Lecture -4-

THE microbial genetics

Dr. Yasir Alabdali

Translation

The genetic code

The mRNA carries the information for the sequence of amino acids in a protein in the form of the Genetic Code in which each occurrence of one of the 64 groups of three nucleotides (triplets or codons) codes for a specific amino acid (or for a stop signal).

The code is almost universal, in all species, although there are occasional minor differences, such as the use of UGA, which is normally a stop codon, to code for tryptophan or cysteine. This indicates that the code as we know it must have originated early in the evolutionary process and then became fixed because of the effect that any change would have on virtually every gene in the cell.

		2nd base							
		T		C		A		G	
1st base	T	TTT	(Phe/F) Phenylalanine	TCT	(Ser/S) Serine	TAT	(Tyr/Y) Tyrosine	TGT	(Cys/C) Cysteine
		TTC	(Phe/F) Phenylalanine	TCC	(Ser/S) Serine	TAC	(Tyr/Y) Tyrosine	TGC	(Cys/C) Cysteine
		TTA	(Leu/L) Leucine	TCA	(Ser/S) Serine	TAA	Ochre (*Stop*)	TGA	Opal (*Stop*)
		TTG	(Leu/L) Leucine	TCG	(Ser/S) Serine	TAG	Amber (*Stop*)	TGG	(Trp/W) Tryptophan
	C	CTT	(Leu/L) Leucine	CCT	(Pro/P) Proline	CAT	(His/H) Histidine	CGT	(Arg/R) Arginine
		CTC	(Leu/L) Leucine	CCC	(Pro/P) Proline	CAC	(His/H) Histidine	CGC	(Arg/R) Arginine
		CTA	(Leu/L) Leucine	CCA	(Pro/P) Proline	CAA	(Gln/Q) Glutamine	CGA	(Arg/R) Arginine
		CTG	(Leu/L) Leucine	CCG	(Pro/P) Proline	CAG	(Gln/Q) Glutamine	CGG	(Arg/R) Arginine
	A	ATT	(Ile/I) Isoleucine	ACT	(Thr/T) Threonine	AAT	(Asn/N) Asparagine	AGT	(Ser/S) Serine
		ATC	(Ile/I) Isoleucine	ACC	(Thr/T) Threonine	AAC	(Asn/N) Asparagine	AGC	(Ser/S) Serine
		ATA	(Ile/I) Isoleucine	ACA	(Thr/T) Threonine	AAA	(Lys/K) Lysine	AGA	(Arg/R) Arginine
		ATG	(Met/M) Methionine	ACG	(Thr/T) Threonine	AAG	(Lys/K) Lysine	AGG	(Arg/R) Arginine
	G	GTT	(Val/V) Valine	GCT	(Ala/A) Alanine	GAT	(Asp/D) Aspartic acid	GGT	(Gly/G) Glycine
		GTC	(Val/V) Valine	GCC	(Ala/A) Alanine	GAC	(Asp/D) Aspartic acid	GGC	(Gly/G) Glycine
		GTA	(Val/V) Valine	GCA	(Ala/A) Alanine	GAA	(Glu/E) Glutamic acid	GGA	(Gly/G) Glycine
		GTG	(Val/V) Valine	GCG	(Ala/A) Alanine	GAG	(Glu/E) Glutamic acid	GGG	(Gly/G) Glycine

nonpolar	polar	basic	acidic	(stop codon)

Ribosomes

Bacterial ribosomes consist of two subunits with sedimentation coefficients of **50S and 30S**, the whole structure being referred to as a 70S ribosome. **The larger (50S) subunit has two RNA molecules (23S and 5S) plus 31 different polypeptides, while the smaller one (30S) contains a single RNA molecule (16S) and 21 polypeptides**. Note that the structure of eukaryotic ribosomes is different in several respects.

In eukaryotes ribosome full (80s) a split to bottom and two small (40s) and large (60s), the unit large contains rRNA (5.28 s), either a small rRNA contains acid type (18s).

The ribosomal RNA molecules form a very stable three-dimensional structure by extensive base-pairing, which allows them to perform a scaffolding role by attachment of the various ribosomal proteins. The role of the rRNA is involved in recognition of the mRNA and in the catalytic events leading to peptide bond formation.

In bacteria, the ribosomes attach to a specific sequence on the mRNA (the ribosome binding site, or RBS, also known as the Shine–Dalgarno sequence). This sequence is partly complementary to the 3′ end of the 16S rRNA, so that binding of the ribosomes can be mediated by hydrogen bonding between the complementary base sequences. This will normally occur as soon as the binding site is available, so the mRNA will start to be translated while it is still being formed.

However, it is not the complete ribosome that initiates these events. Ribosomes that are not involved in translation dissociate into their constituent 50S and 30S subunits.

For translation to start, a 30S subunit binds to a ribosome binding site, and an initiator tRNA associates with an adjacent initiation codon (usually AUG, but sometimes GUG or even less commonly CUG). The 50S subunit can then attach to this initiation complex, and the process of translation can get underway. Since the mRNA is read in consecutive groups of three (with no punctuation), it could code for three completely different proteins, depending on where it starts, i.e. there are three potential reading frames. The position of the ribosome binding site and the initiation codon determines the reading frame. As the addition or deletion of a single base will change the reading frame and the coding property of the subsequent message is totally different.

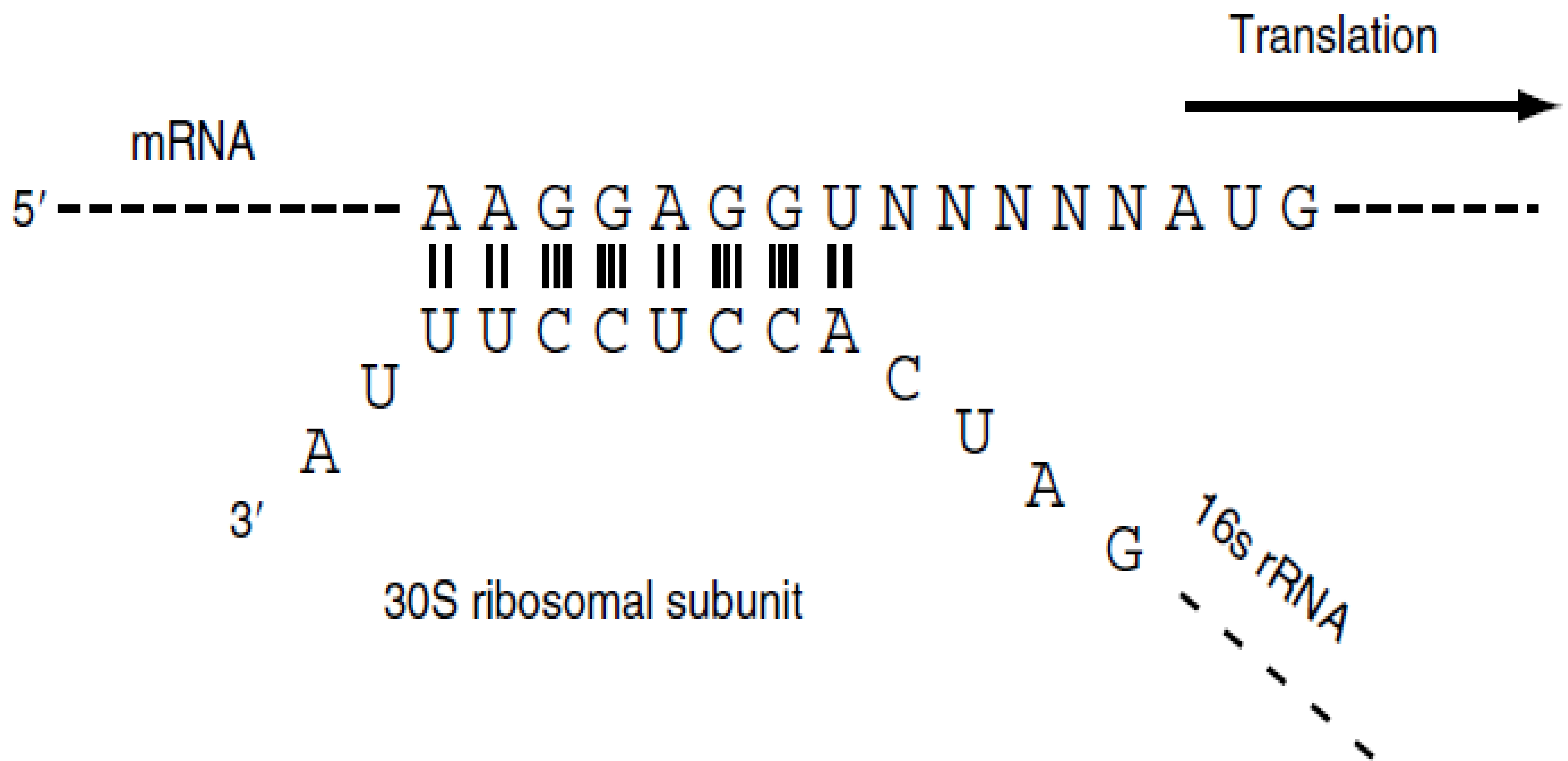

Figure 1.19 Ribosome binding site (Shine–Dalgarno sequence)

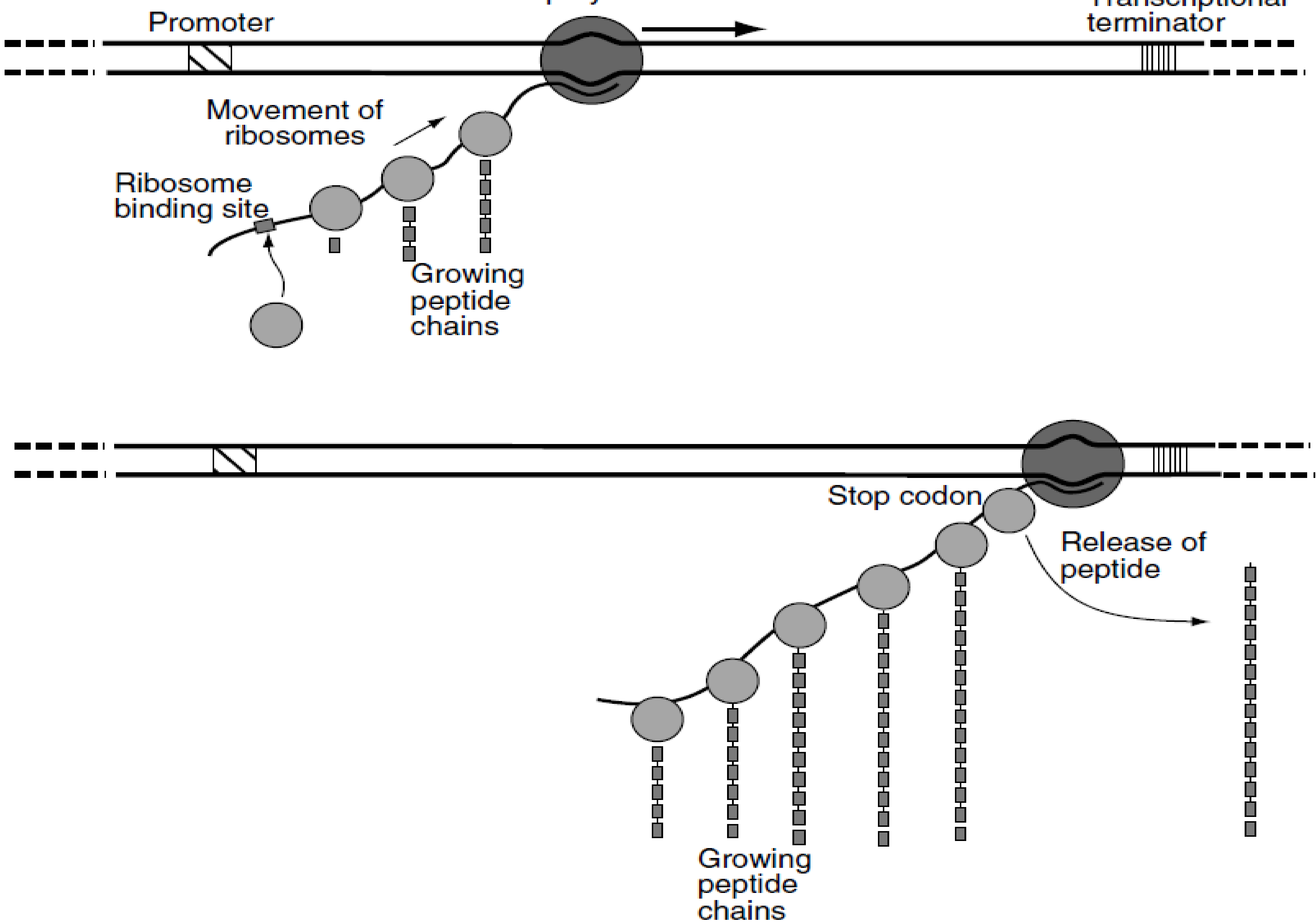

Figure 1.20 Translation of mRNA

Transfer RNA

Recognition of each triplet codon is mediated by small RNA molecules known as transfer RNA (tRNA). There is at least one tRNA species specific for each amino acid. However, they are all quite similar in their structure, consisting of a single RNA chain of 75–100 nucleotides folded back on itself in a form usually depicted as a **cloverleaf structure**; the actual three-dimensional structure is more complex and compact than this simplified two-dimensional diagram. Two parts of this molecule have clear functions in protein synthesis: **the acceptor arm**, formed by base-pairing of the 5′ and 3′ terminal regions, provides the site for attachment of an amino acid (by acylation of the 3′ end), and **the anticodon arm** which contains the bases (the anticodon) that recognize the triplet codon in the mRNA by base-pairing.

The appropriate amino acid is added to the tRNA by a specific enzyme (one of a number of **aminoacyl tRNA synthetases**) which has a crucial dual specificity:

it is capable of recognizing a single tRNA species and also the correct amino acid with which that tRNA should be charged. Thus for example the codon UGG (which codes for tryptophan) will be recognized by a specific tRNA (designated tRNAtrp). This tRNA will be recognized by the **tryptophanyl tRNA synthetase**. This therefore ensures that the tRNA is charged with the appropriate amino acid.

So there are three separate elements to the specificity of this process: **codon– anticodon interaction**, **recognition of the specific tRNA by the aminoacyl tRNA synthetase** and **recognition by the enzyme of the appropriate amino acid.** Since tRNA molecules are all basically quite similar, and some amino acids (such as isoleucine and valine) are also similar to one another, **it would not be surprising if mistakes** were made occasionally. The low frequency of such errors (it has been estimated that one protein molecule in a thousand contains one incorrect amino acid) is due to the existence of an **editing mechanism** whereby the **synthetase** is able to cleave the amino acid from an incorrectly charged tRNA molecule.

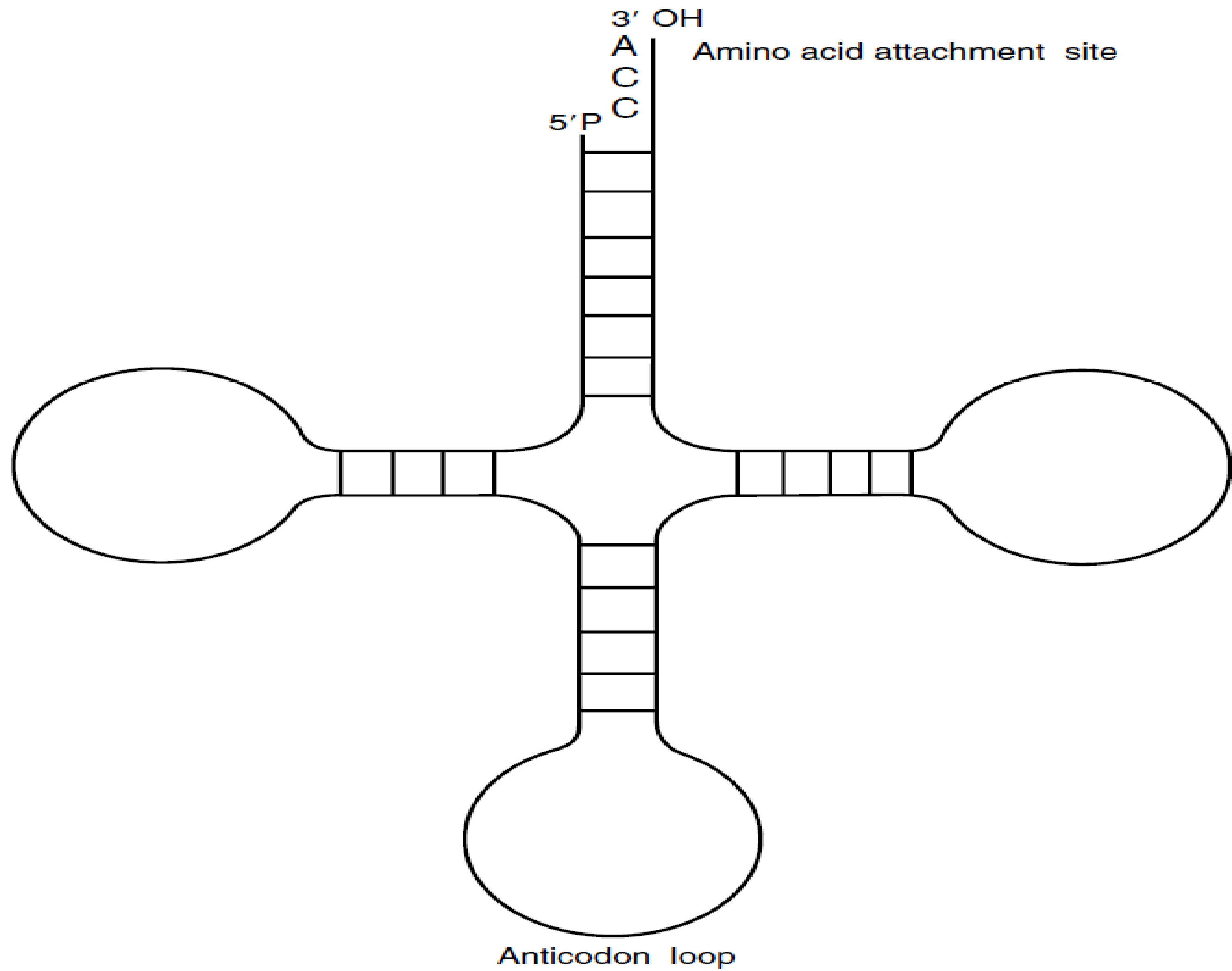

Figure 1.21 Diagrammatic structure of transfer RNA

61 different tRNA species:

one for each of the 64 codons - the three stop codons

For many of the amino acids there are indeed multiple tRNA species with different codon specificities.

Some of these tRNA molecules are present at comparatively low levels in the cell, which would indicate that there could be a difficulty in translating that particular codon. **This can be correlated to some extent with the frequency of occurrence of particular codons** (codon usage): those codons that require a rare tRNA species tend also to occur less commonly, at least in highly expressed genes.

For many tRNA molecules, the codon–anticodon recognition is not absolutely precise; in particular, there is some latitude allowed in the matching of the third base of the codon. A rather complex set of rules (**the wobble hypothesis**) **has been developed to account for the extent of allowable mismatching**. So some tRNA molecules are able to recognize more than one codon. The number of tRNA species required for recognition of the complete set of codons is thus considerably less than 61 (commonly between 30 and 40).

Mechanism of protein synthesis

In bacteria, the initiation codon is recognized by a specific tRNA molecule, **tRNAfMet**. After this tRNA molecule is charged with methionine, the amino acid is modified, to **N-formylmethionine**. Aminoacylated tRNA molecules normally bind to a site on the ribosome known as the A site (Acceptor), while their anticodon region pairs with the mRNA. Only after peptide bond formation is the tRNA able to move to a second site on the ribosome, the P (Peptide) site. The fMet-tRNAfMet (i.e. the tRNAfMet charged with formylmethionine) is unique in being able to enter the P site directly.

The tRNAfMet anticodon recognizes (forms base pairs with) the start codon on the mRNA, in association with binding of a 30S ribosome subunit to the nearby ribosome binding site. The 50S ribosome subunit then joins the complex. The charged tRNA corresponding to the second codon then enters the A site on the ribosome and peptide bond formation occurs by transfer of the fMet residue to the second amino acid. **The tRNAfMet, now uncharged**, is released, and the ribosome moves one codon along the mRNA, which is accompanied by movement of the second tRNA molecule (**now charged with a dipeptide**) from the A site to the P site; this step is known as translocation. The A site is thus free to accept the charged tRNA corresponding to the third codon. The cycle of peptide bond formation, translocation, and binding of a further aminoacylated tRNA requires several additional **non-ribosomal proteins (elongation factors)** and is accompanied by the **hydrolysis of GTP.**

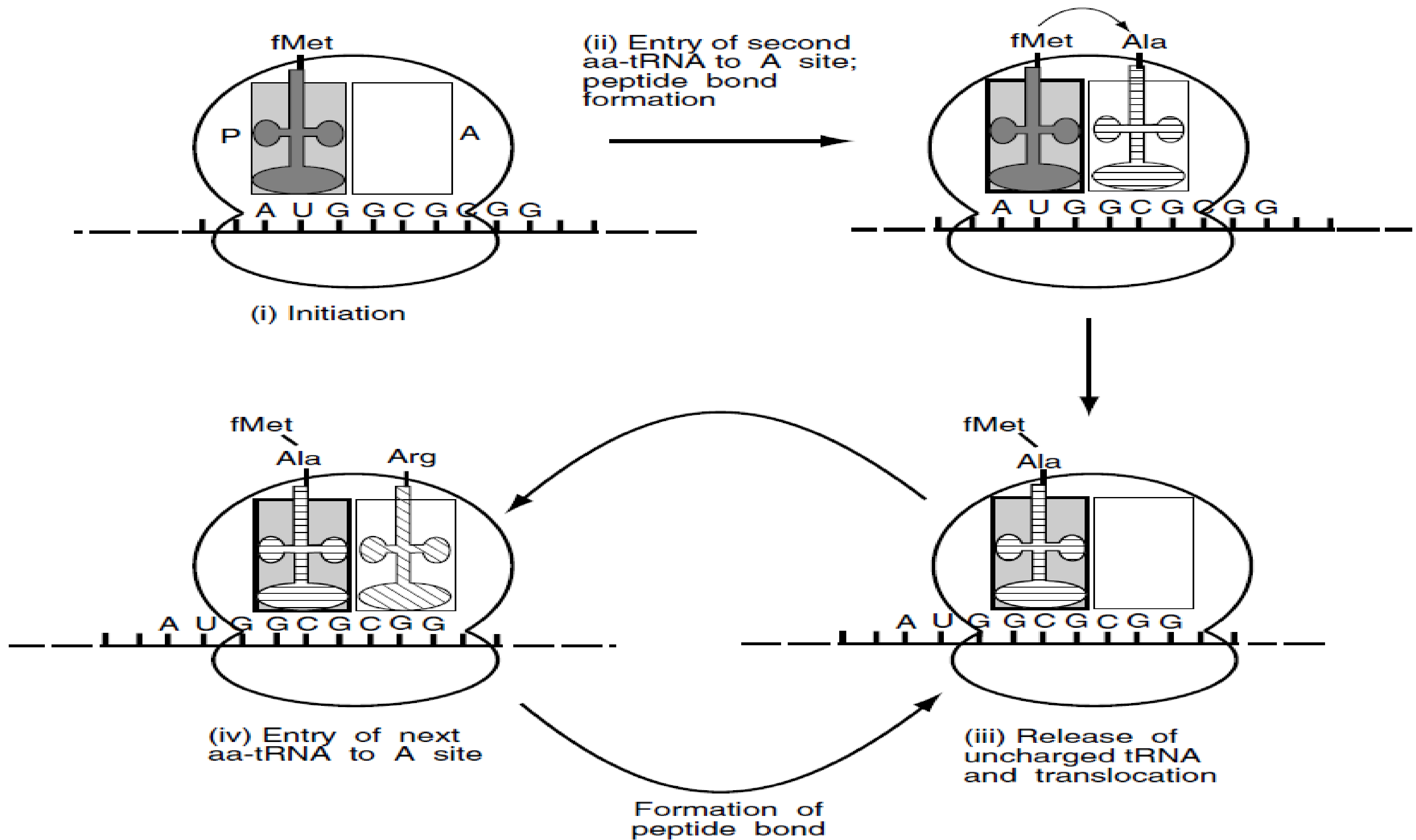

Figure 1.22 Outline of the mechanism of protein synthesis. (i) The initiation complex is formed by the binding of a 30S ribosomal subunit to the mRNA, followed by the initiating mRNA and a 50S subunit. (ii) The second tRNA enters the A site and the first peptide bond is formed by transfer of the N-terminal fMet to the free amino group of the second amino acid. (iii) The uncharged initiating tRNA is released and the ribosome moves along the mRNA. The second tRNA is now in the P site and the A site is free for the next tRNA. (iv) Entry of the next tRNA to the A site, ready for peptide bond formation as in step ii. Steps iii and iv are then repeated, until a stop codon is reached

When the ribosome has moved far enough, the ribosome binding site on the mRNA is exposed again and another ribosome can attach to it. A single mRNA molecule will therefore carry a number of ribosomes actively translating the sequence. Each ribosome moves along the mRNA until a stop codon is reached. The absence of a corresponding tRNA species capable of recognizing this codon causes translation to stop at this point. **The polypeptide chain is then released with the aid of proteins known as release factors and the ribosome dissociates from the mRNA.**

The process needs (Translocation) factors help is elongation factor such as Ef-TS, Ef-TS, Ef-G is the energy you need to bring a new tRNA by an amino acid according to founding in genetic code on the site (A) and thus continue the process of elongation chain (polypeptide) by adding amino acids, respectively, as codes are on the mRNA, possible that the translation is the presence of more than ribosome one at the same time relate to mRNA is the known (polyribosome) or (polysome) is located in the prokaryotic and the eukaryotic cells.

Post-translational events

The formation of a biologically-active product involves several further steps.

First of all, the protein has to fold correctly. There are three conformational levels in addition to the **primary structure** (which is the amino acid sequence itself). The **secondary structure** is the spatial arrangement of successive amino acids which may form regular structures such as α**-helices or β-sheets**, which are stabilized by **non-covalent interactions (such as hydrogen bonds)**. Different regions of the protein will adopt different secondary structures, separated by turns or less defined loops.

The various elements of secondary structure are in turn folded together to form the **tertiary structure.** This conformation is stabilized by non-covalent interactions and also by covalent disulphydryl bridges between cysteine residues.

The tertiary structure may include two or more semi-autonomous regions known as **domains**. Many proteins consist of several (identical or different) polypeptide chains; the way in which these polypeptides are associated constitutes the **quaternary structure.**

Furthermore, the folding of the polypeptide is not entirely spontaneous. Cells contain proteins known as **molecular chaperones** which assist in obtaining the correct conformation of proteins, for example by interacting with the nascent polypeptide to prevent it from adopting an incorrect conformation until the complete protein is produced. Molecular chaperones can also play an important role in the refolding of denatured proteins, which can provide a degree of protection against heat and other stress conditions. **Some of these molecular chaperones are specifically produced under conditions that lead to the accumulation of denatured protein and are known as heat-shock proteins or stress proteins.** A further role of molecular chaperones is concerned with the assembly of polypeptide subunits into multimeric proteins or larger structures; for example, the **assembly of bacteriophage heads may require the action of molecular chaperones.**

Secretion

Many bacterial proteins have functions that require them to be present on the surface of the cell or in the extracellular environment. The first barrier to protein export is the cytoplasmic membrane and the most common mechanism for transport of **proteins across this membrane** is known as **the general secretory pathway (GSP, sometimes called the Sec-dependent pathway**). All proteins that utilize this system have a **specific sequence at the N-terminus** which targets the protein to this pathway and which is cleaved during transport. In Gram-positive bacteria, this mechanism is sufficient for export of proteins to the cell surface or to the surrounding medium. However, Gram-negative bacteria also have an outer membrane and the GSP by itself will deliver proteins, not to the outside of the cell, but to the region known as the periplasm, between the cytoplasmic and outer membranes.

Although Gram-negative bacteria are less prolific than Gram-positive bacteria in secreting proteins, they do have important secretion mechanisms. The most common of these, **the Type II mechanism**, is dependent on **the GSP for transport of proteins to the periplasm and then uses a specific multiprotein complex to transport the protein across the outer membrane**. **Type V secretion systems** are similar, in that they are dependent on the Sec machinery for transport to the periplasm, but the translocated protein has a specialized C-terminal sequence which is able to insert into the outer membrane to form a pore. The N-terminal sequence of the same protein is able to pass through this pore, after which it is cleaved from the C-terminal sequence releasing it into the extracellular environment.

These proteins are often called '**autotransporters**' because they mediate transport through the outer membrane themselves and do not require additional machinery.

Most **Gram-negative bacteria** also possess a number of separate secretion systems which are not dependent on the GSP and which are, therefore, termed **Sec-independent**. **These include Type I, Type III and Type IV secretion pathways.**

The **Type I pathway**, is relatively simple, consisting of just three proteins, and targets proteins with a specific 60-amino acid secretion signal at the carboxy terminus. The Type IV secretory apparatus bears many similarities with the conjugal plasmid transfer systems and can be used by a bacterium to introduce proteins into eukaryotic cells. Perhaps the most remarkable secretion system is the Type III pathway, which has been likened to a molecular syringe and which is used to inject proteins directly into the cytosol of eukaryotic cells.

Secretion is only triggered by direct contact between the bacterium and the host cell. Consequently, this pathway is widely used by pathogens to introduce effector molecules into host cells, thus subverting the normal function of the cell to the benefit of the bacterium.

Other post-translational modifications

In addition to the events described above, proteins may undergo a wide range of additional post-translational modifications, such as glycosylation, biotinylation, addition of lipids and proteolytic cleavage. The full range of these is too complex to be covered here, but the outcome is that the final structure can be influenced very strongly by the nature of the cell itself. Since the post-translational events may be essential in obtaining a product with full biological activity, the difficulty in obtaining accurate post-translational modification can severely affect the outcome of attempts to obtain functional gene expression in heterologous hosts.

Gene organization

In bacteria, genes with related functions are often (but not always) located together in a group known as an **operon** .

An **operon** has a single promoter and is transcribed into a single polycistronic mRNA molecule, which carries the information for several proteins. This group of genes will therefore be coordinately controlled: growth of the cell under the appropriate conditions will induce all the genes in the operon simultaneously.

After the ribosome has translated the first cistron in an operon, it may **dissociate**, in which case translation of the next cistron will require attachment of ribosomes to another binding site adjacent to the initiation codon of the second cistron.

In some cases the start codon for the second gene is very close to the stop codon of the first (in fact the sequences may actually overlap). If this occurs, then after the first polypeptide has been released, the ribosome may start translating again at the nearby start codon **without dissociating.**

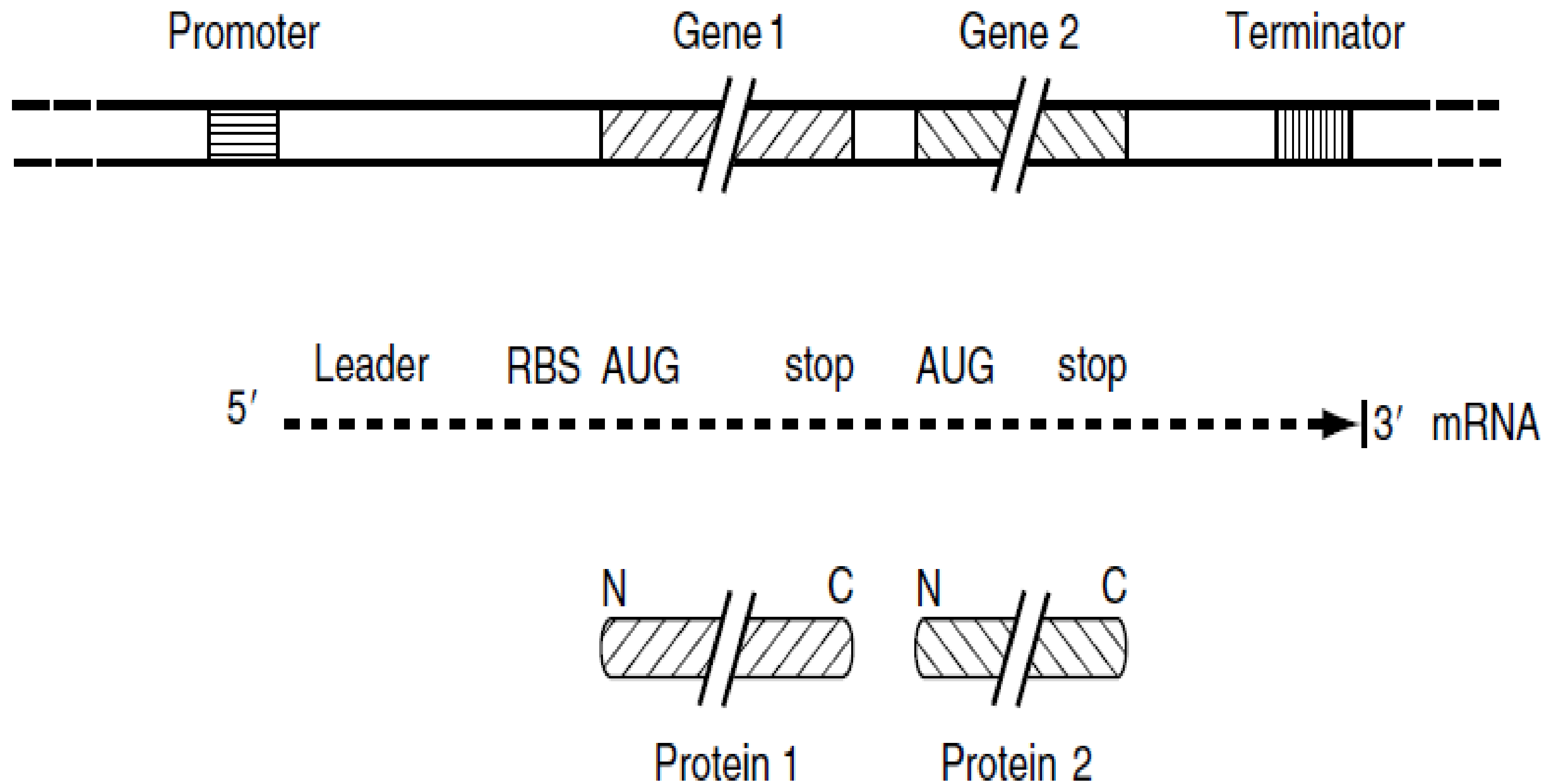

Figure 1.23 Structure and expression of a bacterial operon. A typical operon is transcribed from a single promoter into a polycistronic mRNA from which several independent polypeptides can be translated

Some major differences between bacteria and eukaryotes are

The mRNA in eukaryotes acts as a true 'messenger', being produced within the nucleus and migrating to the cytoplasm for translation to occur**.**

In bacteria, transcription and translation **occur in the same compartment** and the ribosomes will attach to the mRNA as soon as a ribosome binding site is available. So, in bacteria, the mRNA is being actively translated while it is still being made.

Most bacterial mRNA is extremely short-lived; it typically has a half-life of a few minutes only, which may be less than the time required for producing or translating it. This can only be achieved by the coupling of transcription and translation.

A further difference related to the mechanism of ribosome binding is that eukaryotic mRNA (in general) codes for a single polypeptide only. The ribosome in eukaryotic cells attaches to the 5′ end of the mRNA and migrates until it reaches the start codon.

A further difference is that in eukaryotes, the initial product of transcription is a precursor of the mRNA. This precursor, which is found only in the nucleus, contains additional sequences (introns) that are removed by a process known as **splicing or processing**. In some cases the final size of the mRNA is less than 10 per cent of that of the original gene. Generally, **bacterial genes do not contain introns**, but there are a few examples of prokaryotic genes (mainly from bacteriophages) that do contain introns.

Finally, **eukaryotic mRNA** is often (but not always) polyadenylated, i.e. it has a run of adenine residues at the 3′ end. The presence of poly-A tails is often used as the basis of procedures for purifying mRNA from eukaryotic cells.

Bacterial mRNA, on the other hand, is not consistently polyadenylated, although a small proportion of bacterial RNA molecules may carry a short oligo-A tail.

References

Molecular Genetics of Bacteria. By Jeremy W. Dale and Simon F. Park2004 , 4 edition, John Wiley & Sons Ltd.

p.: 26-35

http://www.google.iq/url?sa=t&rct=j&q=&esrc=s&source=web&cd=4&ved=0ahUKEwjFzsnz7rzSAhUC3iwKHRI9BhgQFggqMAM&url=http%3A%2F%2Fshahroodut.ac.ir%2Ffa%2Fdownload.php%3Fid%3D1111127349&usg=AFQjCNH3Q6IyrUwkBEqLYW_wTj1iyJj3QQ&bvm=bv.148747831,d.bGg