RNA world

In prebiotic world, RNA thought to have filled two distinct roles:
1. an information carrying role because of RNA's ability (in principle) to self-replicate,
2. a catalytic role, because of RNA's ability to form complicated 3D shapes.

Over time, DNA replaced RNA in its first role, while proteins replaced RNA in its second role.
RNA classification

Messenger RNA:
- Carry genetic information,
- Structure less important.

Non-coding RNA:
- Functional,
- Structure is important.
Cellular functions of RNA

Genetic Functions:
- Messenger RNA
- Viroids
- Transfer RNA

Enzymatic functions:
- Splicing (snRNA)
- RNA Maturation (ribonuclease P)
- Ribosomic RNA
- Guide RNA (snoRNA)
RNA structure and function

- RNAs have a 3D structure,
- This 3D structure allows complex functions,
- The variety of RNA structures allows the specific recognition of a wide range of ligands,
- Some molecules target these RNA structures (antibiotics, antimitotics, antiviruses):

Linezolid  
Doxycyclin  
Sparsomycin  
Chloramphenicol
RNA vs DNA: Chemical nature

- 2′-OH group attached to sugar (instead of 2′-H): *more polar*
- Substitution of thymine by uracile = suppression of group 5-CH3

Small modifications => big effects
RNA vs DNA: Modification of the local and global geometry

- **Local conformation**
  - RNA favorite: C2’endo
  - DNA favorite: C3’endo

- **Global conformation**
  - RNA favorite:
    - ![RNA conformation image]
  - DNA favorite:
    - ![DNA conformation image]
RNA vs DNA: Consequence of the modification of the geometry

Small furrow is flat

Big furrow is deep
RNA vs DNA: RNA-Protein and DNA-Protein interactions are different

**DNA-Protein:** Secondary structure elements insert in big furrow

**RNA-Protein** interaction are more specific. Usually using less structured regions.
RNA vs DNA: Last (?) differences

- RNA is a short linear molecule
  DNA long ≠ RNA short

- RNA are usually single stranded
  ADN double stranded ≠ ARN single stranded

- « turnover » relatively fast
  ADN stable ≠ ARN versatile
Base pairing in RNAs

- As in DNA, bases can interact through hydrogen bonds.

- Beside the two canonical base-pairs, RNA structure allows “Wooble” base-pairs.

- A-U and G-C are “isosterus” while G-U induce a distortion of the backbone.
RNA secondary structure

The secondary structure is the ensemble of base-pairs of the structure.
RNA secondary structure

**Central assumption:** RNA secondary structure forms before the tertiary structure.

Secondary structure prediction is an important step toward 3D structure prediction.
RNA secondary structure

The secondary structure can be very complex. Usually most of it can be drawn on a plane.

Few “irregularities” remain.

Non-canonical base-pairs

Pseudo-knot (crossing interaction)

Base triplets (Not on the picture)
RNA secondary structure representations

Circular

Brackets

..(((((((...((…)))))...(((….))))).))))

Classical

Dot plot
RNA secondary structure prediction using comparative methods

The secondary structure can be predicted from the alignment of homologous sequences. Base-pairs are identified through compensatory mutations.

97% of the base pairs predicted by comparative analysis in rRNAs have been confirmed later in the crystal structure.
RNA secondary structure elements

- Hairpin loop
- Bulge loop
- Interior loop
- Multibranched loop
- Stem region
- Unconstrained single-stranded region
Loop decomposition

Base pairs?

Stacking pairs!!

Hairpin loop

Bulge loop

Interior loop

Multibranched loop

Unconstrained single-stranded region
RNA secondary structure description

A secondary structure can be decomposed into a sequence of loops:

- **Hairpin**
- **Stack**
- **Bulge**
- **Internal loop**
- **Multi-loop**

● ● : Sequence neighbors  ● ● ● : Spatial neighbors
Stacking base pairs

Base stacking interactions between the pi orbitals of the bases' aromatic rings contribute to stability. GC stacking interactions with adjacent bases tend to be more favorable.

Note: Stacking energy are orientated.

\[
\begin{align*}
5' & - \text{CG} - 3' \\
3' & - \text{GC} - 5' \\
5' & - \text{GC} - 3' \\
3' & - \text{CG} - 5'
\end{align*}
\]
Suite of tools to predict and analyze the secondary structure of RNA. Here, we will overview:

• RNAfold: The main program that is used to predict the secondary structure from sequence data alone.
• RNAdistance: A program to compare secondary structures.
**RNAfold**

Command line: RNAfold -p -d2 --noLP < test_sequenc.fa

Text output:
```
GGGCUAUUAGCUAGCUGCUGUAGGUGAGGCGCACCUCUGAUAAGGGUGAGGUGACUGAUUCGAAUUGGACGCAUAGCCCCA
```

Graphic Output:
Base pair distance:
$ $ RNAdistance -DP
Input structure; @ to quit
.....,.....1.....,.....2
(((((((.....)))),))))
..((((((.....)))),)))
P: 6

Number of base pairs you need to remove and add to go from one structure to the other.

WARNING: Use only with structures with same length.

RNA distance implements many other metrics.