1. Introduction

1.1 Book Overview

After the general introduction, RNA motifs and their functional implications are explained with different specific tables and examples in Chapter 2. Experimental (Chap. 3) and theoretical (Chap. 4) standard techniques to identify RNA have been kept, as in the pre-runner of this book, as these techniques in the field remain the most powerful tool to identify new instances of regulatory RNA motifs. However, we suspect that these may change in the future with large-scale array techniques in functional genomics becoming more and more powerful. In a similar way, network models of the cell are also becoming increasingly popular and may soon provide solid tools of their own to identify new regulatory RNA networks. Dynamics as well as medical implications of RNA motifs (Chap. 5) have experienced interesting new contributions during the past 3 years and become increasingly more important. Similarly we note that unexpected new hot topics (Chap. 6) of research in the field have come up, such as translational silencing and IRES mediated translation initiation.

The interested student is heartily welcomed to this book; however, probably he or she will be on an advanced level, for example doing a thesis in the field. For the specialist the book should be a well-suited reference, summary and incentive with specific chapters covering different angles of interest on regulatory RNA elements.

1.2 RNA, the Underestimated Molecule

RNA has long been underestimated in its capacities, originally considered an uninteresting contaminant of the more important proteins. Torbjörn Caspersson was among the first during the late 1930s to suggest from microscopic studies that RNA has an important role in the cytosol of the eukaryotic cell. Using chemical approaches, Jean Brachet (reviewed in Brachet 1987) confirmed Caspersson's findings and found in cytosolic ribonucleoprotein particles a first hint of the ribosome. Only in the 1950s did radiolabelling studies show that amino acids were assembled sequentially on ribosomes to form proteins. In 1955 Crick formulated the adaptor hypothesis, in which translation was thought to occur through the mediation of transfer RNA (tRNA) adaptor molecules. At about this time Paul

Zamecnik and Mahlon Hoagland started to discover tRNA: during protein synthesis C^{14}-labelled amino acids became transiently bound to a low molecular mass fraction of RNA, soluble RNA or sRNA (Holley et al. 1965). In 1958, including his adaptor hypothesis, Francis Crick formulated the central dogma of molecular biology, the flux of information from DNA via RNA to proteins (reviewed in Crick 1970).

Sidney Brenner and coworkers (1960) were the first to describe mRNA as "an unstable intermediate carrying information from genes to ribosomes for protein synthesis". In 1961 Jacob and Monod published their seminal paper on the existence of mRNA and operons and explained how the transcription of operons was regulated. The involvement of RNA in the synthesis of proteins was established in 1963 by Watson. Hall and Spiegelman in 1964 showed sequence complementarity of T2-DNA and T2-specific RNA. As a result of these studies, the exciting discovery of the genetic code began (Crick 1965). This culminated in the RNA being established as the messenger, the transient carrier of genetic information. As a further surprise, in 1975 Thomas Cech and colleagues (reviewed in Cech 1993) showed that RNA could also be catalytically active in addition to carrying genetic information. In view of these additional capabilities, RNA is now generally considered to have been critical for the evolution of life before the genetic code had evolved (reviewed in Gesteland et al. 1999): RNA could thus have performed dual functions in ancestral cells for which DNA and proteins are now used. DNA stores more stable information than RNA and proteins can be more catalytically active than RNA, but only RNA can do both. Reverse transcriptase found in retroviridae (Temin 1972), explained the emergence of DNA as a more solid permanent form of storage but challenged the "central dogma of molecular biology".

Increasing evidence supports the RNA world. Splicing, the metabolism of precursor mRNA involving the excision of internal sequences is another surprise in this respect. Initial studies (by the groups from Phillip Sharp and Richard Roberts; Chow et al. 1977, Berget et al. 1977) compared mRNA and DNA from adenovirus, and found viral DNA segments which were not contained in the mature mRNA. Subsequently such gene interruptions were found in ovalbumin and ß-globin genes (Crick 1979; Chambon 1981; Perry 1981). The RNA world theory is further supported by the observation of many nucleotide cofactors such as NADH, FAD found in basic metabolic enzymes which may be considered as a vestige from the transition of an RNA-dominated world to a protein-dominated one. Moreover, many new different types of ribozymes are now known which catalyze a large variety of reactions. Catalytic ribonucleoprotein particles in which the active RNA molecule is the principal part are another supporting observation such as RNAse P (Altman et al. 1993; Warnecke et al. 1996; Gesteland et al. 1999). The double capacity of the RNA molecule, both to carry information and to be catalytically active, is the foundation for its many functions and uses in the living cell.

1.3 Examining Regulatory RNA Elements

Information storage and transport are required for many tasks of the cell. RNA is an optimal molecule for this. Moreover, RNA metabolism creates many different levels where information may be stored, transported or specifically released in the cell. For such regulatory steps and other specific functions, specific structures in the RNA, so-called RNA elements or motifs, have evolved (several ones in RNA are shown in Fig. 1.1; see below for details). They allow specific and controlled release of information in the cell from an RNA containing such an element (Dandekar and Hentze, 1995). An RNA element (sometimes also called RNA motif, regulatory element, binding site or simply RNA signal, each term stressing different aspects of the function in which the RNA segment is critically involved) is a small segment of an RNA molecule which is essentially required for a specific interaction of the RNA with the enviroment. This includes interaction with itself (autocatalytic cleavages, for instance) or other binding partners (RNAs and RNA "recognizing", i.e. binding proteins). Essential features of the RNA segment are specific nucleotides and the RNA structure. RNA structure is formed by base-pairing of different parts of the RNA molecule to each other. There are two possibilities: direct base pairing forms an RNA structure, this is called secondary structure. If there are higher-level interactions, this is called tertiary structure. Examples for tertiary structure are interactions between two secondary structures, or between a secondary structure and an additional single-stranded region. Tertiary interactions often happen between nucleotides separated by long (100 or more base pairs) distances, however, shorter tertiary interactions are also known (see Fig. 1.2, right).

As an introductory RNA illustrating the combination of sequence, secondary structure and further structural features, the rev-response core element (RRE) from human immuno deficiency virus (HIV) is shown in Fig. 1.2. The nucleotide sequence of the core motif where rev protein binds the RNA is shown in full. Besides the primary sequence, base pairing between both strands is apparent, forming a stem. Further features of the secondary structure are nucleotides not paired but "bulged out", the A and U on the right side of the stem. The motif itself is part of a more complex structure. The whole rev RNA structure encompasses 234 nucleotides and some of the additional secondary structure, composed of RNA stems and single-stranded RNA loops are sketched. Opposing base-paired RNA strands form stems, shown as parallel lines; unpaired single-stranded RNA forms "loops", shown as open circles (the complete element includes further sequences close to the core consensus structure shown).

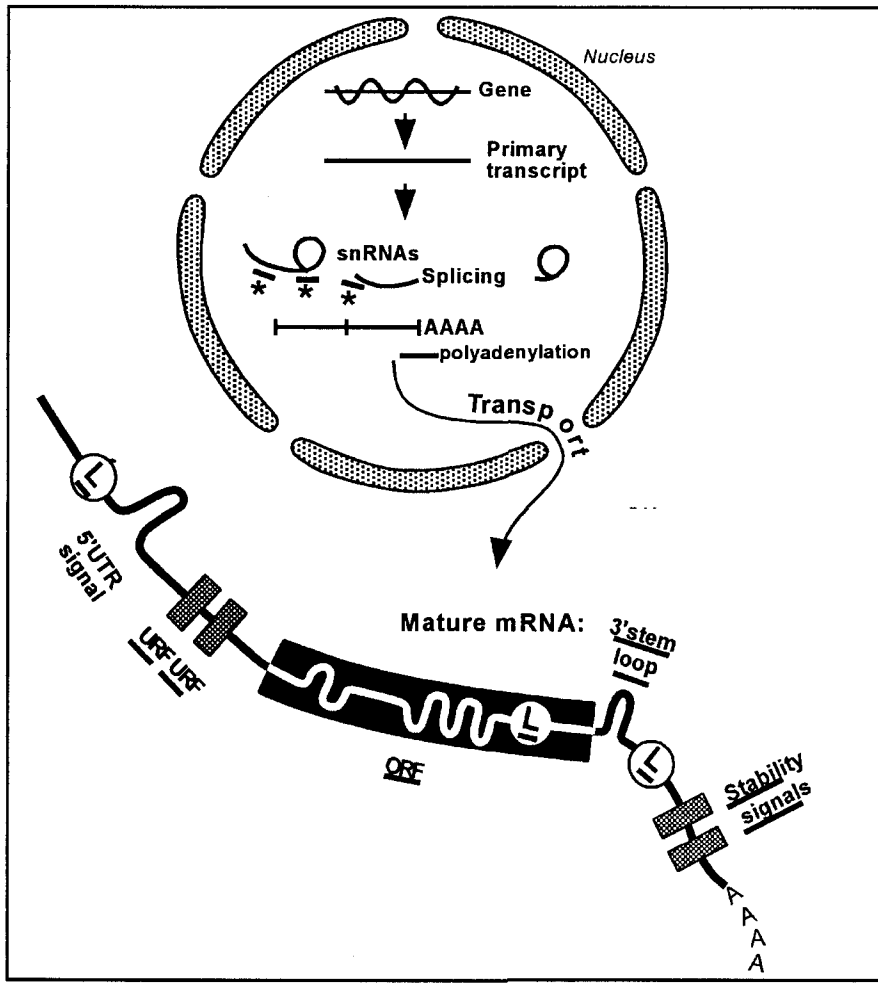

Fig. 1.1. Regulatory motifs involved in messenger RNA metabolism.
The figure summarizes regulatory motifs (indicated by *short solid lines*). *Top* Regulatory motifs involved in processing the mRNA precursor. After the gene is transcribed the primary transcript undergoes splicing, which requires several RNA-RNA interactions and recognition of RNA motifs. Schematically indicated by * in the nucleus interactions between snRNA and mRNA precursor at the 5' splice site, 3' splice site and the branch point are shown; not shown are numerous motifs present in the snRNAs involvedin the splicing process. After splicing the exons are assembled and polyadenylated (AAAA). This process as well as the subsequent transport from the nucleus into the cytoplasm again involves different RNA motifs on the mRNA. Motifs in the mature mRNA (*bottom part*) include: localization signals (*L*) in the 5'UTR and 3'UTR; stem loop structures in 5' UTR, open reading frame and 3' UTR; short upstream reading frames (uORF) in the 5' UTR; regulatory signals in the open reading frame encoded in its secondary structure and primary sequence exploiting the degeneracy of the genetic code and involving also localization signals; regulatory signals in the 3' UTR include stem-loop structures, further localization signals, sequences involved in mRNA stability and poly(A) tail (AAAA) formation

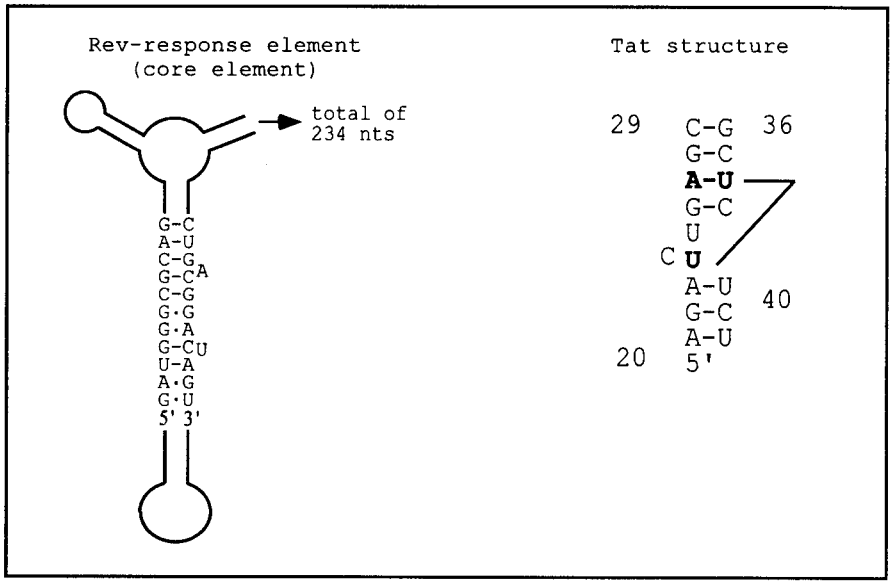

Fig. 1.2. Schematic diagramm depicting RNA motifs.HIV rev serves to illustrate a simple example. The core motif where rev protein binds the RNA is shown in full. Base-pairing between the 5' and 3' strands leads to the formation of a stem. Double stranded regions are drawn as *parallel lines* and unpaired single-stranded regions are drawn as *loops*. Watson Crick base pairs between A and U as well as between C and G are shown as -; all other hydrogen bonding interactions between different nucleotides are shown as throughout the book. A specific feature of the secondary structure are nucleotides (the A und U on the right side of the stem) which are not part of the base-paired stem but are " bulged out" of this RNA stem structure.The HIV rev core motif itself is part of a more complex structure, the complete rev element, which encompasses 234 nucleotides. The complete structure also forms a complex tertiary structure by long distance interactions of the RNA.
The Tat structure on the right shows a stem which is formed by base pairing between RNA strands. It illustrates a very simple tertiary interaction, in which three nucleotides,shown in bold are connected because a uracil contacts an A-U Watson-Crick base pair. Further details on the particular motifs shown can be found in Malim and Cullem (1994).

Besides this secondary structure, the complete structure also forms tertiary interactions in which distal parts of the RNA interact with each other. For comparison, the *trans*-activation response element (TAR) from HIV on the right shows another regulatory RNA structure. Again, a stem is formed between base-pairing RNA strands; however, this RNA element also illustrates a very simple tertiary interaction. Apart from a standard Watson-Crick base pair between A and U, the three nucleotides shown in bold are connected by a tertiary interaction indicated by the lines forming a V on the right. Further details on the particular motifs shown are described in Cullen (1994). As evident from the figure and also the case for catalytic RNA structures, the function of the element depends both on critical sequence features and nucleotides as well as specific base pairing and

http://www.springer.com/978-3-540-41701-9

RNA Motifs and Regulatory Elements
Dandekar, Th. (Ed.)
2002, X, 234 p., Hardcover
ISBN: 978-3-540-41701-9