EteRNA Dictionary

Welcome to the Eterniverse![1]

I have collected definitions from players, devs, biology books, science papers, wikipedia and the web in general. It is meant as a help to avoid misunderstandings between us and make it a bit easier for the newer eterna members - and who knows, perhaps also for scientists to understand some of the EteRNA slang. I will put in a few pictures as to guide understanding. This is just like my lab strategy guide, a work in progress. Feel welcome to send me corrections and definitions that is missing. Especially thanks to Quasispecies for letting me put in his collection of definitions.

For more terms and different explanations also see the Eterna Dictionary in our WIKI.

Eli Fisker

Started June 2011, latest update Dec 2015

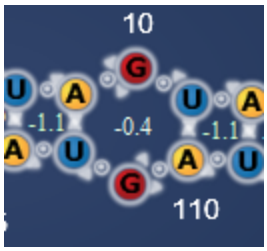
As this dictionary keeps growing I have put in some alphabetical short cuts.

[1] Eterniverse is what Edward Lane calls the eterna universe. Intro idea by Starryjess.

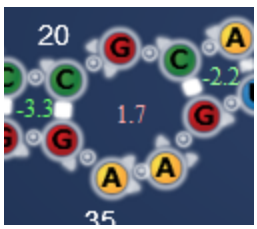
0-1 bulge - It has a nucleotide that sticks out of the stack on one side, but none on the other.

Explanation and picture by Hoglahoo

1-1 loop - only one unbonded nucleotide opposite only one unbonded nucleotide. Explanation: The 2 red G's are the unbonded nucleotides. 1-1 loop is the same as an internal loop (definition from Merryskies, Clollin suggested this term put in). These two red boost points help the loop stay together.

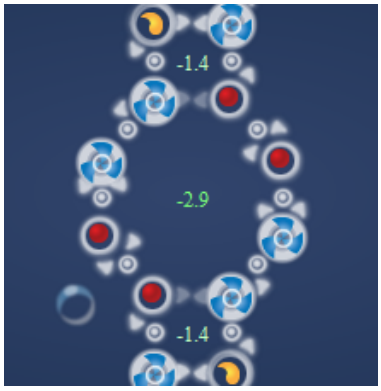


1-2 loop



The same as a bulge (cut from Merryskies Human astrovirus, added on suggestion from Clollin)

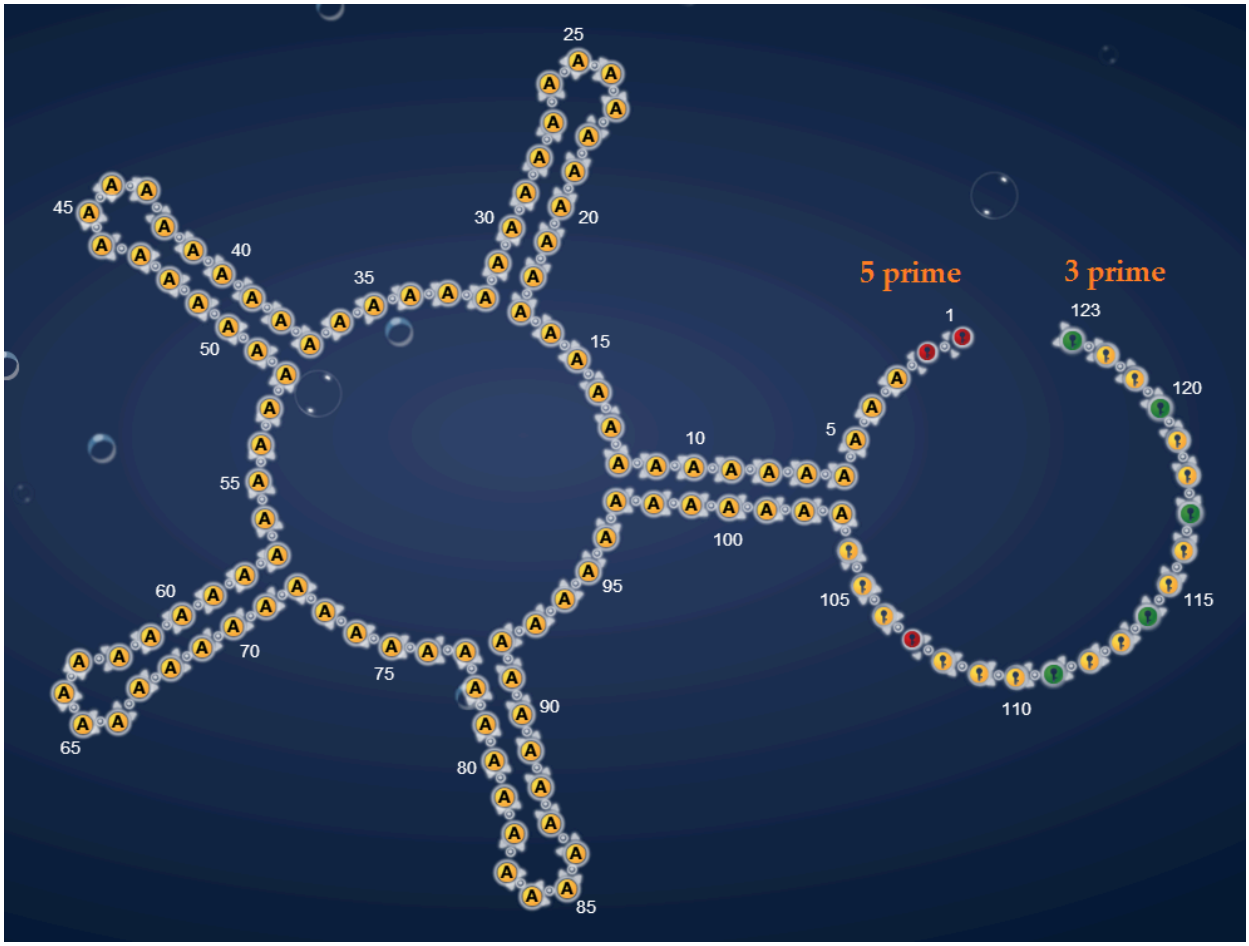
2-2 loop



There are some variations on how to stabilize a 2-2 loop. But in general, the higher the negative energy inside, the more stability it add to the puzzle. We still haven't found a way to stabilize 2-2 loops in lab yet.

5 prime end and 3 prime end - reading direction of RNA

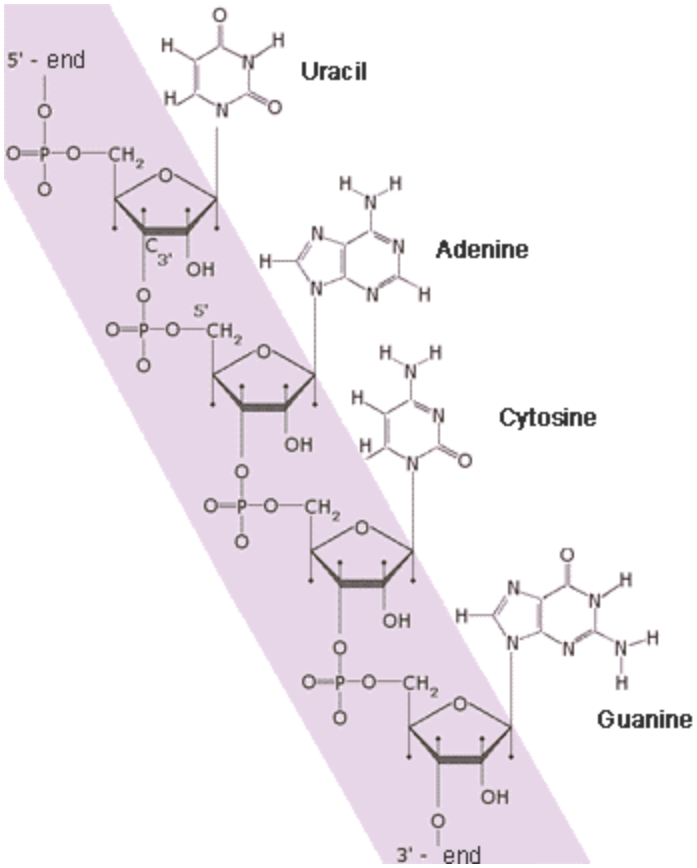
ElNando888: 5' is used to refer to the start of the sequence (base 1), 3' for the end of the sequence. It indicates the order of the sequence (5 April, 2013)



The prime numbers comes from which carbon on the sucker part of the base, the attachment is made. 5 prime means that the connection is made on the fifth carbon and 3 prime means that connection is made on the third carbon.

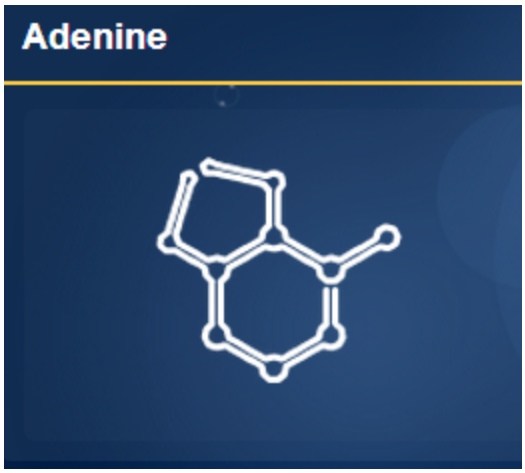
This video by Andrew Douch is doing some fine explanation on the topic.

Quasispecies explains: The two ends of an RNA molecule are not equivalent – it has two distinct ends, or termini, labeled 5’ and 3’ (“five prime” and “three prime”). Customarily, RNA is numbered starting at the 5’ end.



(Quasispecies' glossery)

Adenine



Base in RNA and DNA

Adjacent pairs - base pairs in stacks being right next to each other.

Antiparallel DNA - DNA runs antiparallel - or as it is formulated in a book: In essence, the two strands lie ‘head to toe’. (Molecular biology - Principles of genome function.)

Aptamer - target binding RNA molecule (Definition from Molecular biology of RNA, Elliot & Ladomery). This means an aptamer is an RNA structure that has the ability to bind up with something not RNA.

“The term aptamer means something like "fitting pieces" (from the Latin word aptus, meaning to fit, and the Greek word meros, meaning piece).”

Nano Tech Wire

Any subsequence of nucleic acid or protein, selected from a large random sequence-pool, used to bind to a specific target molecule

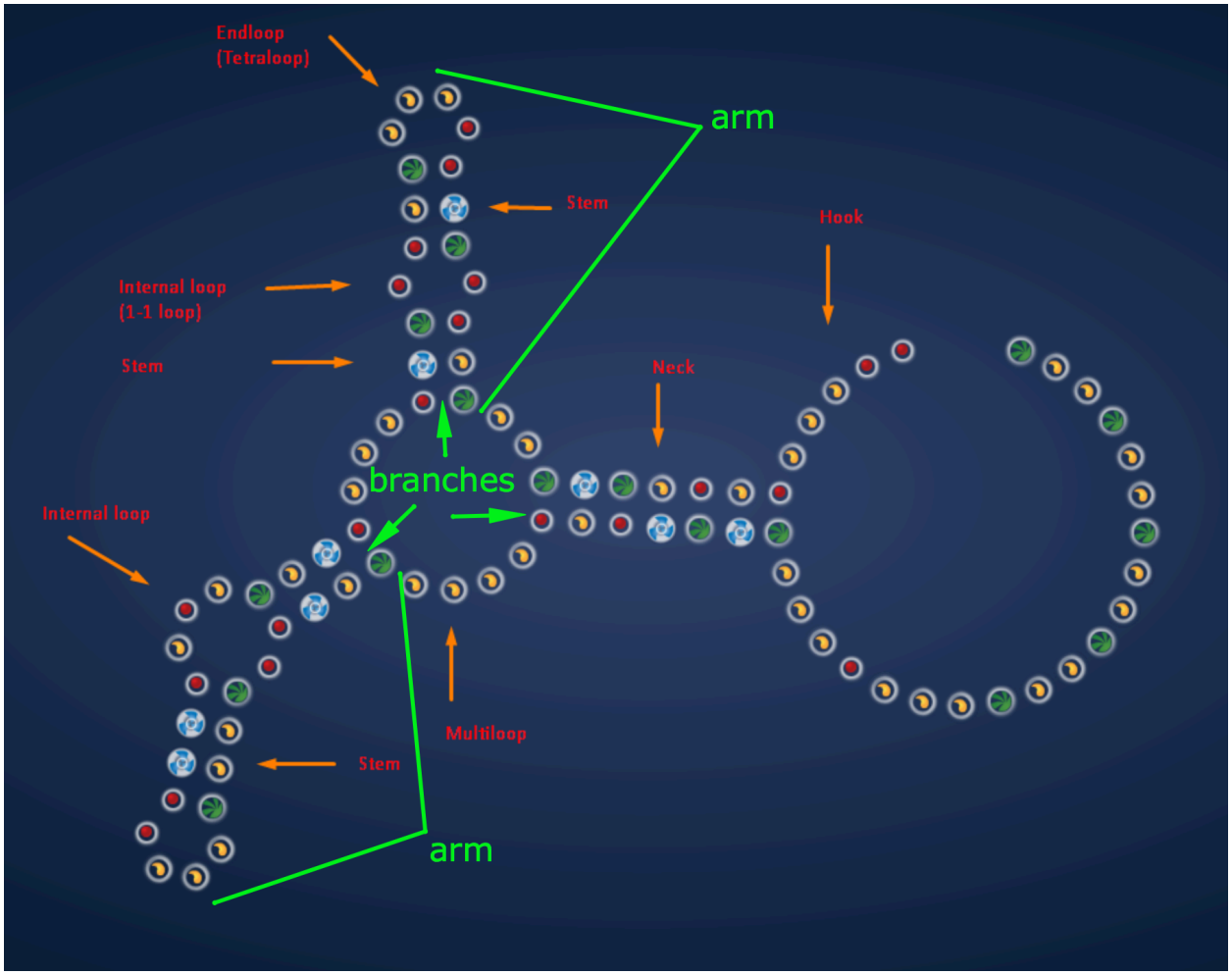
Wiktionary

The aptamer in our designs, is the loop structure with special nucleotide sequence, that is doing the binding up work to an outside molecule.

Or as Hogla said: The ring with the flashing red dot in the middle.

Thanks to Mat for reminding me that we didn’t had definitions on aptamers.

Arm – the same as leg. Meaning the whole arm in a design, eg. two stems and the elbow bulge or internal loop between them.



Thx to Machinelves for adding the arm and branches terminology to the picture.

Asymmetric multiloop - multiloop with different numbers of nucleotides between the multiloop arms

ATP - is the energy currency of all organisms on our planet.

Gerhard Gottschalk, Discover the world of microbes

Backbone - Tom has a very fine description here.

Backing pair - Backing pair - not the closing pair but the next one. (Same meaning as Mat's term Next pair.

Definition by jandersonlee

Bacteria - The goal of every bacterium is to become bacteria.

Stanley Falkow

Barcode, hairpin

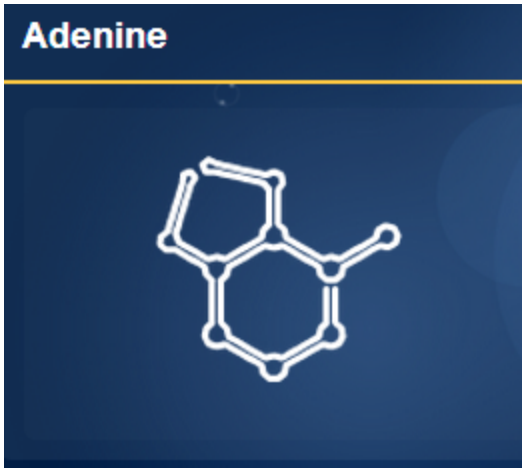
Special hairpin that is present in our cloud labs. Go see this explanation from the devs. There is also a section on it here.

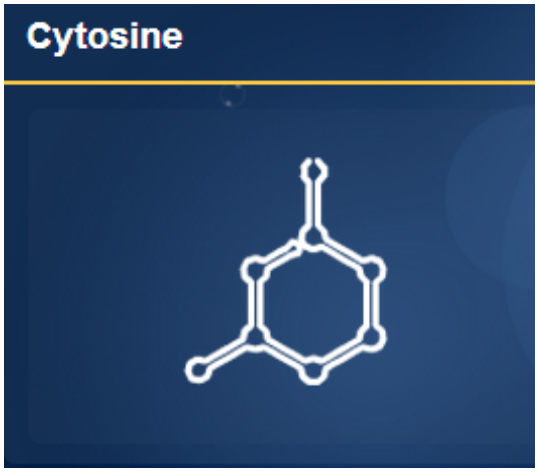
The hairpin barcode identifies the design. the rest of the design can be the same as others.

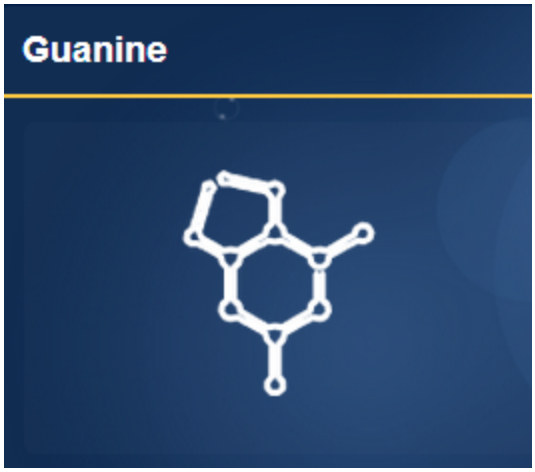
Hoglahoo

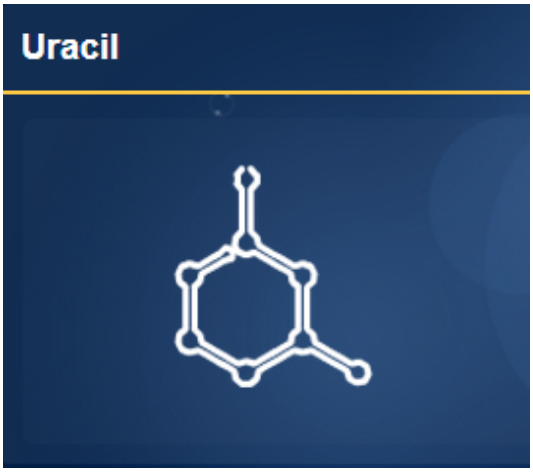
Base = nucleotide (Edward Lane)

Chemical structure of the bases.







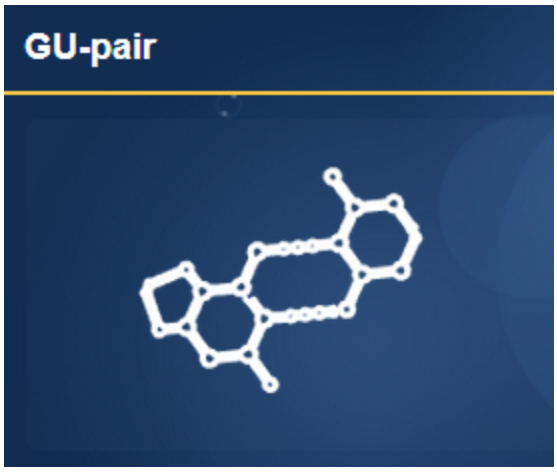


base pair - two nucleotides bonded to each other. There are three types of base pairs:

The GC-pair are strongest - of one guanine nucleotide and one cytosine nucleotide

The AU-pair is the middle strength base pair - consists of one adenine and one uracil

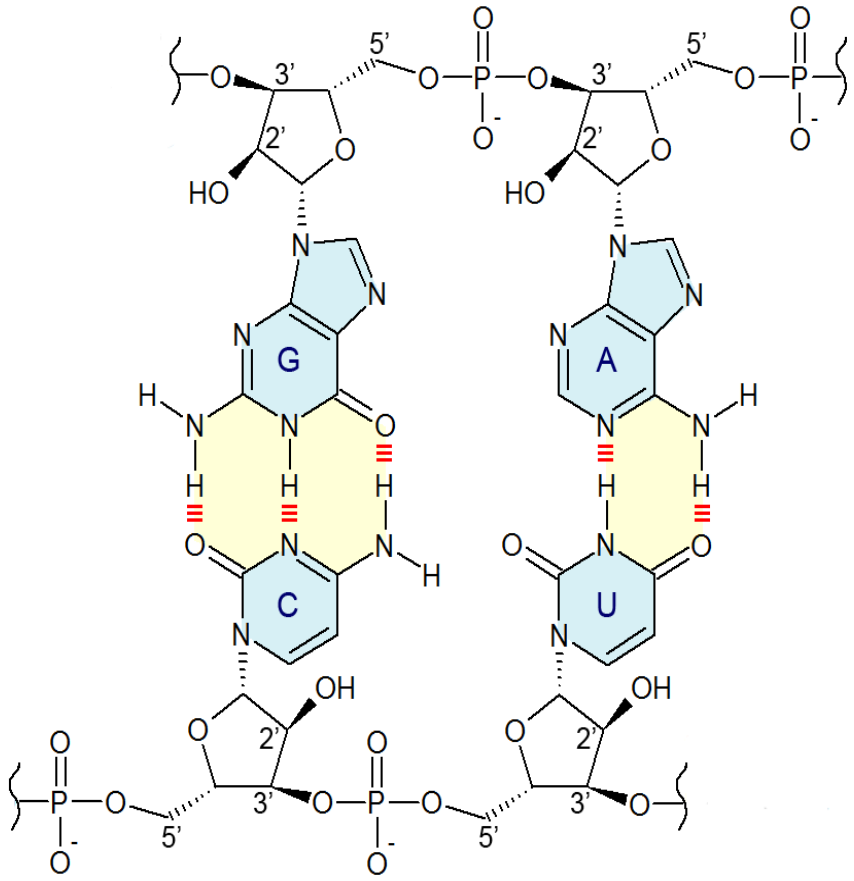
The GU-pair is the weakest - consist of one guanine and one uracil



In general a base avoids itself (Jareth Evers)

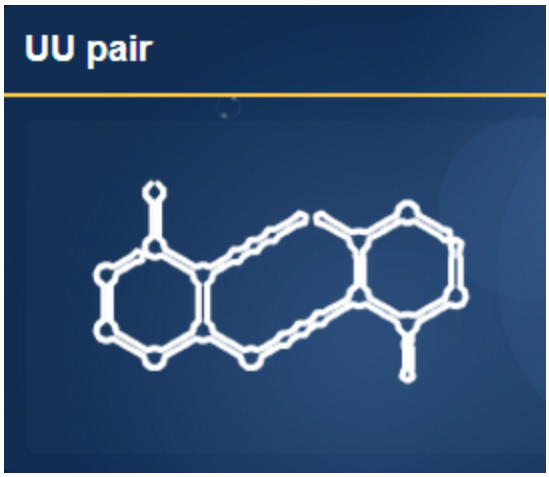
A base pair is two round things next to each other that are supposed to bond. (Hoglahoo)

Base pairing – The combination of H-bonding [h= hydrogen] and stacking interactions that causes complementary sequences to associate with one another.



(Quasispecies' glossery)

Base Pairing, unconventional



I also want to bring up the unconventional base pairs. Here is an outline from the intro I wrote to Rfam (collection of Natural RNA) some time back. I was fascinated as I saw very different base pairing in Rfam and I got lucky that I accidentally stumbled upon an explanation in one of the

books I was reading.

GC and AU base pairs are known as canonical or Watson-Crick base pairing. Those are the strong bindings. One of the non-canonical base pairs we already know is GU. Here is an intro to the others.

“By contrast, the other eight possible pairings (AA, AC, AG, CC, CU, GG, and UU) are much weaker - energetically close to no pairing at all. In the case of the strong pairings, the two surfaces fit together, so to speak, whereas in the case of the weak pairings, there is a mismatch between the respective molecular shapes.”

George Church, Regenesis, page 30

There are 29 possible base pairs that share at least two hydrogen bonds. The most commonly occurring noncanonical base pairs are the sheared GA, GA imino, AU reverse Hoogsteen, and GU and AC wobble pairs.
[The Molecular Interactions That Stabilize RNA Tertiary Structure: RNA Motifs, Patterns, and Networks]

Base triplets - interactions involves more than two nucleotide residues.

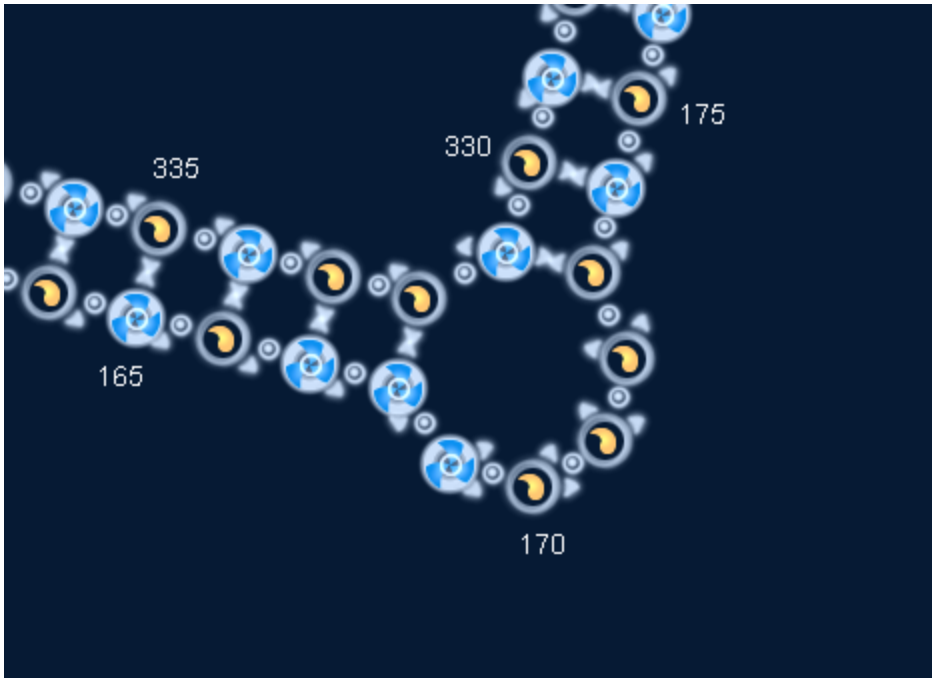
RNAMotifScan: automatic identification..., Rhiju’s science paper collection.

Biopunks - John Dwyer’s term for non-scientists like Foldit and Eterna players dabbling in science. (It is meant as a compliment)

Blocking point

Blocking point is when you have bulge closed by a AU and the bulge is all A, the U wants to bond with one or more of the unbonded A's in the bulge, so you change that A to a U and block the bonding and it works. (Definition by Iroppy)

Or put real short, also by Iroppy: Blocking stops a bond from forming



Nucleotide 169 is a blocking point, if not there the puzzle would have red spot here.

Recently I discovered that a blocking point also can be a C. Here is an example from a puzzle. When the puzzle is solved, if this blocking point (base 10) gets removed, the puzzle becomes unstable again:

Blocking points prevent the structure from sliding. For more on blocking points, go see the blocking section in Tips and tricks.

Boost points - could also be called the power of the loops. (Phrase by MarkPhilip)

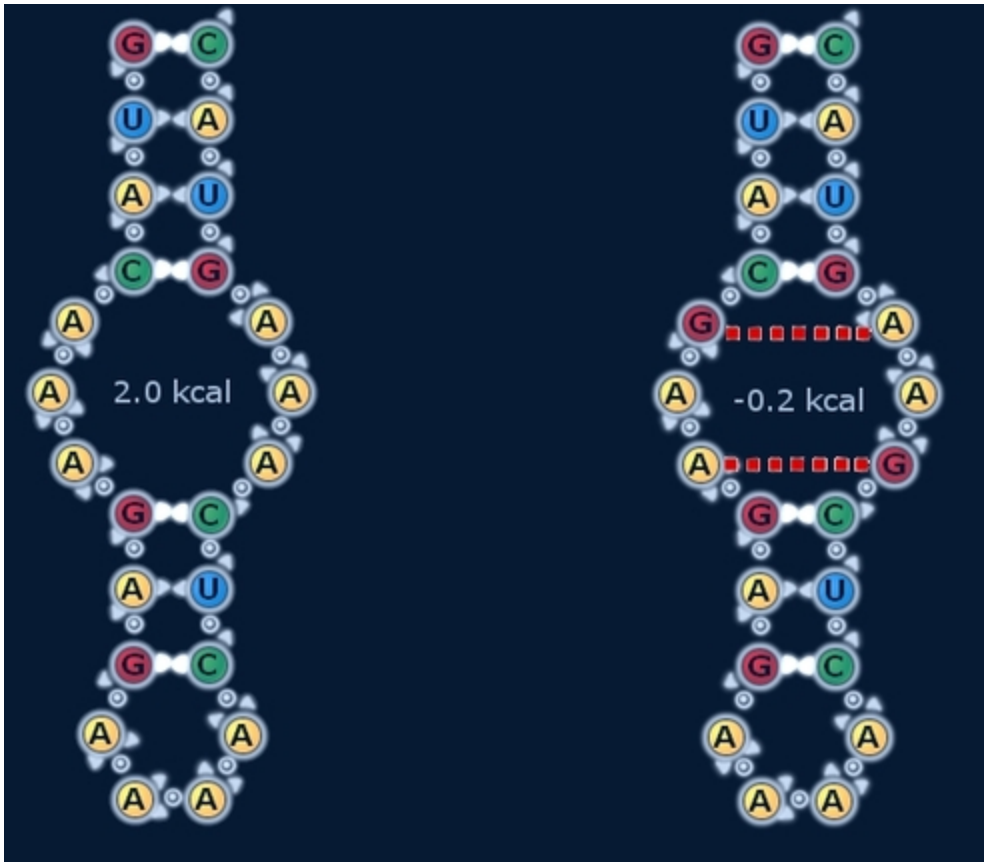
Boost points is an expression Lroppy uses for placing red or green nucleotides in multiloop ring, end loop or hook area, in a spot where it raises the level of negative energy inside and stabilizes the structure. (See also blocking and catalyst points)

Mattduncan asked if there are a way to spot which color of boost points to use and where. Here is my advice:

Mostly boost points are 1 red G's for a loop on one stem or two G's for a loop between two stems. For multi loops boost points can be G's and C's. U's are only rarely used as boosts and mostly when two relatively big loops are close together, with only one bond between them. Triloops should only rarely be blocked with a blocking point - boost can be G, C or U. For 2-2 loops there are special boost settings. See 2-2 loops.

Boosting – Boosting a loop means give the loop a lower free energy. (Explanation Hoglahoo) Which is good as it stabilises a loop.

Changing the default A-A mismatch in a loop to a G-A (or other) mismatch, lowering the energy of the loop in EteRNA. This is how EteRNA attempts to account for the formation of non-canonical base pairs in loops.



(Quasispecies' glossery)

If a different than A unpaired base makes it work then it boosts - that actually seems like a better

definition for boost

Hoglahoo

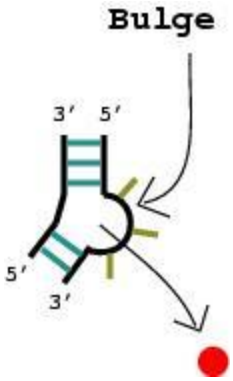
(if at boost spot - the bases just in the opening of the loop to either side)

Specification between boosting and mismatch
biggestlegoheroicafanever: hoglahoo, is boosting an scientific term, or is it game slang for eteRNA?
hoglahoo: it's a slang term for eterna - someone told me that lroppy coined the term :)
biggestlegoheroicafanever: Okay, so boosting is street RNA slang.
hoglahoo: yeah, it's grand theft rnauto
hoglahoo: 'boost three loops before the cops arrive!'
Eli Fisker: mismatch is the scientific term
hoglahoo: sort of - but mismatch doesn't imply changing the energy of a loop like boosting does
hoglahoo: I prefer to think of the mismatch affecting the loop now - but for some reason, the idea of boosting seems simple and helpful
hoglahoo: maybe it's good to say the mismatch is the only part of the loop that can boost it
hoglahoo: except in tetraloops (there's always an exception?) (12 Feb 2014)

For more on boosting go see the boosting section in Tips and tricks.

Bp - scientist abbreviation for base pair

Bulge - A protruding part; an outward curve or swelling. (The free online dictionary)



(from: http://www.biomath.nyu.edu/?q=rag/create_tree)

1-nt loop means only 1 nucleotide is unpaired, making those little pentagon shaped bulges

Vors

Bulge and anti-syn pair - see Freywa's article on bent bulges.

Cap (also called 5 prime cap) What gets put on messenger RNA.

"What is the function of capping? Capping turns out stabilizing mRNA"

Qiang Zhou
Molecular and Cell biology 110, Fall 2008, Berkerley,
lecture, RNA processing: Part lll

It has several other functions too.

Capillary Electrophoresis - is essentially a super powerful gel electrophoresis.

Eterna dev

See an explanation of gel electrophoresis and Tom's explanation about capillary electrophoresis here.

Capillary Electrophoresis builds onto Sanger sequencing. Here is a small video about Sanger sequencing.

The Sanger method has been modified to improve its efficiency. Fluorescent labeled analogs are spiked into the experiment instead of radioactive nucleotides. Samples travel by capillary electrophoresis to a detection area within a DNA sequencing machine where a laser excites the fluorophores producing fluorescence emission that correspond to the base calls.

(Bioinformatics and functional genomics, Jonathan Pevsner, page 545)

Catalyst points

The definition of a catalyst point is a U or C that doesn't seem to lower the energy of a loop in the way a boost point would, yet when placed in a loop, helps solve the puzzle. (Definition Starryjess)

See Starryjess post on this topic under catalyst points in Tips and tricks guide

cDNA - Stands for complementary DNA

- a DNA that is complementary to a given RNA which serves as a template for synthesis of the DNA in the presence of reverse transcriptase

Merriam-Webster

Technique used in relation to the RNA in our lab.

Clamping

As for clamping, that's a common term that (among other things), means limiting a measurement to a certain range. So, for example, if the nominal range is 0.0 .. 0.1, a measurement of 1.3 would be "clamped" at 1.0.

(Explanation Omei)

Cloud lab

Our new lab - with lots of labs. I have a small intro to it here.

Codon - a group of three bases in mRNA (messenger RNA) that codes for a single amino acid.

From genes to genomes

Complementarity – Complementary RNA bases are bases that can pair with one another – typically A and U or C and G. When two strands associate, one is oriented 5'-> 3' while the other is oriented 3'->5'. The sequence 5'-AUCG-3' is complementary to 3'-UAGC-5'.

(Quasispecies' glossary)

Complementary strand - a nucleic acid strand that will pair with a given single strand of DNA (and RNA)

From genes to genomes

Conformation - scientist word for shape.

Covalent bond - The force that links nonmetals together in molecules.

Matthew W. Stoltzfus

C-trick - see Freywa's post

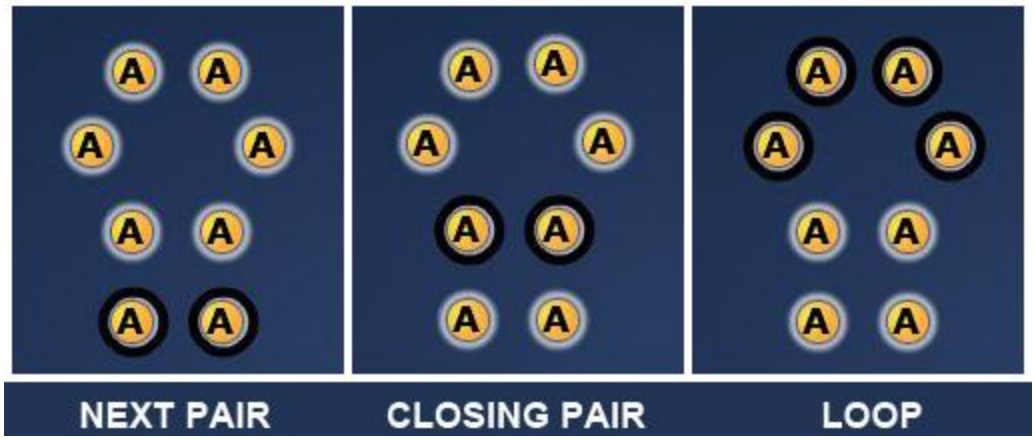
C-stifling - term by Freywa to describe the use of excess use of C's in loops to block off the single nucleotides in a loop from pairing up elsewhere.

Christmas tree - Eterna slang for overuse of **G**uanine (red) and **C**ytosine (green) bases. Or simply Christmas 3's as Starryjess coined it. :)

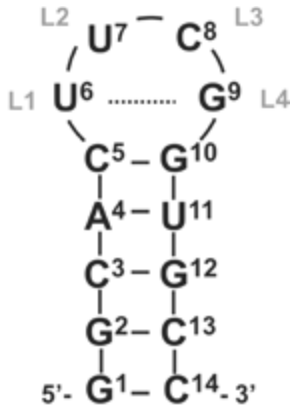
Also see Dimension9's post

Closing base pair (same as closing pair)

Mat made a nice picture explanation



Closing pair - The pair that appears at the end of a stack. C5 and G10 constitute the closing pair in this hairpin. (Same as closing base pair)



(Quasispecies' glossary)

Coaxial stacking - Two of the RNA helices form a coaxial stack, meaning that one helix sits on top of the other, and they are collinear.

(Explanation from paper Jnicol found)

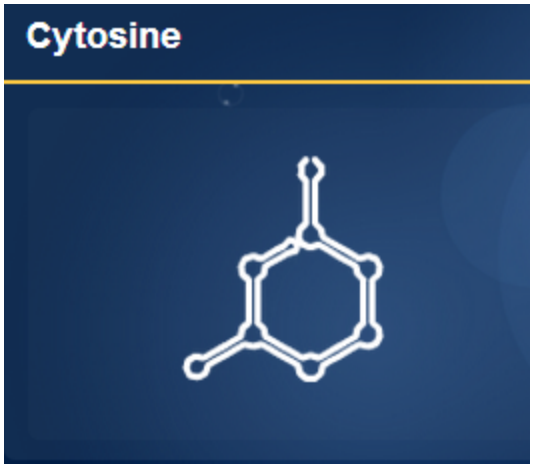
Codon - A section of DNA (three nucleotide pairs in length) or RNA (three nucleotides in length) that codes for a single amino acid. (Definition from EverythingBio.com)

Corner loop - bulge of any size (Term coined by Freywa)

Cub scout - Eterna slang for overuse of **A**denine (yellow) and **U**racil (blue) bases.

Also see Dimension9's post

Cytosine

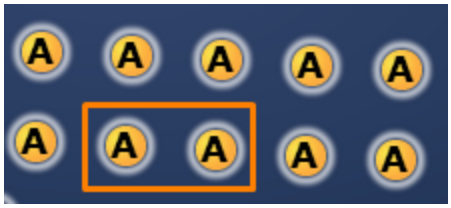


Dangling ends - also what we call hook - the two single stranded sequences at beginning and end of a lab design.

Dev/s - development team

Dev chat - scheduled chat with the developers

Dinucleotide - Scientist word for two bases in line in a strand. Example



Dot plot - Jee has a great explanation here

For more on how to use dot plot, see Lab guide for new players.

Dot bracket structure - same as shape notation.

This means periods represent loops, and the parentheses base pairs.

Starry gave this example:

Looks like this: ((....)) is two pairs and a hairpin loop - a loop 4 nucleotides big

DMS

DMS stands for dimethyl sulfate. Basically DMS is footprinting or fingerprinting if you like, of the RNA structure. It is a chemical check of which bases are paired as they should be and which are not. Rhiju describes it as doing following: DMS reacts only with unpaired A and C nucleotides.

Here is a science paper, the abstract do a good job of explaining what DMS is.

E Coli - The lab rat of the bacterial world

Jesse Noar, BacterioFiles Micro Edition 122 - Coliforms Consume Caffeine Compulsively

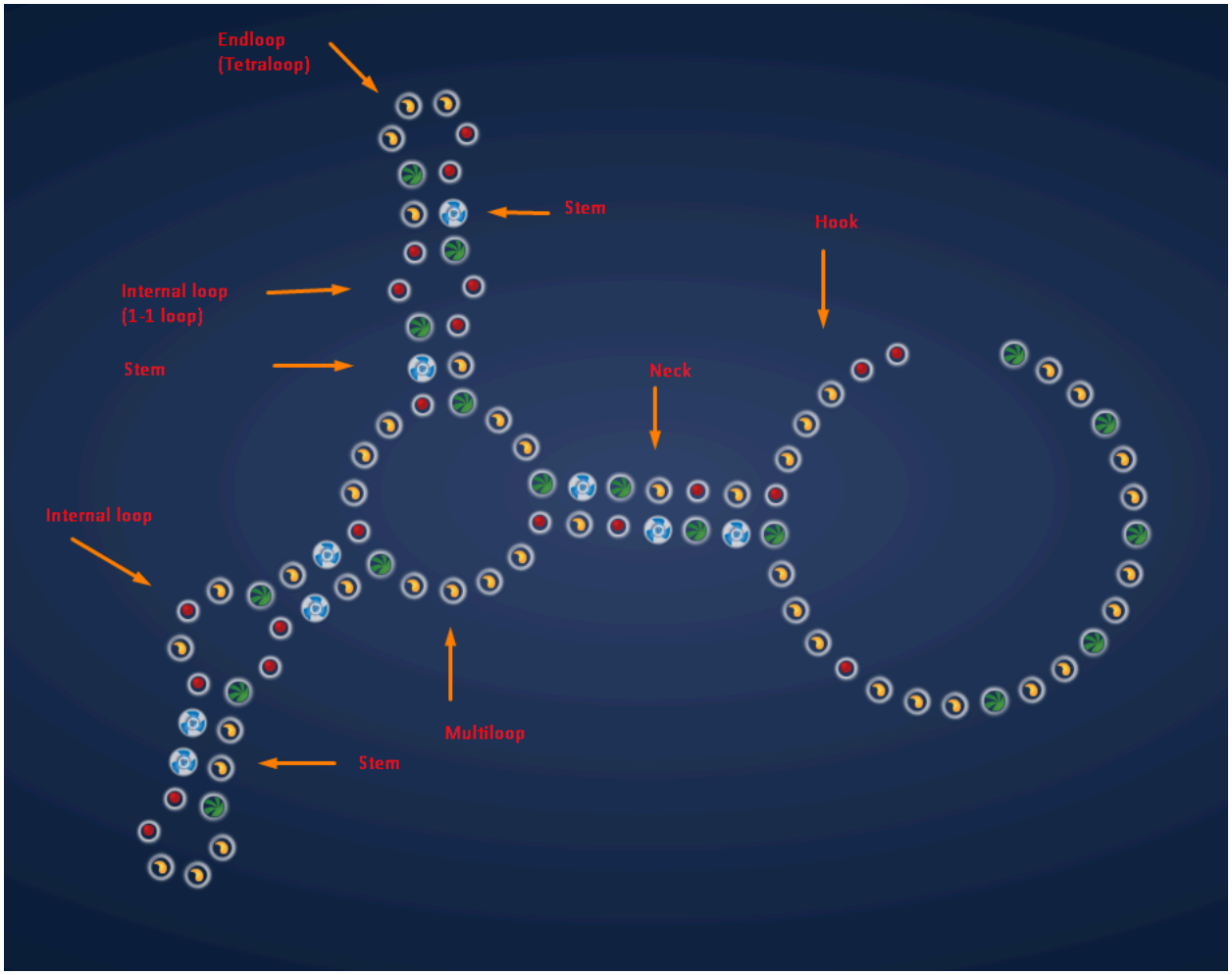
Electrophoresis - separation of macromolecules (DNA, RNA or protein) by application of an electric field, usually across an agarose or acrylamide gel.

From Genes to genomes

Gel electrophoresis is a method for separation and analysis of macromolecules (DNA, RNA and proteins) and their fragments, based on their size and charge.

Wikipedia

Elements



Element - also what scientists call a motif

End loops - Mat coined the term Ends for a loop at the ending of a stem. I added the word loop to specify the meaning of the term. On suggestion from Starryjees, I will mention the scientific term for end loops too, they are called hairpin loops. Sometimes I have called them outer loops too.

Energy - Natural and target shape:

Two base pairs together is called a quad. The energy of the quad, changes based on which base pairs gets grouped together. Loops also have a certain amount of energy. Adding all the energies up, will give the puzzles total energy. That total energy is what you can see in the top left corner of the puzzle, and there is an energy for both target and natural shape.

Timeroot gave a fine explanation of the energy in natural shape versus:

So the target shape has one energy.
The natural one has a different.
You want to bring the target shape's energy down to the natural shape's energy
...or vice versa, bringing the natural shape's energy up..
Because once the target energy is below the other shape's energy, the target shape is now the lowest-energy configuration, and the RNA will take this shape you want.
If you make small changes, like swapping a pair, or adding something in a loop, you might see the energy level change.
You can also switch the native mode and see how the energy changed there.
Anything that brings the two closer together is good. :-)

A chemical reaction with a negative free energy change should occur spontaneously, given enough time.

Molecular biology - Principles of Genome Function, Oxford, 2010, page 103

Energy, difference
Difference in Free Energy between Natural and Target modes, must be > 0 in order to solve the puzzle. (12 Dec 2013)

RedSpah

Ensemble defect - The ensemble defect is the average number of incorrectly paired nucleotides at equilibrium over the ensemble of (possible) secondary structures(=shape). Definition Jeehyung
See also ensemble diversity

Ensemble diversity - When RNA folds, often it shifts around various different shape. Both ensemble diversity, defect is a measure to indicate how much time the RNA stays in the actual "target" shape. Definition Jeehyung

We want RNA to stay in target shape, so it will fold correct. Therefore we want ensemble diversity to be low.

Entropy - Entropy is a measure of disorder in a chemical reaction. The universe likes to be disordered rather than ordered. That is why space is mostly empty with only small collections of atoms in stars. It's like when people fill a bus, they never sit next to another person unless they have to. Atoms won't bond with other atoms unless they have to. Definition from Davenpoc

I also like the definition from Wikipedia:
Entropy is a measure of disorder, or more precisely unpredictability...

...English text has fairly low entropy. In other words, it is fairly predictable. Even if we don't know

exactly what is going to come next, we can be fairly certain that, for example, there will be many more e's than z's, or that the combination 'qu' will be much more common than any other combination with a 'q' in it and the combination 'th' will be more common than any of them.

Entropy is the measurement of disorder - how chaotic something is. From youtube video: Gibbs free energy

I like this explanation I found in Essentials of Biology, third edition, 2011:
In order to overcome the natural tendency toward disorder, energy input is required, just as it requires energy to organize a messy room.

We want RNA to fold correct, so we want a low amount of chaos. Therefore entropy should be low.

Epigenetics - RNA switches are not the only thing that can turn on and off genes.
I like this description I found about epigenetics in the article A Thing or Two About Twins.

“If you think of our DNA as an immense piano keyboard and our genes as keys—each key symbolizing a segment of DNA responsible for a particular note, or trait, and all the keys combining to make us who we are—then epigenetic processes determine when and how each key can be struck, changing the tune being played.”

Also an australien cancer study came up with an interesting find. Turns out there are some epigenetic "masterswitches"

"In this study, we identified 35 domains including 251 genes. While the genes may seem to be functionally unrelated, their coordinated regulation in the cancer genome suggests the presence of epigenetic 'master controllers' that can switch on or off very large regions of DNA."

Eterna ensemble algorithm (EternaBot)

Our own bot, made from patterns found by Eterna players and programmed by Eterna developers.

Different modes of the bot have been run for the lab designs.

Conventional mode: The conventional bot runs 2 very simple strategy - GC pairs must be 60% and dotplot must be clean. It was designed to test if EteRNA players' non conventional strategy is playing a role in our bot's performance. And it does :]

Explanation from Jee

EteRNAddiction - term coined by Jieux. In case of EteRNAddiction, go see EteRNAaddicts.

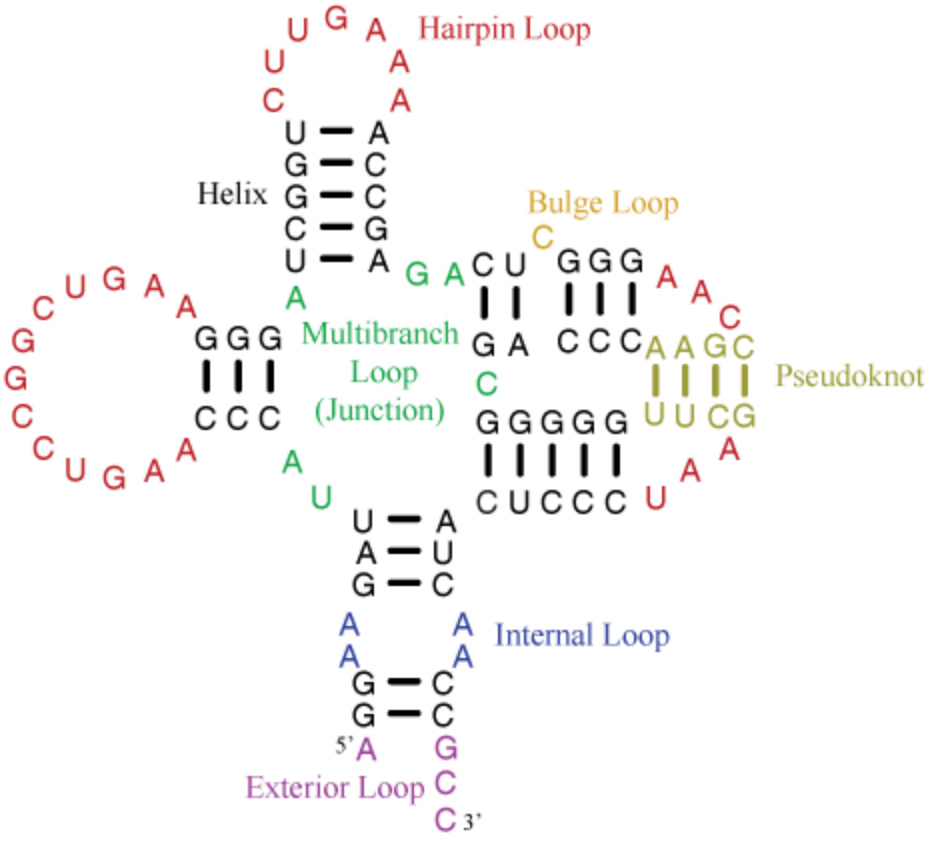
Eterna Bot - Bot made based on player lab strategies. It is a RNA structure prediction algorithm. Here is a link to our bot, here is a link to an intro and read more about it in these two forum posts.

Why do player strategies work better than existing algorithmic models?
How eterna ensamble algorithm works

Eternaists - Adrien Treuille's name for us eterna players

Eterniverse - term for the entire eterna community and the Eterna game. (Term coined by Edward Lane)

Exteriour loop - open loop at the end of the design.



http://x3dna.org/articles/exterior-loop-in-rna-secondary-structure

Thx to Omei for sharing.

FASTA format - an output format for nucleic acid or protein sequences. (Bioinformatics and Functional Genomics, page 896)

JandersonLee has made a script FASTA Lab Submissions v0.01 that will give you a lab designs data in FASTA format, which is a standardized search format, that many outside RNA structure prediction tools, uses for input. If you input the main lab ID in JandersonLee's script, it will output all the designs in the lab at once in correct FASTA format, ready to dump in an outside tool. For an example on how to use the script, see here.

FASTA is also: The first widely used algorithm for database similarity searching. The program looks for optimal local alignments by scanning the sequence for small matches called "words." Initially, the scores of segments in which there are multiple word hits are calculated ("init1"). Later the scores of several segments may be summed to generate an "initn" score... (Bioinformatics and Functional Genomics, page 896)

For further explanation about the FASTA format, see Wikipedia

Free energy - The wiki has a brilliant explanation on free negative energy. Also see this fine explanation Ksteppe have given. Understanding Free Energy (using Legos). The legos was Janelle's idea.

See this video explanation.

The word "free" is a bit of a misnomer. Think of it as available energy (4 May, 2014)

Nando

FMN - is the name of the target molecule that RNA binds to. Jeehyung (This is for the switch puzzles)

The small molecule FMN is added to the solution and that impacts on how the whole thing (switch) binds slightly.

Explanation Edward Lane

FMN is a molecule is called flavin mononucleotide. It's a molecule that is present in all of our cells. All living cells. It acts as a redox cofactor, meaning it is an electron donor or acceptor in many enzymatic reactions. It is just a favorite metabolite or small molecule biochemists play with. There is a RNA sequence that are known that specific bind with FMN. (guess that is our

aptamer, Eli)

Rhiju

Folding kinetics - You know how the transcription process goes, right? Well, the RNA starts folding before it's finished being synthesized. The first foldings may be stable enough to influence the end result, whatever the MFE might be

EINando888

FRABASE - chatlog lesson on how to use it by Nando

More introduction on FRABASE

GC-heat - Using GC-pairs in multiloops to max the negative energy inside

Gene - Ultra short explanation from Gene to protein: Dna helix opens up, transcription enzyme mates with opening, rna fragment produced, rna fragment runs thru ribosome protein, resulting in protein the dna gene codes for. It is the one-gene one protein paradigm.

Rickyjames

GC-line

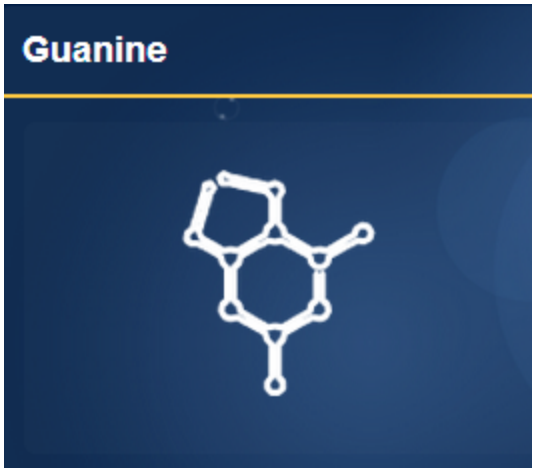
I noticed in puzzles that a rare time that 3 GC-pairs in line could be necessary, to stabilize a puzzle.

The GC-line trick even worked in the neck area in one of our lab puzzles with a really short neck. Elsewhere than in the neck in lab puzzles, three GC-pairs will not be good. (The term GC-line is coined by Starryjess)

GU-pairs - the geometry of GU pairs is different than AU or GC, so GU pairs might help an RNA adopt a conformation that allows it to do something interesting.

Quasispecies

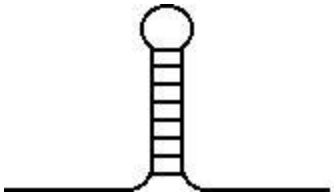
Guanine



Guide RNA - RNA that guides an enzyme to find the correct spot on a DNA sequence.

Read about such an RNA in the article A Powerful New Way to Edit DNA

Hairpin - lollipop-shaped structure (expression from Wikipedia)



(from http://www.emunix.emich.edu/~rwinning/genetics/transcr2.htm)

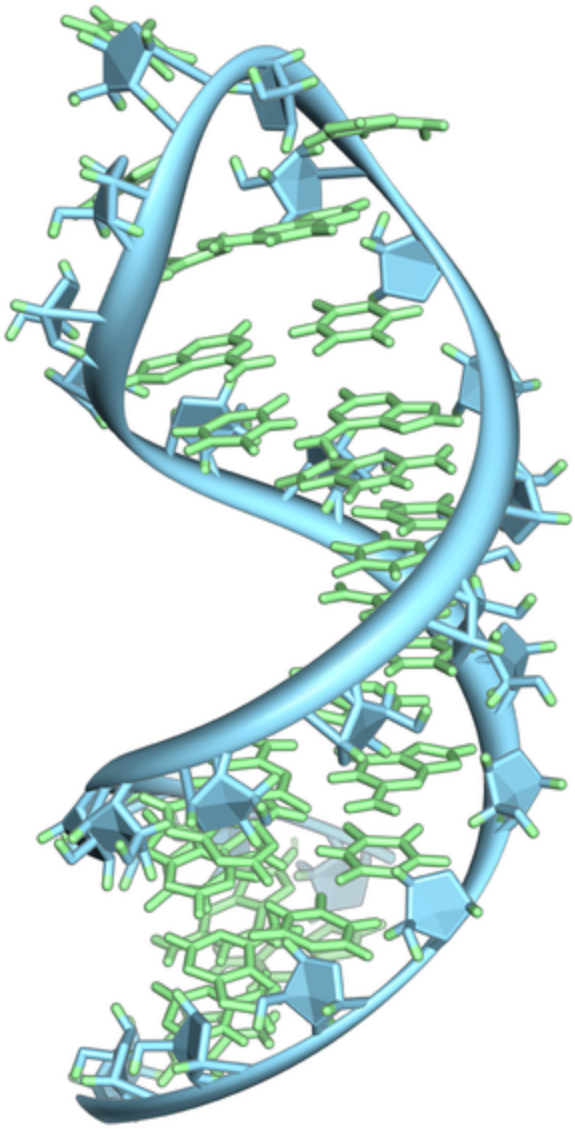
Hairpin in a hairpin - A term Drake coined and that have been causing me headache since.

Here is Hogla's explanation:
hoglahoo: I think drake made some long speeches about that
vors: it's like Inception
hoglahoo: I think it is based on the tendency of the hairpins + the big loop to open up into one large, single hairpin
hoglahoo: you have to destroy the large hairpin to let the other two form

Hammerhead ribozyme - is an RNA that cuts itself.

Quasispecies

Helix – The sugar-phosphate backbone of RNA (highlighted in mauve in the illustration under "things to remember") causes the chain of bases to turn slightly.



(From Quasispecies glossery)

Hidden instabilities - See Quasispecies explanation

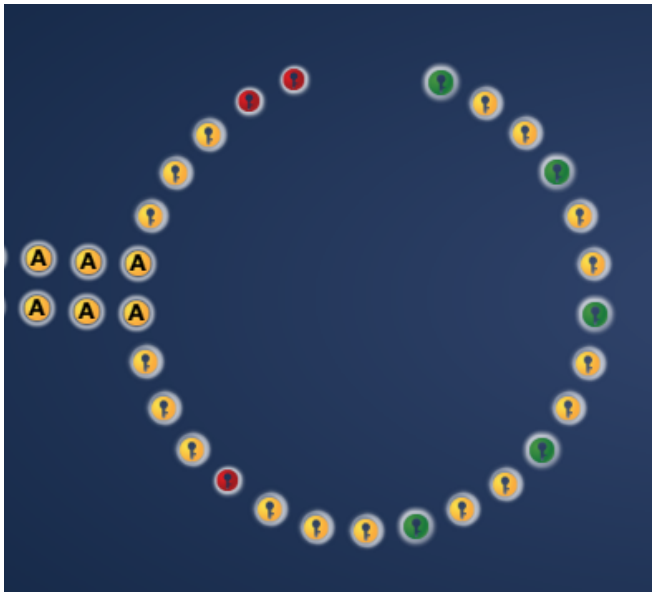
High energy - the same as low negative energy

Hoogsteen base pairing - I found this page that shows what makes a pairing a hoogsteen pairing.

Hook – the part of the design with the opening, dangeling ends, which looks a bit like a fishing hook. This is for lack of a better word for this part of the puzzle As Jeehyung says, those are specific sequences that we need for the lab synthesis. The new big lab no longer requires these

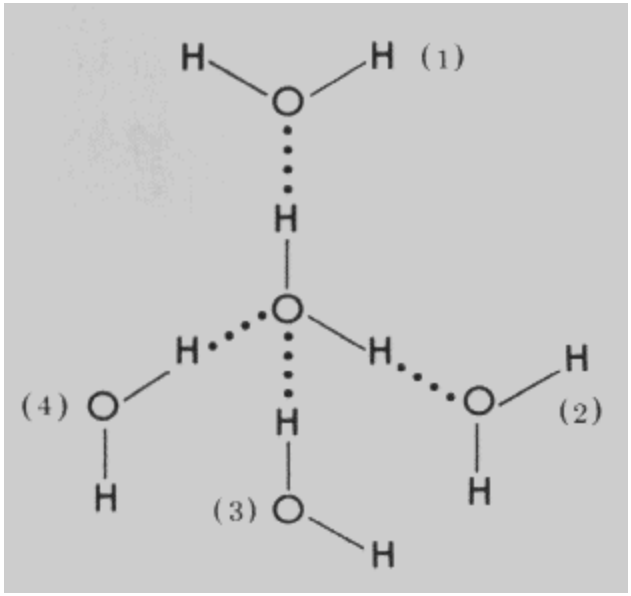
extra sequences.

Thank to Mat for adding that the term Hook was coined by Dimension9.
The locked nucleotides are the hook area.



For more about it see here.

Hydrogen bonding (H-bonding) – Hydrogens attached to certain elements (like nitrogen and oxygen) can be “shared” to varying extents. The dashed lines in the figure below show hydrogen bonding in water. In RNA, hydrogen bonding between bases allows base pairing.



(Quasispecies' glossary)

For more on hydrogen bonding see this introductory video.

Hydrogen bonds are responsible for specific base-pair formation in the DNA double helix.
(Biochemistry, seventh edition, page 8)

Illumina sequencing - a method our lab has shifted to, to enable us getting a huge amount of lab designs and data compared to earlier. Learn more about it here.

Independent stability - a sequence that naturally folds into a given shape, that shape has independent stability. Now maybe you want to throw a bunch of other shapes around it, but not directly attached to it and now it misfolds. It is still independently stable because if those other shapes are made stable, and that first shape is untouched, it will also end up stable. It is only misfolding because the other shapes are grabbing at it, not because it won't hold up on its own.
Explanation by Hoglahoo

Inverse RNA folding - means designing RNA sequences that fold into a given structure. Definition from INFO-RNA - a fast approach to inverse RNA folding, Rhiju's science paper collection.

This is exactly what we are working with.

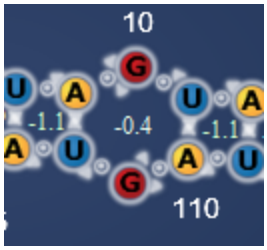
In situ means in the (natural) environment

***In vivo* -** occurring within a living organism, definition from Mosby's Dental Dictionary

In vitro - observable in test tube, definition from Miller-Keane Encyclopedia and Dictionary of Medicine, Nursing, and Allied Health

Or shorter as Minja Lee said: *in vitro* means "in the lab"

Internal loops tend to have the same number of unpaired nucleotides on each side of the stem, where as a bulge is characterized by an imbalance in the number of unpaired bases (Structural Bioinformatics, 2009)



Internal loops: 1-1

Mismatches other than G-G in 1-1 loops

> The C-A, U-U, C-U, and G-A mismatches have not received much attention due to their higher energy relative to the G-G mismatch.
>
> (Quasispecies' glossary)

Internal loops: 2-2

All-A flanked by non GC pairs

These tend to appear as extensively paired in lab results. It is worth noting that all designs have placed GC pairs adjacent to AA / AA mismatches in the 2-2 loop.

5'-GCUAAAGC-3' 5'-GCUAAAGC-3'
3'-CGAAAUCG-5' & 3'-CGAAAUCG-5'

> NMR spectra of these two sequences indicate that the AU pairs adjacent to the 2-2 loop may not form Watson-Crick pairs. If the flanking sequences behave differently, it might be interesting to see if the mismatches do so as well.

All-C

A destabilizing sequence that has not been attempted in lab yet. At lower pH, C-C hydrogen bonding occurs (from protonation of C) and stability increases.

All-U

NMR shows that this loop can be stabilized by U-U hydrogen bonding, regardless of the adjacent base pairs. It would be interesting to see if an all-U loop is scored as "paired" or "unpaired" in SHAPE.

CU and CA Mismatches

These would likely behave very differently depending on their orientation and the sequence surrounding them.

5'-AC-3' 5'-UC-3' 5'-CU-3'

5’-CA-3’ 5’-CU-3’ 5’-UC-3’

These mismatches apparently lack at least one type of base pairing (H-bonding through the imino hydrogen). Other forms of base pairing cannot be ruled out, however, and success in lab is not guranteed. The all-A loop has fared poorly in lab, but it also appears to lack imino hydrogen pairing when examined by NMR.

5’ - GCUCAAGC - 3’
3’ - CGAACUCG - 5’

Unexpected signals arise in the NMR spectrum of this duplex, possibly due to protonated
forms of A and C. Using the mismatch and the same flanking pairs may prove interesting.

5’ - GGACUUCGGUCC - 3’
3’ - CCUGGCUUCAGG - 3’

The duplex below forms a tight helix despite the tandem mismatch being flanked by GU pairs. It would be interesting to see if a similar loop sequence in a player design is scored as “paired” or “unpaired” in SHAPE.

Asymmetric 2-2 Loops

There are many possible sequences, and the overwhelming majority are untested.

Asymmetric internal loops

Lower energy with U-U mismatches

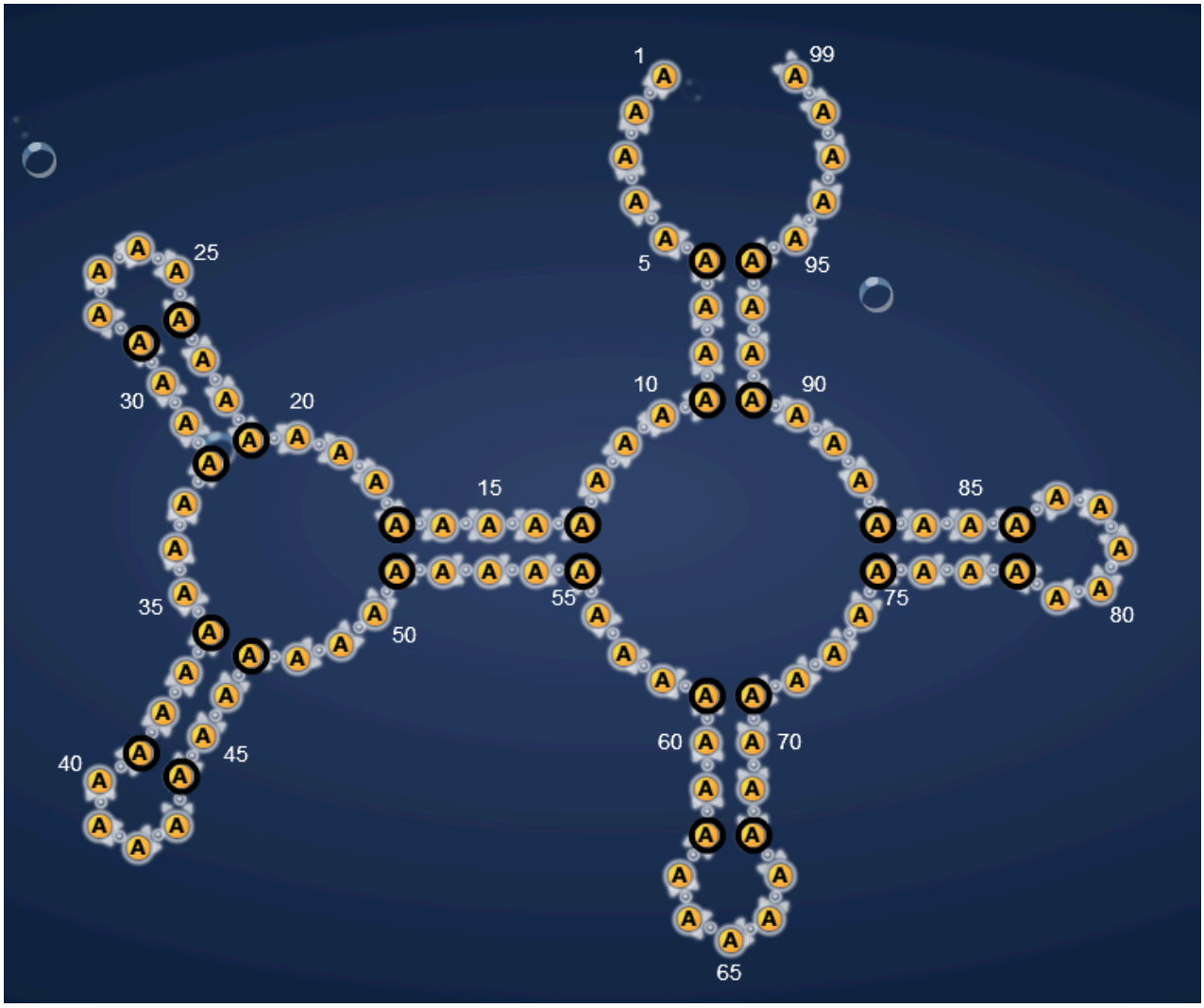
The first mismatched bases in a loop can significantly alter its stability. The conventional approach to “boosting” employs G-A mismatches. U-U mismatches do not receive as large of an energy bonus in EteRNA, but are fairly common in nature.

(Quasispecies’ glossary)

“Junk” DNA - interesting article explaining “junk” DNA.

Junction

Here is a picture with the junctions are highlighted.



Kcal - energy measure.

Kcal is a unit, the number you see on the top left is a measure of the Gibbs free energy of the structure. A predicted one, to be precise. When free energy is positive, the reaction needs external energy to occur, when the free energy is negative, it means the reaction (here, the folding) is spontaneous and the more negative this value goes, the more stable the molecule is supposed to be. In order to solve puzzles, you just need any value that is below 0, you gain nothing by trying to make it as negative as possible

Nando

Labbing - EteRNA slang for doing lab work (term coined by Lroppy)

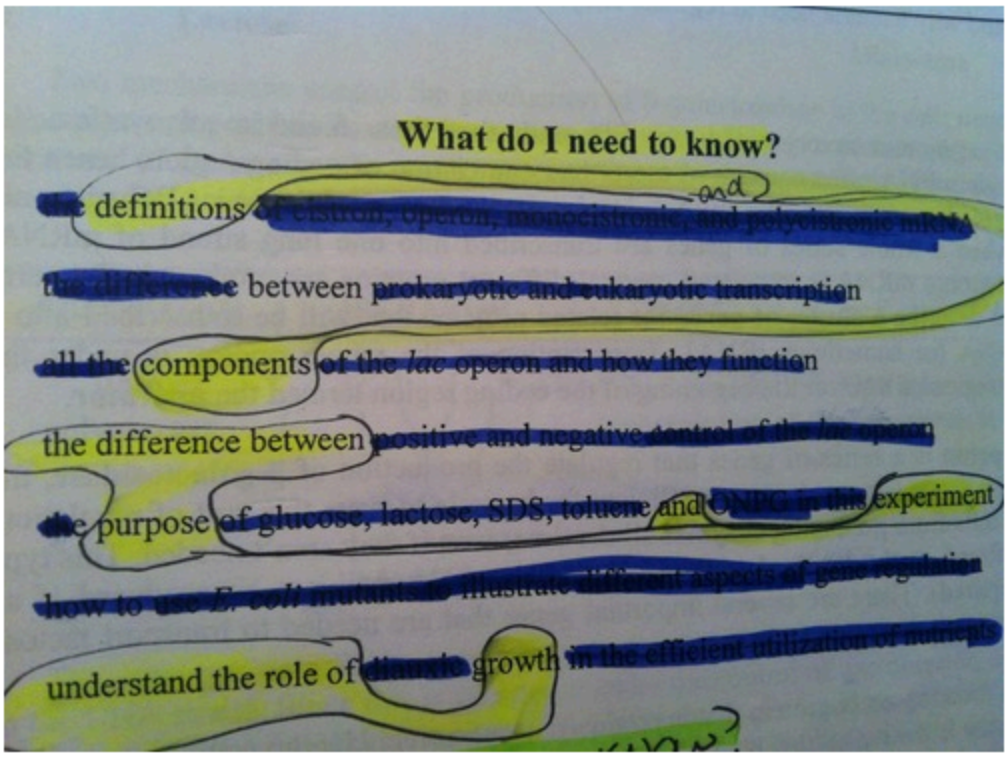
Labcoats - friendly nickname for our scientists and the developers of the game - not just lab

staff.

Lab fever - Doing hours of work to get labs submitted before the deatline. (Expression coined by Lroppy)

Lab manual

2ND APR 2013 | 4 NOTES



creating poetry out of my Microbiology Lab Manual

What Do I Need to Know?

Definitions
And
Components

The difference between
Purpose and experiment

Understand the role of growth

For EteRNA lab manuals, see here:

Quick guide to the lab
Lab guide for new players
My strategy guide to the lab
Ways to deal with lab voting - in huge labs
Ways to deal with lab designing - in huge labs
Switch lab guide

For more lab stuff and guides, see my profile, if you haven't found it already.

Leg - synonym with arm. Can mean two stems with an kneecap of a bulge or internal loop inbetween. Consisting of more stems.

Definition put in on suggestion from Stlnegril9

Ligand - A molecule that binds to a protein with a high degree of specificity. Examples are the substrate of an enzyme and a hormone binding to a cell receptor. (Oxford dictionary of biology)

Loops - official definition

Low energy - the same as high negative energy.

Main design - everything in the design except the neck and hook area. (my definition)

Melting temperature - the temperature at which the two strands of a nucleic acid molecule separate (denature).

From genes to genomes

Melt plot - See my post on the topic.

MFE - Minimum Free Energy. This is identical to the structure and energy that Eterna shows in "natural" mode. The usefulness of MFE frequency is that often there are alternate structures possible that have free energies that are very competitive with the MFE (i.e. within a few kcal/mol). For example, the NUPACK algorithm tries to widen the gap between the MFE and the misfolds as much as possible. (alan.robot)

ribonucleic: mfe is the minimum energy a molecule needs to be stable in nature
bigbloto: my understanding is "every system seeks to achieve a minimum of free energy"
bigbloto: mine is rob from wikipedia :) so for given sequence, rna fold into the shape that provide mfe
ribonucleic: mine is from rnafold

Miseq sequencer

Sequencer: a device for determining the order of occurrence of amino acids in a protein or of bases in a nucleic acid.

Merrian-Webster.

An apparatus for determining the order of constituents in a biological polymer, usually DNA or protein. In this sense also called sequenator. (Free online dictionary)

For more about our new sequencer, see here.

Mismatch - When bases opposite to each other can not form a pair, it is called a mismatch. They result in unpaired bases, which form loops. The bigger the loop, the more energy it takes to hold together. Some mismatches are very efficient at splitting the strands from each other, encouraging loops to form. Loops starting with such mismatches take significantly less energy to hold. One such mismatch is the G-A (red - yellow) mismatch.

Drake

Our definition of boost for loops is a mismatch that yields lower energy than an A-A mismatch does.

Hoglahoo

Misshape - a single shape that appears in a misfold but not in target.

Hoglahoo

Microarrays - this is a word you will see often when you read science papers related to RNA research. I had no idea what it was in the beginning. But I recently found an article from Scientific american that does a nice job explaining a bit about it and how it can be used. Check in this full online number. The article "The magic of microarrays" starts on page 44.

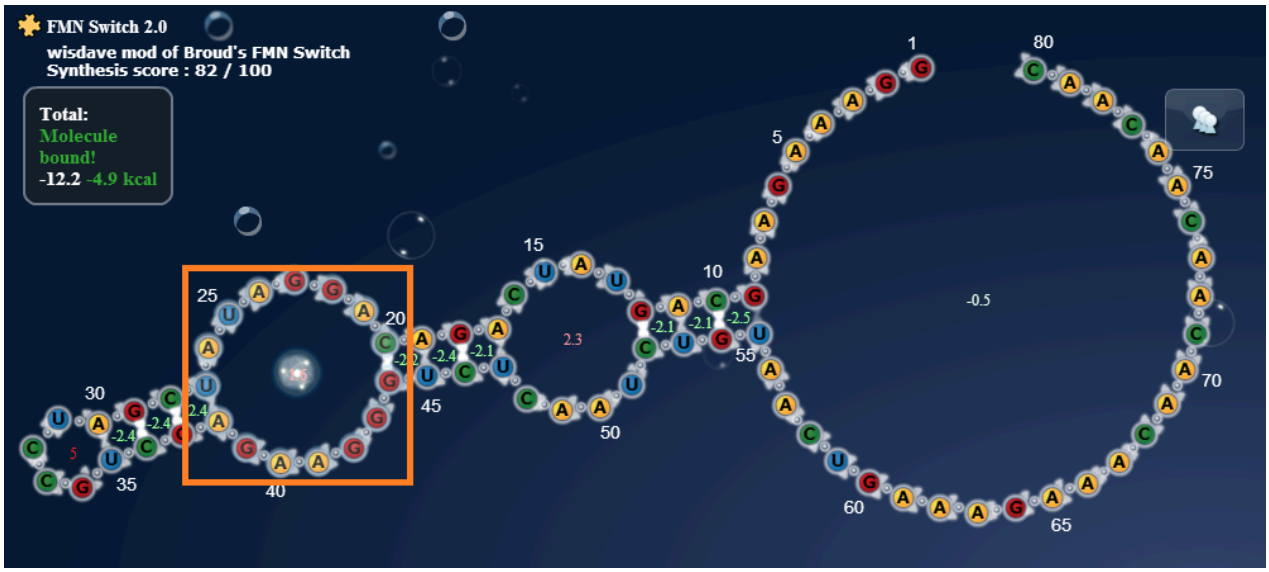
Microarray - a large set of DNA spots immobilized on a glass slide. Used for comparative and differential studies of genomes and transcriptomes.

From Genes to genomes

Micro RNA (miRNA) - small non-translated RNA molecules, involved in regulation of gene expression.

From genes to genomes

Molecule-bound shape - the RNA switch in the molecule bound target shape. (Definition Eli + Jee) The orange box mark the aptamer, the glowing thing inside is the molecule.



Also read Jnicol's fine explanation about [introducing the molecule] in a switch puzzle.

Molecule/no-molecule

In molecule-bound state, means RNA switch is in the second shape where it binds up to the FMN molecule. In other words, the switch is turned on.

In no-molecule state, the RNA is in the first shape and not bound up to anything. In other words, the switch is turned of.

Term put in on suggestion from Stlnegril9

Motif - scientist word for pattern/element.

Motif - a conserved sequence within a family of proteins indicating a specific function.

From genes to genomes

MS2 - I have collected some notes for the MS2 coming switch lab.

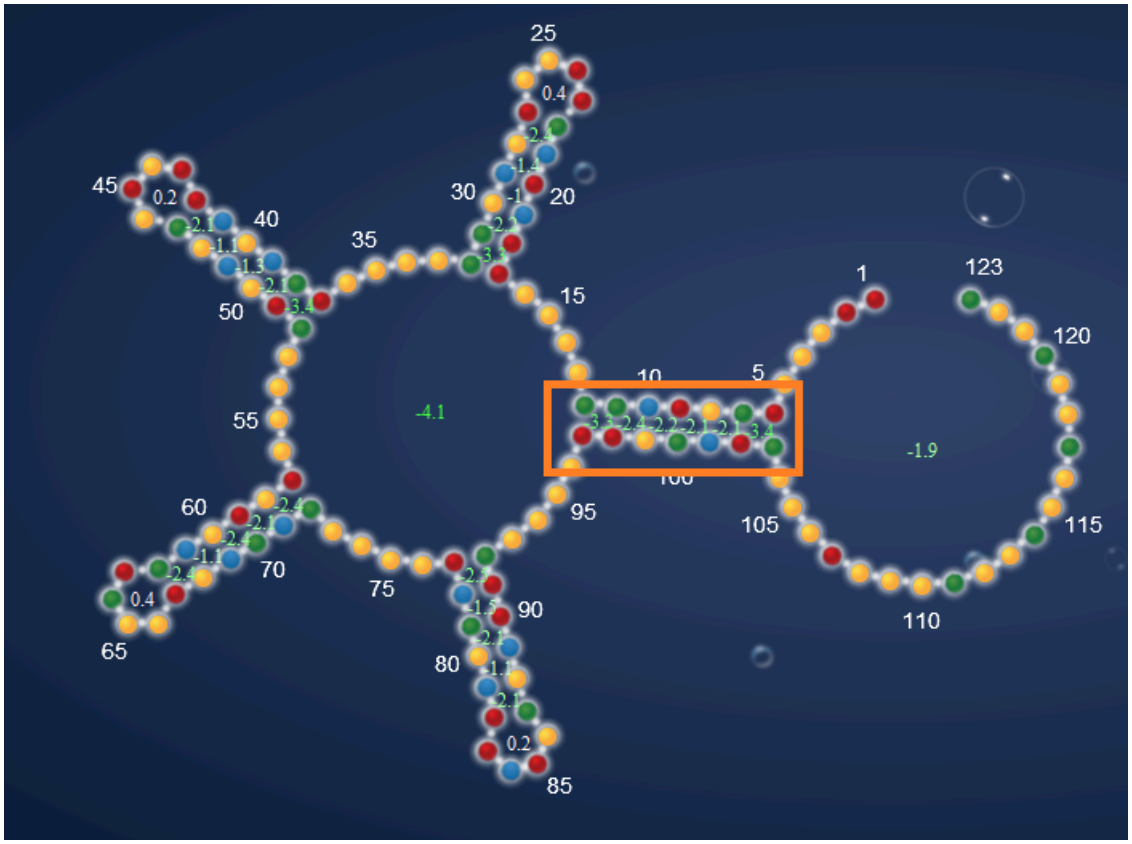
[Notes on the MS2 lab]

Multi branch loop - Scientist word for multi loop

Multi loop – a multiloop must have more than 2 stacks (put in by suggestion from Edward Lane)

Native mode - same as natural mode.

Neck (neck area) The neck area is where the beginning and the end of the RNA sequence ends in a string. Thanks to Ding for coining the term Neck. As Mat said: She thought 104 : The Star looked like a turtle.



My definition is: The neck area, is a stem, that behave quite different to the other stems. The neck area allows more repetitive patterns than in the rest of the stems. Also the neck is allowed to have very high total negative energy or very low total negative energy, compared with the energy distribution in the rest of the design. I found out that there tend to be a relative even energy distribution, in the high scoring designs. The neck area is the only area in the design that is allowed to break it. For more on this, see the Lab Guide Extension on my profile page, with

the headline Neck area.

Also see JandersonLee's fine explanation here.

Negative energy – what is needed for RNA to fold spontaneously, without adding of energy. If a puzzles overall energy is positive, it can't fold. The wiki has a brilliant explanation on free negative energy. Also go see this very fine explanation by Ksteppe on Understanding Free Energy using Legos.

Next pair = the base pair next after the closing base pair. Term coined by Mat747.

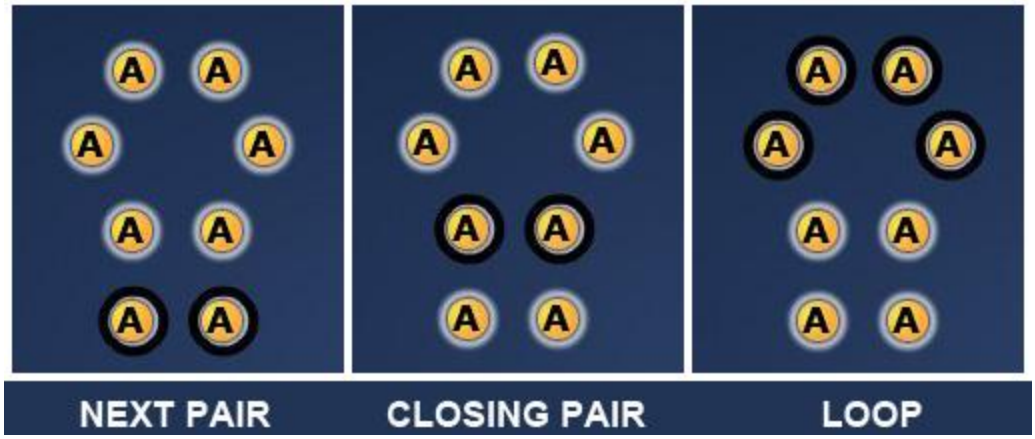
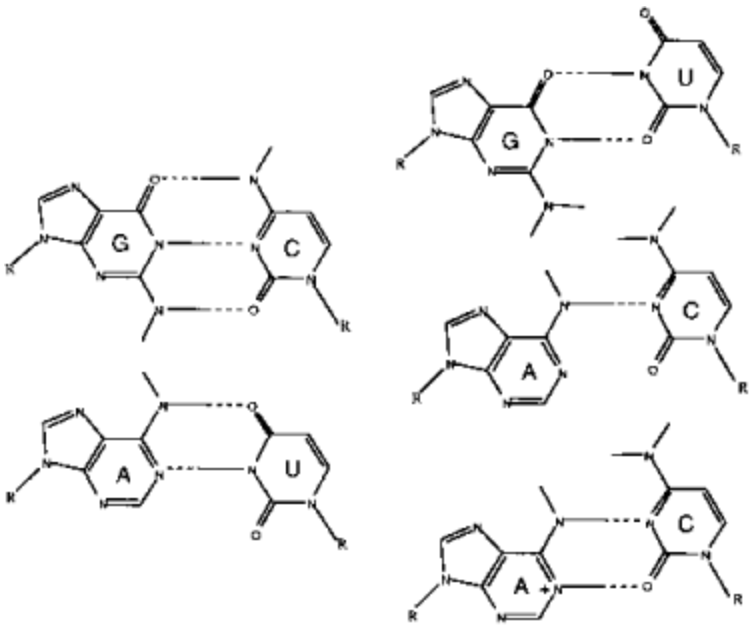


Image by Mat747

Non-canonical base pairs – Any base pair that is not a Watson-Crick base pair, including the pairing between bases that would ordinarily be mismatched. Below left are the two Watson-Crick pairs in RNA. Below right are three of the possible non-canonical pairs in RNA. Others are possible.



(Quasispecies' glossary)

Also see the UU-pair puzzle for an explanation.

Nt - abbreviation for nucleotide (Edward Lane)

Nucleotide = base (Edward Lane)

Nucleotide is a molecule. Bio lego building block for making RNA and DNA. Often abbreviated as nt in game conversation.

Nupack - Rna structure prediction tool.

Optical illusion - see Dimension9's post

Optical melt experiments - see Mseetin's fine explanation here.

Oligonucleotide - a short nucleic-acid chain usually consisting of up to approximately 20 nucleotides. Merriam-Webster

- a short piece of nucleic acid, usually with less than twenty nucleotides, used in genetic engineering and research. (Your dictionary)

Pattern - Eterna word for scientist word motif

Pentagon bulge - 5 nucleotide bulge, the same as a 1-nt bulge (Pentagon bulge is Iroppy's term)

Pentaloops

GNRNA pentaloops (R = A/G N = any nucleotide)

> This sequence adopts a structure similar to that of the GNRA tetraloop. The sequences GAAAA, GCAUA, and GAAGA have been structurally characterized.

GCUMA pentaloops (M = A/C)

> This sequence adopts a structure similar to members of the UUCG and AGNN tetraloop families.

Other pentaloops

> 5' – A-UGUAU-U – 3'
> This sequence appears in the *Tetrahymena* group I intron and shows a few interesting features – namely, a closing AU pair and an initial U-U mismatch.
>
> (Quasispecies' glossary)

PIP - Picture in picture. This is our new double screen view of the switch puzzles. Open this mode when in the switch puzzle by clicking this window symbol.

PCR - method in use for detection of specific nucleic acid sequences.

(Definition from "Protein detection using proximity-dependent DNA ligation assays, Rhiju's science paper collection)

PCR stands for polymerase chain reaction

In essence, PCR serves but one function, which is to amplify a relatively short segment of DNA many times, even if it initially forms only a minute part of a very complex mixture.

From Genes to genomes, page 109)

Here is a short video explaining the basic principles of PCR

Here is some further explanation

PDB - Protein data base file. See how to use it for Eterna purposes here.

Player statistics - several of our players keeps statistics of things. Check them out here.

Pressurized versus unpressurized designs - here is a page with something written about it.

Primary structure - the structure of RNa described with a line of nucleotide letters

Protein

“Proteins are the molecular machines that control just about every process in the body, from replicating DNA to fighting infection. They are, in a way, nature’s robots. Just as a robot must be precisely constructed to do its job so too a protein must adopt a precise structure to work correctly. The staggering thing is that a protein molecule has more possible structures available to it than there are atoms in the universe. Nevertheless, when a protein molecule is made it folds up into its correct structure in fractions of a second. This is sort of like throwing all the pieces of a car in the air and expecting them to form a fully functioning machine upon landing.”

Dr. Mark Lorch

Proteins are really the chemist of the cell.

Ben McFarland

Proteins do almost everything important inside of cells. They are enzymes and enzymes are biological catalysts, they make reactions happen, that would happen at such a slow rate without this catalyst that life would be impossible.

Biology 1A, Amino acids and protein structure, Berkeley, 2007

For more about proteins, see the videos on the basics in my Educational resources. Or read this fine article about the workings of proteins.

Protocol = scientist word for baking recipe - that have to be followed if you want to be sure to get the same results. If you are not following protocol, then you are experimenting.

Pseudoknots - crossing base pairs.

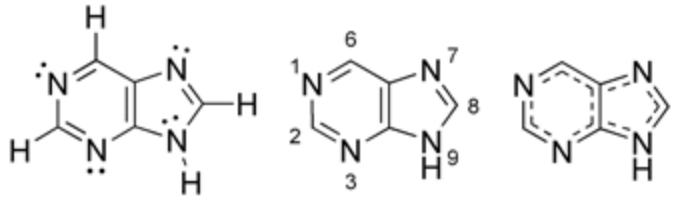
RNAMotifScan: automatic identification..., Rhiju’s science paper collection.

A pseudoknot in RNA is a bit like that, it looks like a knot, but if you could pull the ends, it would make a straight line. (For more from the explaining conversation see here.)

EINando888

There is a explanation about pseudoknots in [this fine paper] about different kinds of RNA structure, that EINando found and shared. (page 3)

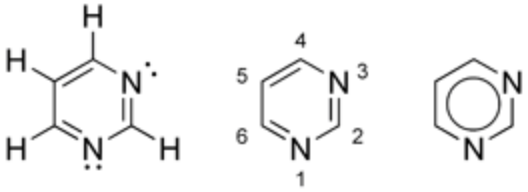
Purine bases – A and G (typically abbreviated “R”). They are based around the structure shown below. Always pair with pyrimidines in Watson-Crick pairs.



(Quasispecies’ glossary)

For more explanation on purines, see [Why do base pairs pairs up?]

Pyrimidine bases – C and U (typically abbreviated “Y”) They are based around the structure shown
below. Always pair with purines in Watson-Crick pairs.



(Quasispecies’ glossary)

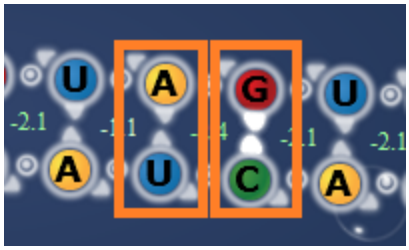
For more explanation on pyrimidines, see [Why do base pairs pairs up?]

Quad – four nucleotides in 2 base pairs

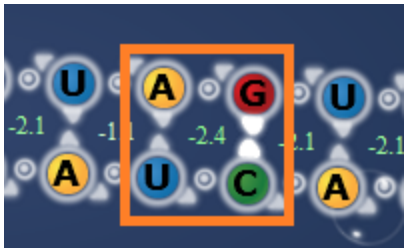
A quad is a stack of two base pairs. The energy in between them will depend on how the pairs are oriented.

Starryjess

Base pairs



Quad



Raw data - With the switch lab, we got a new sort of data. If you want to know how to read them, go see this post + this section on capillary electrophoresis. Also see Pablos fine explanation in his presentation of his project, in this first monthly live developers meeting.

Replication - the synthesis of a copy of a DNA molecule.

From genes to genomes

Retroviruses - Found a really interesting article We Are Viral From the Beginning about our potential viral origin.

Rfam - Database over natural occurring RNA's. Here is a link to Rfam and one to an introduction to it.

Ribosomes - the protein synthesis factories of the cells.

Gerhard Gottschalk, Discover the world of microbes

Riboswitch (RNA switch) - Naturally occurring RNA aptamers which can bind target molecules but can also switch between different conformations depending on whether they are bound to their ligand or not.

Molecular biology of RNA, Elliot and Ladomery

For more on Riboswitches, check here.

Ribozyme (catalytic RNA) Any RNA molecule that can catalyze changes to its own molecular structure.

(Oxford Dictionary of Biology, Sixth edition)

Enzymes are most usually proteins. Ribozymes are basically RNA's that has enzymatic activity. An enzyme acts as a catalyst. That means it speeds up a chemical reaction and makes it both faster and more energetically favorable.

"**RNase**"—a protein that deliberately and specifically cuts RNA molecules and thereby regulates their availability and activity. (Source)

Roadkill plot - scientist slang for really messy and bad dotplot. (Term used by Alan.Robot, sent by Mat)

RMBD - The RNA mapping database. Read more about it here.

RNA

RNA is a nucleic acid that in many ways is similar to DNA, but with key chemical differences that give it far greater functional diversity.

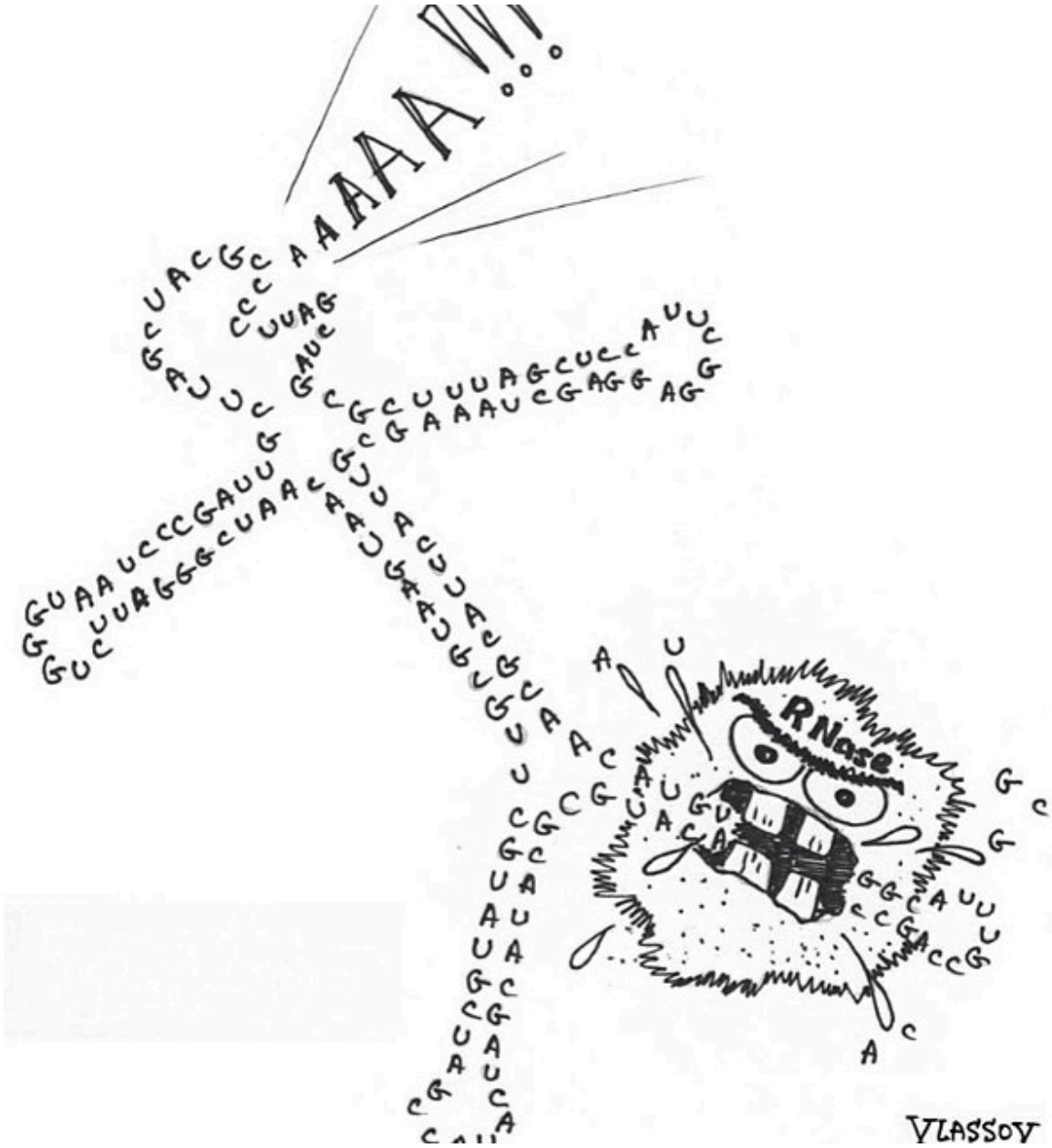
Molecular biology - Principles of Genome Function, Oxford, 2010, page 51

For more on RNA go see my collection of educational videos on RNA and DNA.

RNA (Ribonucleic acid) Equally as important as DNA, it is less stable but more versatile - to the extent that RNA molecules can even catalyze their own replication. The RNA World hypothesis suggest that RNA preceded DNA as the repository of genetic information.

Molecular biology of RNA, Elliot and Ladomery

RNase - Enzyme that degrades RNA.



RNA fold - The precise conformation, or fold, of an RNA molecule is determined by its nucleotide base sequence. Together, the fold and nucleotide composition determine the functional properties of the RNA.

Molecular biology - Principles of Genome Function, Oxford, 2010, page 54

“When studied in the laboratory, RNA bends into so many convolutions that it is nearly impossible to tease out which folds are worthy of scientific inquiry and which can safely be ignored.”

Unusual suspects

RNA polymerase - an enzyme that synthesizes RNA, generally using a DNA template

From genes to genomes

Enzyme which uses a DNA as template to make an RNA copy.

Molecular biology of RNA, Elliot and Ladomery

RNA switches (Riboswitches) - ...controls gene expression throughout the bacterial world.

Molecular biology - Principles of Genome Function, Oxford, 2010, page 59

This short article Synthetic RNA switches explains in simple words what RNA switches are and what they are capable of. Also check out the science daily article that is linked in there. It will give you more details.

Check out this podcast. It is a Yale lecture that gives a nice intro to RNA switches, what they are and what they can do. Riboswitches: Nature's ancient turn on.

This abstract from a science paper also gives a neat overview of RNA switches potential use.

RNA world hypothesis - (working on finding an article for that)

RNAi - RNA interference (**RNAi**) is a process within living cells that moderates the activity of their genes. Two types of small ribonucleic acid (RNA) molecules – microRNA (miRNA) and small interfering RNA (siRNA) – are central to RNA interference.

From Wikipedia

Nova have made a brilliant video explaining RNAi and why it is so promising for medical purposes.

Also found an article from Scientific American DNA Computer works in human cells explaining RNAi.

RNAi (RNA interference) is a potent and highly specific genesilincing phenomemon that is initiated or triggered by 19-23 nt dsRNA's known as small interfering RNA. siRNA's can downregulate the genes responsible for serious diseases and are therefore very promising for target-based drug discovery and development.

From Rhiju's science paper list, Branched RNA nanostructures for RNA interference, Nakashima

Secondary structure - structure in 2D

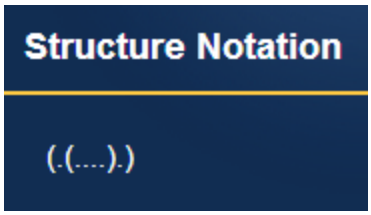
SHAPE = selective hydroxyl acetylation by primer extension. For an intro about it, see this abstract that Drake shared in lab. Thanks to Macclark52 for sharing the shape definition in chat.

Shape notation - the same as dot bracket structure - see example here and the explanation in the eterna Wiki

(.(....).)

Can be dumped into the puzzlemaker to recreate the puzzle and change in it.

Here as seen in game:



Shape data - Shape data is the chemical mapping of how the RNA bases are paired up after the RNA is synthesised. Yellow bases are the unpaired bases and blue base pairs are those that paired up. (Click the yellow/blue picture of the design, when you are viewing the lab results) Shape data is a help when reading the results we get back from lab. For how to read the advanced shape data with continuous colors, see this answer by Jee. Here is an update on the SHAPE data for switch lab. I have written a new and better intro to SHAPE data and on reading the lab results.

SHAPE Probe - You will often hear the words SHAPE probe mentioned.

A probe is a single strand of bases. It will bind up to a complementary sequence. In our case the SHAPE probe will bind to the aptamer.

It can be made to bind to a specific sequence only or to bind more broadly.

Shape reagent

stlnegril9: SHAPE reagent is the 'probing" mechanism, in my simplified understanding

stlnegril9: of whether an NT paired or did not... something like that
Nando: reagent is just a small molecule, barely bigger than a base
Nando: SHAPE is the technique, different chemical agents (called reagents) can be used
(1 July 2013)

SHAPE threshold

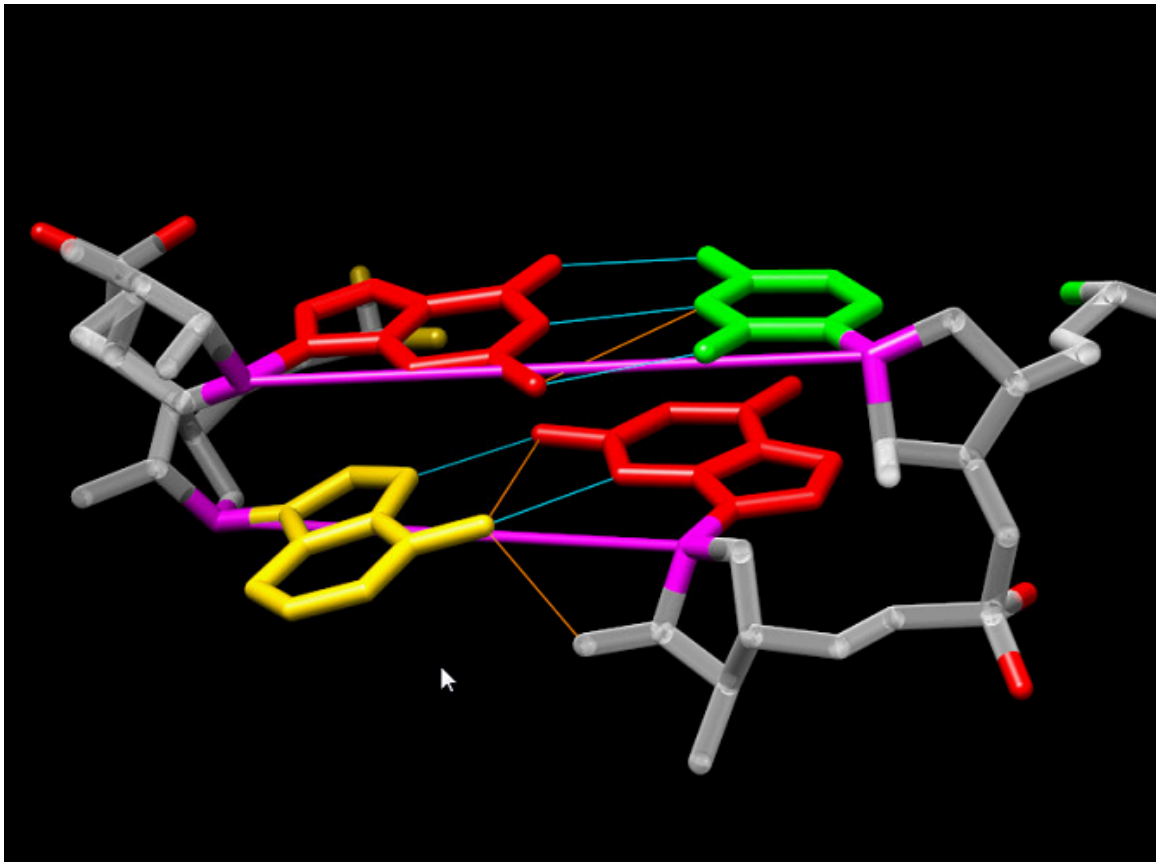
Read more about it here.

Also search for threshold in the forum. This will bring you up some posts on it.

Sheared GA

Sheared GA is a base pairing between a Guanine and an Adenine.

Here is a picture that Nando showed me of the difference between how a GC pair lines up, compared to how a sheared GA shows up. He explains that by definition, scientists only count the canonical pairs GC and AU and the wobble pair GU, for stem. The rest counts as loop, even though it actually looks like stem. Both in SHAPE data and chimera.



Notice how the GC-pair bases line up along the magenta line, where the sheared GA bases doesn't.

Nando: There are a few other "famous" non-canonical pairs, the "Reversed Hoogsteen AU" and the "GA imino" for instance.

You can see more about them here at Nando's blog

Shiftable puzzles:

They can make the puzzle easier or harder to solve if you choose to use them to change the shape

Hoglahoo

Signal-to-noise-ratio

A signal-to-noise-ratio report comes along with our lab data in the Stanford RNA mapping database. Here is an explanation by Omei of what we can use it for:

Basically, any series of measurements can be thought of as being composed of two parts -- the "signal" (which is the value we're actually trying to measure) and the "noise" (which is made of all the things that we're not really interested in, but which affect the data anyway).

We generally can't figure out exactly what is signal and what is noise. but if we make some assumptions about what is producing the noise, we can make a reasonable guess as to what percentage of the overall information in the data is due to signal and what percentage is due to noise. That is what the S/N ratio is. A S/N ration of 1.0 means that we estimate that the noise is having just as much affect on the data as the signal is. Clearly, that makes it pretty hard to interpret that data. Ideally, all the S/N ratios would be very large, so large we could say that the data represented "truth", but that is often not the case. So the S/N ratio for a synthesized design just gives us a rough idea of how much weight we can put on it when using it in support of a particular theory.

Single shape lab - static lab (not moving like switch lab)

Sequencing - is the primary way of characterizing a macromolecule, whether it be determining the order of amino acids in a protein, or of bases in a nucleic acid.

From genes to genomes

siRNA - Small interfering RNA.

Also known as RNAi.

Smurfing - Puzzle solving technique which involves painting most or all single bases in a puzzle blue to prevent them from mispairing with something else. Can be seen in some of Wawan's 2-2 and 1-2 puzzles and Tesla's puzzlets. For smurfing tutorials, see Teaching and tutorial puzzles.

Solvent effects - A part of thermodynamic stability comes from stabilizing through interaction with water and other soluble stuff in your test tube/cell.

Definition by Bravehp

Stack – same as stem - base pairs lined up beside each other

Non-GC closing pairs

> The closing pair of a stack exerts a significant influence on loop energy, and GC pairs are often preferable for this role. Most successful player designs close stacks with GC pairs. That said, would be useful to know the circumstances in which other pairs are equivalent or superior.

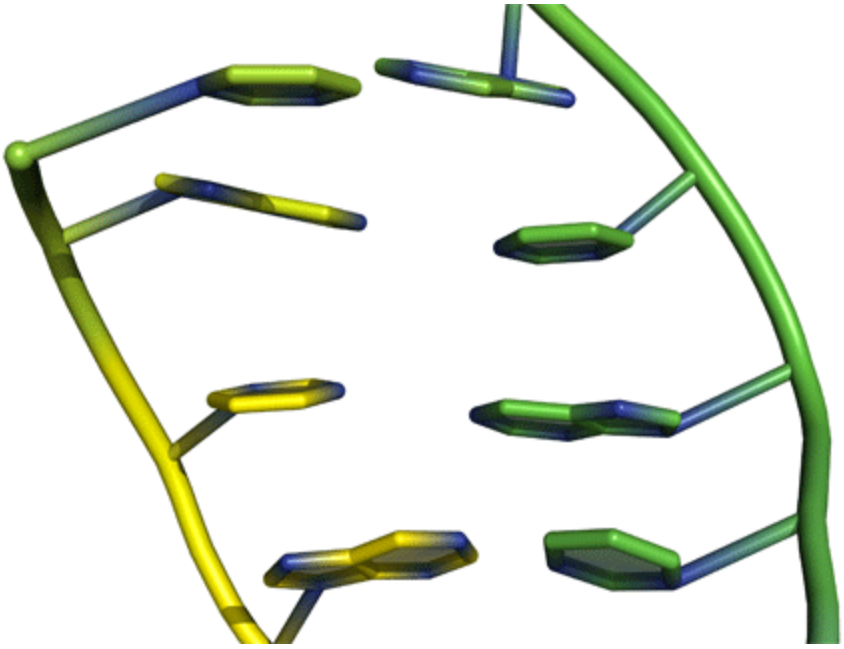
GU pairs

> The geometry of a GU pair in an RNA helix differs from that of Watson-Crick pairs, weakening stacking interactions 5' of guanine. Stacks closed by GU pairs should ideally orient the pair so that the first base of the loop occurs on the 5' side of guanine. Few designs have placed GU pairs within a stack, and none have been synthesized. Internal GU pairs do occur in nature, however.

Lonely AU pairs

> The strength of the GC pair makes it a natural choice for isolated base pairs; however, this is not always the case in nature. In an unusual RNA motif called the "Lonepair Triloop" a 3-nt loop is closed by an isolated base pair, usually an AU or non-canonical pair. Though this structure is frequently stabilized by a variety of outside influences (proteins, coaxial stacking with other helices, etc), it raises the question of whether isolated AU pairs can form under other conditions as well.
>
> (Quasispecies' glossery)

Stacking – The bases of RNA are flat, and their tops and bottoms are relatively "greasy". Like droplets of oil in water, the greasy parts clump together. This causes the bases to stack on top of one another.



(Quasispecies' glossery)

Static lab (not moving like switch lab) - single shape lab

Stem – same as leg or stack. The scientists calls them stem most of the time. Basically it is a stack of base pairs. Anything with 2 base pairs or more is a stem

Strand – nucleotides in a row – like strawberries on a straw. It takes two strands to make a stem.

A stem is two lines of nucleotides or a line of base pairs on a row. Stem - from one junction to another. A stem stops as soon as a bulge or loop begins.

Structure notation - same as shape notation and dot bracket.

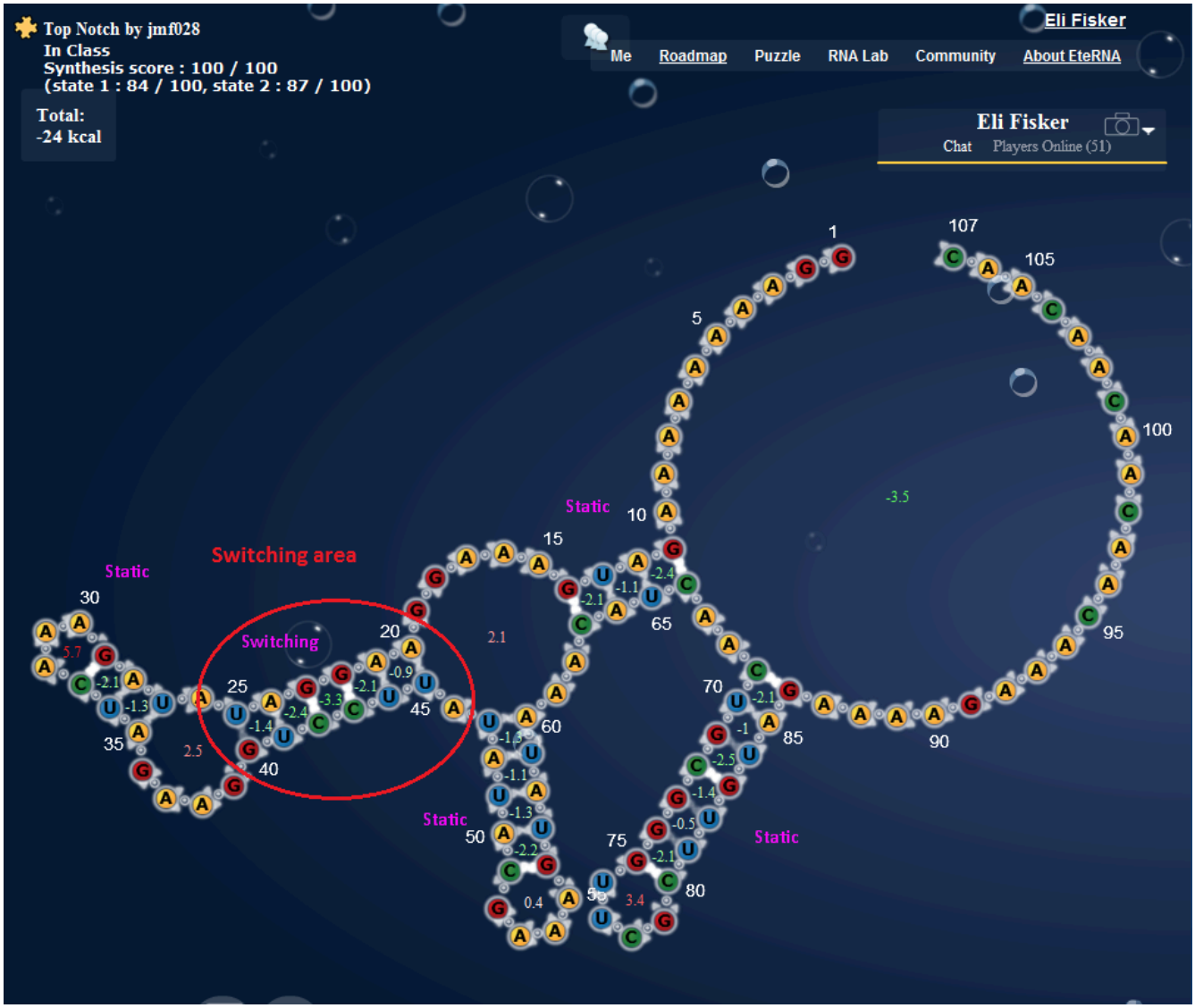
This means periods represent loops, and the parentheses base pairs.

Starry gave this example:

Looks like this: ((....)) is two pairs and a hairpin loop - a loop 4 nucleotides big

Substrate - What every self respecting enzyme must have. (Berkeley, Biology 1A, spring 2012, lecture 30)

Switching area - The part of the switch the two states are moving in relation to each other. However in some switch labs the whole of the two states move compare to each other. So generally what I think about as the switch area, is the area around the aptamer which is a switch hotspot.



Switch Puzzle

A switch puzzle is a puzzle which requires you to design a sequence to stabilize two or three different states at the same time, with one or more states getting a little bonus negative energy to help it fold.

(Hoglahoo)

Symmetric and asymmetric tandem mismatches – Symmetric tandem mismatches include any sequence that fits the pattern:

5’-XY-3’

3’-YX-5’

The lowest-energy 2-2 loops contain GU and GA mismatches of this type. The larger set of asymmetric mismatches remains largely unexplored in player lab designs.

(Quasispecies' glossery)

Synthesis - putting together, from greek, source: Foundations in microbiology

Tertiary structure - structure in 3D. Check out this paper that Rhiju mentions about 3D structure.

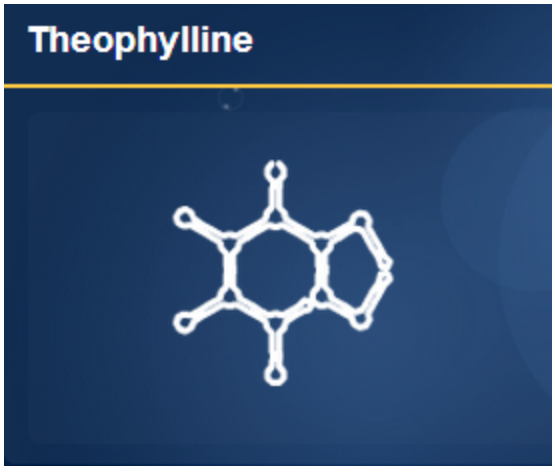
Theophylline (aptamer)

One of the aptamers we are playing with in lab.

- ...is used as a drug to treat asthma and other lung diseases) can form a part of a fifty-five-nucleotide-long stretch of RNA that can have two different morphologies and two different functional states depending on the concentration of theophylline.

George Church, Regenesis page 32

Theophylline molecule



You can read a little more about theophylline here.

Theophylline puzzle

Researcher explanation:
So this puzzle is a riboswitch that we're using to control genes in yeast. It has two state, one which binds a target molecule and the other it takes a form which allows it to fold around and cleave itself, when it cleaves it turns a gene off. so you can turn a gene on and off by adding/removing the target molecule. we've been building these, so far, to just turn a fluorescent

protein on -- makes the yeast colonies glow green when the target is present.

Brent

You can read a bit more about it here in [the developer chat] we had the 15 november 2012.

Thymine

DNA base - corresponds with the base U in RNA

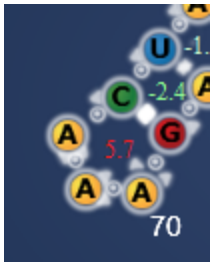
Transcription (DNA transcription) - Copying of one strand of DNA into a complementary (matching) RNA sequence by the enzyme RNA polymerase.

Molecular biology of the Cell

Tri-, tetra-, penta-, hexa-, etc… loops – Hairpin loops that contain the number of unpaired bases indicated by the prefix.

(Quasispecies' glossery)

Tri Loop – end loop which besides it's closing base pair has three nucleotides. (the three yellow A's)



For tricks see under Triloop tricks under the headline Loops in Tips and tricks guide

Triloops

Non-GC closing pairs.

Currently, the only way to accomplish this in EteRNA is to place *two* consecutive AU pairs in the stack preceeding the triloop. EteRNA predicts that a sequence such as 5' – GA-AAA-UC – 3' will form a pentaloop instead of the desired triloop. This is a quirk of the parameters used. In nature, triloops like this can and do form.

5' – A-UCU-U – 3'

An x-ray structure of the *Tetrahymena* group I intron (PDB: 1GID) shows a UCU triloop with a closing AU pair. Unfortunately, the entire hairpin cannot be recreated exactly in EteRNA because a GC pair precedes the closing pair.

5' –U-UUU-A – 3'

Free energy of this loop was measured as 4.86 kcal – nearly 2 kcal less than the value in EteRNA.

Non-AAA triloops

Not all triloops are equivalent. The nature of the closing pair and unpaired bases can cause their their energy and the orientation to vary considerably. All-C loops are less-stable than most other loop sequences. All-U-loops tend to be more stable. The sequences below have all been measured to have loop energies < 5 kcal, and might be worth trying. See (1) for more loops to try.
5' – C/G-CUU-G/C – 3'
5' – C/G-UUU-G/C – 3'
5' – C/G-GAA-G/C – 3'

(Quasispecies' glossery)

Tetraloops – outer loop which besides it's closing base pair has four nucleotides (the four yellow A's)

Tetraloops

CUUG and UNCG loops (N = any nucleotide)

> These classes of tetraloops are very common in nature. EteRNA doesn't grant this sequence the same energy bonus as a GNRA tetraloop, likely discouraging its use in lab.

GU closing pairs

> A GU pair preceding a GNRA tetraloop can be found in many natural RNAs.

Other GNRA tetraloops (R = A/G N = any nucleotide)

> The energy bonus may not be as great for some loop sequences, but other members of the GNRA family might merit use.

(Quasispecies' glossery)

Total energy
The energy total is the total of all the energies of the various shapes in the puzzle (the loops, pair quads, etc). Stronger bonds result in lower energies [4:00 PM]

Hoglahoo

Transfer RNA (tRNA) is the information adapter molecule. It is the direct interface between the amino-acid sequence of a protein and the information in DNA. Therefore, it decodes the information in DNA. The different tRNA molecules have between 75 and 95 nucleotides. Transfer RNA from all organisms have a similar structure; indeed, a human tRNA can function in yeast cells.

Rolf M. Flügel, Chirality and life, chapter 4, 2011

tRNA are also called amino acid adapters, as they pick up amino acids for translating messenger RNA code into protein code. I like that term.[2]

UI - often heard in chat. Means User Interface. Basically it this is what see on your computer screen. And it is a term typically used by developers, game designers, programmers and gamers.

[2] RNA: Structures, roles and transcription - Dr. Highsmith, July 22, 2011. San francisco Dental student store 2011's official podcast.

Um - means micromolar. Wiktionary describes it as “a concentration of one millionth of a mole per litre.

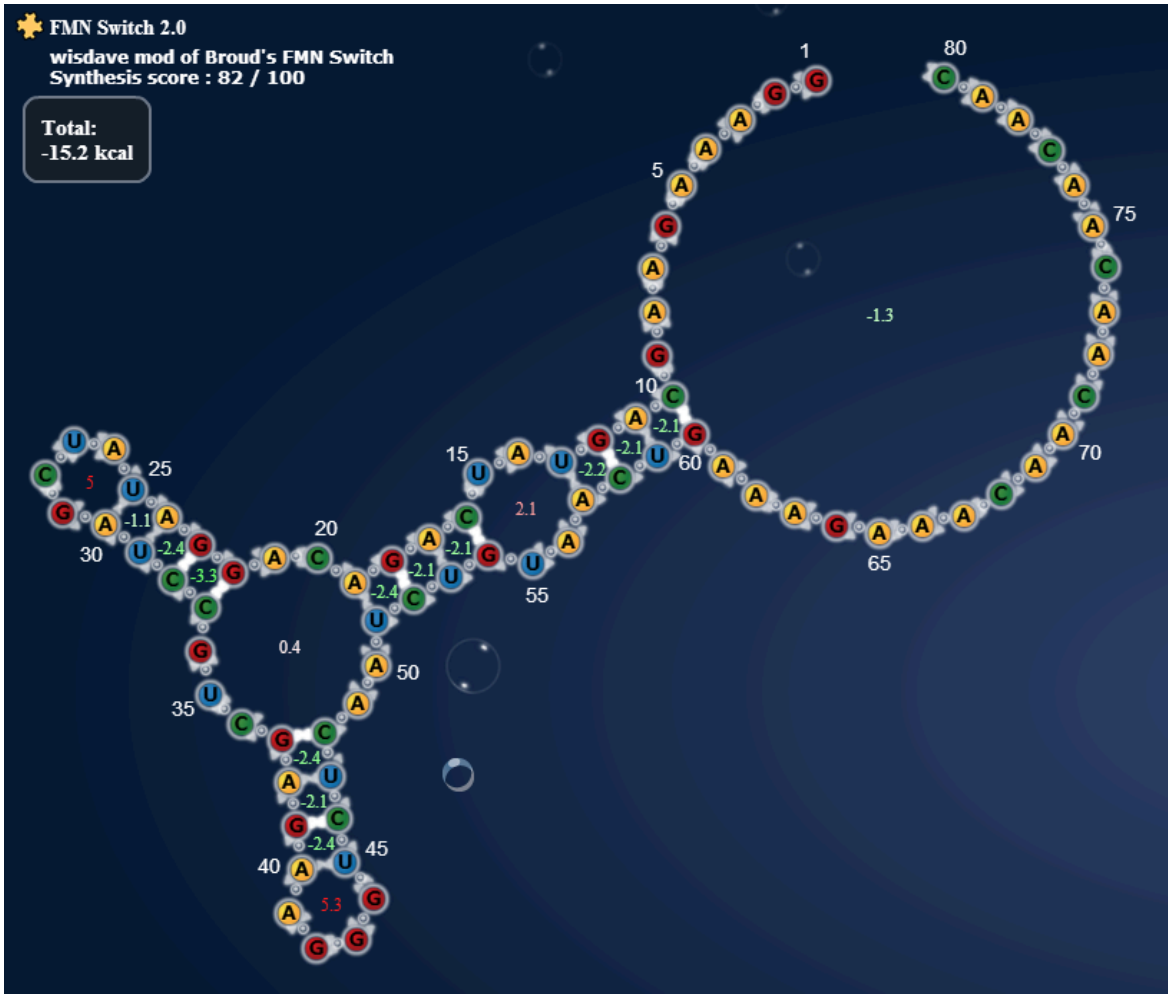
I asked Quasispecies what UM meant in relation to Aptamers. He says: The RNA molecule is shifting between multiple structures, but only one (or a few) can bind FMN. For binding to occur, a molecule of FMN needs to encounter an RNA molecule while the RNA exists in the right structure. At the given concentration (the midpoint), 50% of the RNAs are binding FMN and 50% are free. The more time the RNA spends in the right structure, the less FMN you need to get this level of binding.

Thanks to Mat for bringing up that we needed an explanation.

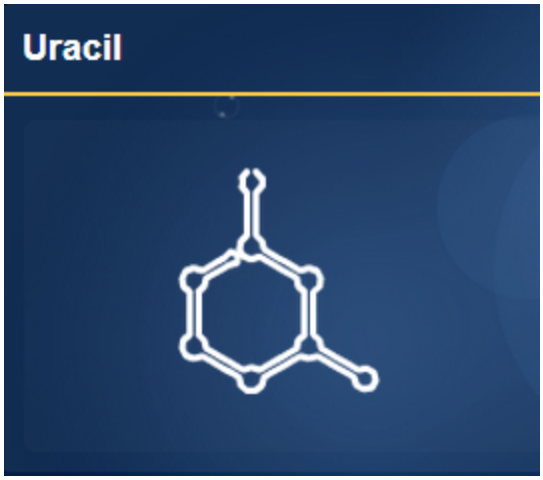
Unsolvable series and about Prevotella in particular - There are also some challenge puzzles that are unsolved in the game. The eterna game says a puzzle is only solved when the energy model it uses says so. Some of the challenge puzzles are based on actual RNA that the energy model says can't be stable.

Hoglahoo

Unbound shape - the unbound hairpin shape in the RNA switch (Definition Eli) (Same as unbound target) The one without the bound molecule. See example below:


FMN Switch 2.0
wisdave mod of Broud's FMN Switch
Synthesis score : 82 / 100
Total:
-15.2 kcal

Uracil


Uracil

RNA base - corresponds with the base Thymine in DNA

Vienna - RNA secondary structure prediction algorithm. For how to use it, check here.

Viruses - Viruses look and often act like little golf balls, just lying around or flying through the air. They aren't able to do much by themselves. But as soon as they have entered a host cell, they start their devilish work.

Gerhard Gottschalk, Discover the world of microbes

Watson-Crick pairs – GC and AU pairs.

(Quasispecies' glossary)

Wobble pair – In the context of EteRNA, GU pairs

(Quasispecies' glossary)

XNA - synthetic DNA. Get an introduction here.

Zigzag - the amount of time it takes to solve a zigzag puzzle increases exponentially with the addition of each zig. Or zag. (Definition Merryskies)

Series of bulges, usually on different sides, are called zigzags. (Definition Drake, from Advanced puzzle solving guide)

Other things worth to know about zigzag. Placing blocking points in zigzags prevents the structure from sliding and just become a long stem instead of a zigzag. Can we make them in lab? Only time will tell.

For different types, see under zigzag in Tips and tricks guide.

Ångstrom - Unit of length used to measure atoms and molecules. Equal to.... 0,1 nanometer. (nm)

Molecular biology of the Cell