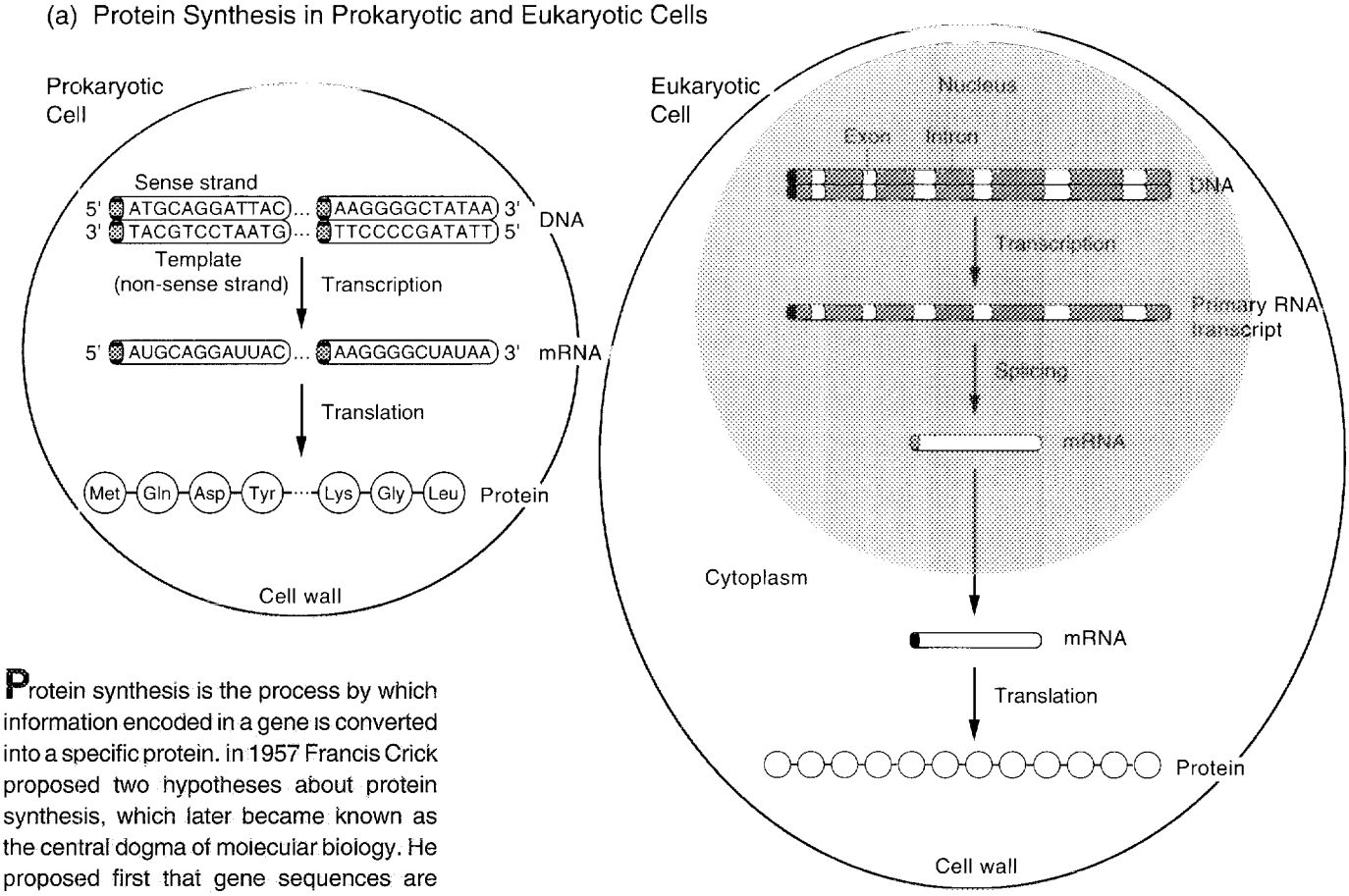
PROTEIN SYNTHESIS

(a) Protein Synthesis in Prokaryotic and Eukaryotic Cells



Protein synthesis is the process by which information encoded in a gene is converted into a specific protein. In 1957 Francis Crick proposed two hypotheses about protein synthesis, which later became known as the central dogma of molecular biology. He proposed first that gene sequences are "collinear" with protein sequences. In other words, the linear arrangement of subunits (deoxyribonucleotides) composing a gene corresponds to the linear arrangement of subunits (amino acids) composing a protein. Second, Crick proposed that a segment of RNA (a ribonucleotide sequence) acts as an intermediate translator between the deoxyribonucleotide sequence and the amino-acid sequence, or, in other words, that genetic information flows from DNA to RNA to protein. Crick had no experimental evidence to support his hypotheses. But very shortly Charles Yanofsky and Seymour Benzer working independently, provided the first evidence in support of the collinearity hypothesis. Their experiments showed that mutations in the genes of *E. coli* and of the T4 bacteriophage produced parallel changes in amino-acid sequences. And as details of protein synthesis were worked out, the role of RNA as an intermediary was also established.

Shown in (a) is an overview of protein synthesis in a prokaryotic cell. In the first stage, called transcription, a DNA segment, a gene, serves as a template for the synthesis of a single-stranded RNA segment called a messenger RNA (mRNA). The base sequence of the mRNA is complementary to the base sequence of one strand of the gene (the template, or "non-sense," strand) and is therefore identical to the base sequence of the other strand of the gene (the "sense" strand). The one exception to the identity is that the base U (uracil) replaces the base T. (Recall that in RNA uracil, rather than thymine, is the base complementary to adenine.)

In the second stage of protein synthesis, called translation, the mRNA serves as the template for the stringing together of amino acids into a protein. The protein is assembled according to the genetic code. That is, the succession of codons (triplets of adjacent ribonucleotides) that compose the mRNA dictates the succession of amino acids that compose the protein. (A listing of codons and corresponding amino acids is presented in "The Genetic Code.") Although transcription and translation are depicted here as if they occurred at different times, translation of a prokaryotic mRNA often begins before its synthesis by transcription is complete.

Also shown in (a) is an overview of protein synthesis in a eukaryotic cell. Unlike prokaryotic genes, most eukaryotic genes are composed of stretches of protein-coding sequences (exons) interrupted by longer stretches of noncoding sequences (introns). Both the exons and introns within a eukaryotic gene are transcribed. The resulting primary transcript is then spliced; that is, each intron is removed and the adjacent exons are linked together.

The shortened RNA is now an mRNA, an RNA that contains only protein-coding sequences. The mRNA leaves the nucleus and in the cytoplasm is translated into a protein according to the genetic code. Thus transcription and translation are of necessity temporally separated in eukaryotic cells.

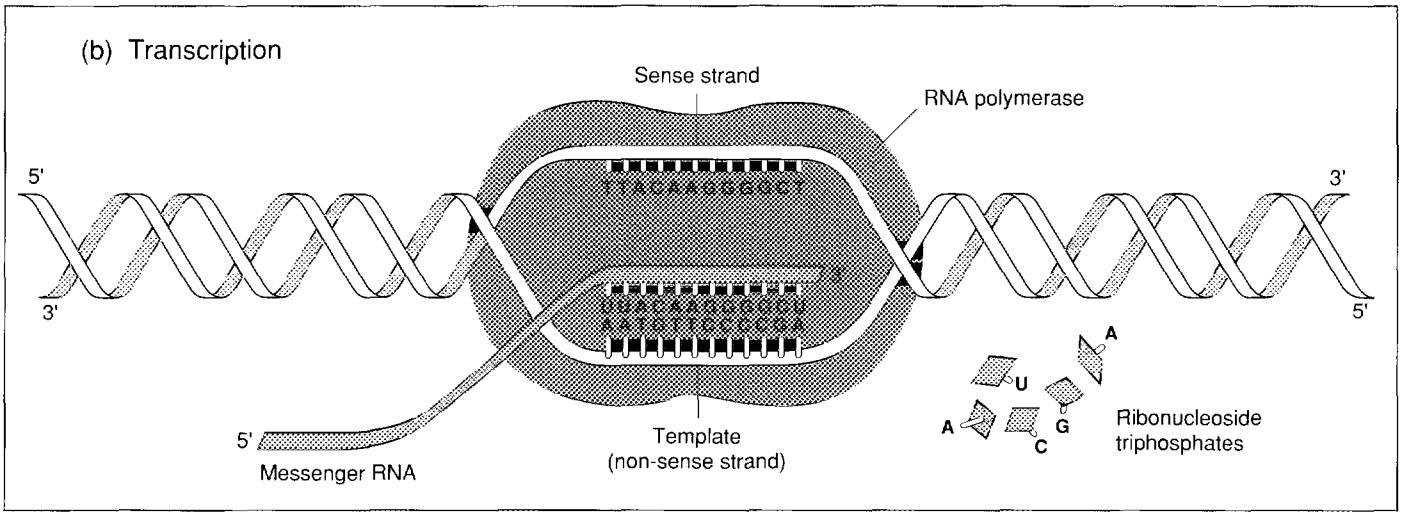
The overviews in (a) illustrate that, as Crick had postulated, genetic information flows from DNA to RNA to protein within both prokaryotic and eukaryotic cells. One important exception to the central dogma is the class of viruses known as retroviruses, of which the AIDS virus is an example. Retroviruses store genetic information in RNA and then convert the information to DNA—a reversal of the usual information flow that is known as reverse transcription.

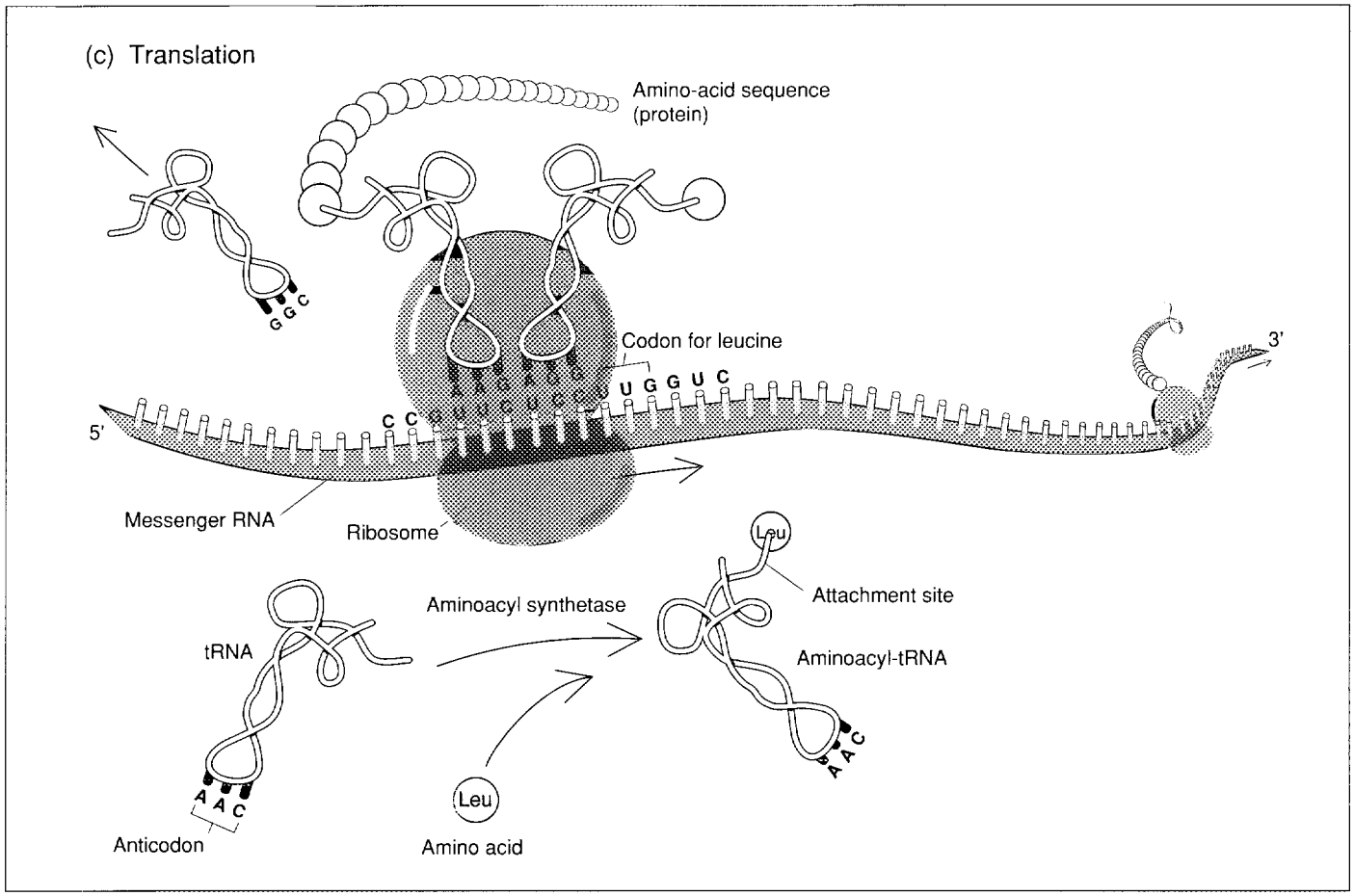
Details of transcription and translation are shown in (b) and (c) respectively. Transcription begins when an enzyme, an RNA polymerase, binds to a particular segment of a gene called the promoter. The double helix then uncoils and separates into two strands, exposing a small number of bases. The RNA polymerase facilitates hydrogen bonding between an exposed base in the template strand and its complementary base in a free ribonucleoside triphosphate (NTP) and then between the next exposed base in the template strand and its complementary base in another free NTP. While the two NTPs are held in proximity by the hydrogen bonds, the RNA polymerase catalyzes the formation of an -O–P–O- bridge between them, thus forming a chain of two covalently linked ribonucleotides. (See "DNA Replication" for details about formation of -O–P–O- bridges.) A third NTP is hydrogen-bonded to the third exposed base in the template strand and is covalently linked to the second ribonucleotide in the chain. The RNA polymerase moves along the template in the 3′-to-5′ direction, continuing to unwind and separate the double helix and to elongate the RNA chain in the 5′-to-3′ direction by catalyzing the addition of successive ribonucleotides. At the same time, the distorted DNA in the wake of the polymerase rewinds. After the gene is fully transcribed, the polymerase separates from the double helix. If the gene transcribed is a eukaryotic gene, the newly minted RNA is spliced and the resulting mRNA enters the cytoplasm through pores in the nuclear membrane.

As shown in (c), translation occurs with the help of transfer RNA molecules (tRNAs) and ribosomes. Each tRNA is a tiny, cloverleaf-shaped molecule that serves as an adapter: At one end it contains a triplet of ribonucleotides (an anticodon) that binds with a complementary codon on the mRNA strand, and at the other end it has an attachment site for a single amino acid. Many varieties of tRNAs exist. An important difference between one tRNA and another is the presence of a different anticodon on the central cloverleaf stem. The number of different anticodons found in the various tRNAs is less than the number of codons in the genetic code. That is so because the base pairing between the third base of the mRNA codon and the first base of the tRNA anticodon can depart from the usual Watson-Crick rules. For example, G can pair with U in addition to C.

Ribosomes are very large molecules composed of ribosomal RNA (rRNA) and approximately fifty different proteins. As a ribosome travels along an mRNA it catalyzes the reactions that lead to synthesis of the protein encoded in the mRNA. Thousands of ribosomes exist within each cell.

Before a tRNA molecule participates in translation, it must be converted to an aminoacyl-tRNA (become attached to the amino acid corresponding to its anticodon). Each of the twenty amino acids found in proteins can be attached to at least one type of tRNA, and most can be attached to several. The binding between tRNA and amino acid is cata-





lyzed by one of a group of enzymes. Those exquisitely specific enzymes, called aminoacyl synthetases, are in fact the agents by which the genetic information in mRNA is decoded.

Translation begins when an aminoacyl-tRNA containing the amino acid methionine and a ribosome bind to an initiation sequence near the 5′ end of the mRNA. The initiation sequence consists of the START codon AUG, to which the aminoacyl-tRNA binds through complementary base pairing. A second aminoacyl-tRNA, which contains an anticodon complementary to the second mRNA codon, binds to the mRNA. Then the amino acid on the first aminoacyl-tRNA is joined by a peptide bond to the amino acid on the second aminoacyl-tRNA, thus creating a chain of two amino acids dangling off the end of the second aminoacyl-tRNA. The process continues as the ribosome moves along the mRNA (in the 5′-to-3′ direction) and as peptide bonds are formed between successive amino acids. When the ribosome reaches a STOP codon within the mRNA, the ribosome detaches from the mRNA, and the completed protein is released into the cytoplasm.

The process of translation is fast: A single ribosome can translate up to fifty ribonucleotides per second. Furthermore, at any one time numerous ribosomes may be traveling along a single mRNA, each producing a molecule of the same protein. Thus a protein needed for diverse tasks within the cell can be quickly and efficiently produced.

Note: Published only recently (in June 1992) was strong evidence that the formation of peptide bonds between amino acids during translation is catalyzed not by some protein enzyme within a ribosome but instead by an RNA component of the ribosome. That news is exciting but not completely unexpected, since the ability of RNA to function as a catalyst in other situations had been demonstrated in the early 1980s. In particular, the primary transcript of a ribosomal-RNA gene of the protozoan *Tetrahymena thermophila* had been shown to effect its own splicing and the catalytic action of an RNA-protein complex that processes the primary transcripts of certain transfer-RNA genes had been ascribed to the RNA component of the complex rather than the protein component.