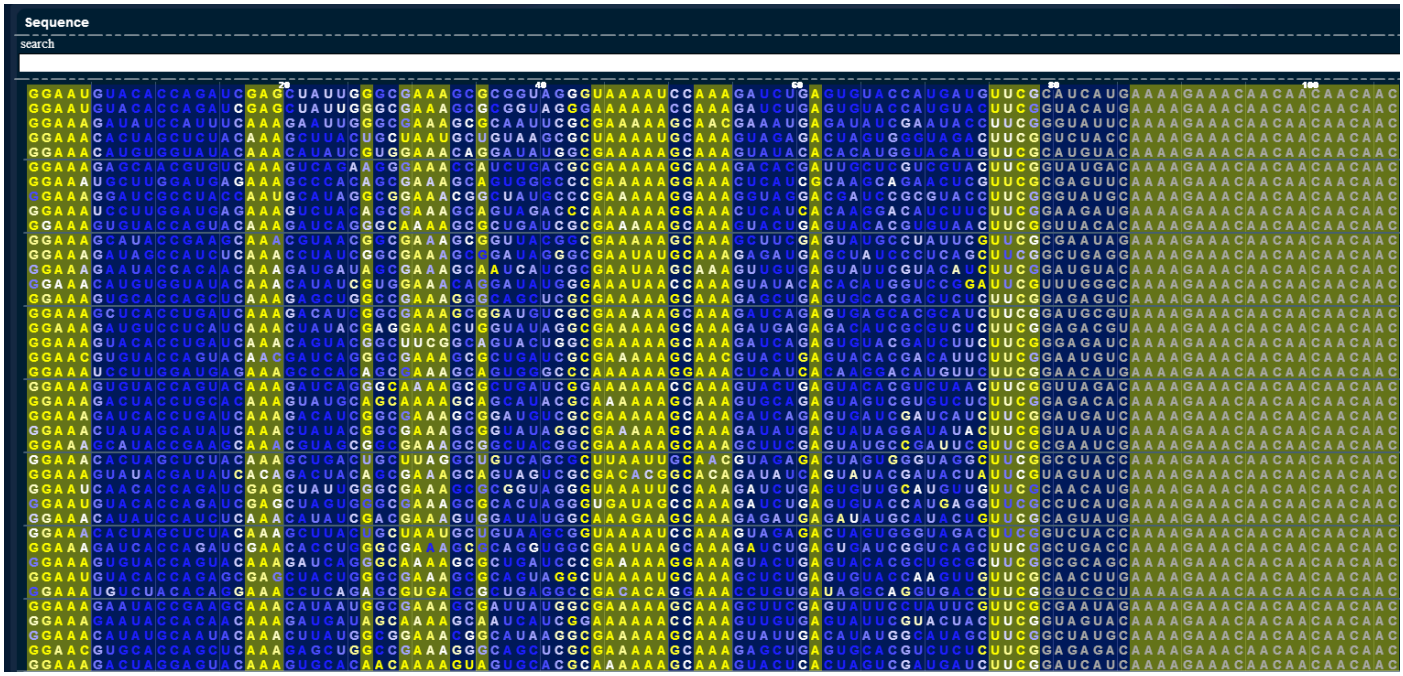


If it looks all blue and yellow with no light blue shades, you need to change the settings in the puzzles. The advanced SHAPE data will give you much more information about which designs to borrow a good spot for fixing a bad spot in an otherwise good lab design.

To set this, open Gear when in a puzzle

Choose Continuous colors




Sequence
search

Different lab designing strategies

I will describe different methods for lab designing. Methods where you will be using information from the data we already have, to make an educated guess for a good design.

Watch the estimate mode

Watch the winners in estimate mode also and notice where they break. The winner from round 1, broke in the bottom 5 nt loop.

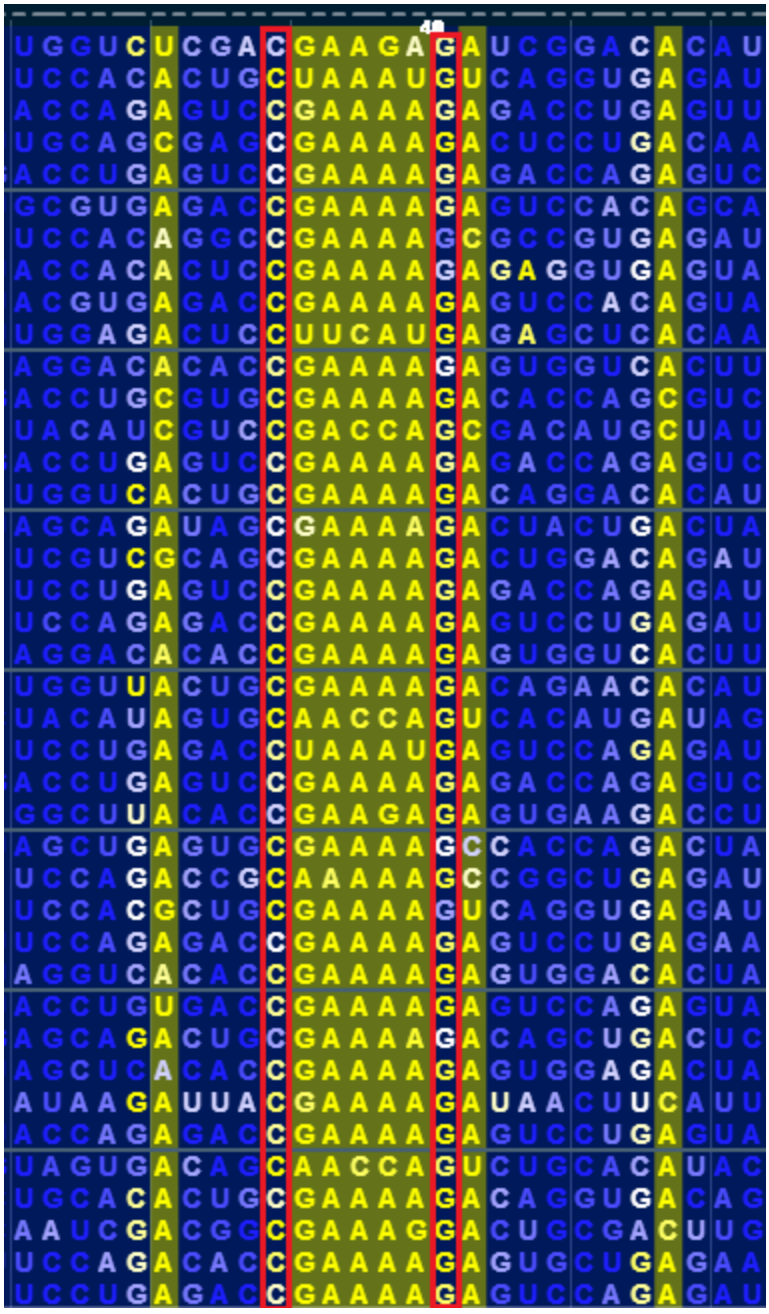
And only very few designs did stay together in base pair 35 and 41. One way to locate the designs which works in this loop to transplant it to the winner is to look through all designs and check their estimate shape.

Now we are coming to the smart part. Though it is generally very useful to see both the estimate and target mode for a good design, looking though all to a working one, might not be on your wish list. Here is how you can use the SHAPE data in the lab interface. We are interested in designs that do well at base 35 and 41.

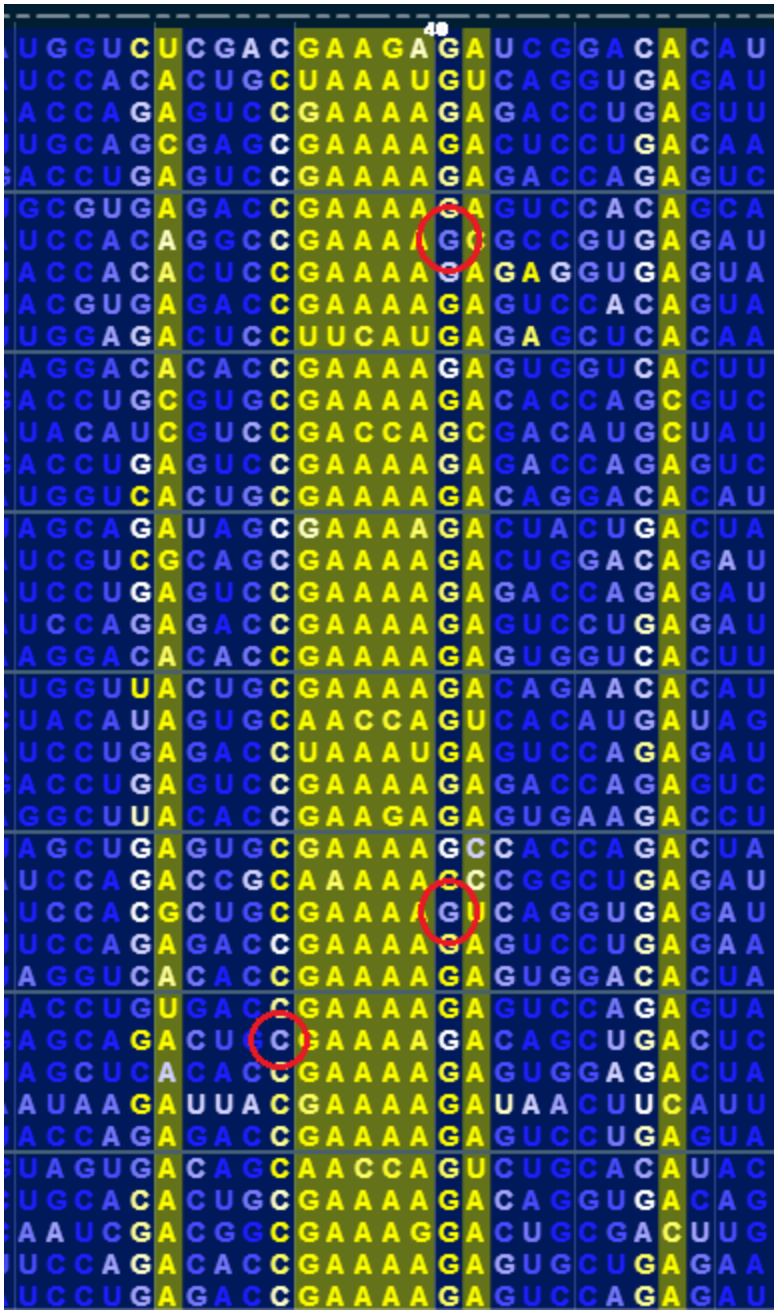
So now I sort all the round 1 designs after how well they scored in synthesis. (if possible it is best to pick the fix from a design that is overall good).

A way to make sense of this blue yellow moon landscape is to watch the numbers on top of the columns. Marked with red below.

Now I have marked the bases 35 and 41. And am on the look for a solution that has two dark blue bases at 34 and 41.



In this round there aren't any, so I try check out the next best. I check out the designs that has one of the base 35 and 41 stable. I have circled them with red below.

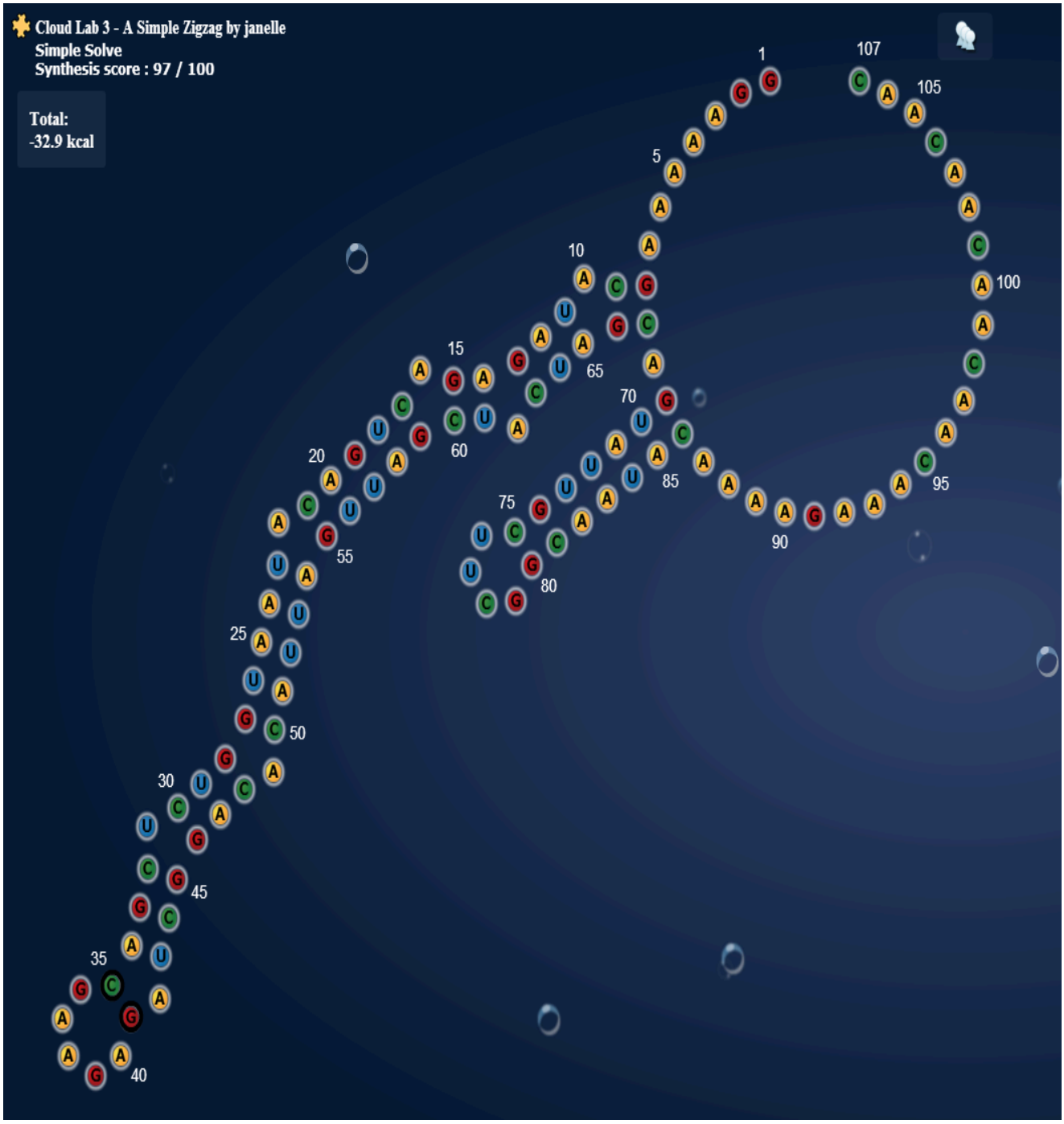


I checked the one with the highest score, the one in top. And indeed it did stay together when I checked the estimate mode. You can check it out yourself.

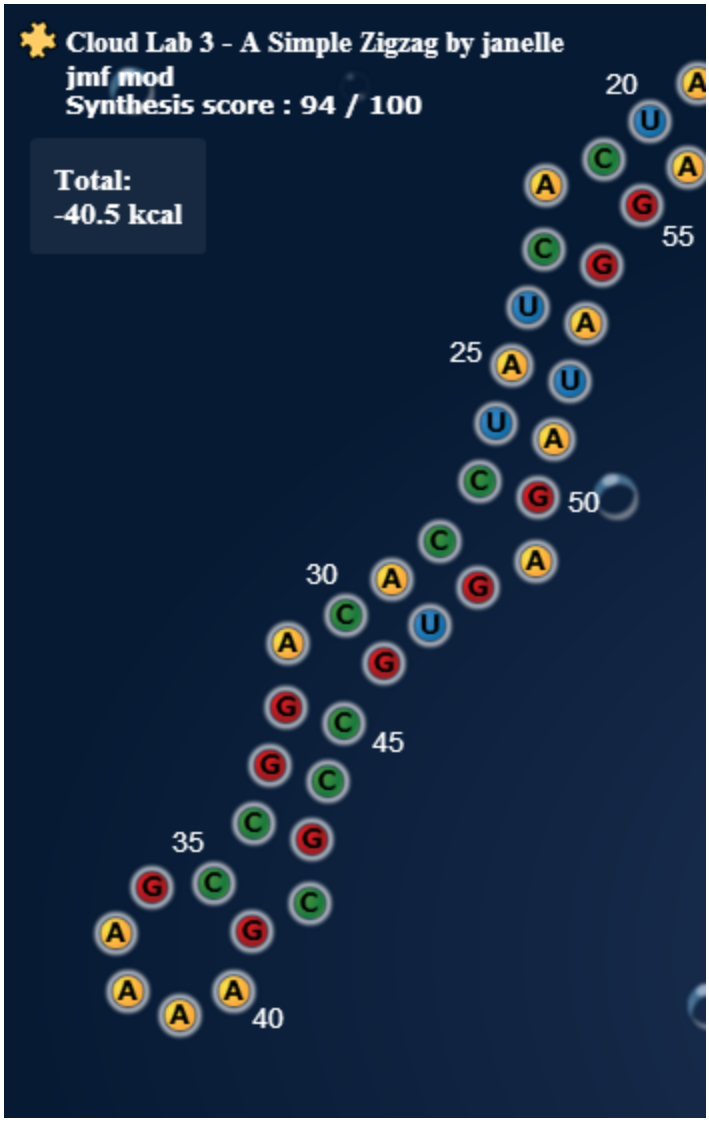
Borrowing whole elements

“Use what already works”.

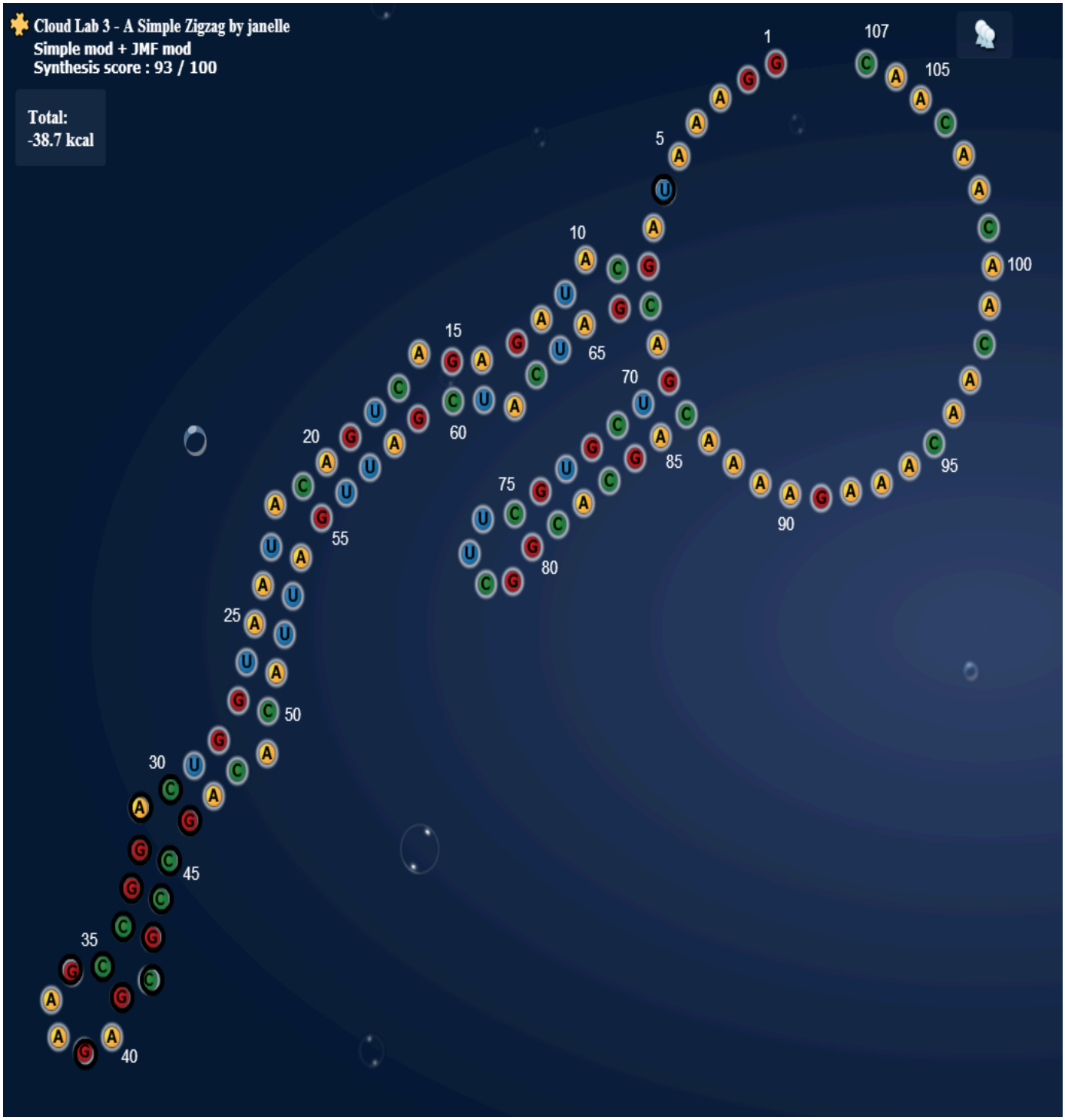
Continued from former section. The round 1 winner in this lab is only unstable at base 35 and 41.



I have located another design that is not breaking, when watched in the estimate mode. As the basepairs following the break in the winner looked quite different to the design with the potential fix, I decided to borrow more than just the base pair 35,41. Borrowing just one base pair, I call a micro fix. Read more about it here.

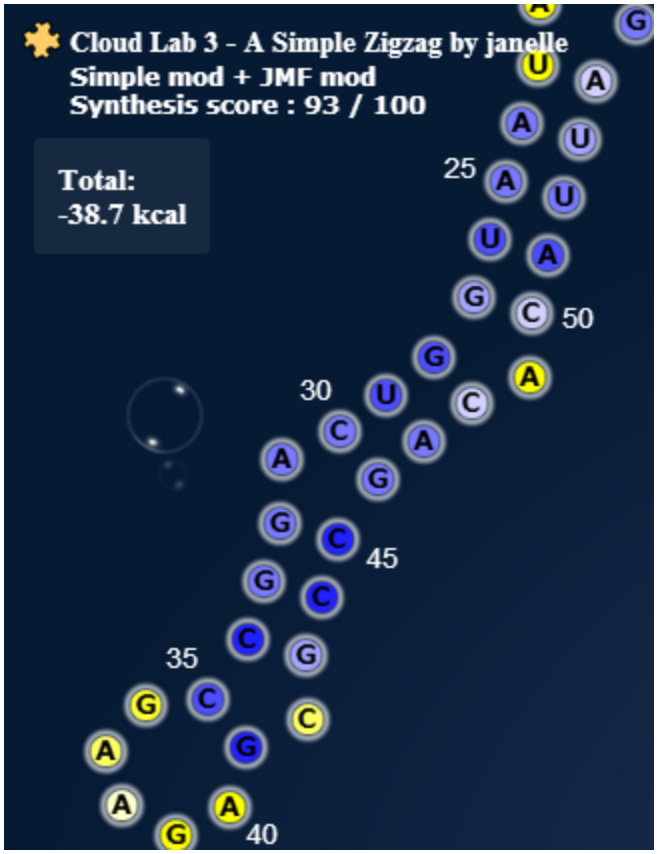


I then borrow the whole element plus a bit more. I have marked the bases I changed in my mod below. I added new constraints too.


Cloud Lab 3 - A Simple Zigzag by janelle
Simple mod + JMF mod
Synthesis score : 93 / 100
Total:
-38.7 kcal

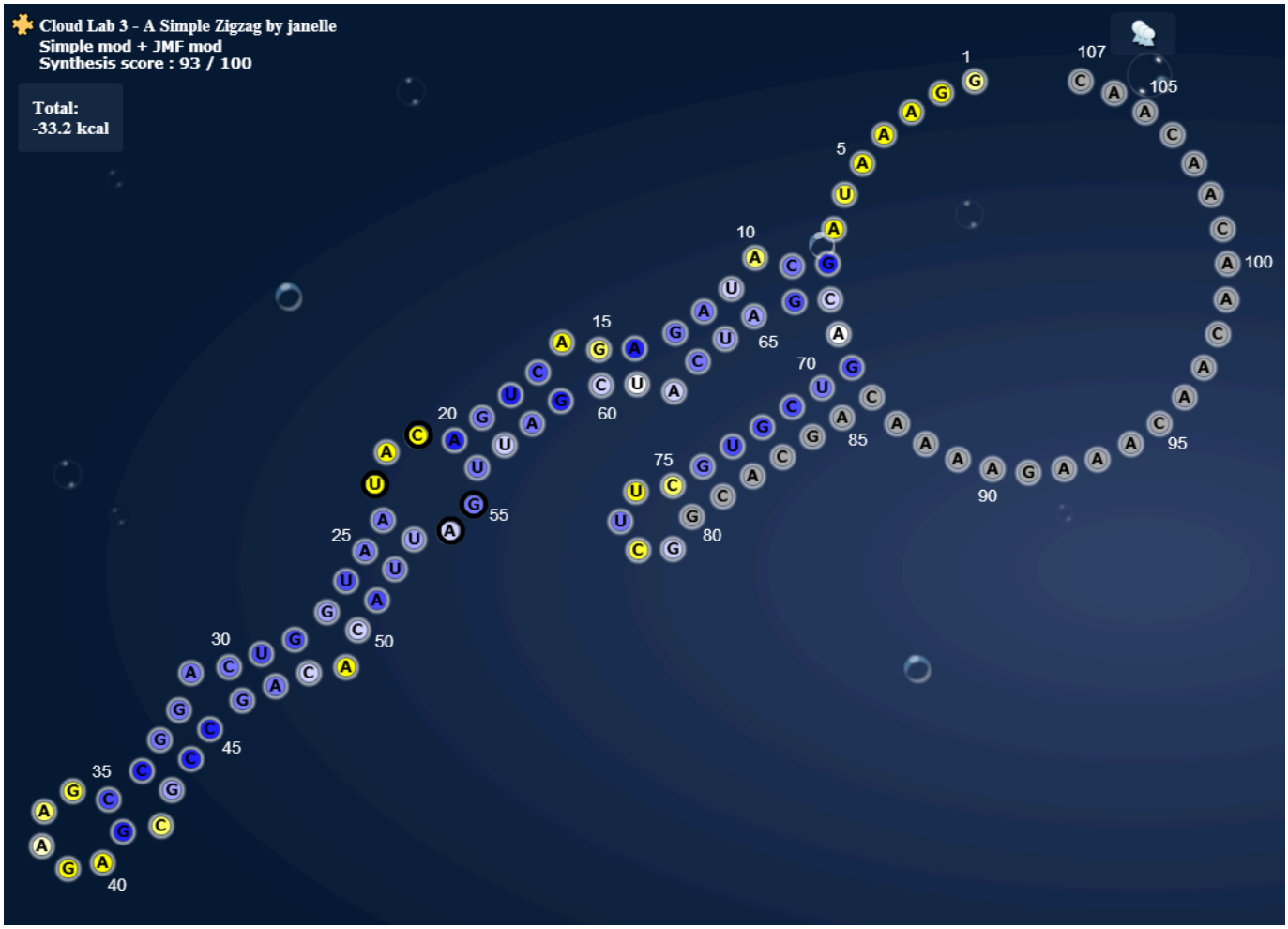
As you can see here, the tip ended up looking better and holding in my fix. But as you also can see, the overall score dropped.

.



My design ended up opening somewhere I had actually not touched. (Breaking at marked bases.)

It might have been because the strong tip end that I added, may have sent instabilities further up in the design.



You can borrow elements in other ways: If this is round 1 in the lab - pick a similar and working element from past labs. That can be a string of same length from an earlier lab. Or a loop area. If you can't get an element of the same size, like a bulge of a certain size, looking at bulges at a size close to and what is successful in those, might help as well.

CSE lab designing

Where possible, I use the CSE tool, to pick out a good element. It sorts the element after how well their SHAPE data score. Use it for strings. It is not set for aptamer use and for loops the SHAPE data is sometimes shady. Eg. a good tetraloop can have blue nucleotides in the loop area and thus score less well in the CSE tool.

If you are making a lab patchwork of elements borrowed from elsewhere, don't pick elements that looks too different. If the design has four long and similar length strings, don't pick an element that fully GC-pairs and the other strings are only just stitched together with a few GC-pairs. Even if each of the elements looks good when viewed individually in in the SHAPE

data or with the CSE tool. For more on the why, read here.

Optimal for round 1 designing, and optimal for improving an element for later round designs

Cloud Lab Data Mining Tool designing

Omei made a new tool that is good for comparing data between designs in the same lab. Read more about it here and check out the manual.

LabDataMiner designing

JandersonLee has made a tool that can be used to find data across different labs. Here is the manual. Here is an example design for how to use the tool.

Mat's cut and paste method

A good cut and paste will often take two basepairs along with it from the element it is borrowing from.

Go see Mat's Lab Design strategy. He describes a good method for fitting elements from different designs together. His method is generally safer than just borrowing a single basepair (making a micro fix). He is the RNA designer with the highest amount of lab winners.

Micro fixes - Or borrowing of whole designs

For second rounds in lab, where we already have results in from first round, I focus on the trouble spots in a good design. Then I look at the other designs and try find some that are working in this specific spot. I used the above mentioned method to isolate them. Here is an example on what I do.

Then I often pick a working a certain spot from another design, in attempt to fix a trouble spot in a good design. Often I only change 1 base pair. I also try watch out for that the fix I take match well into the design I put it into. It is generally more certain to work, if the surrounding it comes from, matches the surroundings where it goes in. So sometimes I prefer not to take the absolute best design, to put the fix into, if the fix looks like a match for the second or third highest scoring design.

- Modify good designs from round 1

I generally look at the highest scoring. I don't always pick the design with the actual highest score. Not all designs are equally easy fixing. When I have noticed the weak areas in some of the high scoring designs, I go and look for designs that do well in that particular area. Let's say there is a breaking base pair in the design I want to fix.

Sometimes the design that has a potential fix, looks very different in the surrounding to that base pair that is breaking in the design I want to move the fix to. This is why it can be sometimes useful picking the 3 or 2 highest scoring design from round 1, instead of the highest scoring. If the fixing base pair and its surrounding two base pair looks more similar to the 2 or 3 highest scoring design, there is a higher likelihood of the fix working, as I am only changing little. The fix might otherwise not work in the changed surroundings.

Note, my micro fix method is not always working. But I think it beats the alternative method of making a random change in the problem area in the hope that it works. Which is of cause very reasonable when there are no usable results available to borrow likely fixes from.

Barcode hairpins

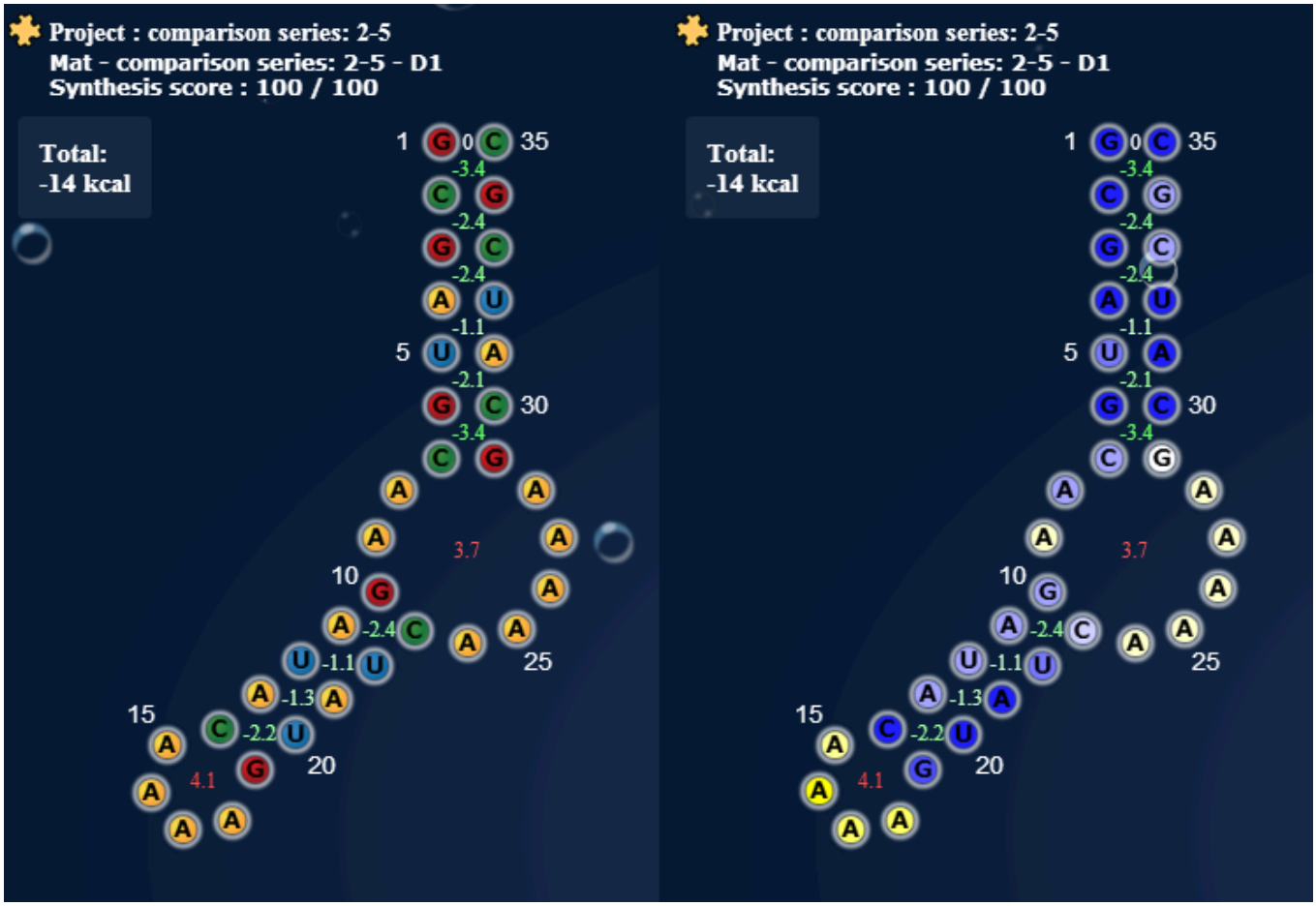
I can see that in several good designs that I made a mod of, the mod went well, but me changing the hairpin barcode, made the design break open there. I'm suggesting copying in hairpin that have showed itself to work in previous good and similar looking designs.

Janderson advice on A constraints

From design description: Yes, it has 5 As in a row, but they tend not to be a problem until 6 or more and breaking up the line seems to cause other problems.

Mat sent me a picture from the new player project lab results, with the comment:

"The funny/interesting thing is the Tetraloop is all A`s."

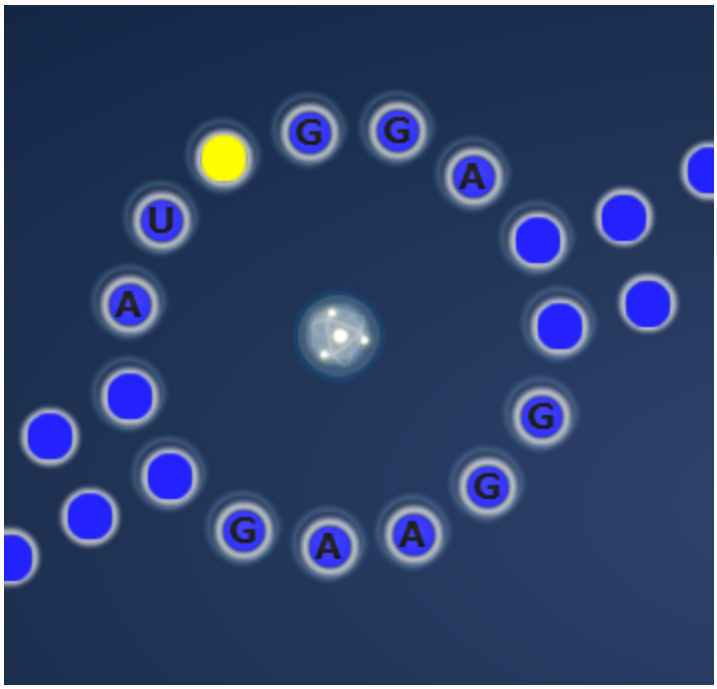

Project : comparison series: 2-5
Mat - comparison series: 2-5 - D1
Synthesis score : 100 / 100
Total:
-14 kcal
Project : comparison series: 2-5
Mat - comparison series: 2-5 - D1
Synthesis score : 100 / 100
Total:
-14 kcal

Aptamers

First a short note on aptamers. An aptamer is a loop structure with special nucleotide sequence, that will bind up to a small molecule if the aptamer folds correct. Here is a picture of one. Locked nucleotides = aptamer sequence. Notice the molecule inside it.

Tom (one of the developers) described what a binding aptamer would look like. Mat and I have added a picture, that shows what a perfect aptamer should look like in the SHAPE data. For more technical details see here.

FMN aptamer - perfect binding



Normally SHAPE data should be yellow in loops and blue in stems. But as the aptamer is binding to a molecule it will turn blue though it is a loop.

I believe that good aptamers can be recycled just as can good necks. It is still early saying what will for sure work with aptamers.

Reuse perfect aptamers and near perfect from earlier designs. I let the aptamer, so it has two basepairs to both sides with it, from the design it originated in. (Some designs will have locked nucleotides in the area, so this approach won't be able in all switch labs. It also depend on what the two states of the switch looks like.)

I asked Tom if aptamers in both switch and static lab would have same binding pattern and he said yes. He also gave following further explanation:

Just to let you know, I made sure that all of the cloud lab RNAs were together probed with FMN. The FMN aptamer is still capable of binding FMN even when it is not in the switch. Also, the nucleotides in the aptamer should have the same reactivities in the presence of FMN, regardless of whether or not they are in a "static" shape or as part of a switch. The exception to this is if the switch is not switching all the way!

Another important point to remember is that FMN should only have an effect on RNAs that contain the FMN aptamer. Non-aptamer nucleotides and RNAs without the aptamer should not be affected by the presence of FMN. This fact is important for our new RNA chemical probing strategy, where we probe many more RNAs at once in a single tube!

The last thing means that all the RNA's, switch and non-switch can be run at the same time. Which is real cool!

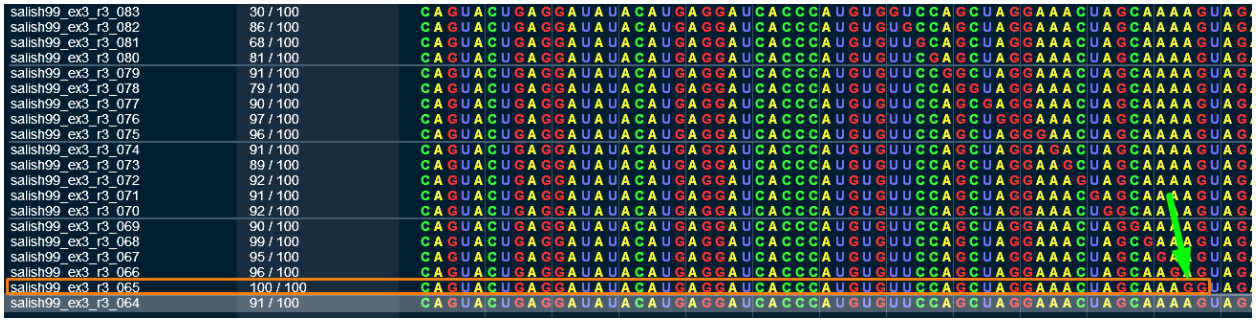
Freestyle lab designing

Where we have no lab designs with a structure close to what I see in a new lab, I go freestyle based on the general rules I see for good designing. You can see some of them in My strategy guide to lab and in Mat's lab design strategy. I try fit my style to what I know works for different elements like long strings, short strings, bulges and so on.

Mutate and map design style

Salish simply systematically mutated his way to a winner by taking last round high scorer and mutating base by base, in the switching area.

Winning design



You can read a bit more about it here.

Omei did something related in the microRNA/reporter lab, round 2. He mutated along the microRNA complementary stretch of bases, so when I flipped my way through his designs it looked like a 1-1 loop walk along up against the microRNA. :)

Melt temp and average stem length

I thought perhaps I could see something from the melt temperature when watching different kind of structures.

Another interesting tendency I see when I compare the melt plot for the winners in different labs is that designs with a stem average below around 2.8 basepairs, seems to have less melt temps of 107 degrees. There are a few labs having 2-2 loops that behaves as exceptions. Designs with longer stems have a higher amount of 107 degree. It is actually very logical. Designs with longer stems gain power by hydrogen bonding and are as such harder to make melting. However the many short stem designs, also have a general high amount of GC-pairs, that are the hardest ones to melt. I think we might be able to use our knowledge about what melt will be typical for a certain design. This is nice knowing when designing.

Is there other lab designing methods?

Know there are many more methods for making and improving lab designs. Read the descriptions on the lab designs to get an idea about how other players are designing and follow the chat discussions about lab designing. Last but not least, watch the lab results from the winning designs. The good thing about lab work, is that it is something that can be learned. You won't make perfect fixes and designs all the time, but being more conscious about why you change what you change, will help you improve.

Different ways of watching lab results

Stable “unstable” regions

Use stable regions from designs that the energy model deems unstable. Read about Janelle’s find on unstable regions that score well in the lab and looks fine in the SHAPE data.

Watching the misfolds...

This is something that Mat made me aware of, back when we were working with presenting the CSE tool idea. See his explanation to me, here.

With our coming more accurate data, this is a method that will have growing relevance. I would like a CSE tool additon, where we will be able to search for an element in the misfolds too. As Mat already realized back then, this will double the amount of data we get per design.

Advice on how to use it for lab designing and puzzle solving

Mat explains: It’s ok to search for element in misfold as long as the designs had scored well. Some players use the lab misfolds as a reference for player puzzles. Players copy lab misfolds and make them into puzzles. That also means that what you see as a fold in the lab, has a potential to work in puzzles too. As the lab and puzzle build on the same energy model and is thus close.

But the lab misfolds doesn't follow the energy model exactly. As the lab result as it is folding is how nature it is nature and not just the energy model. But it still can provide fine inspiration and potential useful solving methods for puzzles. The better you are able to solve puzzles, the better for your lab work as it will give you the ability to choose between more options. This is in particular true for the switches.

Also if you are watching how a puzzle misfold and makes a certain element in the native mode, that can give you the key to how to solve a similar element, when it is present in a target shape

Then use it for making player puzzles.

It made me wonder how accurate the data on the misfolds were. So I asked Rhiju. Find his explanation here.

Using the misfolds

True Forms

This is an expansion of this section. Mat was explaining me a part in his Lab designing strategy, when he opened my eyes to a whole new perspective on the lab data.

(Chatlog extract, from 5 nov, 2012)

Mat: I did want to show you something using the puzzle maker.

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

First is a tetraloop

Mat: Now if you seen that shape in the results, it looks like if had formed the correct shape.

(((.....)))

Mat: now what about this tetraloop

Mat: that the structure i wanted you to see
Mat: one that was a correct tetraloop
and the other wasn`t

Mat: "which one of those formed correctly" as a tetraloop
Mat: now if you were looking at lab results
Mat: which one of those element would you pick, if you were looking at tetraloops results
Eli: Normally I would pick the tetraloop size one. Unless I could find misfolded hexaloops
Eli: Which I probably wouldn't have thought about. Until now
Mat: now would you say that it is important that a elements formed correctly when it come to the lab
Eli: Do you mean it is important it is correct already when it get in. You want to read the analysis from the misfolded shape data too? That we can use that too, to get the best elements?
Mat: I give closing pairs more importance because I want the elements that form correctly
Eli: Awesome
Eli: I think this is very well thought
Mat: if it didn't form correctly it is not the correct shape
Eli: and hard to see at first
Mat: yes very easy just to look at the energetic alone
Eli: yes, and forget about shapes forming
Eli: And I think it is amazing that you just "doubled" our data, because the correctly formed shapes, even in misfolds, are valuable prey too.

See more about this idea that Mat shared in Computationally Selected Elements under the section Extra penalties, on page 8. There are picture examples in there.