7.3 RNA EDITING (4)

RNA editing, also known as RNA modification, is a molecular process through which cells can alter specific nucleotide sequences within an RNA molecule after its synthesis by RNA polymerase. This phenomenon occurs in all living organisms and is one of the most evolutionarily conserved properties of RNAs. It involves the insertion, deletion, and base substitution of nucleotides within the RNA molecule, impacting the activity, localisation, and stability of RNAs, and is associated with various human diseases. RNA editing can be broadly categorised into two main types: substitution editing and insertion/deletion editing. **Substitution Editing** includes two main types, A to I editing and C to U editing. Adenosine to Inosine **(A-to-I) editing** is catalysed by adenosine deaminase acting on RNA (ADAR) enzymes. A-to-I editing is prevalent in vertebrates and alters the coding potential and splicing of mRNAs by deaminating adenosine to inosine, which is interpreted as guanosine by the ribosome. ADAR enzymes recognise double-stranded RNA (dsRNA) structures within pre-mRNAs. They deaminate adenosine residues to inosine, affecting the coding sequence and splicing of the RNA. This process is regulated by the binding of ADAR to specific RNA sequences and secondary structures. Cytidine to Uridine **(C-to-U) editing** is mediated by cytidine deaminase enzymes, which convert cytidine to uridine. In C-to-U Editing, initially the binding of cytidine deaminase occurs to specific RNA sequences, where it catalyses the deamination of cytidine to uridine. The process is often regulated by RNA-binding proteins and secondary RNA structures that facilitate the access of the editing enzyme to its substrate. A notable example is the editing of apolipoprotein B (apoB) mRNA in humans, producing different protein isoforms in the liver and intestines. The gene encoding this protein synthesizes a 4563-amino-acid polypeptide called Apo B100, which is produced in liver cells and released into the bloodstream to facilitate lipid transportation throughout the body. A related protein, Apo B48, is synthesized by intestinal cells. This protein is only 2152 amino acids long and is produced from an edited version of the mRNA for the full-length protein. In intestinal cells, mRNA is altered by the deamination of cytosine (catalyzed by cytidine deaminase), converting it into uracil. This alteration changes a CAA codon present at position 2153, specifying glutamine, into a UAA codon. This alteration causes translation to halt, resulting in the production of a truncated protein. In the case of A → I editing, deamination of adenosine leads to the formation of inosine, a process catalysed by enzymes called Adenosine Deaminases Acting on RNA (ADARs). Inosine (I) mimics guanosine behaviorally: it preferentially base pairs with cytosine, and when present in a codon, it is translated during protein synthesis as if it were G.

The second common type of editing is **Insertion-Deletion Editing.** This type of editing is common in mitochondrial and plastid RNAs of certain organisms, such as the kinetoplast protozoan *Trypanosoma brucei*. Here, the editosome complex binds to pre-mRNA in conjunction with gRNAs. The gRNAs guide the insertion or deletion of uridines by base-pairing with complementary sequences in the mRNA. The editosom then catalyzes the necessary insertions or deletions to align the mRNA with the gRNA template. A similar mechanism is also present in prokaryotic CRISPR-cas editing, which is a common immune mechanism of bacteria to degrade the foreign DNA and has recently emerged as a powerful molecular tool.

Box 7.1 CRISPR-Cas System

CRISPR-Cas, which stands for "Clustered Regularly Interspaced Short Palindromic Repeats" and CRISPR-associated proteins, is a groundbreaking technology in genomics due to its simplicity and effectiveness in gene editing. The origins of CRISPR date back to 1987 when Yoshizumi Ishino at Osaka University accidentally discovered these unique DNA sequences. What made CRISPR intriguing was the organisation of its repeated sequences, which were not consecutive like most DNA repeats. This peculiar arrangement puzzled scientists until later research revealed its significant role in bacterial immunity. In 1993, researchers studying Mycobacterium tuberculosis identified the diversity of these repeats and used them to develop a genotyping method called spoligotyping. By 2000, Francisco Mojica's team at the University of Alicante in Spain had identified these interrupted repeats in 20 microbial species and coined the term CRISPR to standardise the nomenclature. The CRISPR-Cas system functions like a personalised defense mechanism for bacteria, analogous to designing a weapon based on parts of an invader's body. When bacteria face viral attacks, they use Cas proteins to capture and integrate short sequences of the viral DNA into their own genome at specific CRISPR loci. This viral DNA is then transcribed into RNA, which guides the Cas proteins to recognise and cut the viral DNA during subsequent attacks, effectively neutralising the threat. This process unfolds in three phases: adaptation, where viral DNA is captured and integrated; biogenesis, where RNA is produced from the integrated DNA; and targeting, where RNA guides the Cas proteins to cleave the viral DNA. CRISPR-Cas systems are classified into different types based on their components and mechanisms. The most commonly used system in research is the Type II CRISPR-Cas, which includes the Cas9 protein. Cas9 is an endonuclease that cuts DNA at precise locations guided by RNA, making it an invaluable tool for genome editing. Cas9, along with guide RNA, can be programmed to target almost any sequence in the genome, allowing for precise modifications. Other types of CRISPR systems, such as Type I and Type III, involve different sets of Cas proteins and have distinct mechanisms for RNA processing and target recognition.

The applications of CRISPR-Cas are vast and transformative, extending across various fields of biology and medicine. In human cells, CRISPR has been used to inactivate specific genes, offering insights into gene function and the development of gene therapies. In agriculture, CRISPR is used to modify crops for better yield, disease resistance, and climate resilience. In the biofuel industry, it has been employed to enhance the production capabilities of yeast and other microorganisms. One of the most promising applications is in vector control, where CRISPR is used to engineer mosquitoes that cannot transmit diseases such as malaria, potentially reducing the prevalence of this deadly disease. CRISPR's versatility extends to RNA editing, allowing scientists to target and modify RNA molecules. This capability is particularly useful in combating RNA viruses like HIV and poliovirus. CRISPR can also be harnessed as an antimicrobial agent, where phages deliver CRISPR components to bacterial cells, leading to targeted cell death and offering a novel approach to tackling antibiotic-resistant bacteria. Additionally, CRISPR-based gene drives can propagate desirable traits through populations with high efficiency, which can be used for purposes such as controlling invasive species or spreading disease resistance. Beyond editing, CRISPR-Cas technology has non-editing

applications in research. For instance, catalytically dead Cas9 (dCas9) can be used to regulate gene expression without altering the DNA sequence. dCas9 fused with fluorescent proteins helps visualise specific genomic loci, providing insights into nuclear organisation. Furthermore, dCas9 can be used for epitope tagging, enabling the isolation and study of specific genomic regions and their associated proteins or RNAs.

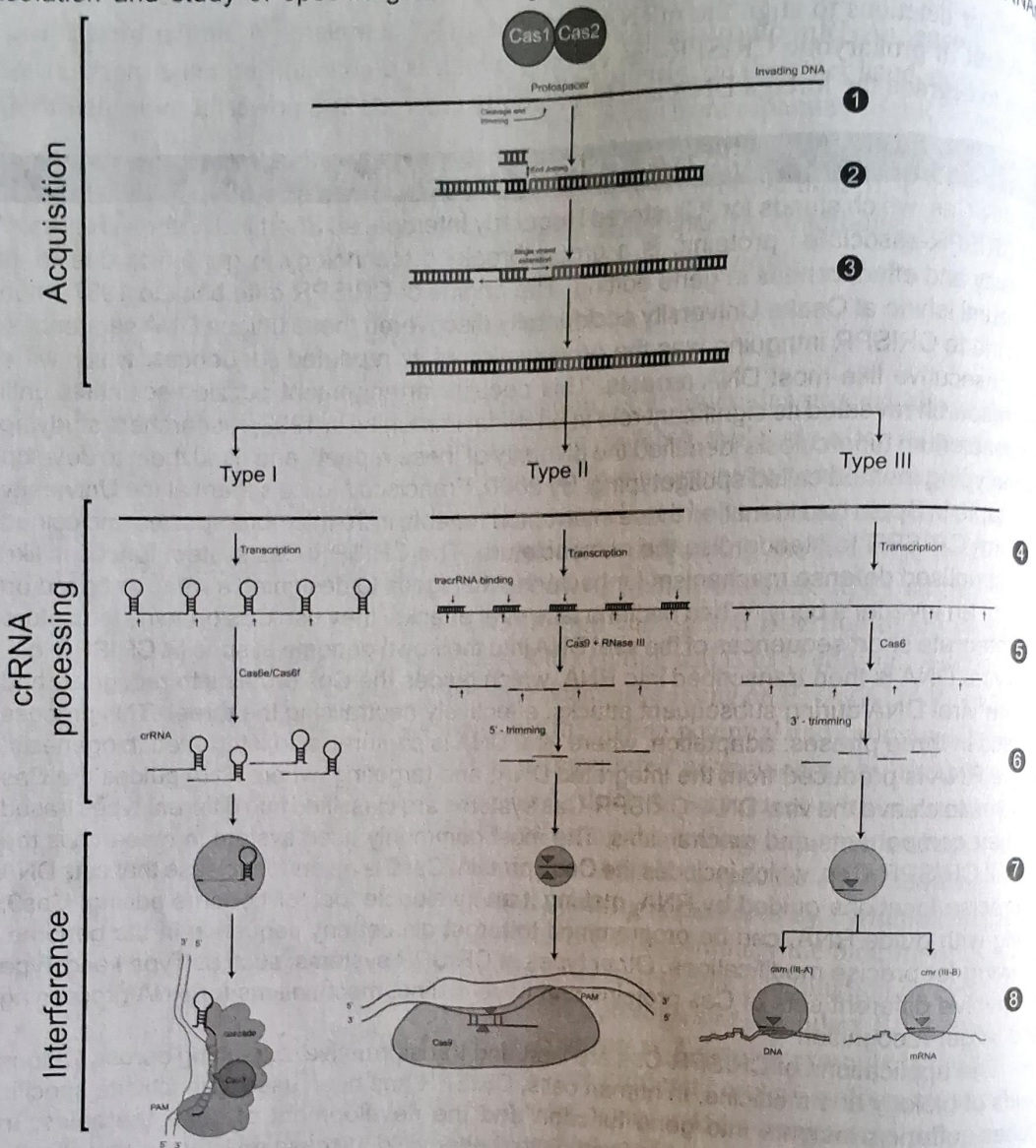

Fig. 7.4. *An overview of mechanism and types of CRISPR-Cas system.*

RNA editing plays a crucial role in regulating gene expression and ensuring the proper function of proteins. RNA editing can modify the coding sequence of mRNAs, leading to changes in the amino acid sequence of proteins. This can affect the activity, interactions, and localisation of the resulting proteins, thus finely tuning cellular functions and responses. In some organisms, extensive RNA editing allows for rapid adaptation to environmental changes. For example, in squids, RNA editing enables the production of protein variants that function optimally under different environmental conditions. RNA editing contributes to the regulation of gene expression by, modulating splicing or altering mRNA localisation and translation. Editing events can create or abolish splice sites, influencing the inclusion or exclusion of

exons in the final mRNA. Also, edited mRNAs may have different localisation signals or translation efficiencies, affecting where and how efficiently proteins are synthesized in the cell. Furthermore, editing can alter the binding sites for RNA-binding proteins and microRNAs, influencing RNA stability, translation, and degradation.

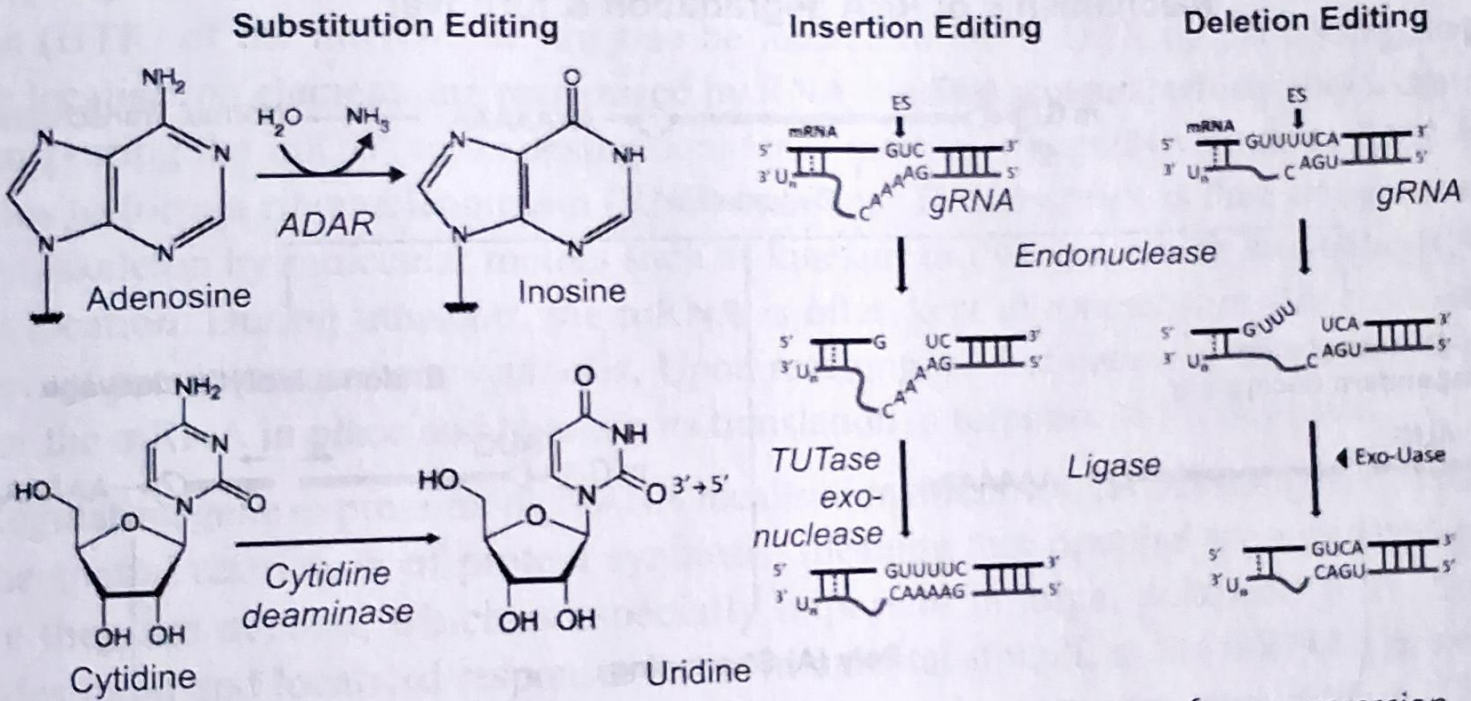

Fig. 7.5. *RNA Editing as a means of post-transcriptional regulation of gene expression.*