The RNA World

E. Westhof

http://www-ibmc.u-strasbg.fr/upr9002/westhof/
Fossils of stromatolites

Decrease in the Number of Meteorite impacts

Organic matter From living organisms

Stabilization of the hydrosphere

RNA WOLRD

Earth Formation

4.5 4.2 4.0 3.8 3.6 3.4
(x 10^9 ans)

3.8
3.6
3.4
3.2
3.0
2.8
2.6
2.4
2.2
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
RNA WOLRD
Global overview of all life

RNA World
The RNA World

• Origin of life / central dogma paradox
  • DNA needs proteins to replicate
  • Proteins coded for by DNA
• RNA can be code and machinery
  • Selex, aptamers
• RNAs are remnants
  • Ancient and Essential
DNA = modified RNA

- Ribonucleoside triphosphates → RNA
- Ribonucleoside diphosphate réductase
- NADP⁺ + Thioredoxine (-SH)₂
- NADPH + H⁺ + Thioredoxine (S-S-)

- Deoxyribonucleoside triphosphates → DNA

RNA = modified DNA
Nucleic acids are negatively charged biopolymers ...
Conformations of RNA

• Primary structure of RNA similar to DNA

RNA, like DNA, can be single or double stranded, linear or circular.

• Unlike DNA, RNA can exhibit different foldings
• Different folds permit the RNAs to carry out specific functions in the cell
Tautomeric forms

Charge delocalization
Protonation possibilities
Always seen

Never seen
Modified bases have different electronic properties
H-bond characteristics

\[ d_{-}N \quad 1,8 - 2 \, \text{Å} \]

\[ d_{-}O \quad 2,8 - 3 \, \text{Å} \]
Horizontal Interactions
Base pairing.
In helices
Complementary Watson-Crick
Vertical interactions: stacking
Stacking forces

- Driving Force: hydrophobic effect.
- Not very specific

Partition in

very polar regions (phosphates) &
less polar ones (exocyclic groups of bases)
History of the many roles of RNA
Central dogma

The flow of genetic information

transcription  translation

DNA → RNA → Protein
DNA notoriety

Rosalind Franklin: Dark Lady of DNA
Book Sheds New Light on a Scientific Landmark

Oct. 6, 2002 -- Early next year, scientific institutions in the United States and Great Britain will mark the 50th anniversary of one of the greatest discoveries in science. In April of 1953, James Watson, Francis Crick and Maurice Wilkins identified the substance of life -- the structure of DNA.

They later shared a Nobel Prize. Their discovery depended heavily on the work of a woman, chemist Rosalind Franklin, whose research was used without her knowledge or permission. Watson’s memoir of the discovery dismisses Franklin as frumpy, hostile and unimaginative. A later work by a friend casts Franklin as a feminist icon, cheated of recognition.
Short history of RNA

- Late 1800’s - 2nd kind of nucleic acid not in the nucleus (rRNA)
- 1920’s - sugar for DNA vs RNA
- 1958 - tRNA (Hoagland)
- 1960’s - mRNA

Relative Amount of RNA in E. coli
- rRNA 80%
- tRNA 15%
- mRNA 5%
“Three” different types of RNA

- mRNA - messenger RNA, specifies order of amino acids during protein synthesis
- tRNA - transfer RNA, during translation mRNA information is interpreted by tRNA
- rRNA – ribosomal RNA, combined with proteins aids tRNA in translation
Ribosomal RNA (rRNA)
Phylogeny of Life using SSU rRNA
Discovery of catalytic RNA

- Early 1980’s:
- Tom Cech & Sidney Altman
- Self-splicing group I introns and
- Ribonuclease P
Group I self-splicing introns
Core structure
Catalytic Core of Group I introns

Transesterification
Second discovery of RNA

- **ncRNAs** - functional RNA molecules rather than proteins; RNA other than mRNA (ex. XIST)
- **RNAi** - RNA interference
- **siRNA** - active molecules in RNA interference; degrades mRNA (act where they originate)
- **miRNAs** - tiny 21–24-nucleotide RNAs; probably acting as translational regulators of protein-coding mRNAs (regulate elsewhere)
Other Roles of RNA

• **stRNA** - Small temporal RNA; (ex. lin-4 and let-7 in *Caenorhabditis elegans*: development

• **snRNA** - Small nuclear RNA; includes spliceosomal RNAs

• **snoRNA** - Small nucleolar RNA; most known snoRNAs are involved in rRNA modification

• **RNA world** - RNA as catalyst
Central dogma

The flow of genetic information

transcription  translation

DNA  RNA  Protein
ncRNA genes

- Genomic dark matter
  - Ignored by gene prediction methods
  - Not in EnsEMBL
  - Computational complexity
- ~10% of human gene count?
RNA interference

<table>
<thead>
<tr>
<th>RNA Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short interfering RNAs (siRNAs)</td>
<td>A class of double-stranded RNAs of 21–22 nucleotides in length, generated from dsRNAs. siRNAs silence genes by promoting the cleavage of mRNAs with exactly complementary sequences, or recruiting inhibitory proteins to, or directing the modification of, DNAs with exactly complementary sequences.</td>
</tr>
<tr>
<td>MicroRNAs (miRNAs)</td>
<td>A class of 19–25-nucleotide, single-stranded RNAs that are encoded in the genomes of most multicellular organisms studied. Some are evolutionarily conserved and are developmentally regulated. They silence certain cellular genes at the stage of protein synthesis.</td>
</tr>
<tr>
<td>Tiny non-coding RNAs (tncRNAs)</td>
<td>A newly discovered class of short, 20–22-nucleotide RNAs that are encoded in the genome of <em>C. elegans</em>. They are not evolutionarily conserved, but some are developmentally regulated. Their function is still unknown.</td>
</tr>
<tr>
<td>Small modulatory RNA (smRNA)</td>
<td>A short, dsRNA, identified earlier this year in mice, that allows the expression of neuron-specific genes only in adult neurons.</td>
</tr>
</tbody>
</table>
Intricate maturation pathways

[Diagram showing the maturation process of microRNAs (miRNAs) from cellular gene to mature miRNA]

- Cellular gene
- Pri-miRNA
- Pre-miRNA
- Mature miRNA
- Drosha
- Dicer
- miRNP
- mRNA degradation
- Translational inhibition

- Exactly complementary
- Partially complementary
Implicated in chromatin remodelling
Cleavage part of DICER
Measuring device for 21 nt
ONE Role of RNA in cells

1. Transcription
   - DNA
   - RNA polymerase
   - mRNA
   - tRNA
   - Polypeptide chain

2. Translation
   - Ribosome
   - Protein
   - Amino acids
   - Anticodon
   - Nucleotides

DNA → RNA → Proteins
Versatility of RNA functions

- Not only messenger RNAs (alternative splicing)
- Number & variety of non-coding RNAs: ribosomal RNAs, snRNAs, snoRNAs, and numerous regulator RNAs.
- Cofactor RNAs: telomerase,
- ...... to be discovered
Properties of RNA molecules

Assemble in double-stranded helices like DNA

Carry GENETIC INFORMATION like DNA

Fold in complex tertiary architectures like proteins

Perform CHEMICAL CATALYSIS like proteins
Catalytic RNAs

- Ribonuclease P
- Self-splicing introns
- Hepatitis delta virus
- The ribosome and peptide bond formation
- The spliceosome (not fully proven yet)
- and ...
Great diversity of RNA functions and architectures
Different levels of organization possibly hierarchically structured.
The bacterial ribosome:

- 270,000 atoms (C, N, O, P)
- 55 proteins
- 3 RNA (4600 nucleotides)
Ribosome active site

H. F. Noller, T. A. Steitz, P. B. Moore, A. Yonath, V. Ramakrishnan

Nissen et al. (2000) Science, 289

Luc Jaeger ©
Only RNA in the reaction site
Comparisons of 3D structures

Comparisons of sequences

| ( ( ( ( (( | UUAGG | GGA | GUUUUA | UCC | AGCGU | CAG-C |
| GCUG | UUAGG | GGA | GUUUUA | UCC | AGCGU | CAG-C |
| GCCG | UUAGG | GGA | GUUUCA | UCC | AGCGA | UGG-C |
| GUUG | UAGG | GGA | GUCUCA | UCC | AGCA | CAA-C |
| GCUG | GAGG | GAA | GC | AA | UUC | AGCA | CAG-C |
| ACUU | CAGU | GGA | GC | AA | UUC | AGCA | GAGAU |
| ACUU | CAGU | GGA | GC | AA | UUC | AGCA | GAGAU |
| GAUG | GAGG | UUG | G | AAA | CAA | UGCA | CAU-C |
| GGGC | CAGG | GGU | G | AAA | ACC | AGCA | GCC-A |
| GGCC | UAGG | UCG | G | AAA | CGG | AGCA | GGU-C |
| GGCC | CAGG | UCG | G | AAA | CGG | AGCA | GGU-C |
| GGCC | CAGG | UCG | G | AAA | CGG | AGCA | GGU-C |
| GGCC | CAGG | UCG | G | AAA | CGG | AGCA | GGU-C |
Biological sequence analysis

Protein easy
RNA hard
Watson-Crick base paired helices

Internal loops (symmetric, asymmetric, bulge)

Hairpin loops

Single-strands junctions
RNA alignments

- RNA sequences are aligned/compared differently because sequence variation in RNA maintain base-pairing patterns

  ![Diagram](image)

- Thus an alignment will exhibit covariation at interacting basepairs

- RNA specifying genes will have conserved regions reflecting common ancestry
Main building block: the RNA double helix held together by Watson-Crick pairs.