NUCLEIC ACIDS (Contd.)
Lecture VI

Eric Mbindo Njunju

Bsc; Msc

Ribonucleic acids (RNAs)

- Ribonucleic acids: Present in nucleus and cytoplasm of eukaryotic cells.
- Also present in prokaryotes.
- Involved in the transfer and expression of genetic information.

Act as primers for DNA formation.

Some RNA act as enzymes as well as coenzymes.

RNA also functions as genetic material for viruses.

Chemical nature of ribonucleic acids

- Like DNAs, RNAs are also poly nucleotides.
- In RNA polymer, purine and pyrimidine nucleotides are linked together through phosphodiester linkage.
- The sugar present in a RNA is ribose.

Mainly three types of RNAs in all prokaryotic and eukaryotic cells.

1. Messenger RNA or m-RNA
2. Transfer RNA or t-RNA
3. Ribosomal RNA or r-RNA.

Differ from each other by size, function and stability.

Messenger RNA

Accounts for 1-5% of cellular RNA.

Structure

1. Majority of mRNA has primary structure. They are single-stranded linear molecules.

Consist of 1000-10,000 nucleotides (Fig).

2. mRNA molecules have free or phosphorylated 3’ and 5’ end.

3. mRNA molecules have different life spans. Their life span ranges from few minutes to days.

4. Eukaryotic mRNA are more stable than prokaryotic mRNA.

5. The mRNA nucleotide sequence is complementary from which it is synthesized or copied.

6. Some eukaryotic mRNA molecules are capped at 5’ end. The cap is methylated GTP (m7GTP).
Some mRNA contain internal methylated nucleotides. Capping protects mRNA from nuclease attack.

At 3′ end of most of eukaryotic mRNA, a polymer of adenylate (poly A) is found as a tail.

Poly A tail protects mRNA from nuclease attack.

8. In prokaryotes 5′ end of mRNA contains a sequence rich in A and G. Such a sequence is known as *Shine-Dalgarno sequence*. Helps attachment of mRNA with ribosome during protein synthesis.

9. Some prokaryotic mRNA has secondary structure. Intra-strand base pairing among complementary bases allows folding of liner molecule. As a result hairpin, or loop like secondary structure is formed (Fig).

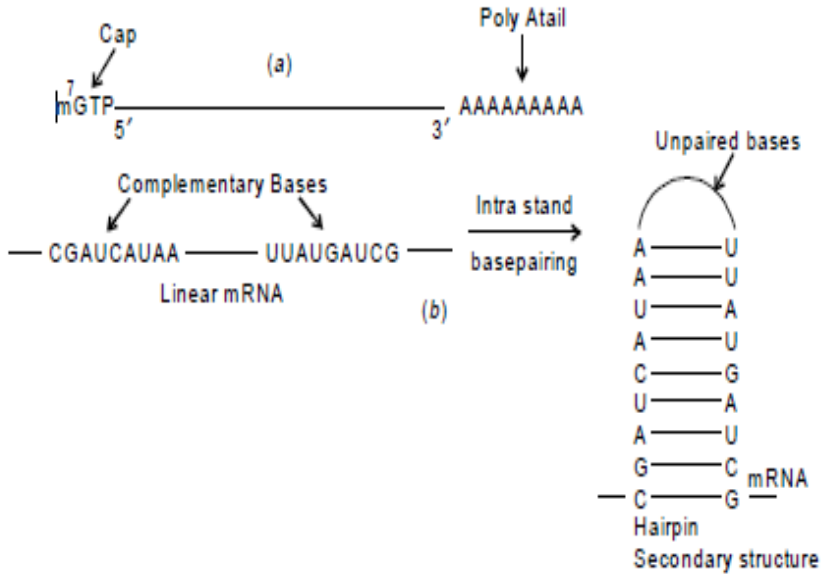


Fig: (*a*) Structure of mRNA

(*b*) Secondary structure formation from linear mRNA molecules

Functions

1. mRNA is direct carrier of genetic information from the nucleus to the cytoplasm.

2. Usually a molecule of mRNA contains information required for the formation of one protein molecule.

3. Genetic information is present in mRNA in the form of genetic code.

4. Some times single mRNA may contain information for the formation of more than one protein.

Transfer RNA

Accounts for 10-15% of total cell RNA.

Structure

Smallest of all the RNAs. Usually they consist of 50-100 nucleotides.

They are single strand molecules. t-RNA molecules contain many unusual bases 7-15 per molecule.

They are methylated adenine, guanine, cytosine and thymine, dihydrouracil, pseudo uridine, isopentenyl adenine etc. These unusual bases are important for binding of t-RNA to ribosomes and interaction of t-RNA with aminoacyl-t-RNA synthetases. About half of the nucleotides in t-RNA are involved in intrachain base pairing. As a result, double helical segments are formed in t-RNA. Further some bases are not involved in the base pairing resulting in loops and arms formation in t-RNA. Thus, folding in primary structure generate secondary structure. Though t-RNAs differ in chain lengths they have some common features with regard to secondary structure.

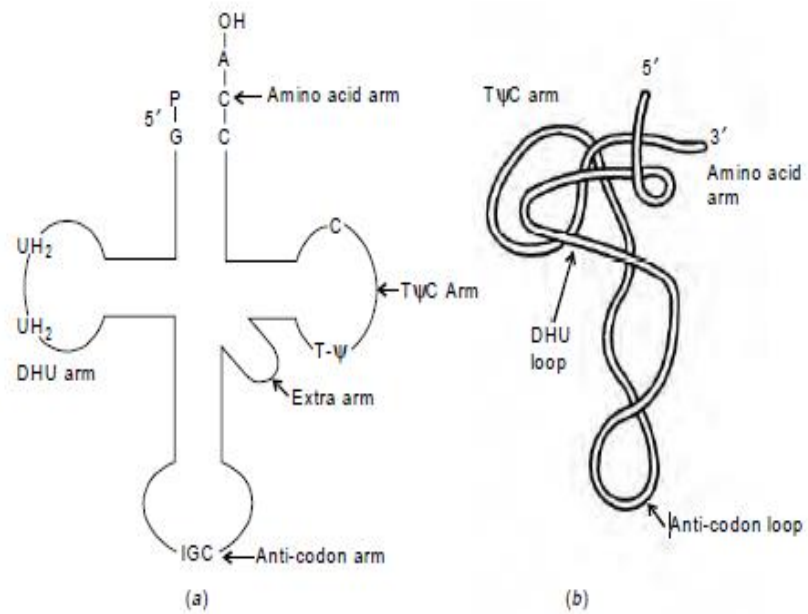


Fig:(*a*) Secondary structure of t-RNA (*b*) Tertiary structure of t-RNA

An amino acid arm where amino acid is attached to 3′-OH of adenosine moiety of t-RNA.

ACC is the common base sequence at this 3′-end.

2. Tφc arm, which contains sequence of ribothymidine-pseudouridine-cytidine. Greek alphabet φ (Psi) stands for pseudo uridine. Thymine and pseudouracil are the two unusual bases found in this arm.

3. An anti-codon arm, which recognizes codon on mRNA.

4. DHU arm, which contains many dihydrouridine (UH2) residues.

5. The 5′ end of t-RNA is phosphorylated and the residue is guanosine.

6. About 75% t-RNA molecules have extra arm. It consists of 3-5 base pairs. It is found between TφC and anti-codon arm.

Tertiary structure of t-RNA

X-ray diffraction analysis indicated complex three-dimensional structure for t-RNA molecule.

Three-dimensional structure of t-RNA looks like an inverted or tilted L. The anti-codon arm is at the tip of the vertical arm of tilted L. The acceptor arm is at the tip of the horizontal arm of tilted L. The D loop and TφC loop are pushed into corner of tilted L (Fig).

Functions

1. Carrier of amino acids to the site of protein synthesis.

2. There is at least one t-RNA molecule to each of the 20 amino acids required for protein synthesis.

3. Eukaryotic t-RNAs are less stable where as prokaryotic t-RNAs are more stable.

Ribosomal RNA

Accounts for 80% of total cellular RNA.

Present in ribosomes.

In ribosomes, r-RNA is found in combination with protein.

It is known as *ribonucleoprotein*.

The length of r-RNA ranges form 100-600 nucleotides. Both prokaryotic and eukaryotic ribosomes contain r-RNA molecules.

r-RNAs differ in sedimentation coefficients (S). There are four types of r-RNAs in eukaryotes. They are 5, 5.8, 18 and 28S r-RNA molecules.

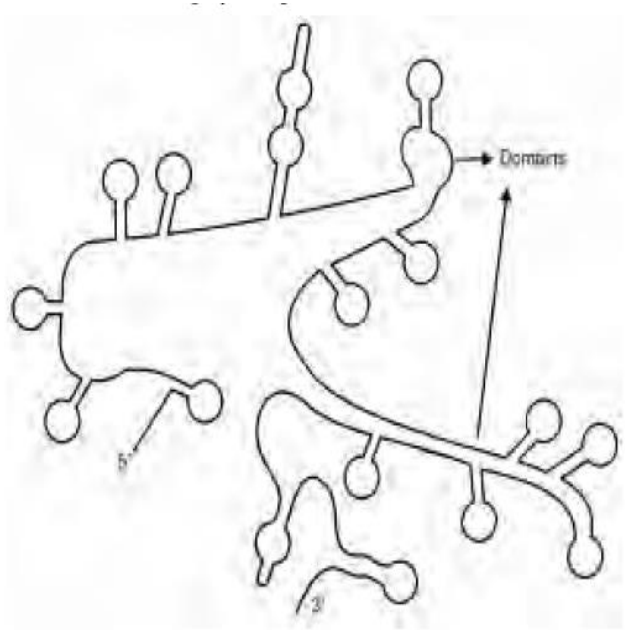
Prokaryotes contain 3 types of r-RNA molecules. They are 5, 16 and 23S r-RNA molecules

Structure

r-RNA molecules have secondary structure. Intra strand base pairing between complementary bases generates double helical segments or loops.

They are known as domains. 16S r-RNA with 1500 nucleotides has four major domains (Fig).

The three-dimensional tertiary structure of r-RNA is highly complex



Fig(*c*): Secondary structure of 16S r-RNA

Functions

1. r-RNAs are required for the formation of ribosomes.
2. 16S RNA is involved in initiation of protein synthesis.

Differences between DNA and RNA

DNA	RNA
1. Sugar moiety is deoxy ribose	Sugar moiety is ribose
2. Uracil, a pyrimidine base is usually absent	Thymine, a pyrimidine base is usually absent

(Contd.)

3. Double-stranded molecules	Single stranded molecules
4. Sum of purine bases is equal to sum of pyrimidine bases A + G = C + T	Sum of purine bases is not equal to sum pyrimidine bases A + G # C + T
5. Resistant to hydrolysis by alkali because of absence of hydroxyl group on 2 carbon atom of deoxyribose	Because of presence of hydroxyl group on 2 carbon atom of ribose RNA is easily hydrolyzed by alkali
6. Bases are not modified	Bases are modified
7. No catalytic activity	Some RNA are catalytically active
8. Only one form or type	More than three types
9. Usually not subjected to degradation in cell	Degraded in the cell by nucleases

Non-Coding RNAs

Other types of RNAs found recently in mammals, yeast and bacteria.

Small RNA molecules. They are so named because they do not code for protein product.

They are often referred to as genes without protein product.

These RNAs may arise from junk DNA, which is an inert part of genome which is of little transcriptional and translational potential.

Some non-coding RNAs serve as molecular chaperones. Some serve as anti-sense molecules and interfere with transcription and translation. They are also involved in genomic imprinting, X-chromosome in activation, germ cell formation, Meiosis, oxidative stress and diseases like cancer.

Human Genome Project (HGP)

Involved in sequencing of whole genome of humans, which is organized as chromosomes.

Two groups:

(*a*) Human Genome Consortium consisted of 16 international centres

(*b*) Celera Genomics of USA engaged in project. It began in 1990 and completed by 2000.

In February 2001, the two teams published results in two separate papers. Completion of human genome project is an extra ordinary achievement of man comparable to that of landing on moon.

Genome used for sequencing is obtained by an elaborate process involving DNA samples from blood of female donors and sperm of male donors. Identity of donor is not disclosed.

Dideoxy method of Sanger-used for sequencing by two groups. Sequencing is done by specially designed high-speed sequencers with little human involvement, which have very high (through) put. Though both groups used dideoxy method for sequencing, they adopted different approach for sequencing. Human genome consortium adopted Top-down approach in which genome is first segregated into smaller segments in a stepwise manner and when pieces are small enough, they are sequenced. After sequencing, these individual pieces are joined together to get chromosome of their origin by back tracking.

Shot-gun procedure or Bottom-up approach is adopted by Celera genomics headed by Venter for sequencing. It is known as whole Genome Shot-gun (WGS) procedure. It involves breaking the genomic DNA into small fragments and sequencing all of them in an unbiased manner. All the sequenced fragments are assembled by matching identifying pairs of sequences among any two fragments.

When sequencing of fragments is completed, its genes or protein coding regions are detected by using computational biology procedures. Total number of genes present in 3.2 billion base pairs containing human genome ranges from 33,000 to 1,50,000. The remaining non-coding DNA is often termed as junk DNA. It is genes-containing array of sequences that determines coordination, communication and functions of the cells which are ultimately responsible for proper health and well being of an individual. Human genome sequence provide some solutions to atleast few medical problems which remained mystery.

Sequencing of human genome allows mapping of disease genes on specific locations on chromosomes. Some disease genes, which are mapped on chromosomes are given on the next slide

Chromosome	Disease genes
Chromosome 1	Gaucher's disease, Breast Cancer
Chromosome 3	Alkaptonuria, Myeloid, Leukamia
Chromosome 6	Celiac disease, Hemochromatosis
Chromosome 7	Cystic fibrosis, Split hand/Foot malformation
Chromosome 8	Lipo protein lipase, Cohen syndrome
Chromosome 9	Fanconi anaemia Type C
Chromosome 10	Wolman disease
Chromosome 11	Ataxia telangiectasia, Wilm's tumour
Chromosome 12	Phenyl ketonuria
Chromosome 13	Wilson's disease, Retinoblastoma
Chromosome 14	Alzheirmer disease, Spastic paraplegia
Chromosome 15	Tay-Sach's disease, Bloom syndrome
Chromosome 16	Fanconi anaemia Type A, Inflammatory bowel disease.
Chromosome 18	Niemann–Pick disease, Colorectal cancer.
Chromosome 20	Polymorphic severe combined immuno deficiency
Chromosome X	Duchenne muscular dystrophy, Adreno leukodystrophy
Chromosome Y	Prostate cancer, Adenocarcinoma

Identification of new disease genes may provide starting point for the development of new diagnostic kits. Further genome sequence enables identification culprit genes involved in diseases whose underlying causes are yet to be elucidated.

Sequence information provides molecular details of signal transduction, differential expression of gene products in various tissues during normal growth, uncontrolled growth in tumour tissues. Human genome sequence allows identification of species

THANK YOU