12

The translation of mRNA: protein synthesis

12.1 AN OVERVIEW OF PROTEIN BIOSYNTHESIS

In the preceding chapters the reader has already encountered the concept that the mRNA (messenger RNA) is an intermediary in the expression of that portion of the genetic information in the DNA that encodes proteins. The present chapter presents a detailed consideration of the process of *translation* of the mRNA. Although this will be prefaced with a summary of the main features of translation, the reader is directed to suitable textbooks of biochemistry (e.g. [1]) for a more elementary account of this topic.

In essence, protein biosynthesis involves translating the information of the sequence of the four different nucleotides of DNA or RNA into a protein sequence with twenty different possible amino acid units. The genetic code provides the conceptual basis of this in the relationship of single amino acids to groups of three nucleotides in the mRNA (triplet codons); whereas tRNA (transfer RNA) provides its physical basis through possessing an anticodon complementary to a mRNA codon, and a specific covalently attached amino acid. Protein biosynthesis involves the successive reading of the codons of the mRNA by the aminoacyl-tRNAs in an ordered manner, and the linking of the amino acids to form a polypeptide chain. This is a complex process and takes place on an elaborate organelle, the ribosome. The direction of growth of the polypeptide chain is from the *N*-terminus to the *C*-terminus [2], and the direction of reading of the mRNA is 5′ → 3′ [3, 4]. Figure 12.1 provides a schematic summary of the interactions of mRNA, tRNA and ribosome. It represents a stage in protein biosynthesis just before the aminoacyl ester bond of the peptidyl-tRNA is broken and the polypeptide chain transferred to the α-amino group of the aminoacyl-tRNA. Also represented in Fig. 12.1 are two sites on the ribosome to which tRNA can bind, the A-site and the P-site, and the integral ribosomal enzymic activity that catalyses the formation of peptide bonds, *peptidyltransferase*.

The overall length of the mRNA and the rate of initiation are usually such that a second ribosome can attach to the mRNA before the first one has completed its polypeptide chain. In fact, several ribosomes are normally found on a given molecule of mRNA, translating different parts of it

R. L. P. Adams et al., *The Biochemistry of the Nucleic Acids*

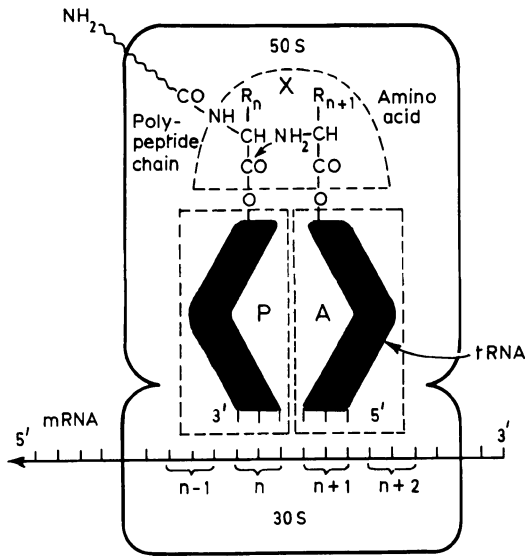

Fig. 12.1 Diagrammatic representation of a prokaryotic ribosome. Two tRNA molecules are bound to the ribosome in response to the mRNA codons designated *n* and *n* + 1. The tRNA bearing the growing polypeptide chain is occupying the peptidyl site (rectangular area, marked P), and the tRNA bearing an amino acid is occupying the aminoacyl site (rectangular area, marked A). The peptidyltransferase centre, where the peptide bond formation is catalysed, is represented by the semicircular area, marked X. Note the exaggeration of the amino acid (1 Å – 0.1 nm) relative to the tRNA (75 Å – 7.5 nm), and the misrepresentation of the shape of the ribosomal subunits (cf. Fig. 12.21).

simultaneously, and such groups of ribosomes are termed *polyribosomes or polysomes* (Fig. 12.2). The size of the polysomes increases with the size of the mRNA: polysomes synthesizing haemoglobin β-chains (M_r = *c.* 16000) contain four to five ribosomes [5], whereas those synthesizing myosin heavy chains (M_r = *c.* 200000) contain about 50–60 ribosomes [6].

The rates of protein synthesis in prokaryotes and eukaryotes appear to be quite similar: values of 15 amino acids s^{-1} for β-galactosidase in *E. coli* [7], and seven amino acids s^{-1} for globin chains in rabbit reticulocytes [8] have been reported. These are, however, much lower than the rates estimated for DNA or RNA synthesis (800 and 50 nucleotides s^{-1}, respectively, in prokaryotes).

It should be mentioned that the synthesis of certain small bacterial peptides occurs in a manner not dependent on mRNA and ribosomes. The reader interested in this subject is directed elsewhere [9].

12.2 THE GENETIC CODE

12.2.1 The standard genetic code

The elucidation of the genetic code represented one of the major breakthroughs in modern biology. Here we shall merely describe the features of the code, as accounts of the history of the code are to be found in a number of reviews (e.g. [10–12]).

The genetic code is a *triplet* code with individual amino acids represented in the mRNA by code words (*codons*) of three nucleotides. Although certain codons also specify initiation and termination signals, the code is uninterrupted, with no 'commas' between codons, and these follow one another in succession and do not overlap. The code was at one time thought to be universal, i.e. each triplet codon had the same meaning, regardless of the species. This assumption derived from comparison between *E. coli* and higher mammals, where initially it was shown that the tRNAs recognized the same codon triplets *in vitro* [13], a result subsequently corroborated by comparison of protein and nucleic acid sequences. It is now known that in certain organisms and in the mitochondria of eukaryotes the genetic code differs from that first established in *E. coli* (section 12.2.3). However, the latter code is so wide-

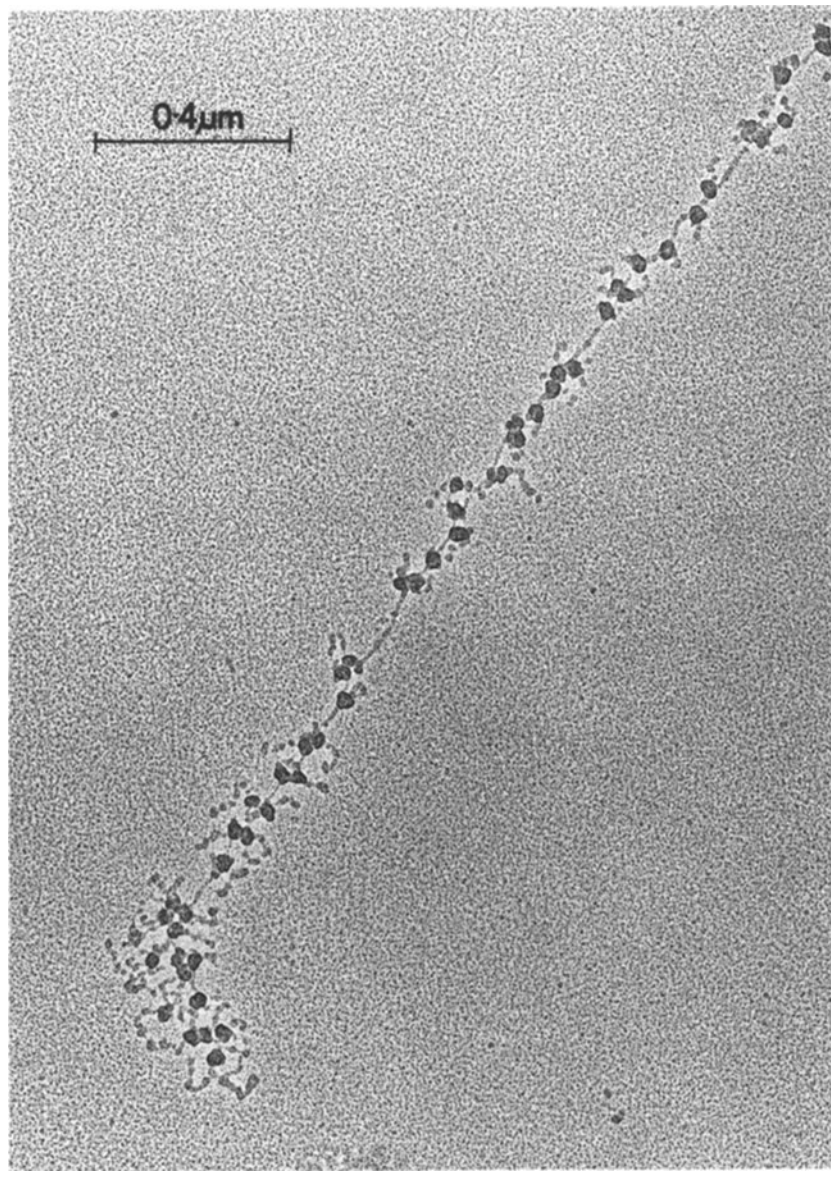

Fig. 12.2 Electron micrograph showing the translation of silk fibroin mRNA on polysomes. The extended fibrous fibroin molecules can be seen emerging from the ribosomes (dark irregular particles). The length of the fibroin molecules increases from the top right to the bottom left of the frame, indicating that this is the 5′ → 3′ direction along the mRNA (courtesy of Dr Steven L. McKnight and Dr Oscar L. Miller, Jr).

spread (it is found in the vast majority of prokaryotes and eukaryotes, animals and plants) that we shall refer to it as the *standard* genetic code. It is presented in Fig. 12.3.

It can be seen from Fig. 12.3 that 20 different amino acids are specified by the genetic code. Other amino acids are found in proteins, but almost all of these are

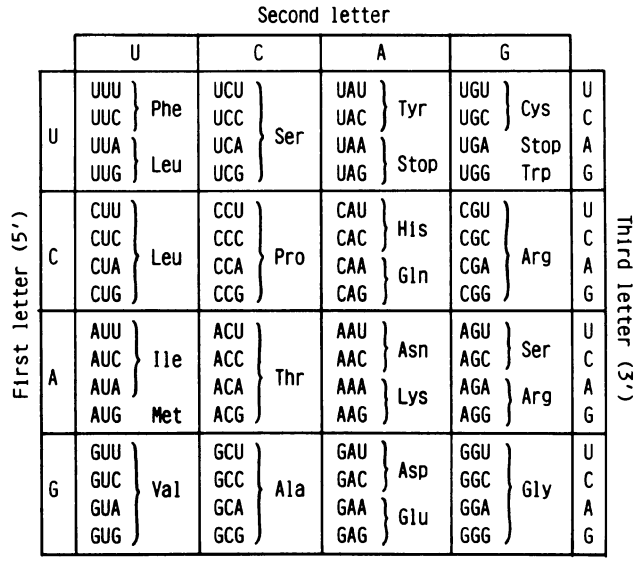

Fig. 12.3 The standard genetic code. Termination codons are indicated as 'Stop'.

generated by post-translational enzymic modification of these 20 genetically-defined amino acids. The one known exception is the rare amino acid, selanocysteine, which is encoded by UGA, otherwise a termination codon (see below). As this is not a complete reassignment of the meaning of the codon, but an alternative translational possibility, thought to depend on the broader context of the mRNA, we shall consider it in section 12.9.6 with other examples of 'suppression' of termination codons.

It is also apparent from Fig. 12.3 that three of the codons (UAA, UAG and UGA, designated 'Stop') do not normally specify an amino acid, but are all signals for the termination of the polypeptide chain. For historical reasons relating to the type of suppressor mutations (see below) that characterized them, UAG and UAA are also referred to as 'amber' and 'ochre', respectively; and UGA is sometimes called 'opal' or 'umber'. Although the termination process does *not* involve tRNA (section 12.4.3), certain mutant tRNAs have been found that can recognize individual termination codons. These tRNAs occur in the 'suppressor' strains of *E. coli*, so called for their ability to suppress particular classes of 'nonsense' mutants. Nucleotide sequence analysis of such a mutant of a $tRNA^{Tyr}$ from an amber suppressor strain showed that its anticodon is changed from 3′-AUG-5′ to 3′-AUC-5′. Thus, it is able to insert tyrosine into a polypeptide chain in response to the termination codon UAG, rather than to the tyrosine codon UAC [14]. The reason that such suppressor mutants are viable, and do not exhibit premature termination of the bulk of normal proteins, is because the mutations occur in the minor, otherwise redundant, representatives of certain pairs of isoaccepting tRNAs.

Although inspection of Fig. 12.3 does not reveal a codon, the sole role of which is to specify the start of a polypeptide chain, the codon AUG fulfills this role as well as that of encoding methionine residues in the body of the polypeptide chain. In *E. coli* the initiation of protein synthesis involves the AUG codon being decoded by a unique species, *N*-formylmethionyl-tRNA (fMet-

tRNA) [15]. The tRNA that inserts the initiating fMet into polypeptide chains, $tRNA_f^{Met}$, has a different nucleotide sequence from the tRNA that inserts Met internally: $tRNA_m^{Met}$ [16, 17]. Both species accept methionine, but only the Met-$tRNA_f$ can then be formylated by a *transformylase* enzyme that has N^{10}-formyltetrahydrofolic acid as a cofactor [18]. Although formylation of Met-$tRNA_f^{Met}$ is normally an absolute requirement for initiation in *E. coli* and many other bacteria, the specific recognition of the initiator tRNA in the initiation process (sections 12.4.1 and 12.5.1) is most certainly influenced by the structure of the tRNA itself. Thus, in eukaryotes the methionyl residue of the initiator tRNA is not formylated, and the initiator tRNA of at least one bacterium can compensate for lack of formylation by undergoing a base substitution [19].

In bacteria AUG, although the main initiation codon, is not the sole one. The usage of GUG and certain other much rarer codons for initiation is discussed in section 12.4.1; but at this point it is important to stress that these minor initiation codons, like AUG, are recognized by fMet-tRNA.

Although fMet or Met initiate polypeptide chains in prokaryotes and eukaryotes, respectively, these amino acids are not necessarily found at the *N*-terminus of the mature protein. In bacteria the formyl group is removed by a deformylase [20] and the methionine group is frequently removed by an aminopeptidase [21], the action of which is governed by the nature of the penultimate residue in the polypeptide chain [22].

12.2.2 The degeneracy of the genetic code

Since many of the 20 amino acids are encoded by more than one triplet (Fig. 12.3), the code is said to be degenerate. Triplets coding for the same amino acid are not distributed at random, but are grouped together so that they generally share the same 5′ and middle base (although there are two separate groups of codons in the cases of Ser and Leu). This has the consequence that mutations producing a change in the base at the 3′-position of the codon often have no effect on the amino acid specified. Furthermore, the different amino acids are segregated to a considerable extent on the basis of chemical similarity (hydrophobicity, hydrophilicity, acidity and basicity). Thus a mutation in the 5′-base of any of the six leucine codons would give a codon specifying another hydrophobic amino acid. Such a change might not impair the function of a particular globular protein if the altered amino acid merely performed a structural role in the hydrophobic core of this protein. It has therefore been argued that the specific arrangement of codons in the genetic code serves to reduce the potentially harmful effect of mutations.

It might have been expected that each degenerate codon would require its own tRNA with a corresponding anticodon. Although there are some such discrete *isoaccepting* tRNAs, recognizing the same amino acid, their number is less than the 61 required if all codon–anticodon interactions involved three standard Watson–Crick base pairs. This situation is viable because there is a degree of latitude or *wobble* in the complementary base-pairing between the base in the 3′-position of the codon and that in the 5′-position of the anticodon. Rules governing the extent of this latitude were proposed by Crick in his *wobble hypothesis* [23]. He considered the stereochemistry of possible non-standard base-pairing that might allow the two pairs of 3′-codon bases that are grouped together in degenerate codons, U and C, and A and G, to be recognized by a common 5′-anticodon base. He found that relatively small movements from the

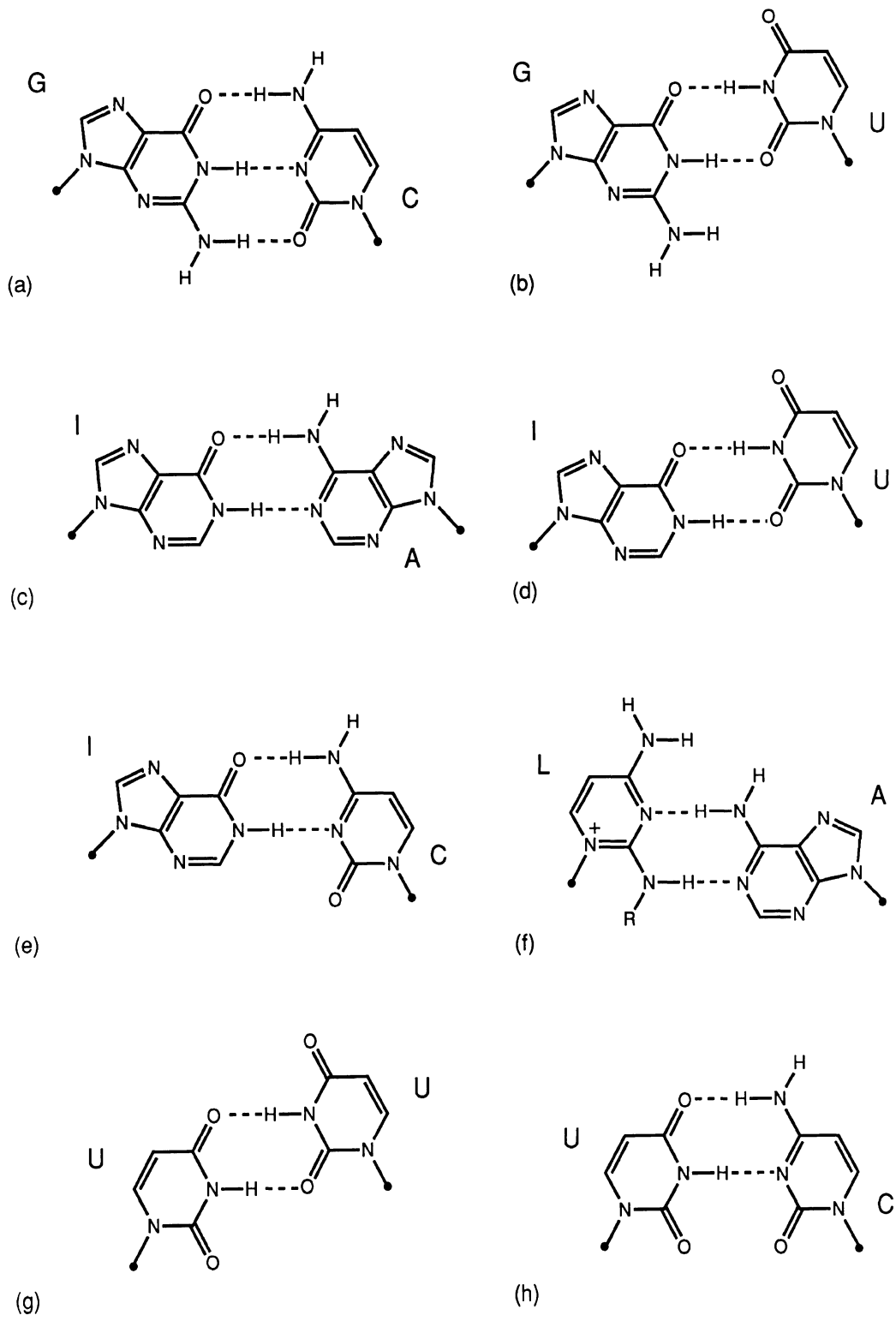

Fig. 12.4 Some proposed patterns of hydrogen bonding involving nucleosides in the 'wobble' position of the anticodon of tRNA. In each case the 5′-nucleoside of the tRNA anticodon is shown on the left, and that of the 3′-nucleoside of the mRNA codon on the right, with the ribose represented as •. (a) and (b) 'wobble' for guanosine; (c), (d) and (e) 'wobble' for inosine; (f) proposed base-pairing for lysidine, R = $-CH_2CH_2CH_2CH_2CH(NH_3^+)COO^-$; (g) and (h) possible interactions of uridine with pyrimidines in mitochondrial tRNAs.

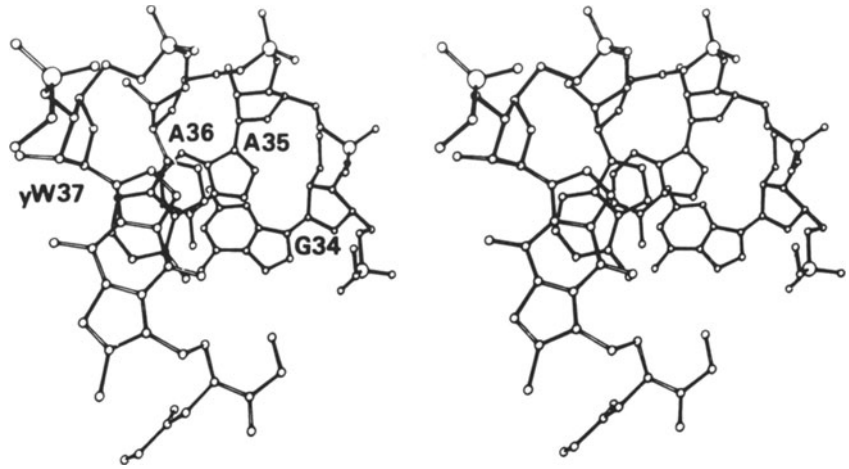

Fig. 12.5 Stereoscopic view of the anticodon bases 34–36 of yeast $tRNA^{Phe}$ and the 3′-hypermodified base yW37 as viewed from the interior of the molecule. It can be seen that the stacking between residues yW37, A36 and A35 is greater than that between A35 and the 'wobble' base G34. (From [49], with permission.)

standard positions might allow the 5′-anticodon base, G, to pair with *both* pyrimidines, U and C, in the 3′-position of the codon; and allow the 5′-anticodon base, U, to pair with *both* purines, A or G, in the 3′-position of the codon. Moreover it had already been found that several tRNA anticodons contained the nucleoside inosine (which pairs only with C in a double helix), and he suggested that this might be able to pair with A, U or C in the 3′-position of the mRNA codon. The hydrogen bonding patterns of these wobble interactions are shown in Fig. 12.4. The elucidation of the three-dimensional structure of tRNA (section 12.3) helps explain how this wobble occurs. It is possible to accommodate a small conformational change into the tertiary structure of the anticodon of yeast $tRNA^{Phe}$, such that the 5′-anticodon wobble base Gm (the hydrogen bonding pattern of which is identical to G) could move into an alternative position that would allow pairing with U rather than C. The reason that such a change is confined to the 5′-anticodon base seems to lie in the fact that the base immediately 3′ to the anticodon, which is always a purine (Fig. 12.7) and frequently heavily modified, has strong base-stacking interactions with the two adjacent bases of the anticodon (Fig. 12.5). This serves to anchor these bases in a helical conformation that only allows standard Watson–Crick base pairs between the anticodon and the two first codon positions.

Although the basic concept of the wobble hypothesis has been validated, it is important to realize that not all its detailed predictions have turned out to be correct. Furthermore, one must consider the rules for the base-pairing of important modified 5′-nucleosides of tRNA which were unknown at the time of the formulation of the wobble hypothesis. Table 12.1 presents a comparison of Crick's proposals with the observed pattern of codon–anticodon interaction in prokaryotes and the cytoplasm of eukaryotes, on the one hand, and mitochondria (which show a unique pattern of interaction), on the other hand. The wobble predictions that

Table 12.1 Predicted and observed 'wobble' base-pairing

5'-Anticodon nucleoside	*3'-Codon nucleoside*		
	Crick's prediction	*Prokaryotes, and eukaryotic cytoplasm*	*Non-plant mitochondria*
Unmodified			
C	G	G	G*
G	C, U	C, U	C, U
A	U	n.f.†	n.f.
U	A, G	n.f.‡	A, U, G, C
Modified			
(C) m^5C, Cm, ac^4C	–	G	n.f.
L	–	A	n.f.
(G) Gm, Q	–	C	C, U§
(A) I	U, C, A	U, C, A	n.f.
(U) mcm^5U, mcm^5s^2U, mnm^5s^2U	–	A	n.f.
mo^5U, cmo^5U, $mcmo^5U$	–	A, G, U	n.f.
$cmnm^5U$	–	n.f.	A, U

–: No prediction was made for the modified 5'-nucleoside (unknown at that time).
n.f.: The 5'-nucleoside is not found in tRNAs of that category.
* Except for that occurring in $tRNA_m^{Met}$, which decodes both A and G.
† Except for isolated examples in which the prediction in probably incorrect (see text).
‡ Except for isolated examples in which all four 3'-nucleosides are decoded (see text).
§ With Q in the 5'-anticodon position of tRNA.

have been substantiated are the ability of the 5'-anticodon nucleoside G to recognize 3'-codon nucleosides U and C (but not A or G), and the 5'-anticodon nucleoside I to recognize 3'-codon nucleosides U, C and A (but not G). Furthermore, with one exception (considered below), 5'-anticodon nucleoside C only recognizes 3'-codon nucleoside G. The problems arise with the wobble predictions for the recognition properties of 5'-anticodon nucleosides A and U, neither of which occur in an unmodified form in the vast majority of prokaryotic or eukaryotic cytoplasmic tRNAs.

The almost complete absence of A from the 5'-position of anticodons has been rationalized on the basis that it is more economical to use G, as all codons with U in the wobble position are in the same family as a codon with C in that position. Likewise in three- or four-codon families, an A in the initial transcript is replaced by an I in the mature tRNA, which can decode three codons. In the rare cases in which A does occur in the wobble position of the anticodon (a $tRNA^{Thr}$ of mycoplasma and a $tRNA^{Arg}$ of yeast mitochondria are examples) the pattern of base-pairing is unclear, however. For example, the $tRNA^{Arg}$ of yeast mitochondria with anticodon 3'-GCA-5' would seem to have to decode the Arg codons in the 5'-CGN-3' family, all of which are represented in the yeast mitochondrial genome. Thus, a suspicion existed that there was some more fundamental reason for the rarity of an A in the wobble position of

tRNA anticodons. This has been confirmed by experiments in which a 3′-CCA-5′ anticodon was engineered into an *E. coli* $tRNA^{Gly}$, which did not then show the predicted discrimination against glycine codons other than 5′-GGU-3′ [24]. It therefore appears likely that the reason that A is avoided in the wobble position of anticodons is that, contrary to prediction, it would wobble.

The situation with U and the 5′-position of anticodons is somewhat similar. In prokaryotes and the cytoplasm of eukaryotes U in the anticodon position of tRNAs is almost always modified. The reason for this would seem to be that the wobble that could occur with unmodified U would exceed the limitation predicted by the hypothesis. In two-codon amino acid families this would cause violation of the genetic code; in four-codon families the extended wobble is generally not efficient enough to be utilized. The basis for this viewpoint is as follows. Unmodified U does occur in mitochondria (and also in mycoplasma [25]), where it can decode all four 3′-codon bases. In mitochondria this has the important consequence that fewer tRNAs are needed to decode four-codon families (section 12.7.2). When the 3′-CCC-5′ anticodon of the aforementioned *E. coli* $tRNA^{Gly}$ was mutated to 3′-CCU-5′, the unmodified wobble U was able to decode Gly codon GGA, but was much less efficient than the mycoplasma $tRNA^{Gly}$ with the same anticodon in translating the codons GGG, GGU or GGC [26]. Thus, other structural features of the tRNA besides the anticodon are required to allow unmodified U to decode effectively all four bases in the third position of the codon. Even the modified derivatives of U employed by the majority of prokaryotes and the eukaryotic cytoplasm do not decode according to the wobble rules: mcm^5s^2U, mcm^5U and mnm^5s^2U decode only A; whereas mo^5U and cmo^5U decode all three of A, G and U (these abbreviations are defined at the front of the book). Structural explanations for the extension or restriction of wobble recognition with some of these modified U residues have been proposed in terms of the adoption of the flexible C3′-endo or the rigid C2′-endo conformation (Fig. 2.16) of the ribose [27, 28]. It is interesting, however, that mitochondria have evolved tRNAs with a derivative of U, $cmnm^5U$, that allow them to decode codon pairs, NMA and NMG, according to the wobble rules, hence minimizing their complement of tRNAs [29]. Because this modified base has not been found in the wobble position of non-mitochondrial tRNAs one suspects that there are other features in the structure of the mitochondrial tRNAs that contribute to the pattern of decoding.

Similar extra-anticodon structural features must be invoked to explain the ability of the C in the 3′-UAC-5′ anticodon of yeast mitochondrial $tRNA_m^{Met}$ to recognize both AUA and AUG [30]. One possible structure for the C:A base-pair that is presumed to form is that found in crystals of an artificial deoxyoligonucleotide [31].

The essence of the wobble hypothesis is the possibility of alternative base-pairing interactions involving the 5′-base of the anticodon. It is evident from the foregoing discussion that a sound structural basis exists for the standard wobble base-pairings of unmodified and modified anticodon bases. What, however, of the decoding of four-codon families described above for the unmodified U found in some mitochondrial and mycoplasma tRNAs? One possibility is that the unique structure of these tRNAs allow a relaxed pattern of base-pairing in which, by rotation of the C4′–C5′ and P–O bonds, U:U and U:C become close enough together to form satisfactory hydrogen bonds (Fig. 12.4) [32]. However, an alter-

native possibility is that there is no interaction in the third position in this situation, and that decoding is by the 'two-out-of-three' mechanism proposed by Lagerkvist [33]. This 'two-out-of-three' hypothesis was actually developed before the situation in mitochondrial tRNAs became apparent, and proposes that in certain circumstances a single tRNA that can recognize the first two codons of a four-codon family has the potential to decode the whole of such a family. The four-codon families for which such 'two-out-of-three' recognition was proposed are UCN, CUN, CCN, CGN, ACN, GUN, GCN and GGN (Fig. 12.3). All of these involve at least one strong G:C base pair in the first two positions. Some experimental evidence to support this hypothesis was obtained from the successful translation of phage MS2 RNA in a cell-free system from *E. coli* in which the individual tRNAs that can normally translate particular codons were omitted. These could be replaced by tRNAs that would be predicted to translate them according to the 'two-out-of-three' hypothesis, but not according to the 'wobble' hypothesis [34]. The efficiency of this process was such that it would be unable to compete with the normal tRNA that can make a 'three-out-of-three' interaction, and would, therefore, be most unlikely to participate in contemporary codon–anticodon interactions in prokaryotes (or eukaryotic cytoplasm). Nevertheless it cannot be excluded that in the peculiar environment of the mitochondrion similar interactions might be involved in the decoding of four-codon families by anticodons containing unmodified U in the wobble position.

12.2.3 Alternative genetic codes

There are two types of divergence from the standard genetic code. Outside mitochondria the most widespread of these is the reassignment of one or two of the three termination codons to an amino acid. Thus, in some ciliated protozoa and algae UAA and UAG code for glutamine rather than acting as chain-termination codons [35], whereas in several species of the prokaryote, *Mycoplasma*, these codons are chain terminators but UGA codes for tryptophan [36]. In *Euplotes octocarinatus*, another ciliate, UGA codes for cysteine, with UAA functioning as a termination codon and UAG apparently not being employed in any capacity [310]. Species of tRNA with appropriate anticodons exist in these species to decode these erstwhile termination codons.

How can these changes be accounted for in evolutionary terms? The first stage is thought to have been the loss of the particular termination codon or codons from the repertoire in use. This could have been facilitated by a small coding capacity in the case of mitochondria, and by pressure against usage of G in the extremely AT-rich genome of *Mycoplasma*. Next, the formation of a new tRNA or a mutation in the anticodon of a tRNA would have been required to allow recognition of the codon as an amino acid when it re-emerged in the mRNA. For example, the mutation of the tryptophan anticodon from 3′-ACC-5′ to 3′-ACU-5′ (with appropriate post-transcriptional base modification) would have allowed it to recognize the codon UGA as well as UGG (Table 12.1). Such a change would have had to be accompanied by a loss of, or alteration in specificity of, the termination factor that had previously recognized the codon. In *Mycoplasma* the loss of RF-2 (section 12.4.3) could have accomplished this.

The second type of divergence from the standard genetic code involves changes in the assignment of codons specifying amino acids. Most of the examples are encountered

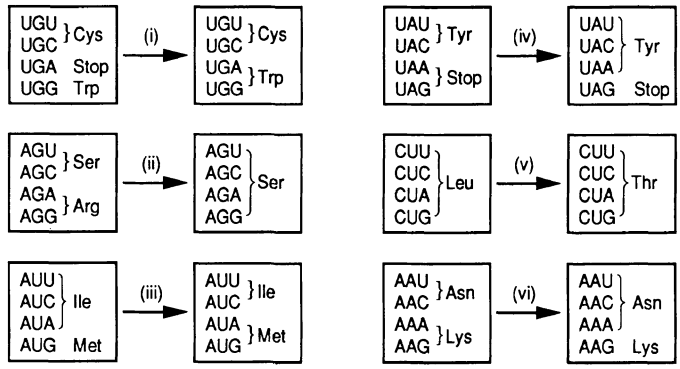

Fig. 12.6 Some deviations from the standard genetic code found in non-plant mitochondria. The figure shows the changes in relation to the other members of the four codon family involved. The changes indicated are found in: (i) all non-plant mitochondria examined; (ii) echinoderms, molluscs, nematodes and platyhelminths; (iii) vertebrates, arthropods, molluscs, nematodes and yeasts; (iv) platyhelminths; (v) yeasts; (vi) echinoderms and platyhelminths. Details of other changes can be found in [37].

in non-plant mitochondria, which will be the focus of consideration, but a non-mitochondrial example is known. In the fungus, *Candida cylindracea*, CUG encodes serine rather than leucine [35]. Some of the changes found in non-plant mitochondrial genetic codes [37] are illustrated in Fig. 12.6. It should be stressed that there is no single mitochondrial genetic code, implying that different mitochondrial genomes have undergone separate evolution. Indeed, plant mitochondria (and chloroplasts) employ the standard genetic code. Examination of the nature of the changes suggest that although some appear quite gratuitous (e.g. Leu to Thr for CUN in *S. cerevisiae*), others can be regarded as simplifying the code into four-member or two-member families (e.g. change from Arg for AGR). This facilitates the decoding of the mitochondrial genome by a smaller number of tRNAs according to the special rules already described in section 12.2.2.

Such changes in the value of individual codons from one amino acid to another can also be envisaged in terms of temporary abandonment of use (all the examples are degenerate, i.e. other codons for the same amino acid remain) followed by reassignment. In the case of a gross change, such as that in which all the codons in yeast mitochondria of the type CUN changed from Leu to Thr, all that would be required would be for a mutation (outwith the anticodon) in the single tRNA that decodes this family to cause a change in aminoacylation specificity (section 12.3.2). In the case of the conversion of the AAA codon from Lys to Asn in echinoderm mitochondria, mutation of the Asn anticodon from 3′-UUG-5′ (which can only decode AAC or AAU) to 3′-UUI-5′ could have accomplished this if accompanied by the loss of the $tRNA^{Lys}$ uniquely decoding AAA. In other cases there has been no change in the anticodon, but a change in the decoding capacity of the tRNA, presumably as a result of mutation elsewhere in the molecule. Thus, the Met-tRNA anticodon 3′-UAC-5′ normally decodes only AUG, but in yeast mitochon-

dria it has acquired the capacity to decode AUA as well (section 12.2.2).

12.2.4 Differential codon usage

Because the genetic code is degenerate, there exist, in principle, alternative choices between 'synonymous' codons for most of the amino acids, for the termination signal, and even for the initiation signal. The usage of alternative initiation (section 12.4.1) and termination codons (section 12.4.3) is discussed later: the present section considers the pattern of choice of amino acid codons and its implications.

The usage of such synonymous amino acid codons is definitely non-random, and differs between prokaryotes and eukaryotes, viruses and their hosts. In some cases, e.g. certain eukaryotic viruses and mycoplasma, codon choice seems to be determined by the pressure to a particular extreme base composition. However, in bacteria (and to a certain extent in yeast) there is evidence that the reason for non-random codon usage is that it can influence the rate of translation. In *E. coli* it was observed that certain codons that were rarely used in the genes for highly expressed proteins (e.g. ribosomal proteins) were found at a significantly greater frequency in the genes for weakly expressed proteins (e.g. the *lac* and other repressors) [38]. Circumstances where such translational, rather than transcriptional, regulation is important are nicely illustrated by the case of DNA primase. The gene for this protein is in the same transcription unit as those for ribosomal protein S21 and the σ factor of RNA polymerase, proteins that are expressed under similar circumstances but approximately 800 and 50 times more strongly, respectively. The difference in codon usage between this gene and its cotranscribed neighbours is most striking [39].

There are two main ideas as to how the pattern of codon usage might regulate translation. The first suggestion, for which the indirect evidence is strongest, is that the speed of translation of synonymous codons varies with the abundance of the corresponding isoaccepting tRNAs (section 12.2.2). A study of the iso-accepting tRNAs in *E. coli* showed a strong correlation between tRNA abundance and codon choice in the genes of highly expressed proteins [40]. A similar correlation between codon usage in highly expressed genes and the relative abundance of iso-accepting tRNAs was found in yeast [41] and for the fibroin mRNA of the silkworm, *B. mori* [42]. Nevertheless, in *E. coli* there is also a difference in usage of codons of the type NMU and NMC, which cannot be explained in terms of abundance of isoaccepting tRNAs as they are both decoded by a tRNA with G or I in the 'wobble' position (Table 12.1). It was observed in such pairs that the usage correlated well with the predicted codon–anticodon interaction energy, those with either maximum (e.g. CGC) or minimum (e.g. AUU) energies being less frequent in the genes for highly expressed proteins than those (e.g. CGU or AUC) with intermediate energies [38]. It was argued that intermediate codon–anticodon interaction energies would produce more efficient decoding by allowing the optimal balance between binding and release of tRNA. Although this latter idea has not found universal acceptance, it is striking that a similar pattern of preference in NMY codons has been reported in yeast [41].

Such correlations between codon usage and gene expression, although highly suggestive, cannot be taken as proof of a causal effect of codon usage on translation. However, strong experimental support has been obtained in both bacteria and yeast. One elegant experiment involved site-directed mutagenesis of

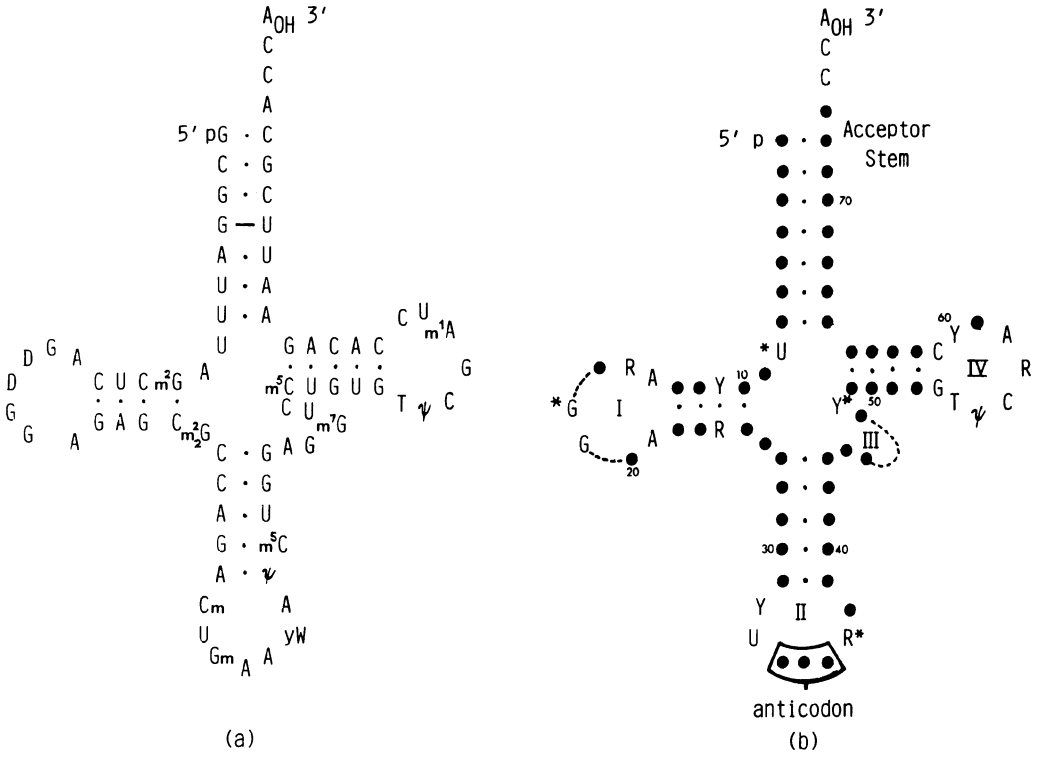

Fig. 12.7 Secondary structure of tRNA. (a) Yeast $tRNA^{Phe}$; (b) generalized structure in which only the invariant or semi-invariant bases are named, the others being represented by filled circles. An asterisk indicates that a base may be modified. Hydrogen bonds in the helical stems are indicated by dots, except for the G:U pair in (a) which is indicated by a line for emphasis. The dotted lines indicate a variable number of nucleotides. R and Y stand for purine and pyrimidine, respectively, other symbols being defined at the beginning of the book. An explanation for the numbering of the variable loops can be found in [45].

rare Leu codons to more highly used codons in the attenuating leader of the leucine operon of *Salmonella typhimurium* (cf. section 10.2.1), and this was found to have a profound effect on the attenuation [43]. In yeast the highly expressed phosphoglycerate kinase gene was mutated so that up to 39% of major codons were replaced by minor ones, and this was found to cause a dramatic decline in the extent of expression [44].

In concluding this discussion of the usage of amino acid codons it should be mentioned that in the small genomes of mitochondria, some codons are completely unrepresented in the repertoire of mRNAs. As mentioned in section 12.2.3, this is one of the prerequisites for the evolution of variant genetic codes.

12.3 THE STRUCTURE AND AMINOACYLATION OF tRNA

12.3.1 The structure of tRNA [27, 45]

The key role of transfer RNA in decoding the genetic information has already been mentioned. To reiterate, each tRNA can covalently attach one specific amino acid to its 3′-end, and each possesses a sequence of

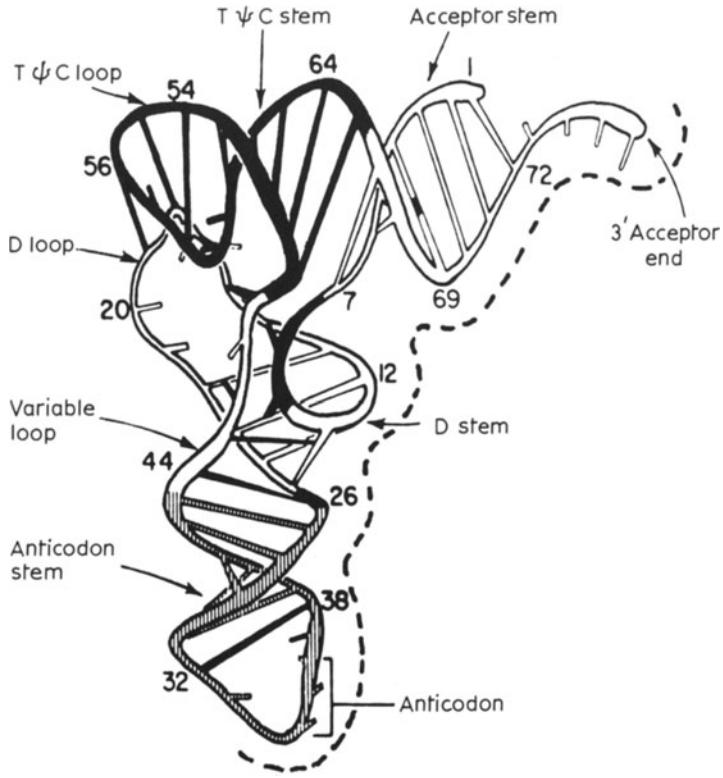

Fig. 12.8 Tertiary structure of yeast $tRNA^{Phe}$. The sugar–phosphate backbone is shown as a coiled tube, the numbers referring to the nucleotide residues starting from the 5′-end. Hydrogen bonding interactions between bases are shown as cross-rungs, tertiary interactions being shaded solid black. Bases not involved in hydrogen bonding are shown as shortened rods attached to the backbone (from [49], with permission).

three bases, the anticodon, complementary to and able to interact by hydrogen-bonded base pairing with a mRNA codon for this, so-called, cognate amino acid.

The tRNAs range in length from 73 to 94 nucleotides and are characterized by a relatively large proportion of modified or non-standard nucleosides (sections 2.1.2 and 11.7.2). Almost all these primary structures (mitochondrial tRNAs are the exception – section 12.7.2) allow themselves to be arranged into a common secondary structure, specific and generalized examples of which are shown in Fig. 12.7(a) and (b), respectively. The common features of these structures are as follows.

1. A stem, the acceptor or aminoacyl stem, containing the 5′ and 3′-extremities of the molecule. It consists of a helix of seven pairs of bases generally making Watson–Crick base pairs (i.e. A:U or G:C) but occasionally (e.g. Fig. 12.7(a)) with a G:U base pair (cf. Fig. 12.4), together with an unpaired sequence of four bases at the end of the 3′-strand of the stem. The last three bases are always CCA and it is to the 3′-terminal adenosine residue that the amino acid is attached.
2. An arm, the dihydrouridine (D) arm, comprising a stem of three or four base pairs together with the D loop (loop I) of eight to eleven nucleotides, some of

which are invariant. Although loop I generally contains one or more dihydrouridine residues (hence one of its common names), several examples are known in which this base is absent.

3. An arm, the anticodon arm, comprising a helical stem of five base pairs together with the anticodon loop (loop II) of seven nucleotides. It is worth pointing out that the base 5′ to the centrally placed anticodon is always U, and that two other bases in the loop are semi-invariant (i.e. restricted to being purines in one case and pyrimidines in the other). The 5′-base of the anticodon, the base in the 'wobble' position, is frequently a modified or non-standard base.
4. An extra arm (III) of extreme variability, ranging from four to 21 nucleotides. This may be either a helical stem together with a loop of three or four nucleotides (13–21 nucleotides in all), or merely a loop of three to five nucleotides.
5. An arm, the TΨC arm, comprising a helical stem of five base pairs and a loop (loop IV) of seven nucleotides.

To date, X-ray crystal structures have been determined for four tRNAs uncomplexed to any protein [27, 45–48]. The structure of yeast $tRNA^{Phe}$, for which greatest resolution has been obtained, is shown in Fig. 12.8. One way of regarding it is as having an L-shape, with two angularly oriented domains: one comprising the TΨC stem and acceptor stem with the 3′-terminal CCA at its furthest extremity; and the other comprising the D-stem, the variable loop (III) and the anticodon arm, with the anticodon loop (II) at its furthest extremity. The manner in which the individual helical stems augment one another to form these two domains of extensive base stacking is perhaps the most fundamentally important feature of the structure. The corner made by the junction of the two domains comprises the D-loop (I) and the TΨC-loop (IV). The dimensions of the domains are similar: approximately 60 Å (6 nm) from the TΨC-loop at the corner to either anticodon or 3′-acceptor end; whereas the distance between these extremities is approximately 75 Å (7.5 nm).

The inner surface of the two domains (broken line in Fig. 12.8) rather belies the description, L-shaped, that applies to the outer surface. It comprises a more or less planar region that in two dimensions can be

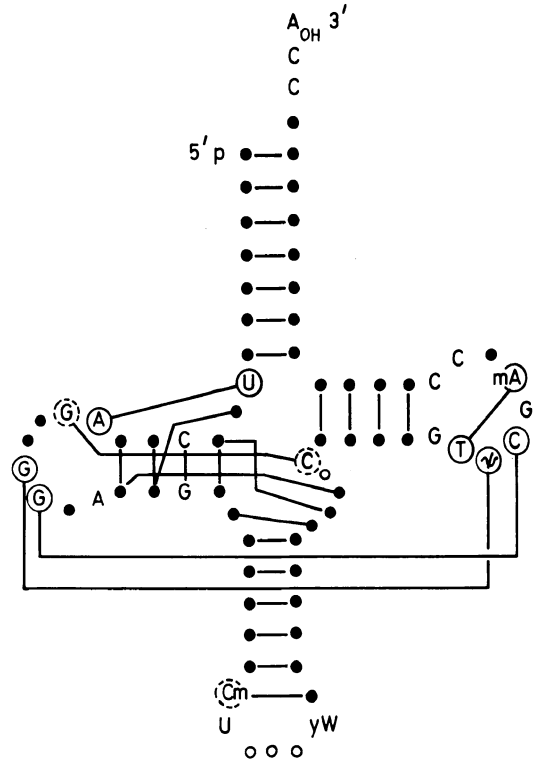

Fig. 12.9 Relationship of tertiary hydrogen bonding interactions to the cloverleaf secondary structure of yeast $tRNA^{Phe}$. The generalized convention of Fig. 12.7(b) has been applied to the structure in Fig. 12.7(a) to highlight the involvement of invariant or semi-invariant bases in the tertiary interactions (extended lines). (After [49], with permission.)

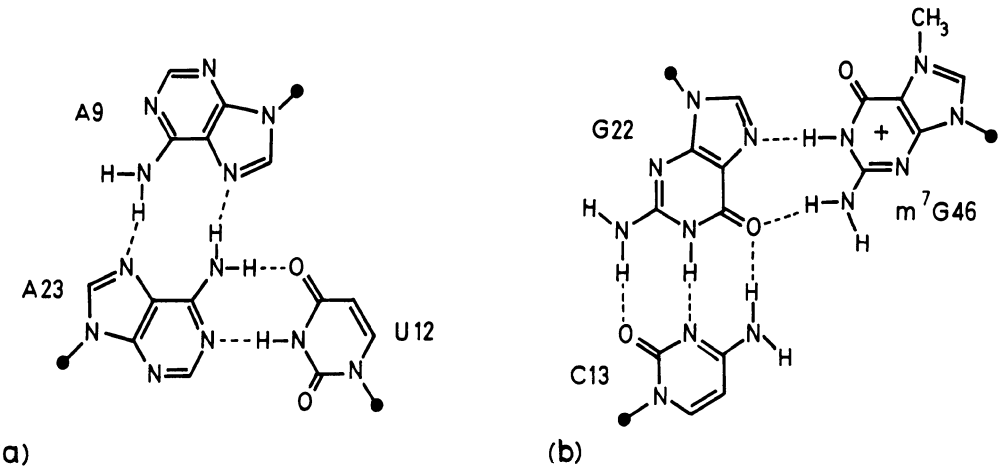

Fig. 12.10 Examples of hydrogen-bonding in yeast $tRNA^{Phe}$ that involve three bases. The numbering is as in Fig. 12.7(b).

regarded as the top of a T, the base of which (at about nucleotide 56) is equivalent to the angle of the L. This inner surface contains many of the variable bases of the D-stem and variable loop (III), which, furthermore, are not involved in any of the tertiary interactions discussed below. They are thus available for interaction with other molecules, particularly the aminoacyl-tRNA synthetase (section 12.3.2).

The tertiary structure of yeast $tRNA^{Phe}$ is maintained by a large number of hydrophobic stacking interactions in the augmented helices, together with additional specific hydrogen-bonding interactions between nucleotide residues that are often widely separated in the secondary structure cloverleaf (Fig. 12.9). Many of these interactions are between invariant or semi-variant bases, suggesting a rationale for their restricted variance and a generality for the yeast $tRNA^{Phe}$ tertiary structure. These hydrogen-bonding interactions are in no way confined to Watson–Crick base pairs, but include a variety of non-standard interactions, some involving three bases (Fig. 12.10). Indeed, some of the interactions involve the sugar-phosphate backbone. Although these hydrogen bonds may at first sight seem esoteric, their occurrence in tRNAs merely reflects the fact that the bases involved are not confined to the relative spatial orientations which they are forced to adopt in a standard RNA A-helix. The potential for the 2′-OH of the ribose to participate in hydrogen bonding may be the basis of the ability of RNA molecules to form such non-helical structures, not possible in the case of DNA.

The general base-stacked two-domain structure stabilized by tertiary interactions described above for yeast $tRNA^{Phe}$ also obtains in yeast $tRNA^{Asp}$ [48], and in yeast and *E. coli* $tRNA_f^{Met}$ (section 12.2.1), suggesting that it is a general feature of tRNAs. (There are certain differences in the conformation of the anticodon loop in the initiator tRNAs which are discussed in section 12.4.1.) There is also evidence indicating that the crystal structure corresponds to the structure adopted in solution [25]. However, as described in the next section, changes in this structure may occur on interaction with proteins.

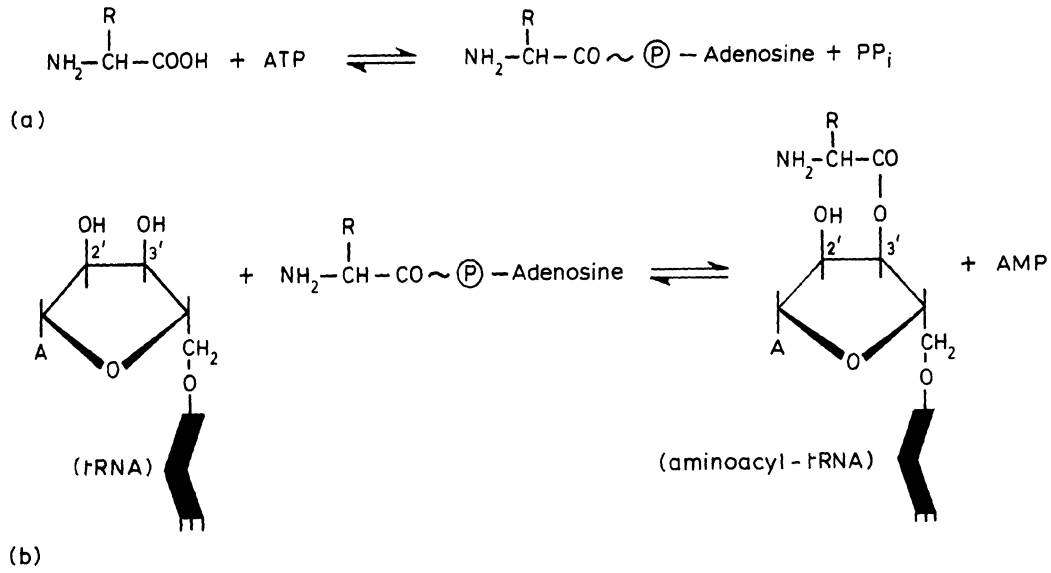

Fig. 12.11 The activation of amino acids and their attachment to tRNA. The reaction occurs in two stages, (a) and (b), both of which are catalysed by the same enzyme, and to which the intermediate is bound (see text). The symbol ~ represents a bond with a relatively high standard free energy of hydrolysis.

12.3.2 The aminoacylation of tRNA

Before a tRNA molecule can act as an adaptor by interacting with its corresponding anticodon in the decoding process, it must first be 'charged' with its cognate amino acid. The enzymes responsible for this process are called *aminoacyl-tRNA synthetases* (*amino acid-tRNA ligases* EC 6.1.1.) and catalyse the reaction illustrated in Fig. 12.11.

The reaction occurs in two stages [50, 51], in the first of which the amino acid is activated by ATP to form an aminoacyl adenylate:

$$\text{ATP} + \text{amino acid}_1 + E_1 \rightleftharpoons$$
$$(\text{amino acid}_1\text{-AMP})E_1 + PP_i$$

The aminoacyl-adenylate complex then reacts with the terminal adenosine moiety of the appropriate tRNA to form an aminoacyl-tRNA:

$$(\text{amino acid}_1\text{-AMP})E_1 + \text{tRNA}_1 \rightleftharpoons$$
$$\text{tRNA}_1\text{-amino acid}_1 + \text{AMP} + E_1$$

Although the amino acid is located at the 3′-OH of the ribose of the terminal adenosine moiety of tRNA during peptide bond formation, the initial point of attachment can be the 2′-OH, the 3′-OH, or either, depending on the amino acid [52]. After attachment, rapid migration between the two positions is possible [53]. The aminoacyl ester linkage has a relatively high standard free energy of hydrolysis, derived from the ATP hydrolysed during its activation. This is important as it provides the necessary energy for the subsequent peptide-bond formation to occur [54].

On the basis of primary structure it is possible to divide the aminoacyl-tRNA synthetases into two distinct classes, which have no discernible relationship between one another. Furthermore, it is observed that the class I synthetases generally catalyse attachment of the amino acid to the 2′-position of the ribose, whereas the class II synthetases generally catalyse attachment to the 3′-position. The evolutionary implication of this – that there were seperate origins for the two classes of synthetase – although startling, has been supported by the finding that the three-dimensional structures of representatives of the two classes are quite distinct [321].

Despite the fact that multiple species of tRNA (isoaccepting tRNAs) exist for a single amino acid (section 12.2.2), there appears to be only one aminoacyl-tRNA synthetase for each amino acid. Even the different initiating and elongating methionyl-tRNAs are recognized by the same enzyme. A fundamental question that arises is what features of these isoaccepting tRNAs are recognized by individual aminoacyl-tRNA synthetases to ensure that no misacylation occurs. Chemical and genetic manipulation of tRNAs has now led to a clearer general understanding of this, and the determination of the X-ray crystal structure of tRNA–synthetase complexes has provided detailed information.

In general, the studies with artificial tRNA constructs [55] have identified two main regions on tRNAs that act as determinants for recognition by synthetases: the anticodon loop and the acceptor stem. Which of these regions provides the major determinant varies from tRNA to tRNA, and it has been suggested that this may reflect the structural classes of aminoacyl-tRNA synthetase, mentioned above [56]. Although not all tRNAs show such clear-cut determinants, the cases of $tRNA^{Ala}$ and $tRNA^{Met}$ will be used to illustrate this point. A G:U base-pair unique to position 3:70 of the anticodon loop of $tRNA^{Ala}$ determines recognition by the alanyl-tRNA synthetase. Mutation of this causes loss of Ala-acceptor activity, and mutation of the corresponding C:G base-pair of $tRNA^{Cys}$ to G:U conveys Ala-acceptor activity to this latter tRNA. In the case of $tRNA^{Met}$, conversion of the cytosine of the 3′-UAC-5′ anticodon to its modified derivative, lysidine (found in $tRNA^{Ile}$), conveys Ile-acceptor activity on the tRNA; whereas conversion of the 3′-CAU-5′ anticodon of the $tRNA^{Val}$ to 3′-UAC-5′ conveys upon it Met-acceptor activity.

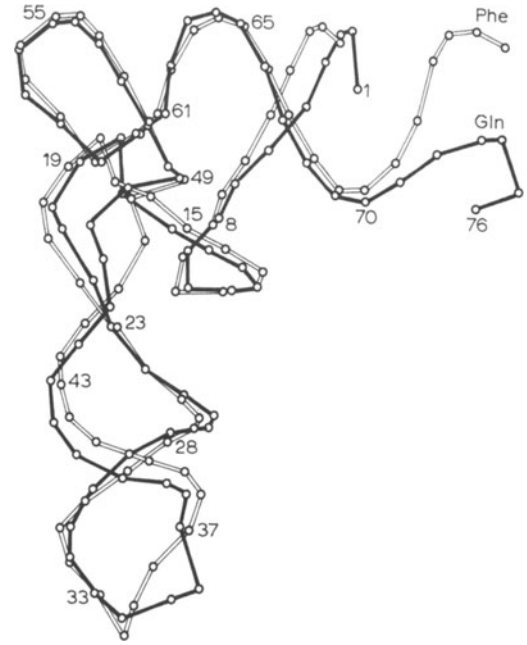

Fig. 12.12 Comparison of the conformation of free yeast $tRNA^{Phe}$ with that of *E. coli* Gln-$tRNA^{Gln}$ in a complex with glutaminyl-tRNA synthetase. The solid lines and open lines represent the path of the backbones of $tRNA^{Gln}$ and $tRNA^{Phe}$, respectively (After [57], with permission.)

The first tRNA–synthetase complex to have its structure determined was that for *E. coli* $tRNA^{Gln}$, a class I enzyme [57]. Both

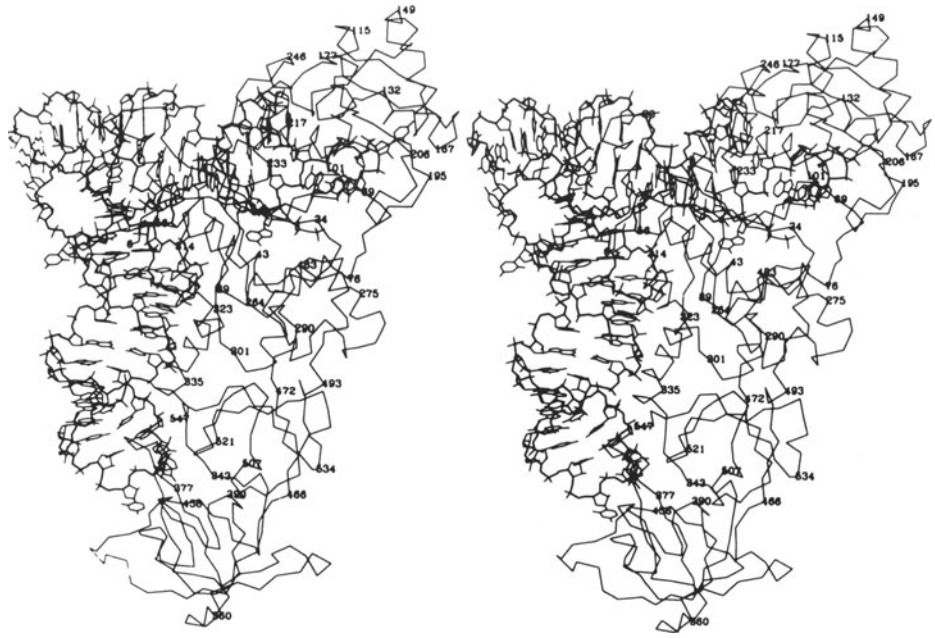

Fig. 12.13 Stereoscopic view of the complex between *E. coli* glutaminyl-tRNA synthetase and Gln-tRNA. (From [57], with permission.)

the central U of the anticodon of this tRNA and its acceptor stem are determinants of synthetase recognition. The striking feature of the X-ray crystal structure is the distortion of both these features of the tRNA compared with $tRNA^{Phe}$ (unfortunately the structure of $tRNA^{Gln}$ alone has not yet been determined) and the interaction of these regions of the tRNA with the synthetase (Figs 12.12 and 12.13). In the complex the base pairing of the acceptor stem is actually broken, explaining why only a weak G:U pair (or a non-pairing C:A) can be tolerated in the 1:72 position. In addition, the observation that this distorted structure is stabilized by a hydrogen bond to the phosphate backbone made by the 2-amino group of the unpaired residue G73 explains the requirement for a guanine residue in this position of the acceptor stem. It is perhaps a little surprising that the major role of the determining residues of the acceptor stem in this tRNA does not lie in direct interaction with the synthetase, the most striking feature of which in this region are amino acid residues that disrupt the helix. However, it should be stressed that there are specific interactions between the enzyme and the tRNA in these regions. For example, the carboxyl group of Asp^{235} of the synthetase makes hydrogen bonding interactions with the the 2-amino group of the guanine of the G3:C70 base-pair, and the importance of this is indicated by the fact that mutations that convert this Asp to Asn or Gly cause misacylation [58]. Furthermore, the anticodon loop, which is also involved in specifying the identity of the tRNA, makes multiple specific interactions with amino acid residues in the protein [322].

The three-dimensional structure of a complex of a class II synthetase with its tRNA has also been determined, that of $tRNA^{Asp}$ [323]. This complex also shows interaction

between the synthetase and the acceptor stem and anticodon loop of the tRNA. Although in this case there is no separation of the acceptor stem of $tRNA^{Asp}$ (this is not surprising as most of the identity determinants are in the anticodon loop) the distortion of the anticodon loop found in $tRNA^{Gln}$ is also observed, and in this case there is a structure for $tRNA^{Asp}$ alone (rather than that of $tRNA^{Phe}$) as a basis for direct comparison.

The recognition of the correct amino acid by the aminoacyl-tRNA synthetase is equally as important for the fidelity of translation as the recognition of the correct tRNA, just described. This is because if an incorrect amino acid is enzymically attached to the tRNA it will be misincorporated into protein on the basis of the codon–anticodon interaction. (This was shown in the classic experiment in which Cys-$tRNA^{Cys}$ was reduced with Raney nickel to give Ala-$tRNA^{Cys}$, which then incorporated alanine in response to codons for cysteine in a cell-free system [59]). There are several pairs of amino acids differing in structure by no more than a single methyl group, and this poses a real problem in discrimination for the synthetases. For example, it was observed that valine bound appreciably to Ile-tRNA synthetase. There is good evidence to support the existence of a 'proofreading' or 'editing' mechanism in those synthetases for which there are inappropriate isosteric or smaller amino acids with which the tRNA may be mischarged. This involves a hydrolytic site on the enzyme, close to, but distinct from, the acylation site, to discharge such inappropriate aminoacyl-tRNAs [60]. The accuracy of recognition at the hydrolytic site can, of course, be no greater than that at the initial acylation site; but by requiring the amino acid to be recognized twice, an error frequency of (e.g.) 1 in 10^2 would be reduced to 1 in 10^4.

12.4 THE EVENTS ON THE BACTERIAL RIBOSOME [61–63]

Bacterial protein synthesis will be described using the traditional two-site model of the ribosome (cf. Fig. 12.1) as a convenient visual aid. There are, however, several different proposals for additional sites on the ribosome which will be discussed subsequently.

12.4.1 Chain initiation [64–66]

In polypeptide chain initiation fMet-tRNA is bound to the initiation codon of the mRNA on the 30S ribosomal subunit and the resulting 30S initiation complex then reacts with the 50S ribosomal subunit to give a 70S initiation complex. This process requires GTP and the initiation factors, IF-1, IF-2 and IF-3, and in polycistronic mRNAs can occur independently at several different initiation sites.

The exact sequence of events in initiation and the precise roles of all the factors is not entirely certain, but that thought most likely is presented in Fig. 12.14. The initiating 30S subunit most probably has bound to it IF-3 and IF-1 (Fig. 12.14(a)) which are involved in generating free ribosomal subunits after polypeptide chain termination (section 12.4.3). Their role in initiation is distinct from this latter as they are required for the formation of a 30S initiation complex (Fig. 12.14(b)) even when 50S subunits are not present. IF-3 (M_r = 21 000) is primarily involved in binding mRNA to the ribosome. Although it is needed for translation of natural mRNAs, there is no absolute requirement for IF-3 in either the AUG-dependent ribosome binding of fMet-tRNA or the translation of artificial polynucleotides such as $AUGA_n$. This suggests that, either directly or indirectly, IF-3 facilitates the

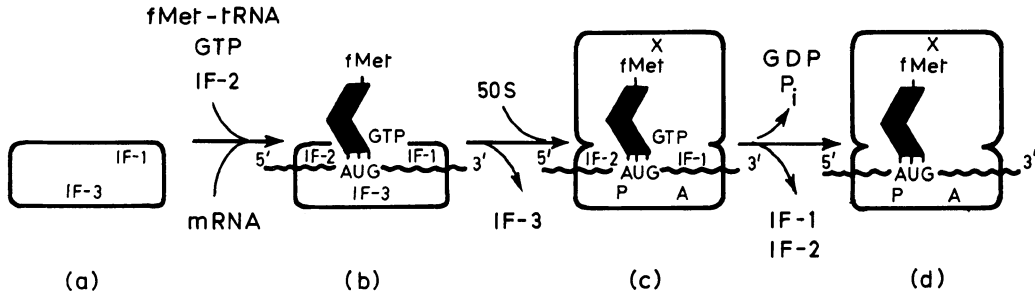

Fig. 12.14 A schematic diagram of prokaryotic polypeptide chain initiation. The ribosome and tRNA are represented as in Fig. 12.1.

16S rRNA	3′	HOAUUCCUCCACUAG. . . . .
		⊢—7±2—⊣
Phage φX174 *A*	5′. . . .	AAUCUUGGAGGCUUUUUUAUG. .
Phage MS2 coat protein	5′. . . .	UCAACCGGAGUUUGAAGCAUG. .
Phage λ *cro*	5′. . . .	AUGUACUAAGGAGUUGUAUG. . .
galE	5′. . . .	CCUAAUGGAGCGAAUUAUG. . . .
β-lactamase	5′. . . .	AUUGAAAAAGGAAGAGUAUG. . .
Ribosomal protein S12	5′. . . .	AAACCAGGAGCUAUUUAAUG. . .
RNA polymerase β	5′. . . .	GCGAGCUGAGGAACCCUAUG. . .
lacI	5′. . . .	CUUCAGGGUGGUGAAUGUG. . . .

Fig. 12.15 Some ribosome binding sites of *E. coli* and bacteriophage mRNAs. The key polypyrimidine region in the 16S rRNA and the bases complementary to this are shown with background shading, and the initiation codons are indicated in bold.

recognition of the untranslated 'leader' sequences that precede the initiating AUGs of natural mRNAs. Such bacterial 'leader' sequences contain short polypurine regions complementary to a polypyrimidine sequence at the 3′-end of the 16S rRNA (Fig. 12.15); and it was suggested by Shine and Dalgarno [67] that base-pairing between these regions is the means by which bacterial ribosomes select the correct AUG codon for initiation. There is now overwhelming evidence (reviewed in [68]) that this hypothesis is correct, although other factors, such as secondary structure, may prevent every potential 'Shine and Dalgarno' sequence being used for initiation. Three known exceptions in which initiation codons lack such a polypurine region in their 'leader' sequence nicely prove the rule, as these occur in mRNAs coding for extremely weakly expressed proteins: the cI repressor protein of phage lambda, and the DNA primase (the particular codon usage pattern of which was mentioned in section 12.2.4) and *trp* repressor of *E. coli*. The polypurine region is generally found at quite a specific distance 5′ to the initiation codon (Fig. 12.5), as would be expected if its function is to bring the initiation codon into the P-site of the ribosome.

The primary role of IF-2 (M_r = 97000) is to bind fMet-tRNA to the ribosome in a reaction that requires GTP and is stimulated by the other initiation factors, especially IF-1. The fMet-tRNA most probably binds to the ribosome after the mRNA has done

so, in contrast to the situation in eukaryotes (section 12.5.1), and most likely in the form of a ternary complex between IF-2, fMet-tRNA and GTP, although such a complex is experimentally much less stable than those formed by evolutionally related tRNA-binding factors. IF-2 must recognize some specific structural feature of the initiator tRNA as it will not interact with aminoacyl-tRNAs, not even Met-tRNA. One such feature may be the unique unpaired end of the acceptor stem, C1:A72, as mutation of this to a C:G base-pair allowed the tRNA to function in elongation [69]. Another, possibly relevant, structural difference of the initiator tRNA from elongator tRNAs is seen in the X-ray crystal structure, where the anticodon loop has an external-facing rather than internal-facing disposition [47], a feature also conserved in eukaryotic initiator tRNA [46].

Once a 30S initiation complex, containing mRNA and fMet-tRNA, has been formed (Fig. 12.14(b)) the 50S ribosomal subunit can associate with it (Fig. 12.14(c)) causing the release of IF-3. The non-hydrolysable analogue of GTP, 5′-guanylylmethylene diphosphonate, has been used to show that hydrolysis of GTP is not required for this step, but GTP hydrolysis is required for the fMet-tRNA to become available for reaction with puromycin. This reactivity with puromycin is used to define occupancy of the P-site, as puromycin is an analogue of aminoacyl-tRNA, and hence binds to the A site. After hydrolysis of GTP, IF-2 and (probably) IF-1 are released. As GTP hydrolysis does not cause relative movement of mRNA and the ribosome, fMet-tRNA must be bound directly at the P-site [70]. The role of the GTP hydrolysis, which is discussed in more detail below (section 12.4.2), cannot, therefore, be the provision of energy for movement of the fMet-tRNA from A-site to P-site.

It will be evident from the foregoing that although IF-1 (M_r = 8000) is absolutely required for initiation, its role is not clearly defined. The fact that it cycles on and off the ribosome during initiation established that it is indeed an initiation factor, rather than a loosely bound ribosomal protein such as S1 (section 12.6). Although IF-1 seems especially to facilitate the action of IF-2, it is clear that it does not have a function analogous to that of EF-Ts in elongation (section 12.4.2).

Lower-M_r subspecies of IF-2 and IF-3 exist, lacking portions of the *N*-termini of the larger species. The smaller form of IF-2 (IF-2β), rather than being a proteolytic fragment of the larger form (IF-2α), is the product of a second initiation on the mRNA [71]. These subspecies appear to be functionally equivalent to their parents.

The crucial role of the Shine and Dalgarno sequence in the selection of the initiation codon helps explain the fact, mentioned in section 12.2.1, that this latter is not always AUG. The most common variant, GUG, is used at about 3–4% the frequency of AUG. There are, in addition, some examples of UUG [68] and one of AUU [72] functioning as initiation codons in natural bacterial mRNAs. The other initiation codons are translationally less efficient than AUG: relative activities of AUG > GUG > UUG have been demonstrated [73]. However, the occurrence of GUG and UUG codons in highly expressed mRNAs confirms that this effect is not large and may be compensated for by a strong Shine and Dalgarno sequence. In the case of the AUA initiation codon, which is only known to occur in the mRNA for IF-3, the low intrinsic activity of this appears to be important for autoregulation by the cellular concentration of IF-3 itself. It has been suggested that the translation of the mRNA for IF-3 is independent of IF-3, and that a low intrinsic initiation activity because

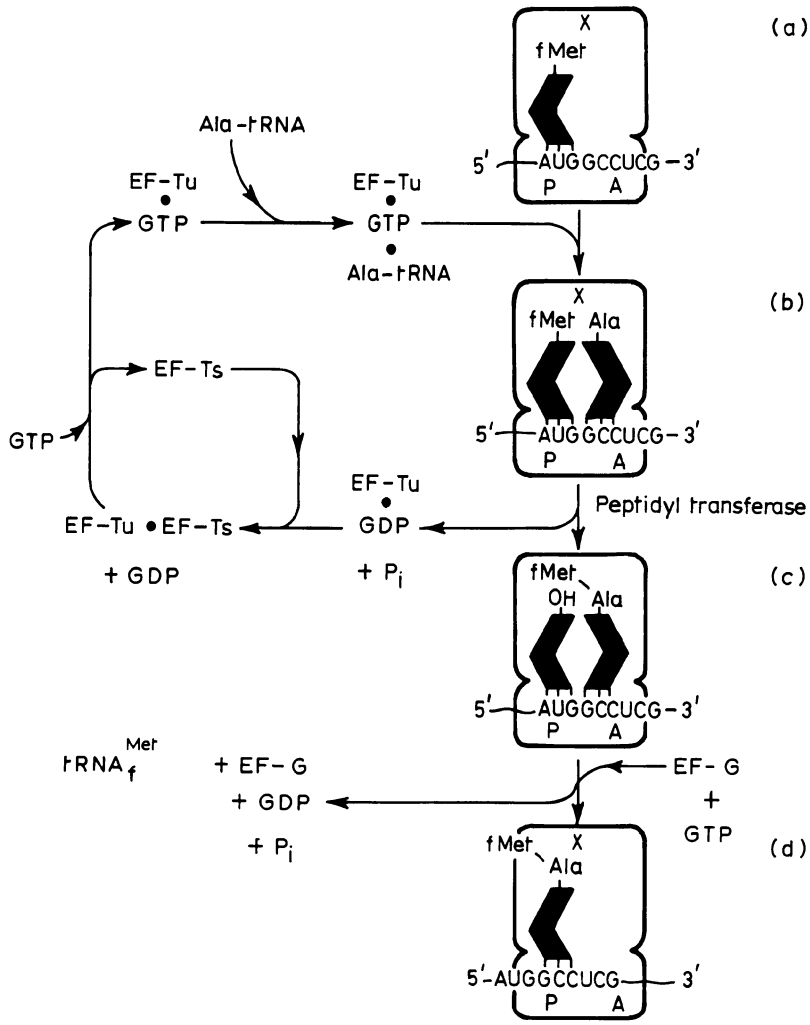

Fig. 12.16 A schematic diagram of prokaryotic polypeptide chain elongation. For convenience the 70S initiation complex of Fig. 12.14(d) has been taken as the starting point, (a), although the scheme applies equally for 70S ribosomes bearing peptidyl-tRNA in the P-site (e.g. (d)). Likewise, the designation of the mRNA triplet in the A-site as coding for Ala, and the third triplet for Ser are purely arbitrary.

of the AUA codon will mean that the mRNA cannot compete for ribosomal subunits unless a low concentration of IF-3 prevents other mRNAs from doing so [74]. Consistent with this idea, mutation of the AUA codon to AUG causes a 40-fold increase in translation and abolishes the autoregulation [75].

Some comment is necessary on the fact that the recognition of the alternative initiation codons, GUG and UUG, involves 'wobble' at the first, rather than the third, position of the codon. It is noteworthy that, in contrast to elongator tRNAs (section 12.2.2), the base 3′ to the anticodon in $tRNA_f^{Met}$ is unmodified. Hence the base-stacking that facilitates fidelity in the first two codon positions of elongator tRNAs is

decreased. It is interesting in this regard that in eukaryotic cytoplasmic $tRNA_f^{Met}$, which does not normally show such first position 'wobble' (section 12.5.1), this base *is* modified.

12.4.2 Chain elongation [76–79]

In polypeptide chain elongation an aminoacyl-tRNA binds to the A-site of the ribosome and reacts with the peptidyl-tRNA (Fig. 12.16) or fMet-tRNA (Fig. 12.14) in the P-site, accepting the growing polypeptide chain. The tRNA is then moved across to the P-site (translocated), with concomitant movement of the mRNA and expulsion of the deacylated tRNA, in order to make the A-site available for the next aminoacyl-tRNA. Elongation requires three soluble factors, EF-Tu, EF-Ts, EF-G, and the hydrolysis of two molecules of GTP.

Elongation factor EF-Tu (M_r = 43000 [126]) is responsible for the ribosomal binding of the aminoacyl-tRNA corresponding to the mRNA codon in the A-site (arbitrarily designated as Ala in Fig. 12.16), before which it forms a soluble ternary complex with the tRNA and GTP. All elongator aminoacyl-tRNAs will form this complex, but fMet-tRNA will not [80]. The non-hydrolysable analogue of GTP, 5′-guanylylmethylene diphosphonate, will allow the aminoacyl-tRNA to bind to the 70S ribosome, but the GTP must be hydrolysed before peptide-bond formation can occur. The GTP hydrolysis is not required for the peptidyl-transferase reaction itself and its possible role is discussed below. The EF-Tu and GDP are released from the ribosome as a complex. In this form the EF-Tu cannot react with GTP or aminoacyl-tRNA, and it is the function of EF-Ts (M_r = 30000) to displace GDP from the EF-Tu.GDP complex. This results in the formation of an EF-Tu.EF-Ts complex, from which the EF-Tu.GTP complex can be regenerated (Fig. 12.16).

Analysis of its primary structure shows that EF-Tu, like IF-2 and EF-G (below), belongs to the family of GTP-binding proteins that include the signal-transducing G-proteins [81]. The X-ray crystal structure of EF-Tu has been determined [82, 83], but, as yet the structure of the complex with aminoacyl-tRNA has not been solved, so that at present only the position at which GTP binds is known. There are two, apparently functionally equivalent forms of EF-Tu in *E. coli*, the products of separate genes (*tufA* and *tufB*), differing only in their *C*-terminal amino acids. EF-Tu is extremely abundant, constituting some 5% of total bacterial cell protein, and occurring in approximately sixfold excess over ribosomes and other elongation factors. The significance of this is not known.

The aminoacyl-tRNA bound in the A-site (Fig. 12.16(b)) can now be linked to the carboxyl group of the fMet or nascent peptide, through the catalytic activity of the intrinsic *peptidyltransferase* centre of the 50S ribosomal subunit. As already mentioned, the thermodynamic free energy for peptide bond formation comes from the hydrolysis of the 'energy-rich' acyl-ester bond of the aminoacyl-tRNA, which in its turn is derived from the ATP hydrolysed during aminoacylation. The fact that GTP and supernatant factors are not required for the transpeptidation was perhaps most convincingly confirmed when it was discovered that, in the presence of ethanol (about 50%), a 3′-hexanucleotide fragment of fMet-tRNA could react with puromycin on the isolated 50S ribosomal subunit [84]. Extension of this 'fragment reaction' to even smaller oligonucleotide fragments has shown that CCA-fMet is the smallest species that can occupy the P-site of the peptidyltransferase, and

that puromycin can be replaced by CA-Gly at the A-site.

The translocation of the peptidyl-tRNA from the A-site to the P-site requires the elongation factor EF-G (M_r = 77000) and GTP. This reaction has been shown to allow movement of the peptidyl end of the tRNA so that it becomes reactive towards puromycin, movement of the mRNA relative to the ribosome, and ejection of the deacylated-tRNA from the P-site. The reaction requires hydrolysis of the GTP, the non-hydrolysable analogue being inactive although it will allow EF-G to bind to ribosomes. The molecular mechanism underlying this process is perhaps the most intriguing and the least understood aspect of protein biosynthesis. It is clearly possible that a large part of the structural complexity of the ribosome, including even the division of the ribosome into subunits, may be a consequence of the need for this specific and concerted movement of macromolecules.

After translocation (Fig. 12.16(d)), one cycle of elongation has been completed (cf. Fig. 12.16(a)). The vacant A-site now contains a new mRNA codon, to which a corresponding aminoacyl-tRNA can bind, starting another round of elongation.

The function of GTP hydrolysis in the elongation reactions was unclear for many years. The earliest ideas were influenced by the biochemical precedents of ATP hydrolysis and the GTP-utilizing phosph*enol*pyruvate carboxykinase, and focused on the provision of energy for movement of molecules: the translocation of peptidyl-tRNA to the P-site in the case of EF-G, and the 'accommodation' of the aminoacyl-tRNA into the A-site in the case of EF-Tu. However, as mentioned above, it is now clear that the GTP-utilizing protein-synthesis factors are members of a larger family of GTPases which includes the G-proteins, in which the function of GTP hydrolysis is generally to terminate the action of the protein by causing its dissociation from its target [81]. Against this context the idea has emerged of a unitary role for GTP hydrolysis in protein biosynthesis to expel the factors (including IF-2) from the ribosome after they have fulfilled their functions. The presence of the factor on the ribosome can be regarded as preventing the next reaction from occurring, and the role of EF-G envisaged in this model is to prime the ribosome for subsequent translocation which does *not*, in itself, require GTP. In the case of EF-Tu, the time taken for the GTP to be hydrolysed after the ternary complex has bound to the ribosome is important in the kinetic 'proofreading' that ensures accuracy of the codon–anticodon interaction (section 12.6.3). However, assertions that the function of GTP hydrolysis is to provide the energy for this proofreading are, in our view, misleading.

As already stated, there are models of the ribosome containing additional sites, and these have implications for models of the elongation cycle. Lake [85] proposed that before occupying the A-site proper, the aminoacyl-tRNA entered through a recognition or R-site, which would allow the anticodon to be proofread for correspondence to the codon. Nierhaus (reviewed in [86]), on the other hand has proposed a model of the ribosome in which the deacylated tRNA is first transferred to an exit or E-site (i.e. after stage (c) in Fig. 12.16), before being expelled from the ribosome. More recently Moazed and Noller [87] obtained evidence for the employment of an additional site during translocation; although one rather different from that proposed by Nierhaus. Chemical footprinting indicated that the translocation occurred in two discrete steps involving first the acceptor end and then the anticodon end of the tRNA. To explain their results they proposed an extra E-site, solely

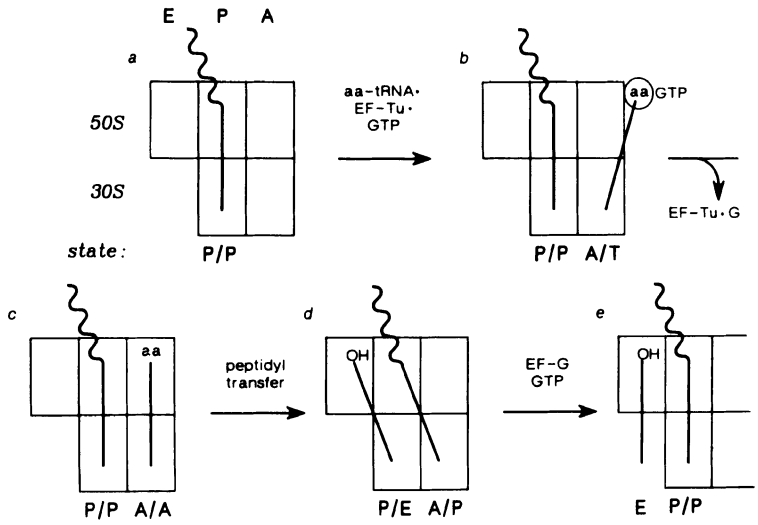

Fig. 12.17 Model for the movement of tRNA during translocation involving hybrid occupancy of A- or P-sites on the 30S subunit, and A-, P- or E-sites on the 50S subunit of the ribosome. In state A/T, no site on the 50S subunit is occupied. (From [87], with permission.)

on the 50S ribosomal subunit, with hybrid occupancy by the tRNA of various 50S and 30S sites occurring during the elongation cycle (Fig. 12.17). It is interesting that there is structural similarity between the RNA in the proposed E-site and that of part of the RNA component of *E. coli* ribonuclease P (section 11.3.4(a)) involved in interaction with the 3′-end of precursor tRNAs [319].

12.4.3 Chain termination [88–90]

In polypeptide chain termination (Fig. 12.18), the ester linkage of the peptidyl-tRNA is hydrolysed in response to one of three termination codons (section 12.2.1) in a reaction involving two of the three release factors, RF-1, RF-2 and RF-3. The deacylated tRNA and the mRNA are expelled from the ribosome in the presence of release factor, RRF, and EF-G, liberating free ribosomal subunits. The subunits will associate to form 70S ribosomes unless prevented from doing so by IF-3 and IF-1 (section 12.4.1).

In contrast to the other 61 codons, the three specific terminators are not read by tRNAs. This was shown using RNA from an amber mutant of bacteriophage R17 in which the first six codons of the coat protein are followed by a stop codon. Only the six appropriate aminoacyl-tRNAs and supernatant proteins were required for release of the hexapeptide. This same system was subsequently used with purified elongation factors and release factors to show that release of the peptide required the peptidyl-tRNA to be at the P-site of the ribosome [91]. To study the factor requirements for termination at all three codons, an assay was developed in which the termination codons could direct the release of fMet-tRNA, previously bound to ribosomes in the presence of the triplet AUG. This led to the resolution of two release factors, RF-1 (M_r = 36000) and RF-2 (M_r = 38000), of different codon specificities:

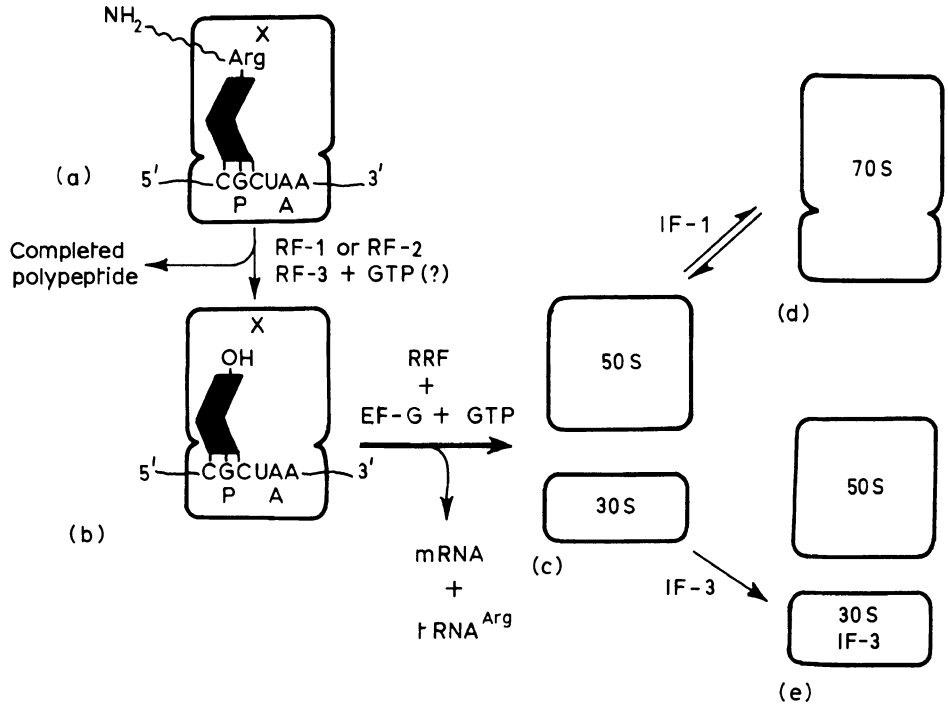

Fig. 12.18 A schematic diagram of prokaryotic polypeptide chain termination. The amino acid designation in the P-site is purely arbitrary. Other possible termination codons in the A-site are UAG and UGA (see text).

RF-1 for UAA or UAG

RF-2 for UAA or UGA

The third release factor, RF-3 (M_r = 46000), is not codon-specific and has no release activity in the absence of the other factors. It enhances the release of polypeptide promoted by the other factors and seems to stimulate both binding and release of these latter from the ribosome. Its activity is increased by GTP, but a requirement for GTP hydrolysis during termination in prokaryotes is not firmly established as it has been reported that GDP can replace GTP in the reaction *in vitro*. In this respect it would be useful to know the primary structure of RF-3 to see whether it is related to those of the known GTPases discussed in section 12.4.2.

There are quite strong grounds for thinking that the actual hydrolysis of the peptidyl ester linkage is catalysed by the peptidyltransferase centre of the ribosome, the reaction specificity of which has been modified by the binding of the release factors. This was suggested by the finding that the peptidyltransferase would catalyse the formation of an ester link to fMet-tRNA or its hexanucleotide fragment (section 12.4.2) if certain alcohols were presented to the ribosome instead of aminoacyl-tRNA. If the hydroxyl groups of an alcohol could replace the α-amino group of an aminoacyl-tRNA as a reactive nucleophile, it seemed possible that the hydroxyl group of water might do likewise. This suggestion was sup-

ported by the fact that a number of antibiotics (e.g. sparsomycin and chloramphenicol) and ionic conditions known to inhibit the peptidyltransferase reaction were also found to inhibit the termination reaction *in vitro*.

More recently a new perspective has been generated on the termination process in the form of a possible role for ribosomal RNA. A ribosomal mutation that specifically suppressed UGA termination codons was found to be the deletion of a cytosine (C_{1054}) from a highly conserved region of the 16S rRNA, and it was suggested that this mutation might prevent codon-specific base-pairing to a proximal 3′-ACU-5′ sequence [92]. There is, as yet, no direct evidence for such an interaction, nor is there any indirect evidence for analogous interactions involving the other two termination codons.

After the release of the peptide, the mRNA and deacylated tRNA are still attached to the ribosome (Fig. 12.18(b)) and must be removed before subunits can be regenerated for another round of protein synthesis. This requires GTP, EF-G and ribosome release factor, RRF (M_r = 18000). Although it might be expected that the primary role of EF-G in this process would be the expulsion of deacylated tRNA, there appear to be no data bearing on this question.

The released ribosomes can be in the form of 70S particles or 30S and 50S subunits. The supply of isolated subunits for reinitiation is controlled by IF-3, acting as an *anti-association factor*, preventing the 50S subunit from associating with the 30S.IF-3 complex [93]. Inactive 70S ribosomes do accumulate in cells, especially when inhibition of initiation results in a relative excess of 30S subunits over IF-3. To regenerate subunits when conditions improve there must be an equilibrium between 70S ribosomes and ribosomal subunits. Initiation factor IF-1 is thought to accelerate the interconversion of these species, without altering the position of its equilibrium [94].

Although all three termination codons are found to occur in mRNAs, the distribution of these is not random. The codon UAG is highly disfavoured, and in a sample of *c.* 800 *E. coli* termination codons the relative frequencies of occurrence, UAA:UAG:UGA, were approximately 10:4:1 [95, 96]. Furthermore, the context of termination codons is also not random, with a tendency for a following U being the most obvious feature of this. These features appear to provide protection against readthrough by natural suppressor tRNAs (section 12.9.6), rather than contributing to the recognition of the codons by termination factors [97, 98].

12.5 THE EVENTS ON THE EUKARYOTIC RIBOSOME [99, 100]

Eukaryotic ribosomes catalyse essentially the same process as prokaryotic ribosomes. Although the details of eukaryotic protein synthesis are less well understood, it is clear that the differences from prokaryotic protein synthesis are relatively minor for elongation and termination, but much greater for initiation. The following discussion relates to the nucleo-cytoplasmic protein-synthesizing system, the protein-synthesizing systems of mitochondria and chloroplasts being dealt with in section 12.7.

12.5.1 Chain initiation [101, 102]

One way in which eukaryotic initiation differs from that in prokaryotes is in the initiating amino acid: instead of fMet, the initiating amino acid is Met. Nevertheless, there is a specific species of methionine tRNA for initiation, distinct from that used in elongation. The eukaryotic initiator

tRNA will be referred to here as $tRNA_f^{Met}$ (and the elongator as $tRNA_m^{Met}$), although it is sometimes designated $tRNA_i^{Met}$. This nomenclature emphasizes the unity in structure between the eukaryotic and prokaryotic initiator tRNAs: despite the fact that there is no transformylase in the eukaryotic cytoplasm, the eukaryotic Met-$tRNA_f$ can be formylated *in vitro* by *E. coli* transformylase [103].

The more fundamental difference in eukaryotic initiation is the mode of selection of the initiating AUG, and it is clear that it is this aspect of initiation, although still incompletely understood, that accounts for a plethora of eukaryotic initiation factors, far exceeding the number in prokaryotes. Instead of the 'Shine and Dalgarno' interaction (eukaryotic 18S rRNA lacks the key CCUCC sequence involved in this), allowing independent internal initiations on a polycistronic mRNA, in eukaryotic initiation there is a different mechanism: the Kozak 'scanning mechanism' [68, 104, 105]. The 40S ribosomal subunit binds to the 5′-end of a monocistronic mRNA and 'scans' along this until it encounters an appropriate (usually the first) AUG initiation codon, when attachment of the 60S ribosomal subunit can occur. In contrast to the situation in prokaryotes, the initiator tRNA binds to the small ribosomal subunit before this attaches to the mRNA [106, 107], and experiments in which the 3′-UAC-5′ anticodon of the $tRNA_f^{Met}$ was mutated to 3′-UCC-5′ have demonstrated that the anticodon is involved in the scanning for the appropriate AUG codon [108].

A model for the mechanism of eukaryotic initiation is presented in Fig. 12.19. The first stage is the formation of a (stable) ternary complex between Met-$tRNA_f^{Met}$, eIF-2 and GTP. Although eIF-2 appears functionally analogous to prokaryotic IF-2, the eukaryotic factor differs from the latter in comprising three subunits (α, β and γ). The ternary complex then binds to a 40S ribosomal subunit bearing eIF-3 and eIF-4C, giving a 43S preinitiation complex. One of the roles of eIF-3 is as an anti-association factor, generating free 40S subunits; but it should be stressed that eIF-3 is an extremely complicated species, consisting of at least seven antigenically discrete polypeptide chains with an aggregate M_r of approximately 700 000 [109].

The next step is a most distinctive feature of eukaryotic initiation, the binding of the 43S preinitiation complex to the 5′-end of the mRNA, involving melting of mRNA secondary structure and recognition of the 5′-cap structure. Four factors are involved at this stage: the three factors eIF-4A, eIF-4E, and p220 (M_r = 220 000), which together make up a complex known as eIF-4F, and a fourth factor, eIF-4B. The factor responsible for promoting unwinding of the secondary structure of mRNA is eIF-4A, which is an ATP-dependent helicase. (At least in mouse, there are two distinct functional genes for eIF-4A, but the significance of this is unclear [110].) The recognition of the 5′-cap structure is the property eIF-4F. Although all three components of the factor are required for this, purification on 'cap' affinity columns has shown that it is eIF-4E (M_r = 24 000) that is the actual cap-binding protein. The factor eIF-4E must play a pivotal role in the regulation of eukaryotic protein biosynthesis, for transfected fibroblasts over-expressing this factor exhibit malignant transformation [111]. The precise role of eIF-4B at this step of initiation has yet to be defined.

The 40S initiation complex, having bound to the mRNA, scans for the correct AUG codon, when the 60S subunit is incorporated, forming the 80S complex. Factor eIF-5 appears to be necessary to promote GTP hydrolysis and the release of the other

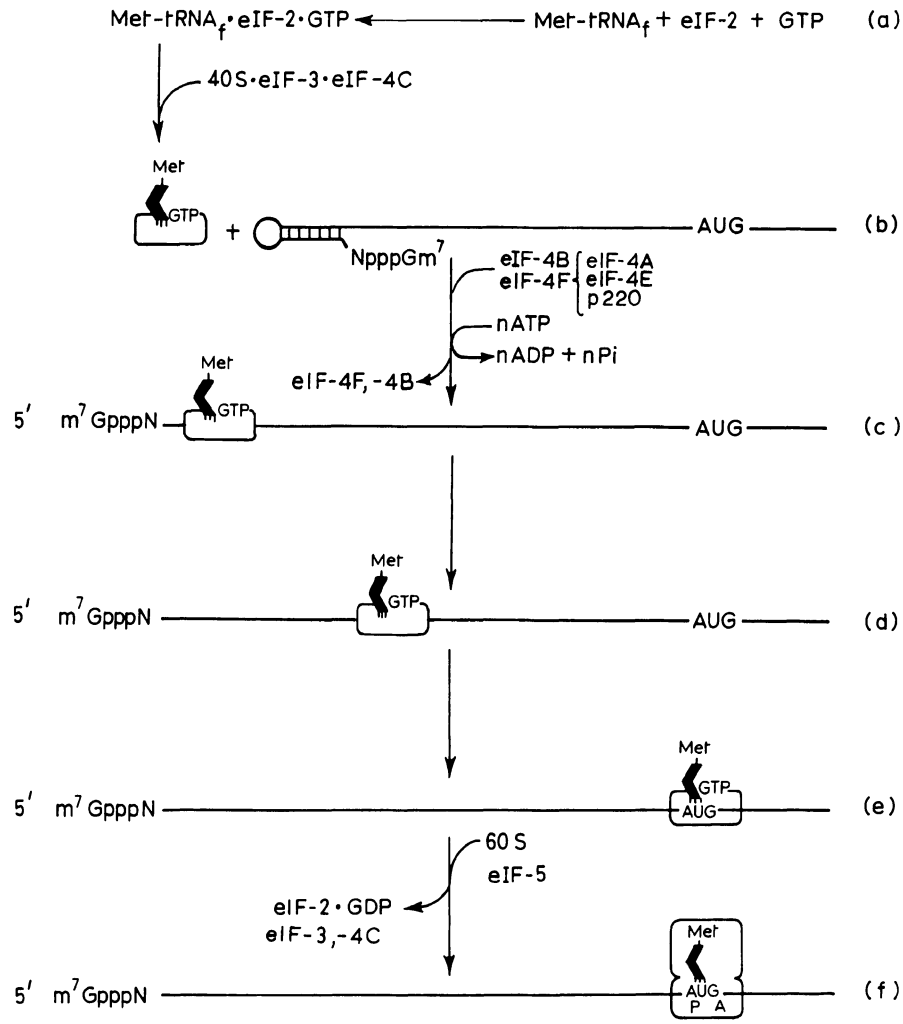

Fig. 12.19 A schematic diagram of eukaryotic polypeptide chain initiation.

factors before this can occur. The eIF-2 released at this stage is complexed to GDP which must be displaced by a factor, eIF-2B (also called eRF and GEF), before it can continue to function (Fig. 12.35). This reaction, which is analogous to that between EF-Tu and EF-Ts (Fig. 12.16), is subject to regulation and is discussed in more detail in section 12.9.5. Another factor, termed eIF-4D, has been implicated in the formation of the first peptide bond, but its role has not been clearly defined and it is not shown in Fig. 12.19. This protein is interesting, however, as it contains a unique modified derivative of lysine, hypusine, which is essential for its biological activity [112].

The key tenets of the original 'scanning model' of Kozak [104], presented above, are attachment of the ribosome to the 5′-end of the mRNA and initiation at the first AUG codon; and the corollary of these is that

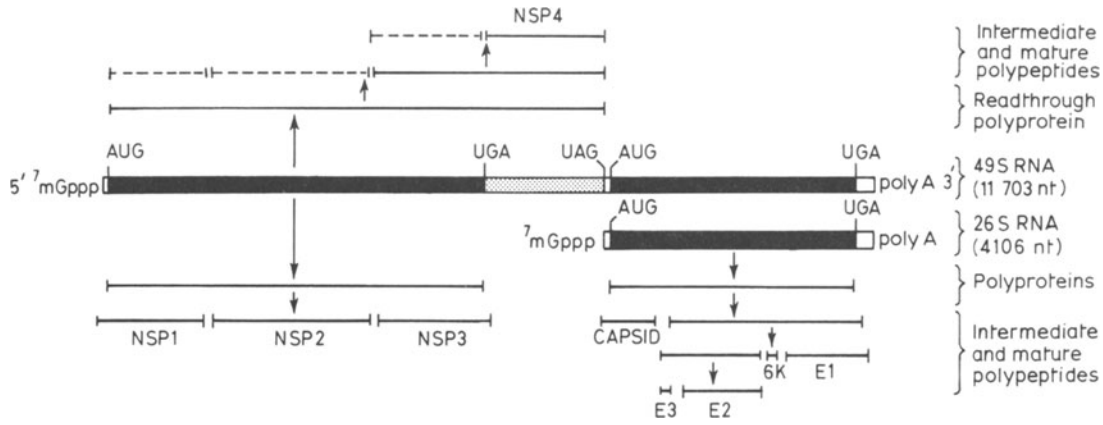

Fig. 12.20 Translational strategy of Sinbis virus [113]. Translation of the 49S genomic RNA occurs only from the first AUG, a second internal initiation site being recognized only on the subgenomic 26S RNA, where it has now become the first AUG. The initial transcripts are polyproteins which are proteolytically processed to give the mature non-structural and structural proteins indicated. Suppression (section 12.9.6) of the first termination codon produces a larger polyprotein from the 49S RNA, the processing of which generates NSP4 (and perhaps more of the other non-structural proteins: indicated by broken lines).

there should be no polycistronic mRNAs with internal initiation, in contrast to the situation in prokaryotes. Certainly, in the vast majority of cases this holds true, and it is striking how many RNA viruses (e.g. [113]) resort to strategies such as the production of polyproteins or the generation of nested mRNA subspecies that allow then to be functionally monocistronic despite having structurally polycistronic genomes (Fig. 12.20). It is true that there are some minor deviations from the original predictions, but these can be accommodated by slight refinement of the scanning model [105]. However, there are also some rare major violations of the predictions that can only be explained by the existence of an alternative mechanism of eukaryotic initiation [114]. We shall consider each of these situations in turn.

In approximately 700 vertebrate mRNAs analysed, the most 5′-terminal AUG is the initiation codon in 90–95% of cases. It was realized that the initiation codons actually used have a similar 'context', the consensus being:

GCCGCCRCC**AUG**G

and it was therefore proposed that this context somehow influenced the selection of the initiation codon, with AUG codons in weak contexts being bypassed. Not all the consensus nucleotides exert equal influence, but the presence of either a purine at position '−3' or a G at position '+4' to the start of the AUG seem to be particularly important. The majority of those 5–10% of mRNAs that did not initiate at the first AUG were found to initiate at the first AUG occurring in a strong context. Moreover, some rare cases where two consecutive AUG codons are used as alternative initiation codons are explicable in terms of this context effect if it is assumed that a

proportion of 43S initiation complexes ignore the first AUG because of a relatively poor context which, nevertheless, does not deviate totally from the consensus. The mode of action of the consensus sequence is unknown, although the occurrence of purines with a periodicity of three suggests that some kind of 'phasing' of the reading-frame may be occurring.

There are two other refinements that have to be made to the 'scanning' model in order to account for cases where initiation does not occur at the first AUG codon in a strong context. One is when the first AUG occurs within 10 nucleotides of the 'cap', in which case it is assumed to be inaccessible to the 43S initiation complex. The second is where the AUG codon from which the protein initiates is preceded by a short open reading-frame. It appears that the first open reading-frame is translated, and that, after termination, continued scanning and re-initiation can occur [115]. The explanation of this phenomenon, which is restricted to short open reading-frames culminating in a termination codon, is unclear. Such preceding short open reading-frames have the effect of decreasing translation of the authentic initiation reading-frame, and Nature appears to have used this as a device to repress the translation of certain proteins. The most studied example of this is the mRNA for the GCN4 regulatory protein of yeast, where four such short open reading-frames are found preceding the initiation codon [116].

There is ample experimental evidence that mutation of AUG initiation codons markedly decreases translation (reviewed in [105]), so that the first reports of initiation at non-AUG codons generated considerable surprise. These, the best cellular examples of which involve CUG codons, appear to be for very weakly expressed proteins [117–119], and therefore seem to represent a further example of how deviations from the optimal situation for operation of the scanning mechanism are used to regulate translation.

In the instances cited above there is no question that the basic mechanism of 'scanning' from the 5′-end of the mRNA still holds. However, there is one case, that of the picornaviruses, where an entirely different mechanism has to be invoked. Picornaviruses, such as polio, are peculiar in having uncapped mRNAs which do not require eIF-4F for their translation; and indeed this is fundamental to the manner in which they subvert the protein synthetic machinery of the cell (section 12.9.5). Although this does not itself preclude scanning from the 5′-end of the mRNA, this appeared unlikely because of the long 5′-leader with multiple AUG codons, some in strong contexts, preceding the actual initiation codon of the polyprotein that polio encodes. There now seems no doubt that there is internal initiation in this case, as it has been shown that a 450-nucleotide segment of this leader, not including the extreme 5′-end, can allow translation of the erstwhile silent second member of laboratory-generated bicistronic mRNAs when inserted in front of the latter [114]. The secondary structure of the internal entry site provided by this 450 nucleotide segment seems to be important, and there is some evidence that this may be recognized by a cellular protein [120], but otherwise the details of the mechanism of such internal initiation are obscure. It is clear, however, that a completely distinct mechanism exists for the complex and fundamental process of initiation of protein synthesis, raising the question whether this could have evolved in picornaviruses. It is perhaps easier to envisage the virus acquiring a mechanism of initiation evolved by a few host mRNAs to enable them to operate the type of translational regulation described in section

12.9.5, and the existence of a functional internal initiation sequence has recently been demonstrated for one host mRNA [121].

12.5.2 Chain elongation [122]

The eukaryotic factors required for binding aminoacyl-tRNA to the ribosome during the elongation cycle are analogous to EF-Tu and EF-Ts of prokaryotes. EF-1α, which is structurally related to EF-Tu [81], forms a ternary complex with aminoacyl-tRNA and GTP; and a second activity, EF-1$\beta\gamma$ (or perhaps EF-1$\beta\gamma\delta$), functionally analogous to the monomeric EF-Ts, promotes release from EF-1α of GDP, and exchange for GTP. Like prokaryotic EF-Tu, eukaryotic EF-1α is an extremely abundant cellular protein [123], and is encoded by two separate genes [124].

The elongation factor involved in the translocation reaction in eukaryotes, EF-2, is structurally and functionally related to EF-G in prokaryotes. One point of interest is the specific inactivation of EF-2 (but not EF-G) by bacterial diphtheria toxin [125], which transfers ADP-ribose from NAD^+ to an unusual modified histidine residue [126] in EF-2. Although there has been a report of an enzyme with an analogous activity to that of diphtheria toxin in eukaryotic cells [127], it is still uncertain whether ADP-ribosylation of EF-2 represents a normal mechanism of cellular regulation.

Eukaryotic, like prokaryotic, ribosomes possess an intrinsic peptidyltransferase activity, which has also been studied using the 'fragment' reaction (section 12.4.2) [128]. Although the eukaryotic peptidyltransferase is inhibited by certain antibiotics (e.g. sparsomycin) that inhibit prokaryotic peptidyltransferase, it is resistant to the action of others (e.g. chloramphenicol).

One surprising feature of eukaryotic elongation is a species difference: yeast and some fungi have a third elongation factor, EF-3, not required by a range of other eukaryotes [122]. Although yeast EF-3 has been purified and its gene cloned, its precise function is as yet unclear.

12.5.3 Chain termination

In eukaryotic termination there is a single factor, RF, to recognize all three termination codons, UAA, UAG and UGA, and hence this is functionally equivalent to both RF-1 and RF-2 of prokaryotes [129]. A second eukaryotic release factor, equivalent to bacterial RF-3, has also been reported and warrants further study in view of the fact that the eukaryotic termination reaction shows a clear requirement for GTP hydrolysis *in vitro*. The amino acid sequence of eukaryotic RF shows no similarity to those of prokaryotic RF-1 and RF-2 (which are quite similar to one another [130]), although, intriguingly, it contains a region homologous to one in the tryptophanyl-tRNA synthetases [131].

Although one imagines that there is a eukaryotic factor analogous to RRF, none has so far been described. Eukaryotic ribosomes are liberated from polysomes as subunits, and a pool of inactive 80S monomers is present in eukaryotic cells [132]. In contrast to prokaryotes there is both an anti-association activity, eIF-3 (section 12.5.1), which binds to 40S subunits, and an 80S subunit dissociation factor, eIF-6 [101].

The usage of the three termination codons in eukaryotes is much more random than in prokaryotes (section 12.4.3), although UAG is still the codon least frequently used. In vertebrates the relative occurrence of UGA:UAA:UGA is approximately 2.5:1.7:1 [133].

12.6 THE RIBOSOME

The focus of this section will be *E. coli*, the ribosomes of which have been the subject of extensive study [134–138]. Further information about eukaryotic ribosomes can be found elsewhere [139, 140].

12.6.1 The structure of the ribosome

(a) *Overall features*

The overall size and shape of ribosomes have been analysed by several different techniques. At the most basic level was the physical separation of the ribosomes into subunits (by removal of the Mg^{2+} that holds them together) and the determination of their sedimentation coefficients in the ultracentrifuge: 30S and 50S in the case of the 70S bacterial ribosome, and 40S and 60S in the case of the 80S eukaryotic ribosome of mammals. (The sizes of the ribosomes of other eukaryotes may differ from this.) Initial small-angle X-ray scattering studies provided values for the overall dimensions of the ribosomes and their subunits, but their gross features have been deduced from electron microscopy of individual negatively stained particles. Although inherent problems in this latter method limit the detail that can be obtained, useful models have been derived from such work, and the one presented in Fig. 12.21 is generally used as a reference for studies on the organization of the ribosomal components [141]. Its most obvious features are a 'cleft' and 'platform' on the 30S subunit, and an elongated 'stalk' protruding to one side of the 'central protuberance' in the 50S subunit.

Refinements have been made in the methodology for interpreting such electron micrographs [142], but the way to more substantial advances would seem to lie through

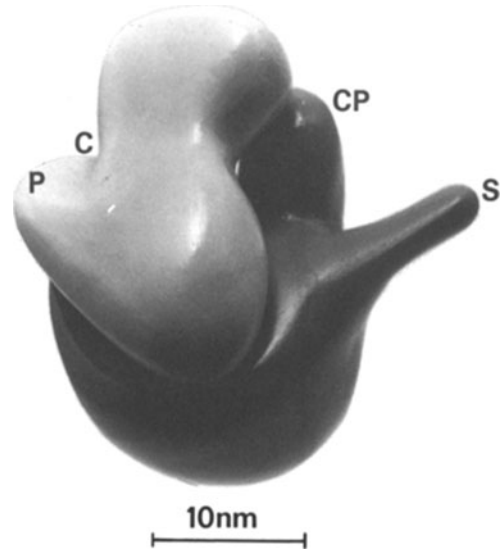

Fig. 12.21 Model of the 70S ribosome of *E. coli* based on electron-microscopic studies by Lake [141]. The 30S subunit is light, and the 50S subunit is dark. The cleft (C) and platform (P) of the 30S subunit, and the stalk (S) and central protuberance of the 50S subunit, referred to in the text, are indicated.

diffraction methods. Three-dimensional crystals of ribosomes which diffract X-rays have been obtained, but these have not yet yielded sufficient resolution to be useful. However, significant progress has been made with three-dimensional image reconstruction of the electron diffraction patterns obtained with two-dimensional crystalline sheets of ribosomes [143]. One striking feature revealed by this technique is a 'tunnel' through the centre of the 50S ribosomal subunit (Fig. 12.22), and a similar feature has been observed in independent studies of eukaryotic 60S ribosomal subunits [144]. This tunnel is large enough to accommodate a nascent peptide of the length (25–40 amino acid residues) that the ribosome is known to protect against proteolytic digestion [145], and it is most likely that this feature represents an exit channel for the nascent peptide.

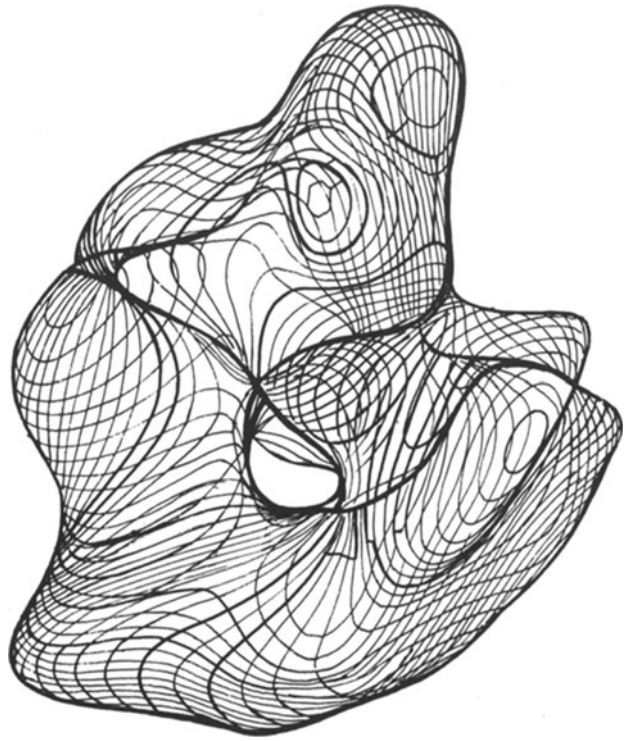

Fig. 12.22 Computer graphic representation of a reconstructed model of the 50S ribosomal subunit of *Bacillus stearothermophilus* obtained from two-dimensional sheets, negatively stained with gold thioglucose (courtesy of Dr A. Yonath).

What of the physical location of other species interacting with the ribosome? The 'length' of the tRNA (75 Å – 7.5 nm) is such that it could be accommodated in the space between the two ribosomal subunits; but this is purely speculation. Immune electron microscopy (see below) suggests that puromycin, and hence the acceptor stem of aminoacyl-tRNA, may interact with the 50S subunit near the base of the 'central protuberance' [146]. Related experiments suggest that the anticodon of the tRNA interacts with the mRNA at the 'platform' of the 30S subunit [147]. It has long been known from nuclease protection experiments that 35–50 nucleotides of mRNA (i.e. 12–17 codons) are in contact with the ribosome. A more recent sophisticated analysis revealed two internal cleavage sites that could represent 'bends' in the mRNA, and have been interpreted in terms of the mRNA 'looping' around the ribosome to re-emerge near its point of entry [148].

(b) *Components*

The composition of the ribosome is approximately 60% RNA and 40% protein. The smaller ribosomal subunit contains a single species of RNA (16S or 18S RNA for 30S and 40S subunits, respectively) together with ribosomal proteins; the larger ribosomal subunit contains a major species of RNA (23S or 28S RNA for 50S and 60S subunits, respectively), an additional small 5S RNA, together with a number of proteins exceeding that in the small subunit (Fig. 12.23). In eukaryotes the portion of the primary

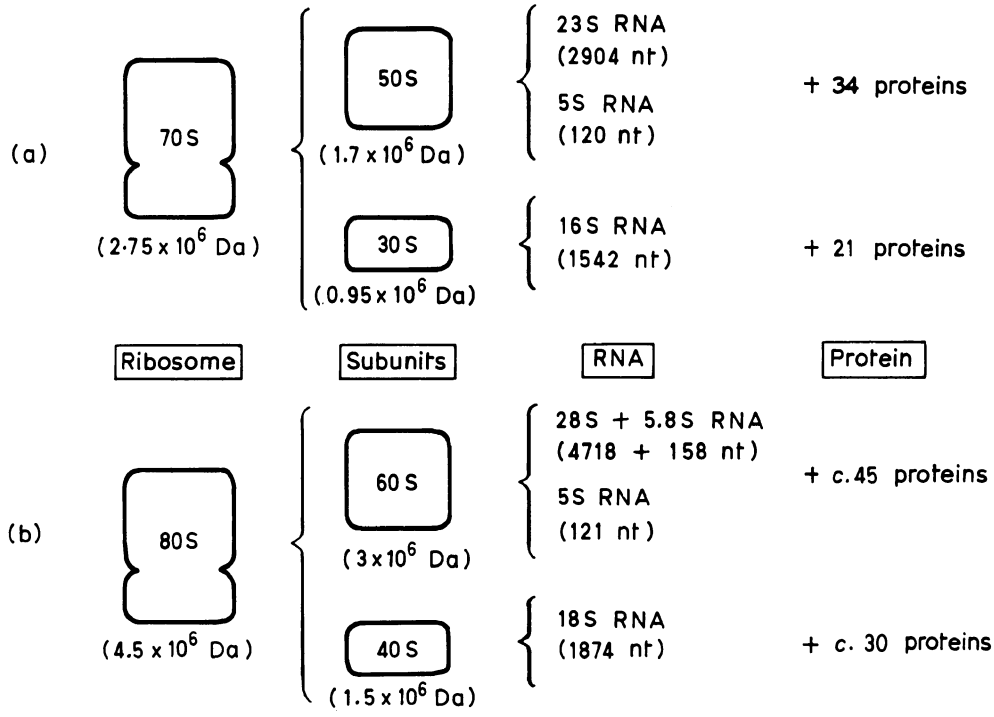

Fig. 12.23 A schematic diagram of the components of the ribosomes of (a) *E. coli* (b) rat. The molecular masses of the particles are the mean of physical determinations. The chemical values for the M_r of 70S, 50S and 30S ribosomal particles from *E. coli*, based solely on RNA and protein content, are 2.3×10^6, 1.45×10^6 and 0.85×10^6. The discrepancy between the two sets of values can be accounted for by the presence of metal ions and spermidine.

transcript that gives rise to the major species of rRNA in the 60S subunit undergoes further nucleolytic cleavage after it has started adopting its secondary structure (section 8.4.3). This results in a 5.8S RNA species hydrogen-bonded to the 28S RNA in the case of mammals. Similar 5′-post-transcriptional processing gives rise to a 2S RNA species in *Drosophila* and a 4.5S RNA species in plant chloroplasts (sections 9.4.2, 9.4.3 and 9.7.2). As discussed further below, from a functional point of view these RNAs are better considered to be integral parts of the larger rRNA species.

The ribosomal RNAs contain a small number of specific modified (largely methylated) nucleotides (section 11.6.3) [149, 150]. In the case of *E. coli*, 10 base methylations have been identified in both 16S and 23S rRNA; and the 23S rRNA also contains three pseudouridine residues and a few ribothymidine residues and ribose methylations. In mammals there are 46 and 71 methylated groups on the 18S and 28S RNA, respectively, most of which involve the 2′-*O* of the ribose moiety. They are predominantly clustered in the 5′-half of 18S rRNA and the 3′-half of 28S rRNA [151]. Eukaryotic rRNAs also contain pseudouridine residues: approximately 37 and 60 for 18S and 28S rRNA, respectively. Neither eukaryotic nor prokaryotic 5S rRNA contains modified nucleotides.

The ribosomes of eukaryotes, prokaryotes,

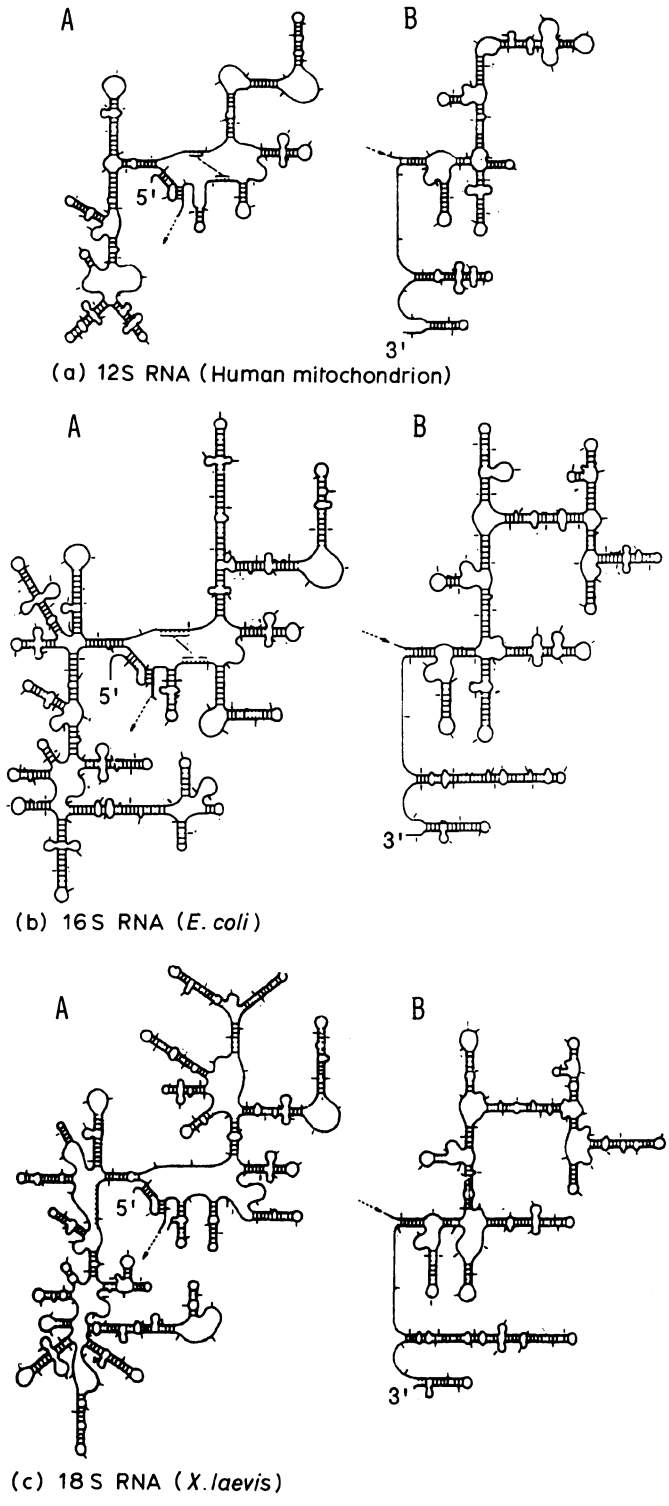

Fig. 12.24 Comparison of the secondary structures of the smaller ribosomal RNA from (a) a human mitochondrion, (b) a prokaryote and (c) a eukaryote. Each structure has been separated into two parts for clarity of presentation. These early proposals have undergone subsequent detailed refinement, but without in any way invalidating the comparative approach. (After [152], with permission.)

mitochondria and chloroplasts catalyse similar reactions. It seems reasonable, therefore, to assume that their RNAs serve similar functions, whatever these might be. In comparing rRNAs that diverge markedly in size (from *c.* 600 to 2000 nucleotides for small subunit rRNAs, and from *c.* 1200 to 4700 nucleotides for large subunit rRNAs [149]) attention has therefore centred on common structural features that may be crucial to such functions. The most evident similarity lies in the secondary structure, in which a common core is present which becomes progressively more elaborated as the rRNA increases in size (Fig. 12.24). It is such comparison of secondary structures that makes it apparent that the 5.8S rRNA of mammalian ribosomes, when hydrogen-bonded to the 28S rRNA, is equivalent to a structural feature present in *E. coli* 23S rRNA [153]. At corresponding regions in the secondary structures of rRNAs from different species (predominantly in single-stranded regions) it is possible to identify a small number of relatively short, highly conserved, regions of primary structure. The significance of these will be discussed in section 12.6.2.

The initial models of a common secondary structure for the rRNAs were primarily based on comparisons of the sequences of rRNAs of different phylogenetic origins. Evidence was obtained from situations where there is poor conservation of primary structure in a feature of proposed secondary structure, which is maintained because of compensatory base changes (e.g. an A → G change in one strand of the RNA is compensated for by a U → C change in the complementary base of the other strand). Experimental confirmation and refinement of the models was then necessary, and was of especial importance where there was either extreme conservation or divergence of primary structure. One approach was to identify regions of single-stranded RNA, which can be done using certain nucleases, chemical reagents and oligonucleotide probes. A particularly extensive study used chemical probes for all four bases, and identified the resulting modified exposed bases by their property of terminating reverse transcription, primed by a battery of oligonucleotides complementary to rRNA [154]. The other, more laborious approach, was identification of double-stranded regions, which involved RNA–RNA cross-linking followed by isolation and identification of base-paired fragments. Such cross-links have been important in building models for the tertiary structure of the rRNAs [155, 156], one of which is shown in Fig. 12.25. Models for the secondary and tertiary structure of 5S rRNA have also been proposed [157].

One feature of the hairpin loops in the rRNA secondary structure (section 2.6) is worth reiterating. This is the frequency with which the loops are found to consist of the tetranucleotides, GNRA or UNCG. Analysis of these has revealed particular structures in which there is a base-pair between the first and fourth nucleotide (G:A and U:G, respectively) [317]. The resulting stability imparted to the loops by these sequences explains their frequent occurrence.

Originally the ribosomal proteins of the ribosomes of *E. coli* were enumerated S1–S21 and L1–L34 for the small and large subunits, respectively. It is still thought that the 30S ribosomal subunit has these 21 distinct proteins, one copy per 30S subunit. However, there has been subsequent revision of the status of the proteins of the 50S subunit. It emerged that the species designated L8 was, in fact, a complex of L7/L12 and L10, and that L26 is identical to S20, there being, on average, 0.2 copies of L26 and 0.8 copies of S20 per 70S ribosome [158]. Furthermore, two additional proteins, A and B, were detected on the ribosome, and as

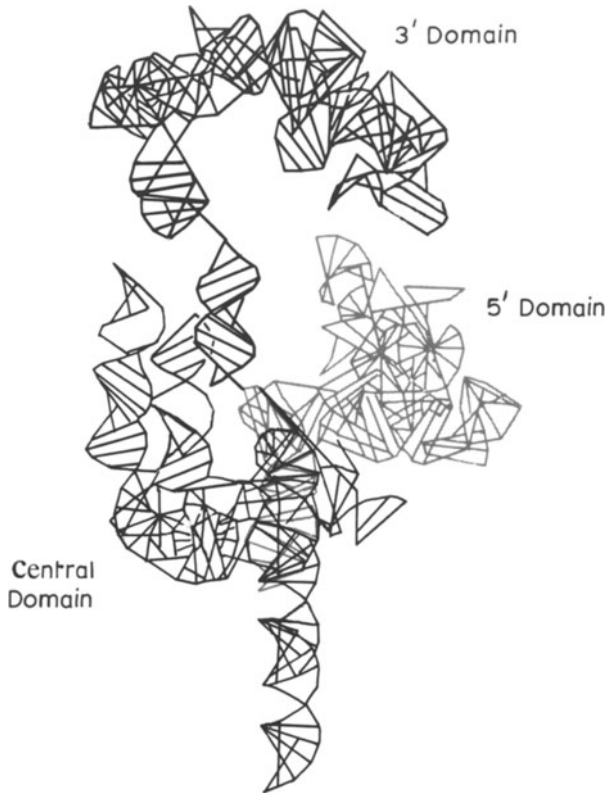

Fig. 12.25 Model proposed by Noller and co-workers for the folding of *E. coli* 16S rRNA. The 3′-domain is at the top of the figure and the 5′-domain at the right. (From [155], with permission.)

these are chemically basic and have genes that map in clusters of other ribosomal proteins (section 12.9.2) they have now been redesignated as ribosomal proteins L35 and L36 [159]. Thus it is best regarded that there are 34 distinct proteins on the 50S subunit. All but two of these proteins are completely dissimilar and are present as single copies. The exceptions are L7 and L12, L7 being the *N*-acetylated form of L12. These two proteins together are present in a total of four copies per 50S subunit. Apart from proteins S1, S6, L7 and L12, the ribosomal proteins are all chemically basic, with values of M_r in the range 9000–35 000. The complete amino acid sequences of all the ribosomal proteins of *E. coli* have been determined. The structures of a couple of ribosomal proteins have been determined by X-ray crystallography and, on the basis of these, possible regions for interaction with rRNA have been proposed [160, 161].

Amino acid sequences are at present available for approximately half the eukaryotic ribosomal proteins, but so far only eight of these have been shown to be statistically significantly related to the sequences of proteins of *E. coli* (to S3, S5, S10, S14, S16, S17, S19 and L6 of the latter). In addition, there are grounds for thinking that certain ribosomal proteins of eukaryotes are functionally homologous to L7/L12 of *E. coli*. These include the size of these proteins, their acidic nature and their occurrence in multiple copies. There is also an indirect primary structure link through mutual relatedness to a corresponding archaebacterial (section 12.7.1) ribosomal protein [162]. Two forms of this acidic ribosomal protein, with similar but distinct primary structures, occur in eukaryotes. A third related, but larger, protein may be functionally equivalent to L10 (section 12.6.2) [163]. All these eukaryotic acidic proteins are phosphorylated to various extents; a modification found in another eukaryotic ribosomal protein (section 12.9.4) but not in prokaryotic ribosomal proteins [164].

(c) *Organization*

We now turn to the question of the organization of RNA and ribosomal proteins in the ribosome. The relative positions of ribosomal proteins to one another have been determined by cross-linking with cleavable bifunctional reagents [165] and by neutron scattering [166]. In the latter, extremely powerful, technique, ribosomes are reconstituted from their components, but with

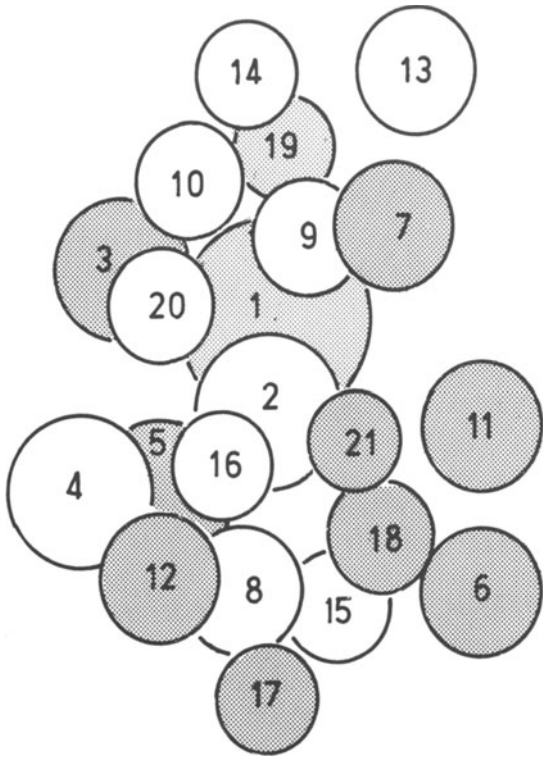

Fig. 12.26 The relative positions of the proteins of the 30S ribosomal subunit of *E. coli* derived from the neutron scattering studies of Moore and co-workers. (From [166], with permission.)

two of their proteins replaced by ones of a neutron density that contrasts with that of the rest of the particle. This is achieved by growing *E. coli* in media with appropriately different proportions of D_2O. The results of such measurements for the proteins of the 30S ribosomal subunit are shown in Fig. 12.26. A more indirect method is 'immune electron-microscopy', although this has the advantage that it relates the positions of the proteins to the gross features of the ribosomes. This method involves electron microscopic examination of ribosomal particles cross-linked by specific antibodies to individual proteins [141, 167]. It is reassuring that there is broad agreement between the results obtained by these different techniques.

To integrate models such as that in Fig. 12.26 with the three-dimensional models for the structure of rRNA both direct and indirect approaches have been used. Studies of the direct interactions between proteins and RNA have shown that a subset of proteins is involved in the primary interaction with RNA during assembly *in vitro*, and probably also *in vivo*. For 16S RNA and 5S RNA these proteins are fairly well defined as S4, S7, S8, S15, S17 and S20; and L5, L18 and L25, respectively. In the case of 23S RNA an original list of 10 proteins (Ll, L2, L3, L4, L6, L13, L16, L20, L23 and L24) was subsequently extended to about half the proteins of the 50S ribosomal subunit. This may reflect a greater complexity of structure, but probably also illustrates the fact that most proteins in the assembled ribosome are in contact with RNA. Thus, using chemical cross-linking reagents [156], it was possible to demonstrate interaction with RNA for proteins other than those involved in the assembly of ribosomes [134, 166]. This approach has provided more precise information than obtained previously by nuclease-protection experiments. The indirect approach to determining the relative positions of RNA and protein has been to compare their electron microscopic locations on the ribosome. Initially electron-microscopic visualization was only possible for regions of rRNA against which antisera could be raised, such as the 5′- and 3′-ends and certain methylated bases. However, after the secondary structure of rRNA had been established it became possible to use biotinylated oligonucleotides complementary to single-stranded regions, and to visualize these after cross-linking with avidin [168].

On the basis of all these different types of information, models of the relative positions

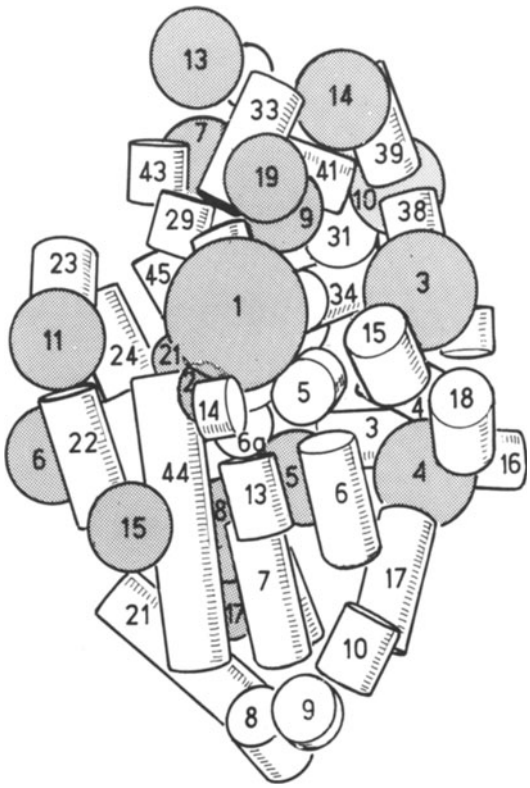

Fig. 12.27 Computer-generated model of the 30S ribosomal subunit of *E. coli* proposed by Brimacombe and co-workers. Helical regions of the rRNA are represented as cylinders and the proteins as spheres, as in Fig. 12.26. (From [156], with permission.)

of RNA and protein in the 30S subunit have been proposed [155, 156], one of which is shown in Fig. 12.27.

12.6.2 Ribosomal structure–function relationships [169, 170]

What is the relationship between the structure of the ribosome already described and its function in catalysing protein biosynthesis? It is evident that some of the components – proteins and portions of the RNA – play only structural roles; moreover bacterial ribosomes with certain individual proteins deleted have been found to be viable [171]. Although it is not yet possible to describe the operation of the ribosome in terms of its constituent parts, it has been possible to identify components of importance to different aspects of ribosomal function. These are present in different domains on the ribosome, and we shall concentrate our discussion on three of these, with emphasis on RNA, rather than on proteins. This is because the experimental evidence is increasingly in line with the evolutionary speculation [172, 173] that RNA has the primary functional role in ribosomes.

(a) *The elongation factor domains*

There is a domain on the 50S subunit involved in the action of EF-G and its intrinsic GTPase activity. It is associated with the 'stalk' of the subunit (Fig. 12.21), which is made up of the elongated L7/L12 tetramer, at the base of which is ribosomal protein L10, and in the vicinity of which is ribosomal protein L11. Proteins L7/L12 are essential for the binding and GTPase activity of EF-G, and L11, although not functionally indispensable, is labelled by affinity analogues of GDP, and has an rRNA binding site (nucleotides 1052–1112) that overlaps the region (1055–1081) to which EF-G can be cross-linked (Fig. 12.28(a)). Strong evidence that this region of 23S RNA is functionally involved in the action of EF-G comes from studies involving the antibiotic thiostrepton, a specific inhibitor of the action of this factor. The natural producer of thiostrepton, *Streptomyces azureus*, unlike *E. coli*, is methylated at nucleotide A_{1067}. Mutation of this same nucleotide from A to a pyrimidine, either under antibiotic selection or by site-directed mutagenesis, conveys resistance to thiostrepton. Furthermore, this region of rRNA (in Domain II of Noller [155]) contains highly conserved portions of primary

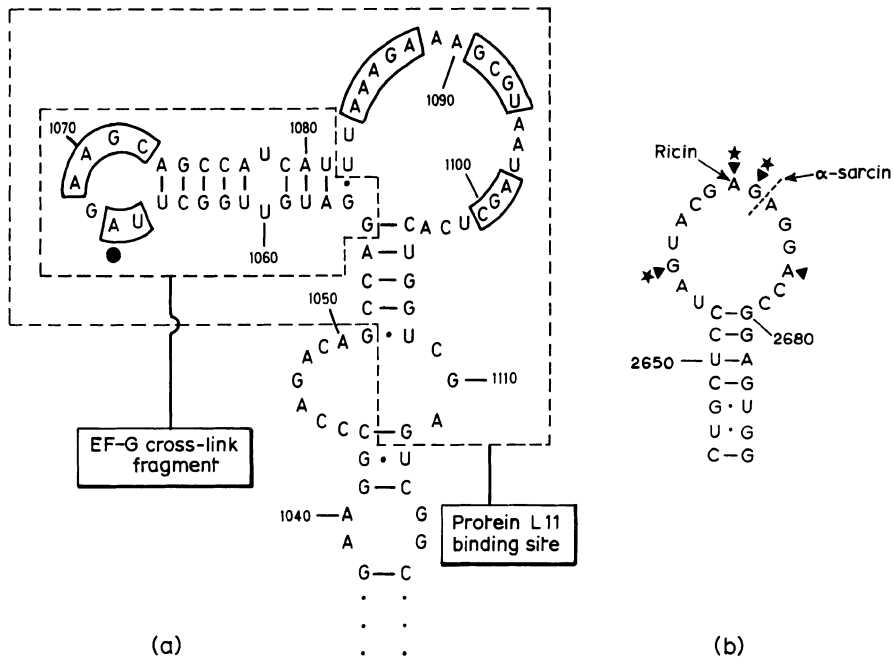

Fig. 12.28 The proposed elongation factor domains of the 50S ribosomal subunit of *E. coli* (after [169] and [155], with permission). (a) Part of Domain II of Noller: ● nucleotide which is methylated in *Streptomyces azureus* and confers resistance to thiostrepton; (b) part of Domain II of Noller: ★, base protected from chemical modification by EF-Tu; ▲, base protected from chemical modification by EF-G. The boxed areas are highly conserved in eubacterial and organelle ribosomes.

structure, consistent with a functional role.

There is a body of evidence suggesting that EF-G and EF-Tu interact with the ribosome in the same region. However, it appears that EF-Tu does not bind to the region just described, consistent with the fact that its action is not inhibited by thiostrepton. It has been shown by RNA 'footprinting' with chemical probes that there is a distinct region of the ribosome where there is overlap between the binding of EF-G and EF-Tu, and this involves the conserved loop in the region of nucleotide 2600 (in Domain VI of Noller [155]). Moreover this loop (Fig. 12.28(b)) is known to be functionally important because it is the target of the cytotoxins α-sarcin (which cleaves between G_{2661} and A_{2662}) and ricin (which removes A_{2660}). One possibility is that these nucleotides are important for triggering the GTPase activities of the two elongation factors to similar effect. In the case of EF-G this would also need to involve the other domain, even though the two regions of 23S rRNA are far apart in both primary and secondary structure.

In focusing on the domains on the ribosome associated with the action of EF-G, one should not neglect the dynamic aspects of the translocation of peptidyl-tRNA from the A- to the P-site. This may involve changes in the conformation of the ribosome as a result of events on the ribosome. It is possible that those changes of conformation that have been detected (by distal changes in protected nucleotides) may involve 'switches' between alternative rRNA secondary structures.

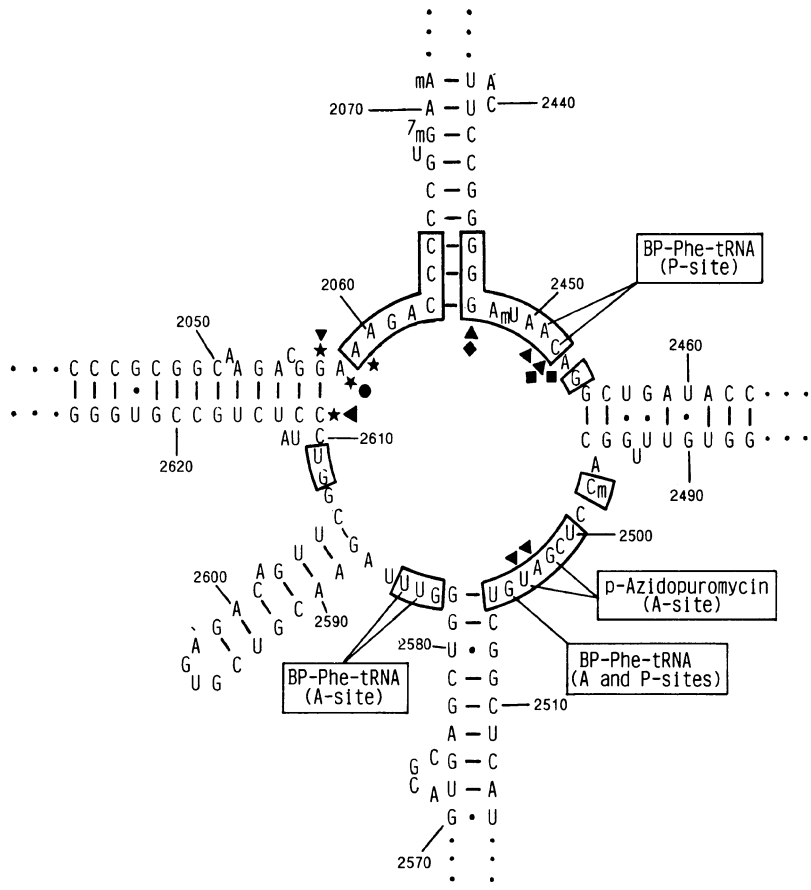

Fig. 12.29 The proposed peptidyltransferase domain of the 50S ribosomal subunit of *E. coli* (after [174], with permission). ★, Site of base change conferring resistance to erythromycin; ▲, site of base change conferring resistance to chloramphenicol; ◆, site of base change conferring resistance to anisomycin; ● nucleotide which is methylated in *Streptomyces erythraeus* and confers resistance to erythromycin; BP-Phe-tRNA, 3-(4′-benzoylphenyl)propionyl-Phe-tRNA. The boxed nucleotides are highly conserved between different organisms.

(b) *The peptidyltransferase domain*

The peptidyltransferase catalytic activity is the most fundamental functional feature of the ribosome, and, before the discovery of ribozymes (section 11.3), it was assumed that it must reside in an individual ribosomal protein. The major tools in probing this activity have been the drugs that specifically inhibit it; but, although several proteins (especially L2, L16 and L27) were implicated in the peptidyltransferase centre by affinity labelling with antibiotic derivatives, none of them satisfied the criteria for catalytic activity. Although, likewise, it has not been possible to demonstrate catalytic activity associated solely with the rRNA, attention became focused on the latter because it was discovered to be the site of point mutations that confer certain mitochondrial ribosomes with resistance to drugs such as chloramphenicol or erythromycin. These

mutations fall in a non-hydrogen-bonded loop (in Domain V of Noller [155]) that forms a junction from which several hairpin stems emanate (Fig. 12.29), and many nucleotides in this junction are conserved. Furthermore, affinity-labelling and 'footprinting' experiments with aminoacyl- or peptidyl-tRNAs or their derivatives allow assignment of parts of the A- and P-sites to conserved nucleotides in this central loop [155, 174, 175]. The key role of the 3′-terminal CCA of tRNAs in the peptidyltransferase 'fragment' reaction (section 12.4.2) suggests that this might base-pair to complementary RNA sequences at the A-and P-sites of the peptidyltransferase centre. There are, however, no totally conserved 3′-UGG-5′ sequences in the 'central loop', and the results of footprinting experiments with tRNAs lacking increasing portions of the CCA sequence suggest that any nucleotides that base-pair with this region would be physically separated from one another in the primary structure and require to be brought together in the tertiary structure [176]. Models for the tertiary structure of 23S rRNA are starting to emerge (e.g. [318]) and should help clarify this question.

(c) *The decoding domain*

This is the domain on the 30S subunit where the codon–anticodon interaction occurs, and this is close to the region of the 'Shine and Dalgarno' interaction between the mRNA and the 3′-end of the 16S RNA (section 12.4.1), which may be modulated by ribosomal protein S21. Although the 'Shine and Dalgarno' interaction is specific to prokaryotes (section 12.4.1), the decoding domain has features that are conserved in prokaryotes and eukaryotes, especially the non-base-paired region from nucleotides 1392–1407 of the 16S RNA of *E. coli* (Fig. 12.30). This region contains a nucleotide (C_{1400}) that can be cross-linked to the modified base, cmo^5U, at the wobble position of $tRNA^{Val}$ when this is bound to the P-site of *E. coli* or yeast ribosomes [177]. Furthermore, resistance to certain antibiotics that cause misreading was associated with methylation or mutation at positions 1405, 1408 and 1409 [178, 179]. Site-directed mutagenesis suggests that C_{1400} is important for ribosome assembly, but that the adjacent base (G_{1401}) is necessary for function [180].

Also in this general area is the highly conserved sequence $m_2^6Am_2^6A$ (nucleotides 1518–1519). The functional importance of this was suggested by the characteristics of an *E. coli* mutant resistant to the antibiotic kasugamycin, which inhibits initiation and causes misreading. The ribosomal RNA from this mutant was unmethylated at these positions [181].

It should be emphasized, however, that not all antibiotics that cause misreading interact with this domain in the 30S subunit. The most thoroughly studied of these is streptomycin, mutations conferring resistance to which occur in both the 900 region (C_{912}) and the 530 region (A_{523} and C_{525}) of 16S RNA, the latter involving a pseudoknot [320]. The role of these latter regions of 16S rRNA may relate to the participation of other species in the control of ribosomal accuracy, as is described in the next section.

12.6.3 Ribosomal optimization of translational accuracy [182]

We have already seen how a mechanism exists to minimize misreading of the genetic message as a result of mischarging of tRNA (section 12.3.2). Misreading can also occur as the result of incorrect codon–anticodon recognition, and reference has been made in section 12.6.2(c), above, to the effects of certain antibiotics on this. Although some mu-

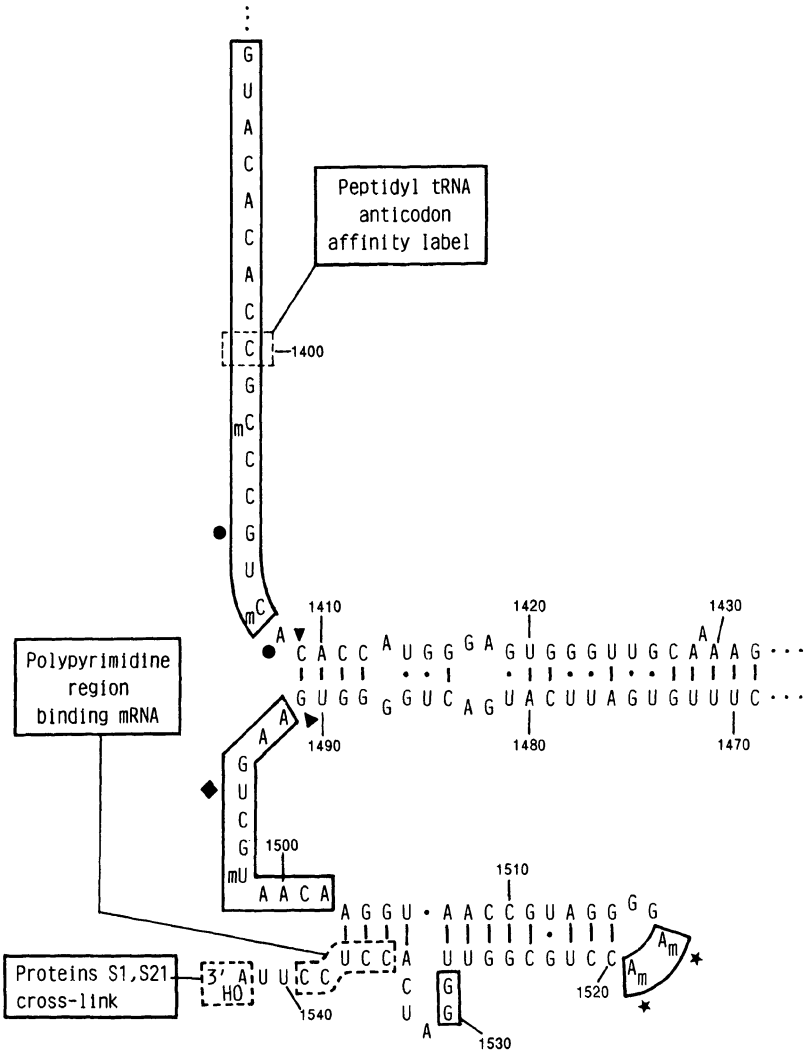

Fig. 12.30 The proposed decoding domain of the 30S ribosomal subunit of *E. coli*. The RNA secondary structure is from [156] and diverges from that of [155] with respect to the hydrogen bonding of the polypyrimidine region at the 3′-end. ★, Nucleotides $m_2^6A_{1518}$ and $m_2^6A_{1519}$, the demethylation of which confers resistance to kasugamycin; ▲, site of base change conferring resistance to paromomycin; ◆, site of base change conferring resistance to hygromycin B; ● nucleotides which are methylated (as m^7G_{1405} and m^1A_{1408}) in organisms producing certain aminoglycoside antibiotics, and confer them with resistance to these. The boxed nucleotides are those not involved in secondary structure that are highly conserved between different organisms.

tations conveying resistance to streptomycin have now been shown to involve RNA, the mutants most pertinent to the question of the accuracy of decoding involve ribosomal proteins. These led to the realization that ribosomes with different error frequencies are possible. Thus, many streptomycin-resistant mutants have alterations in ribosomal protein S12 that confer greater accuracy to the ribosomes; whereas revert-

ants from streptomycin-dependence (also involving S12) have mutations in ribosomal protein S4 that result in increased misreading [183]. Indeed, a role for ribosomal protein S12 in modulating translational accuracy is indicated by the fact that ribosomes lacking S12 are more accurate than those of the wild-type [184].

The accuracy of decoding found *in vivo* is difficult to reconcile with the energy difference between cognate and non-cognate codon–anticodon interactions. Although recognition may involve more than the anticodon of the tRNA (there is a suppressor tRNA with an unaltered anticodon, but a base change in loop I [185]) it seems that the accuracy of the ribosome is achieved by the multiplicative effect of several steps of 'kinetic' proofreading (cf. section 12.3.2), as proposed by Hopfield [186] and Ninio [187]. There is evidence that the first step of proofreading (some workers prefer to call this stage 'recognition', and reserve 'proofreading' for the subsequent checking step) occurs in the binding of the ternary complex, aminoacyl-tRNA.EF-Tu.GTP, when the complex either binds to the ribosome, with hydrolysis of GTP, or dissociates; whereas the second proofreading is of the bound aminoacyl-tRNA.EF-Tu.GDP, which can either form a peptide bond, with liberation of EF-Tu.GDP, or dissociate from the ribosome [188]. In effect, the recognition of the cognate aminoacyl-tRNA by the ribosome is a question of whether the aminoacyl-tRNA dissociates from the ribosome before the GTP is hydrolysed or the EF-Tu.GDP is released: the time taken for GTP hydrolysis provides a 'kinetic yardstick' for proofreading. It has been suggested that the hydrolysis of GTP during the action of EF-G is also involved in proofreading, but, although there are some data that are consistent with this view, it cannot be regarded as proven [182].

The general view of the nature of ribosomal accuracy described above made it possible to resolve an apparent paradox. This was that, if it is possible to select for ribosomes that have increased accuracy, why has Nature not already done this? The most reasonable answer (and the one that is thought to be correct) is that perfect accuracy is only possible if translation is infinitely slow, and hence the actual error frequency is the best compromise between accuracy and speed. However, studies *in vitro* showed that ribosomes of different accuracy exhibited no differences in their maximum speed of translation. This paradox was resolved when it was realized that what was relevant to kinetic proofreading was not the maximum rate that ribosomes could achieve with saturating ternary complex, but the rate at which the ternary complex could interact with the ribosome. When the values of K_m were measured for the interaction of the ternary complex with ribosomes it was found that the K_m increased (and hence the rate of cellular protein synthesis would be expected to decrease) as the accuracy of the ribosome increased [182, 188].

12.7 OTHER PROTEIN-SYNTHESIZING SYSTEMS

12.7.1 Archaebacteria [189, 190]

In the discussion of protein-synthesizing systems so far, a distinction has been made between the nucleo-cytoplasmic system of eukaryotes and the system found in those prokaryotes represented by the commonly studied bacteria (e.g. *E. coli*) and their viruses. There is good reason, however, to think that the prokaryotes comprise not one, but two, kingdoms. These are the *eubacteria*, into which kingdom the majority of bacteria

fall, and the *archaebacteria*, a kingdom containing methanogens, extreme halophiles and certain thermoacidophiles. It is appropriate to discuss briefly the protein-synthesizing system of archaebacteria; not only because this shows certain differences from eubacteria and eukaryotes, but because analysis of the archaebacterial translation apparatus has provided the strongest evidence that the archaebacteria do in fact constitute a separate kingdom.

Archaebacteria have certain structural characteristics that distinguish them from eubacteria. Their cell walls (where these are found) lack peptidoglycan, and their lipids are unique in consisting of branch-chain fatty acids, and being linked to glycerol by an ether, rather than an ester, link. This, in itself, would hardly be sufficient to establish them as a separate kingdom. The evidence that does so comes from phylogenetic studies on ribosomal RNA. Woese and Fox [191] showed that the sequence divergence between the smaller (16S or 18S) rRNAs of eukaryotes and eubacteria clearly distinguished these two kingdoms. (This work was actually done by analysis of oligonucleotides before complete sequences were available.) When the rRNAs of the different archaebacteria were examined it was clear that they were related to one another, but were no more closely related to eubacteria than they were to eukaryotes.

The protein-synthesizing systems of different archaebacteria are not all identical, but, as a generalization, they can be said to show some similarities to those of eubacteria, some to those of eukaryotes, and to have some unique features. Although archaebacterial ribosomes have a sediment coefficient of 70S, rather than 80S, they have an additional morphological feature (a 'bill') also present in eukaryotic ribosomes. Most striking, however, is the sensitivity of archaebacterial ribosomes to antibiotics. Thus, certain archaebacterial ribosomes are insensitive to some antibiotics (chloramphenicol and kanamycin) that had previously been regarded as specific for prokaryotic ribosomes, and sensitive to at least one inhibitor (anisomycin) that had been regarded as specific for eukaryotic ribosomes. The position of archaebacteria in relation to eukaryotes and eubacteria is also apparent from the sequences of their ribosomal proteins [192, 193]. Although the genes for these are organized in clusters similar to those in *E. coli*, only some of the genes in these clusters encode proteins that are related to eubacterial ribosomal proteins. Some are related to eukaryotic ribosomal proteins, and others appear to be unique (although a relationship to eukaryotic proteins not yet characterized cannot be excluded at present). Two proteins of particular interest are the equivalents of *E. coli* L10 and L7/L12, the components of the 'stalk' on the large ribosomal subunit important for translocase function (section 12.6.2(a)). Although it is difficult to see a relationship between the eukaryotic and eubacterial proteins here, the similarity of the archaebacterial proteins to both eubacterial and eukaryotic proteins establishes this and argues for a functionally conserved role.

The sequences of the elongation factors from archaebacteria bear much greater similarity to those of eukaryotes than they do to those of *E. coli*. This is reflected in the fact that, as in eukaryotes, the translocase from archaebacteria has the modified histidine residue that can be ADP-ribosylated by diphtheria toxin, and an EF-Tu/EF-1α equivalent that is insensitive to the antibiotics, pulvomycin and kinomycin. Another apparent difference from eubacteria is the lack of formylation of the methionyl residue on the initiator tRNA, although further characterization of the initiation process is lacking at present.

12.7.2 Mitochondria [194–196]

The majority of mitochondrial proteins are encoded by mRNAs transcribed from nuclear genes, and these mRNAs are translated on 80S ribosomes in the cytosol, the resulting proteins being transported into the mitochondrion (section 12.8). The mitochondrion has its own genome (section 3.4.1) and its own machinery for transcription (section 8.5) and translation. Although the size and coding capacity of mitochondria vary quite widely between species (from *c.* 17 kb in man, to *c.* 2500 kb in some plants) most mitochondria encode their own rRNAs and tRNAs, together with a variable subset of the proteins of the electron-transfer and oxidative phosphorylation systems. The amount of coding DNA in the larger mitochondrial genomes is little greater than that in the smaller ones, as much of the extra DNA in the former is simple-sequence DNA (chapter 3).

It is generally thought that the precursors of mitochondria were aerobic bacteria that were engulfed by, and evolved in a symbiotic relationship with, an anaerobic host cell. Comparisons of rRNA sequences indicate that mitochondria are most closely related to purple photosynthetic bacteria [197]. This relationship between eubacteria and mitochondria is seen in certain functional aspects of the translational machinery, although there are unique features to the mitochondrial translational apparatus [198]. Mitochondrial ribosomes, other than those of plants, differ from the ribosomes of prokaryotes and eukaryotes in lacking the otherwise essential 5S rRNA, vestiges of which can be seen in the rRNA of the large subunit, which may subserve its function. Mitochondrial rRNAs show fewer secondary modifications (methylations, pseudouridine substitutions) than other rRNAs. As already mentioned (cf. Fig. 12.29), the sensitivity of mitochondrial ribosomes to antibiotics shows an extensive similarity to that of eubacteria. The mitochondrial ribosomal proteins (encoded in the nucleus) are not well characterized, but appear to exceed those of bacteria in number. The mitochondrial elongation factor responsible for binding aminoacyl-tRNA is functionally interchangeable with *E. coli* EF-Tu, but the translocase is not interchangeable with EF-G. Mitochondria, like eubacteria, use fMet-tRNA to initiate protein synthesis [199], and there is some evidence that the mitochondrial initiation factors resemble those of prokaryotes [198]. However, the smaller rRNAs from mitochondria lack the bacterial polypyrimidine 3′-sequence; and in any case there is no opportunity for 'Shine and Dalgarno'-type base-pairing to mammalian mitochondrial mRNAs that start directly, or almost directly, with AUG. The unique post-transcriptional editing of protist mitochondrial mRNAs has been described in section 11.8.

Perhaps the most striking features of the mitochondrial translational system (again, with the exception of higher plants) are the deviations from the standard genetic code (section 12.2.3) and the unique manner in which the codons are read by their tRNAs. This seems to be a consequence of the restricted set of mitochondrial tRNAs. The actual number varies between species (mammals have 22, yeast 25), but the number is less than the 32 required to read the standard code (Fig. 12.3) according to the wobble hypothesis (Table 12.1). In the case of mammalian mitochondria (although not in those with larger genomes) there are apparently no separate genes for $tRNA_f^{Met}$ and $tRNA_m^{Met}$, even though both Met-tRNA and fMet-tRNA are found. It is assumed that the single gene gives rise to both tRNA species through differential post-transcriptional modification. As already

described, these restricted sets of tRNAs are able to decode the mitochondrial genome by a different set of wobble rules (Table 12.1). However, there are some protists (*Tetrahymena pyriformis*, *Trypanosoma brucei*, *Chlamydomonas reinhardtii*) in which the number of tRNAs encoded by the mitochondrial genome is too small even for this. In these cases it must be assumed that the tRNAs are encoded in the nucleus and imported from the cytoplasm.

The mitochondrial tRNAs do not conform to the general structural pattern described in section 12.3.1 and summarized in Fig. 12.7, and it is possible that these deviations are important for their unique decoding properties. In particular the mitochondrial tRNAs violate some or all of the following general features previously enumerated: the conservation of the GTΨCRA sequence in loop IV, the constant seven-nucleotide length of loop IV, the pattern of conserved bases in loop I. In individual cases even more striking violations are observed: the mammalian mitochondrial $tRNA^{Ser}$ completely lacks loop I and its stem, and in several tRNAs from *Caenorhabditis elegans* loop IV and its stem have been replaced by a loop of variable size. Nevertheless, these tRNAs may adopt a conformation similar to that of yeast $tRNA^{Phe}$ (Fig. 12.8) [200].

12.7.3 Chloroplasts [201–203]

The protein-synthetic apparatus of chloroplasts resembles that of bacteria even more clearly than does that of mitochondria. Indeed, rRNA sequence comparison (cf. section 12.7.1) shows that chloroplasts are specifically related to cyanobacteria, with which they also share similarities in the structure of their chlorophylls and carotinoids [197]. The genomes of chloroplasts (120–200 kb) contain much more genetic information than those of mitochondria, and the apparent lack of a corresponding intense pressure for space has left them with a full complement of normal-sized rRNAs (including 5S rRNA). Although the standard genetic code is employed in chloroplasts, there are not quite enough tRNAs (31 in tobacco) to decode the mRNAs according to the standard pattern of observed wobble shown in Table 12.1. Unmodified Us are employed in certain anticodons to allow reading of all members of certain four-codon families [37, 202].

The chloroplast genome encodes other components of its protein-synthetic apparatus, including an initiation factor (IF-1), an elongation factor (EF-Tu, although this is absent in some cases) and somewhat less than half the ribosomal proteins. These latter show an average of 44% identity with clearly analogous proteins of *E. coli*, although there is one example of a chloroplast ribosomal protein with no eubacterial homologue [203]. The polypyrimidine sequence capable of 'Shine and Dalgarno' base-pairing is present at the 3′-end of chloroplast 16S RNA, but some chloroplast mRNAs seem to lack canonical Shine and Dalgarno sequences. If initiation codon selection involves a mechanism similar to that in eubacteria it may perhaps operate in a modified form.

12.8 THE CONTROL OF THE CELLULAR LOCATION OF THE PRODUCTS OF TRANSLATION [204]

Proteins synthesized on the cytoplasmic ribosomes of eukaryotic cells can be either retained in the cytoplasm, transferred to subcellular organelles, inserted into the plasma membrane, or secreted. Although, of course, prokaryotes do not have subcellular organelles, the other possibilities

mentioned are also available to them, together with sequestration in the periplasmic space. The molecular mechanisms governing these possible fates of proteins involve recognition of features in their amino acid sequences – *targetting signals*. There are two main classes of such signals: signals that direct the translocation of proteins across a membrane separating two compartments, and signals that cause the retention of a protein in a particular cellular compartment. Some of the signals consist of a conserved amino acid sequence, but more frequently the targetting sequence can only be described in terms of its general chemical nature and position within the protein.

12.8.1 Translocation across the endoplasmic reticulum or the bacterial inner membrane [204–209]

Much work has centred on the execution of the signal for the translocation of proteins across the membrane of the endoplasmic reticulum of eukaryotes, or across the inner membrane of bacteria. This pathway is taken by secreted proteins, which have been the main focus of study, but it is also the common first step in the processes leading to compartmentation of other proteins (see below). Ribosomes synthesizing proteins destined for secretion (unlike the majority of ribosomes synthesizing proteins to be retained intracellularly) are located on the membranes of the rough endoplasmic reticulum of eukaryotes. The proteins are extruded through the membrane as they are synthesized and pass from the cisternae of the endoplasmic reticulum, via the Golgi apparatus, to secretory vacuoles. The chemical basis for the initial membrane translocation, and that occurring in prokaryotes, lies in a largely hydrophobic peptide, the 'signal peptide', at or near the *N*-terminus of the nascent protein [210]. In most cases the signal peptide is an extension of the *N*-terminus of the mature protein, approximately 15–30 amino acids in length, and is later removed by proteolytic cleavage. (In some cases the signal peptide is internal, and in some cases it is not cleaved.) The core of this signal peptide is predominantly hydrophobic, but there are usually basic or neutral residues at the extreme *N*-terminus.

It seems reasonable to suppose that the hydrophobic nature of the signal peptide is necessary for it to cross the membrane. However, in order for this hydrophobic region to interact with the membrane it must remain exposed, and this means that the nascent protein must not be allowed to adopt its native conformation before the interaction can occur. It is now known that there are two mechanisms to achieve this in prokaryotes and lower eukaryotes (such as yeast), one of which is similar to the single well-studied mechanism that operates in higher eukaryotes. In higher eukaryotes the signal peptide interacts with an 11S signal recognition particle (SRP), which contains six polypeptide chains and a 7S RNA species, 260 nucleotides in length, to which the *Alu* family of repeated DNA sequences is related (Fig. 7.36, section 7.7.4(c)). This interaction, which involves the 54 kDa subunit of SRP (SRP54), may temporarily arrest protein synthesis, but, in any case, maintains the exposure of the signal peptide until the SRP interacts with a specific receptor ('SRP receptor' or 'docking protein') on the inner face of the membrane. This allows the ribosome to attach to the membrane, after which the SRP and its receptor are released, translation is resumed, and the nascent peptide is translocated through a specific channel [311] into the intracisternal space. The signal peptide is, in most cases, cleaved by a peptidase as it emerges on the luminal side of the endoplasmic reticulum (Fig. 12.31).

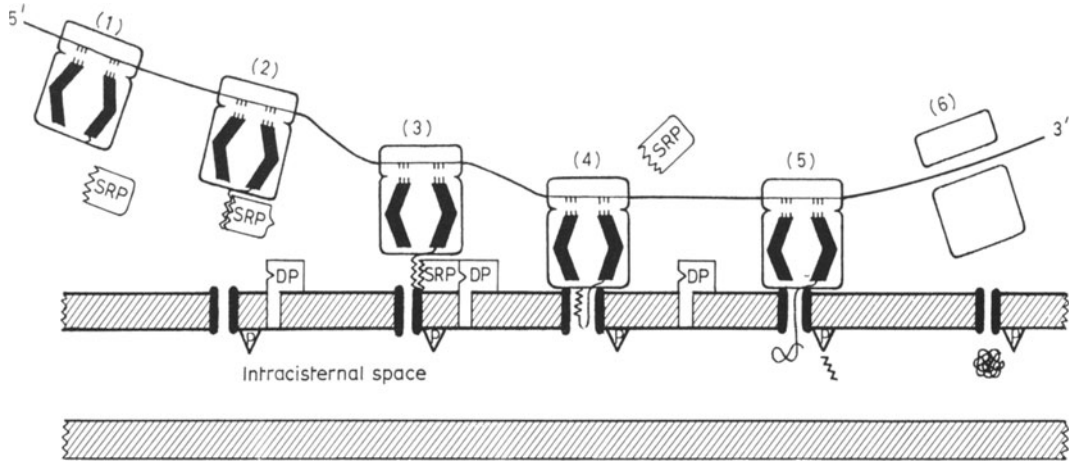

Fig. 12.31 Model for the segregation of secretory proteins into the lumen of the endoplasmic reticulum. SRP, signal recognition particle; DP, docking protein (signal recognition particle receptor); P signal peptidase. For further details see text.

There is evidence that GTP (and presumably its hydrolysis) is involved in the release of the SRP and its receptor [211], and SRP54 and the SRP receptor contain conserved motifs found in other GTPases [208]. It would be consistent with the precedents set by many other GTPases (including the elongation factors of protein synthesis) if the role of the GTP hydrolysis were to disengage these proteins after they had fulfilled their function. By analogy with EF-Tu, there have also been speculations that the time taken for GTP hydrolysis to occur may act as a 'kinetic standard' to allow dissociation of the *N*-terminal peptides of erroneously bound proteins, not destined for translocation across the membrane.

The SRP of prokaryotes escaped detection for a considerable time, even though it was known that bacterial signal peptides were functionally interchangeable with those of eukaryotes. The overall structure of bacterial SRP is still unclear, but it contains a functional counterpart of 7S RNA in the essential and abundant bacterial 4.5S RNA (the function of which had remained elusive for many years), and a structural counterpart to SRP54 in the product of the *ffh* gene. Furthermore, a structural counterpart to the eukaryotic SRP receptor exists in the product of the *ftsY* gene [209]. One of the reasons that the SRP escaped detection for so long in bacteria was that the genetic approach that was applied to the problem only revealed components of the second mechanism for preventing protein folding (see below). However, it did lead to the identification of integral membrane proteins necessary for the process of translocation, and thus common to both mechanisms [212]. There is evidence that two of these, the products of genes *secE* (*prlG*) and *secY* (*prlA*), constitute the actual translocase, but the functions of the other two, the products of *secD* and *secF*, remain to be elucidated. Much less is currently known about the membrane translocase proteins in eukaryotes, and there may be some differences expected here, because in prokaryotes, but not in eukaryotes, the interaction of the negatively charged inner surface of the bacterial cell membrane with a positively charged residue

near the *N*-terminus of the signal peptide is absolutely required for attachment to occur [213]. It also appears that the signal peptidase in eukaryotes is more complex than the general signal peptidase in prokaryotes, although their specificity is similar [214]. This cleavage specificity cannot be characterized simply, but involves recognition of an amino acid with a small side-chain in a relatively hydrophilic environment [210].

The second bacterial mechanism to prevent protein folding before translocation (also found in yeast) involves proteins termed 'molecular chaperones' or 'chaperonins'. Some of these (the GroE protein product of the *groEL*, *groES* genes, and the product of *dnaK*) are heat-shock proteins, which have a more general role in binding partially denatured proteins. However, one, the product of the *secB* gene, is specific for the secretory pathway, and transfers the nascent peptide to SecA, a peripheral membrane protein with ATPase activity [212]. The proteins protected from folding by this mechanism, unlike those protected by the mechanism involving SRP, pass through the membrane post-translationally, rather than co-translationally. The relative roles of the two pathways to the membrane still remains to be resolved. It seems likely that the one involving SRP is the more ancient, and it has been suggested that this pathway is still absolutely required for certain proteins (yet to be identified), although other proteins have evolved the ability to use the second pathway [215]. If this is so, it should be possible to identify features of the signal peptide that confer the ability to use the second pathway.

Finally it should be mentioned that a small minority of secretory proteins do not appear to possess a signal peptide of the type required by the mechanism described above, and it has been suggested that an alternative mechanism must operate in these cases [216].

12.8.2 Fate of non-secretory proteins after translocation across the membrane of the endoplasmic reticulum

Proteins destined for secretion are thought to pass by default from the cisternae of the endoplasmic reticulum to the secretory vacuoles. However, proteins destined for other locations have additional signals to direct them to these or retain them in them.

(a) *Retention as membrane proteins [217]*

Certain plasma membrane proteins do not complete passage through the membrane of the endoplasmic reticulum, but remain embedded in it. The eventual fusion of membrane derived from the endoplasmic reticulum with the plasma membrane results in the internally facing *N*-termini of such proteins becoming exposed to the external surface of the plasma membrane. What stops such membrane proteins being secreted is a hydrophobic sequence at or near the *C*-terminus, the 'stop transfer' sequence, that anchors the protein into the membrane. This is especially well illustrated in the case of the membrane-bound and secreted forms of immunoglobulin heavy chain (Fig. 11.38), which differ only in approximately 20 amino acid residues at their *C*-termini [218]; and this *C*-terminal segment has been shown to convey membrane location when genetically 'grafted' onto a protein that is normally secreted [219].

It should be stressed that there are other orientations of proteins in the plasma membrane: external orientation of the *C*-terminus

or multiple loops back and forth through the membrane. However, proteins with such orientations are inserted into the membrane in specialized manners involving internal signal sequences, for a description of which the reader is directed elsewhere [217, 220].

(b) *Retention in the lumen of the endoplasmic reticulum [221]*

The lumen of the endoplasmic reticulum contains a number of resident soluble proteins (e.g. protein disulphide isomerase, prolyl-4-hydroxylase and others that interact with newly synthesized proteins). These are prevented from being secreted by a specific *C*-terminal sequence, which in animal cells is KDEL, and in yeast is generally HDEL. It is thought that this sequence is bound by a membrane receptor in a salvage compartment, on route to the Golgi, and this receptor recycles the proteins to the endoplasmic reticulum.

(c) *Transfer to the lysosomes [222]*

Proteins such as hydrolases that are resident in the lysosomes of mammalian cells follow the secretory pathway as far as the Golgi. In the *cis* Golgi an unknown feature of these proteins is recognized by *N*-acetylglucosaminyl-1-phosphotransferase, which modifies *N*-linked oligosaccharides on the proteins, and further modification by other enzymes generates manose 6-phosphate groups. These latter constitute the signal that directs the proteins, bound by specific receptors, to the lysosomes.

A different mechanism to the above operates in yeast, and, although the details have not yet been completely elucidated, this involves the action of a protein kinase [312].

12.8.3 Translocation into the mitochondrion and direction to submitochondrial locations [223–225]

Although mitochondria encode some of their own proteins (section 12.7.2) the vast majority are encoded in the nucleus and synthesized on free cytoplasmic ribosomes. Irrespective of their ultimate submitochondrial location, most of these bear a targetting signal that directs them post-translationally to a translocation system which leads through the outer and inner mitochondrial membranes (at a point of contact between the two) into the mitochondrial matrix. Proteins of the outer and inner mitochondrial membrane have additional 'stop transfer' signals that arrest their transfer before they reach the matrix. Proteins of the intermembrane space are re-exported to that compartment on reaching the matrix.

The mitochondrial targetting sequence takes the form of an *N*-terminal extension of the mature protein, but differs from the signal peptide described in 12.8.1 in being more variable in length (*c.* 12 to > 70 residues) and being much more polar in nature. An important property of the targetting sequence is a potential for forming a positively charged amphiphilic α-helix. Soluble cytoplasmic chaperonins (including a family of 70-kDa heat-shock proteins, cytoplasmic hsp70) prevent the mitochondrial protein from folding completely after translation. The protein is then transferred from the cytoplasmic hsp70 to a receptor on the outer mitochondrial membrane, with the concomitant hydrolysis of ATP. Receptor components have been identified [224, 225], but as yeast bearing mutants in these are viable it is assumed that there are alternative receptor systems. The receptor is thought to transfer the protein to the actual transfer pore, one component of which

would appear to be an essential integral outer-membrane protein (ISP42 in yeast, MOM38 in *Neurospora*). An electric potential across the inner membrane is required for transfer of the protein to the mitochondrial matrix, although this is of opposite polarity to that required for periplasmic sequestration in bacteria. Within the matrix the protein first interacts with one of the essential components of the translocation machinery, mitochondrial hsp70. It is then transferred to a second essential matrix chaperonin, mitochondrial hsp60 (a protein structurally related to GroEL of bacteria), with hydrolysis of ATP. There is evidence that hsp60 is required for correct folding of matrix proteins. The final step for matrix proteins is the cleavage of the signal peptide by a protease (MAS protease), which has two non-identical subunits.

It appears that the mechanism for the re-export of proteins to the intermembrane space is basically similar to that employed in bacterial periplasmic sequestration: the relationship of the mitochondrial matrix to the intermembrane space is similar to that of the bacterial cytoplasm to the periplasmic space, and the directionality of the membrane polarity is, of course, the same. In fact, the presumed evolutionary pre-existence of this translocation system allows one to rationalize the initial transfer of proteins to the matrix before transfer to the intermembrane space. A second, hydrophobic, portion of the signal sequence is required for this latter process, and this is finally removed by a protease, perhaps inner-membrane protease I, which is structurally related to bacterial signal peptidase.

Before concluding this section it should be stressed that the general mechanism described above does not apply to all mitochondrial proteins, some of which have specialized individual means of translocation. One example of such an exception is cytochrome *c*, which does not have an *N*-terminal extension and is inserted directly into the intermembrane space, rather than first passing through the matrix.

12.8.4 Translocation across the chloroplast membrane [226, 227]

The problem of translocation to different compartments of the chloroplast has not received the same intense study accorded to the mitochondrion, but the picture emerging is not dissimilar. Initial targetting to the chloroplast is conveyed by an *N*-terminal extension somewhat resembling the mitochondrial one, although with the difference that it is not predicted to form an amphiphilic α-helix. This signal delivers the protein through the double membrane of the chloroplast envelope into the stroma. There is evidence for a requirement for ATP for import into the stroma, but not a membrane potential, and a stromal signal peptidase has been described. From the stroma certain proteins need to be transported further into the lumen of the thylakoids, and this process formally resembles that of mitochondrial transport from the matrix to the intermembrane space. It, too, requires a second, hydrophobic, signal sequence, and, indeed, certain proteins synthesized in the stroma on chloroplast ribosomes have just this single targetting signal.

12.8.5 Translocation to microbodies [228]

Microbodies, which include peroxisomes, glyoxisomes and glycosomes, are structurally related organelles which share the possession of a β-oxidation system. Microbody proteins are synthesized on cytoplasmic ribosomes and their transport is directed by

targetting signals that do not undergo subsequent cleavage. In the case of some of these proteins this targetting signal is a *C*-terminal tripeptide sequence, SKL. However, other proteins appear to carry a different, but as yet undefined, internal targetting signal.

12.8.6 Translocation across the nuclear membrane [229, 230, 313]

At present two methods of post-translational translocation of proteins across the nuclear membrane have been identified. Small proteins pass through the nuclear pores by simple diffusion, and are then concentrated in the nucleus by specific binding to other nuclear proteins. Large proteins, however, require a targetting signal. The targetting signal, which may be as short as a pentapeptide, does not consist of a precise sequence, but contains three or more basic amino acids (lysine being more frequent than arginine). An example of such a sequence is PKKKRKVE of SV40 large T antigen, which can cause nuclear translocation when artificially introduced into proteins normally resident in the cytoplasm. The nuclear targetting sequence does not appear to have to be located at any particular position in the protein, provided that it is exposed. The translocation appears to be a two-step process: initial rapid binding to nuclear pores, probably involving an acidic receptor region, followed by slower ATP-dependent transfer through the pores.

12.9 THE REGULATION OF TRANSLATION

The regulation of protein synthesis by altering the rate of transcription of mRNA has been described in chapter 10. In eukaryotes the sites of synthesis and usage of mRNA are physically separated, and this, and the relative stability of most eukaryotic mRNAs, provide scope for the translational control of protein synthesis. Moreover, even in prokaryotes, examples of translational regulation are to be found. The following account of translational control classifies these examples primarily in terms of the mechanisms involved. The post-transcriptional editing of mRNAs (section 11.8), and regulation of translation by antisense RNAs (section 10.7) and by codon usage (section 12.2.4) have already been described.

12.9.1 mRNA secondary structure

(a) *RNA bacteriophages [231, 232]*

The influence of mRNA secondary structure on translation differs in prokaryotes and eukaryotes because of their different modes of initiation. In the small bacteriophages, such as MS2 and $Q\beta$, there is a striking example of the role of the secondary structure of their mRNA in the differential expression of different cistrons of a polycistronic mRNA. These single-stranded RNA phages direct the translation of four proteins from their structural RNA, and these include the coat protein and the viral subunit of the replicase. During infection, or translation *in vitro*, there is much greater synthesis of coat protein than of replicase; and it was shown that destruction of the secondary structure of RNA by mild formaldehyde treatment or by heating redressed the rate of synthesis of replicase to that of coat protein. The proposed secondary structure for MS2 RNA [233] shows the initiation site of the replicase subunit hydrogen-bonded to part of the coat protein cistron, the initiation site of the latter being well exposed (Fig. 12.32). Thus, it appears that initiation of replicase trans-

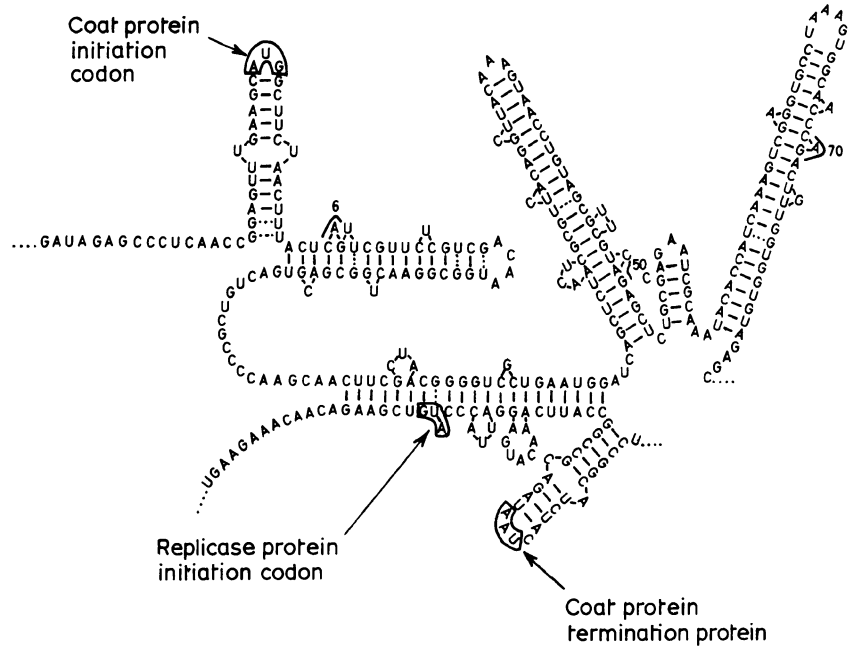

Fig. 12.32 Model for the secondary structure of part of the coat protein cistron of bacteriophage MS2. The positions of the coat protein initiation and termination codons are indicated, as well as the initiation codon for the replicase. The numbers 6, 50 and 70 indicate codons which have been mutated in the related bacteriophage f2 to produce premature termination (see text). (After [233], with permission.)

lation is restricted because it is only accessible when the coat protein cistron is being translated (which would temporally disrupt this hydrogen bonding). Experimental support for this model came from a study of chain-termination mutations in the coat protein gene. An amber mutation at the codon specifying amino acid position 6 of the coat protein severely repressed expression of the cistron for the replicase subunit, but this effect was abolished if the phage RNA was treated with formaldehyde. In contrast, amber mutations at the codons specifying the 50th, 54th and 70th amino acids did not have this effect. It can be seen from inspection of Fig. 12.32 that translation of only the first six codons of the coat protein cistron would not break the putative hydrogen bonds to the replicase subunit initiation site, but that translation to the 50th or subsequent mutant termination codons would. It now appears that the expression of a fourth, minor, protein of MS2 (the lysis protein) is also constrained by the secondary structure (and not by a requirement for frameshifting, as originally proposed).

(b) *Eukaryotic mRNAs*

Although initiation of eukaryotic protein synthesis involves the RNA helicase action of eIF-4A (section 12.5.1), there is experimental evidence that the translational efficiency of mRNAs decreases if artificial stem–loop structures are inserted in the 5′-untranslated region [234]. It is possible that a difference in secondary structure between different mRNAs could affect their translation if eIF-4A were limiting. Examples

might be the α- and β-globin mRNAs, where a difference in translation was eliminated by addition of eIF-4A and eIF-4B [235], and reovirus mRNAs, which are translated to different extents during infection [236]. It has recently been suggested that such secondary structure may be a reason for the translational repression of maternal mRNAs of *Xenopus* oocytes; and evidence has been obtained that at fertilization, when translation of maternal mRNAs occurs, oocytes suddenly become capable of translating artificial mRNAs with extensive secondary structure [237]. In this respect it is interesting that an inhibitor of eIF-4F has been reported in unfertilized sea urchin eggs, which also contain translationally repressed maternal mRNA [238].

12.9.2 RNA–protein interaction

Another aspect of the role of mRNA secondary structure in translation control is that it can provide specific sites for the binding of regulatory proteins, and there are several well studied examples of this in both prokaryotes and eukaryotes.

(a) *Bacteriophage coat protein*

In the case of the translation of the cistrons of bacteriophage RNAs, discussed above, it is thought that the decline in synthesis of the replicase at later times of infection is caused by the binding of the coat protein (as the concentration of this increases) to a region of secondary structure containing the ribosome-binding site [239].

(b) *Autogenous regulation of bacterial ribosomal protein synthesis [240, 241]*

A more intensively studied prokaryotic example, and one involving cellular rather than phage mRNAs, is the autogenous control of the synthesis of bacterial ribosomal proteins. The syntheses of the ribosomal proteins of *E. coli* are closely co-ordinated with one another and with that of rRNA. There are approximately 25 transcriptional units (also referred to as operons) for the 54 ribosomal protein genes, some containing several ribosomal protein genes, whereas others contain only one. Genes for other proteins involved in macromolecular synthesis are also found in these units, some of which are shown in Fig. 12.33. When transducing phages or plasmids were used to increase the copy number of the ribosomal protein genes, without a concomitant increase in the number of rRNA genes, it was found that there was no increase in synthesis of the corresponding ribosomal proteins. This indicated that some form of feedback repression was occurring, and it transpired that a single protein in each transcription unit was responsible for repressing the synthesis of the other members of the unit *in vivo*, as indicated in Fig. 12.33. Experiments *in vitro* (which show consistent, if not identical, results to those *in vivo*) demonstrated that the target of the regulation was the translation of the ribosomal protein mRNA, and that the repressor proteins also inhibited their own translation.

A possible mechanism for the regulation was suggested by the fact that many of the proteins responsible for the feedback inhibition (e.g. S4, S7 and S8) had been shown to bind to rRNA early in the assembly of ribosomes (section 12.6.1). Thus, it was proposed that these proteins were also able to recognize a site on the polycistronic mRNA that resembled their rRNA-binding site, and the binding to which would block translation. Binding to this site would be of lower affinity than that to rRNA so as not to interfere with assembly of ribosomes, but would occur when the ribosomal protein was

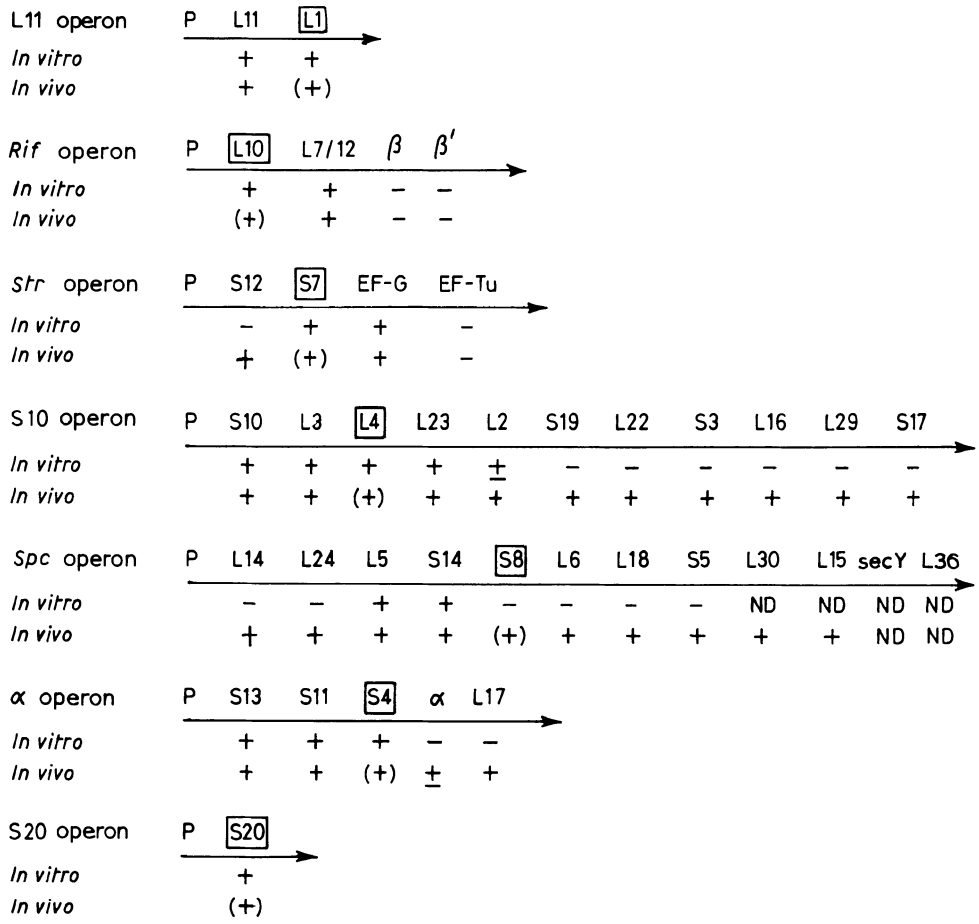

Fig. 12.33 Autogenous regulation of ribosomal protein synthesis in *E. coli*. Each operon is indicated by an arrow, the direction of which is that of its transcription. The promoters are indicated by P and the individual genes by the names of their protein products. Regulatory ribosomal proteins are boxed, and the effects of these on the synthesis of proteins in the same operon *in vivo* and *in vitro* are indicated (+, inhibition; −, no effect; (+), presumed to inhibit *in vivo*; ND, not determined). The autogenous regulation of the synthesis of proteins L14, L24 and S12 operates by retroregulation, involving nucleolytic cleavage of the mRNA [324]. The non-ribosomal proteins in the operons are α, β, and β', subunits of RNA polymerase, and secY, a component of the protein export machinery. (After [240], with permission.)

produced in excess of the available rRNA. As the cotranscribed 16S and 23S rRNAs would regulate the uptake of all rRNA-binding proteins in the same way, the feedback of the excess of these would explain how the translation of the polycistronic mRNAs from different transcription units was also co-ordinately regulated. Extensive mutagenic studies have established that the target of the translational feedback is generally a single site, located near the translational start of the first gene of the transcriptional unit (the *str* and *spc* operons are exceptional: Fig. 12.33). The binding sites on rRNA of certain ribosomal proteins have been determined (section 12.6.1) and,

in the case of S8, L1 and L10, similar secondary structures can be predicted for their regulatory binding sites on the mRNA [241]. A more complex structure, a pseudoknot (Fig. 2.27), has been demonstrated as the recognition site for ribosomal protein S4 in the region of the S13 mRNA initiation codon, and it is assumed, although not yet established, that the S4 binding site on rRNA has a similar tertiary structure. Indeed, in this case, the deduction of the structure of the mRNA has provided an impetus to studies on the more complex tertiary structure of rRNA.

This proposed mechanism of translational regulation (analogues of which appear to operate on the mRNAs for a range of other RNA-binding proteins) raises the question of how the binding of a ribosomal protein to a single site on a polycistronic mRNA can prevent translation starting at separate initiation sites on the same molecule, and why certain members of the transcription units (e.g. EF-Tu and the β and β' subunits of RNA polymerase) are not subject to this feedback regulation (Fig. 12.33). The most likely explanation is that the initiation sites of all but the first mRNA are masked by secondary structure and only become accessible when the first mRNA is translated (cf. the coat protein and replicase genes of the RNA phages – section 12.9.1). This would explain those proteins that escape feedback regulation (EF-Tu etc.) in terms of exposed translational initiation regions, and explain the lack of repression *in vitro* for certain distal mRNA cistrons in terms of exposure through nuclease attack *in vitro*.

(c) *Ferritin mRNA [242]*

The best characterized example of the regulation of a eukaryotic mRNA by interaction with a protein is that of ferritin mRNA. Ferritin is an iron-storage protein, the translation of the mRNA for which increases when the concentration of iron increases. A secondary structure element in the 5′-untranslated region of the mRNA has been demonstrated by mutagenic experiments to be sufficient and necessary for the regulation; and this element has been termed the iron-responsive element (IRE). A 90-kDa protein that specifically binds this element *in vitro* has been identified, and this represses translation of the mRNA, which may also lead to its degradation. It is assumed that *in vivo* the binding of this protein to the IRE is regulated by the concentration of available iron. There is suggestive evidence that this regulation may be mediated by the iron-porphyrin, haemin, which may oxidize free sulphydryl groups on the protein [243, 244]. This may be a general mechanism for iron-binding proteins, as an IRE has also been identified in the 5′-untranslated region of the mRNA for the transferrin receptor [245].

12.9.3 Degradation of mRNA and modulation of its half-life [246, 247]

The destruction of a mRNA will obviously prevent its translation. Although the induction of ribonuclease is used as a regulatory mechanism in certain circumstances, the structural features of individual mRNAs can influence their half-lives, and it is this aspect that will be considered first.

(a) *The role of the poly(A) tail of the mRNA [248, 249]*

There is a large body of evidence indicating that the 3′-poly(A) segment of mRNAs influences their degradation, with the loss of this segment or its reduction to a certain minimum size leaving the mRNA vulnerable to exonuclease action. Thus, the poly(A)

sequences of cytoplasmic mRNAs become progressively shorter with time; histone mRNAs, which lack poly(A), have shorter half-lives than most eukaryotic mRNAs; artificially deadenylated mRNAs are translated less efficiently than normal mRNAs *in vivo* and *in vitro*; and histone mRNAs have been stabilized by artificially polyadenylating them. Although initial polyadenylation occurs in the nucleus, it is possible for the length of the poly(A) segment of mRNAs to be extended in the cytoplasm [250]. This occurs during oogenesis and embryogenesis in certain maternal mRNAs that contain U-rich regions in their 3′-untranslated regions. This is thought to selectively increase the translation of these mRNAs.

Poly(A)-containing mRNA decays exponentially, and this implies that the susceptibility of such mRNA to degradation does not increase with decreasing size of poly(A). Rather, there is evidence that when the size of the poly(A) segment falls below a certain threshold value the mRNA suddenly becomes extremely susceptible to degradation. An approach to understanding the mechanism of this effect has come through the characterization of a specific poly(A)-binding protein containing a four-times tandemly repeated RNA-binding domain [251]. It is the association of this protein with the 3′-poly(A) segment that is thought to protect it from nuclease attack, and the critical poly(A) size below which rapid degradation of mRNA occurs is thought to be the minimum size for binding of the protein.

(b) *The AU-rich instability sequence [252]*

Apart from the cases in which the extent of polyadenylation is increased in the cytoplasm, the 3′-poly(A) segment would appear to be a common determinant of the half-life of eukaryotic mRNAs. Nevertheless, there must be some structural feature of different mRNAs to explain the differences in their half-lives. In the case of some very short-lived mRNAs, e.g. that for the granulocyte–monocyte colony stimulating factor, this has been shown to be an AU-rich component of the 3′-untranslated region of the mRNA. Several other short-lived mRNAs (including the cellular oncogenes, *fos* and *myc*) have a similar sequence, and transfer of this to a stable mRNA, such as globin, dramatically decreases its half-life.

At present nothing is known of the way in which this AU-rich sequence conveys instability on the mRNA. One suggestion is that it competes with the 3′-poly(A) segment for the poly(A)-binding protein [251], although a different protein that specifically binds AU-rich sequences has been described [253]. A further intriguing aspect of the instability of these mRNAs is that their half-lives can be dramatically increased by various mitogenic stimuli [252], and in view of the potential oncogenic nature of the products of some of these mRNAs there is considerable interest in elucidating the mechanism of this regulation.

(c) *Interferon and the degradation of viral RNA [254–257]*

Animal cells infected with viruses produce interferon, a family of glycoproteins which act on adjacent uninfected cells, enabling them to resist infection by inhibiting the replicative cycle of the virus. Interferons are species-specific but are effective against a wide spectrum of animal viruses. (They also have inhibitory effects on cell growth which may be of importance in uninfected cells.) Although in this section we are only concerned with the ability of interferon to inhibit translation of viral mRNAs, it appears probable that the transcription is also inhibited. The common signal for the induction of

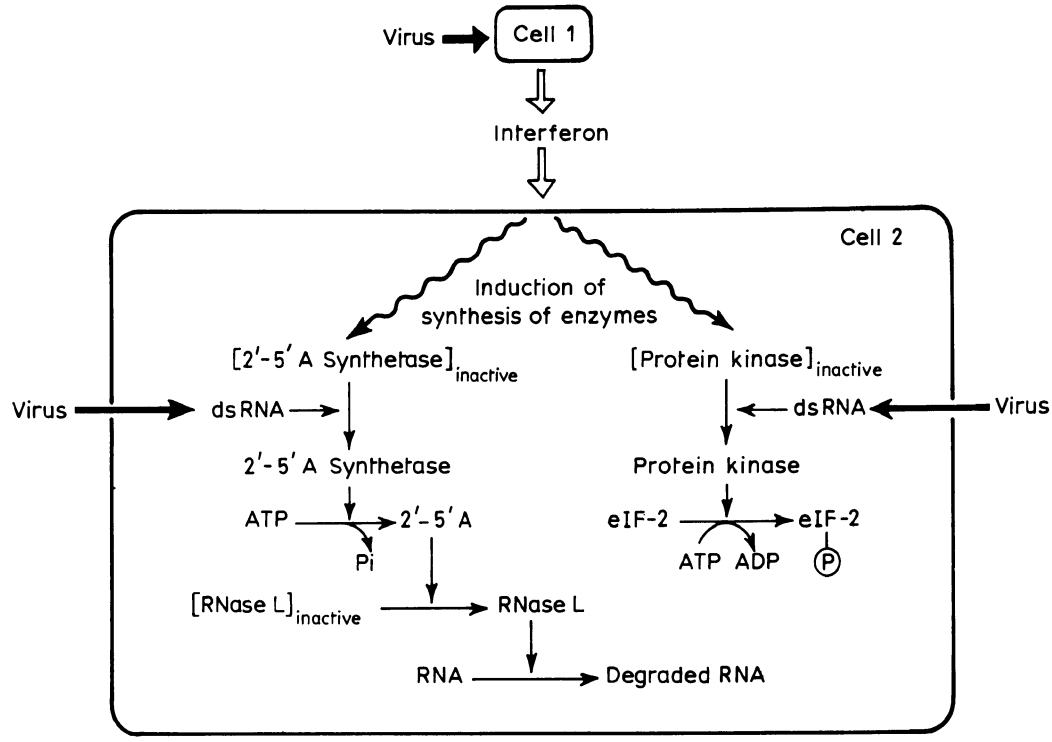

Fig. 12.34 Pathways by which double-stranded RNA and interferon, resulting from infection of cells by viruses, may cause the inhibition of viral protein synthesis. 2′–5′A is an abbreviation for ppp(A2′p5′A2′p5′A)$_n$. For other details, see text.

interferon by different viruses is thought to be double-stranded RNA, which is also the signal that activates the gene-products of interferon that counter the translation of viral mRNA in the 'protected' cell. One of these gene-products is a protein kinase, and this is discussed in section 12.9.4. The other is an enzyme called 2–5A synthetase, which catalyses the synthesis of the trinucleotide, pppA2′p5′A2′p5′A (Fig. 12.34). This nucleotide, in turn, activates the relatively non-specific ribonuclease L (section 4.3.1(e)).

Although it can be seen how this nuclease could prevent the translation of viral mRNA, it is not immediately obvious how host mRNA could escape degradation. Nor is the interferon-induced protein kinase any more specific in its mode of action (section 12.9.4). Nevertheless, there is a considerable body of evidence that both these mechanisms can be important for the translational inhibition *in vivo*, their relative importance depending on the virus. In the case of the action of 2–5A synthetase, the importance of which appears restricted to certain picornaviruses, it has been suggested that the 2′-5′-oligoadenylate required for the activation of the ribonuclease is rapidly degraded as it diffuses away from its site of synthesis, the double-stranded RNA [254]. It is argued that its action would be largely restricted to this viral double-stranded RNA, thought to consist of replicative intermediates in the case of these RNA viruses. (In the case of DNA viruses, overlapping mRNA

transcripts are thought to be the source of double-stranded RNA.) However, it is more difficult to extend this idea to the action of the protein kinase. An alternative way of resolving the problem is possible if one accepts that the inhibition need not, in fact, be specific. The cell, unlike the virus, may be able to compensate for the effect of the nuclease by increasing its synthesis of RNA, and by resuming protein synthesis on dephosphorylation of the cellular target protein after the virus has been destroyed.

(d) *Herpes simplex virus and the degradation of host mRNA [258, 259]*

A situation that appears to be the converse of that described above is found in the case of the inhibition of host protein synthesis by the large DNA virus, herpes simplex type-1. It has been demonstrated genetically that the product of viral gene UL41 is responsible for the early phase of the shut-off of host protein synthesis. However, if this non-essential gene is deleted there is an increase in the half-life of virus mRNA as well as host mRNA. Thus, it would appear that the ribonuclease responsible for the degradation (there are some indications that this is a cellular enzyme, rather than the product of gene UL41 itself) is not intrinsically specific for host mRNA. In this case it may be the concomitant selective shut-off of host transcription that produces an apparent specific effect on translation.

(e) *Autoregulation of the stability of tubulin mRNAs [260]*

The co-ordinated regulation of the α- and β-tubulins (the principal subunits of the microtubules) involves the regulation of the degradation of their mRNAs by a mechanism that is, to date, unique to them. It was demonstrated that the synthesis of tubulin was inhibited when there was a build up of unpolymerized tubulin monomers, and that this involved degradation of their mRNAs. In the case of the mRNA for β-tubulin, it was found that this regulation was conveyed by the section of the mRNA encoding the four *N*-terminal amino acids, MREI. Mutations in this region that did not alter the coding potential did not destroy the regulation, whereas the opposite result was obtained with those that did. It appears that, in some way yet to be determined, the interaction of unpolymerized tubulin monomers with this peptide activates nucleolytic degradation of the mRNA. Although it might be expected that other proteins with the same *N*-terminal sequence might bring into question the specificity of this mechanism, no other example of an *N*-terminal MREI sequence has yet been found. The *N*-terminal sequence of α-tubulin, MREC, is similar enough to expect that it is also recognized by unpolymerized tubulin, although this has not yet been demonstrated.

12.9.4 Protein phosphorylation

(a) *Phosphorylation of eIF-2 [261–263]*

Initial studies of the phosphorylation of eIF-2, performed in reticulocyte lysates, will be summarized before considering the more general but similar phenomenon in cells responding to interferon. Reticulocyte lysates provide a very active protein-synthesizing system, but if haem is not added there is a rapid cessation of protein synthesis. (Haem is rapidly converted in aqueous solution into an oxidized form, haemin, which is what is actually used in such studies.) Although one can rationalize this effect in terms of the co-ordination of the synthesis of the predominant reticulocyte protein, globin, with the availability of its prosthetic

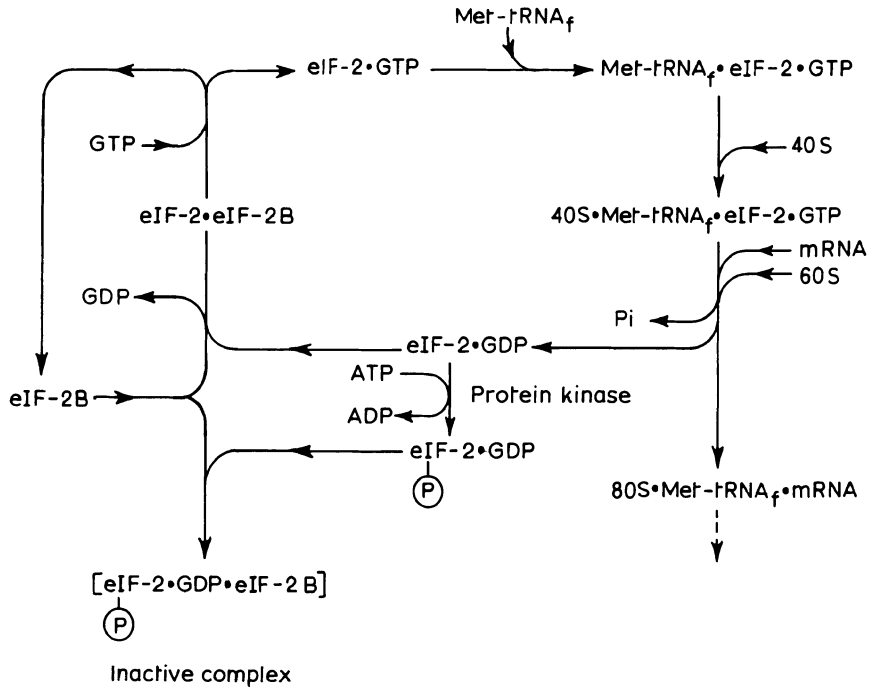

Fig. 12.35 A scheme for the role of factors eIF–2 and eIF–2B in the initiation of eukaryotic protein synthesis, and the effect of phosphorylation of the α-subunit of eIF2. The point at which the phosphorylation of eIF-2 has been indicated is purely arbitrary.

group, haem, it must be emphasized that, in fact, the synthesis of all reticulocyte proteins is similarly affected. It was demonstrated that the step of protein synthesis affected in reticulocytes deprived of haem was the binding of Met-tRNA$_f$ to the 40S ribosomal subunit, the reaction that requires eIF-2 (Fig. 12.19). An inhibitor (known as HCR, haem-controlled repressor; or HRI, haem-regulated inhibitor) was purified from the lysates and shown to be a cytoplasmic protein, pre-existing in an inactive form in normal cells. It was eventually found to be a specific protein kinase that phosphorylates the α-subunit of eIF-2. This has the effect of preventing the liberation of GDP from the eIF-2 released from the ribosome, although eIF-2B still forms a complex with the released species (Fig. 12.35). Although the formation of the complex is readily reversible, phosphorylated eIF-2 has a much greater affinity for eIF-2B than the unphosphorylated factor; in effect sequestering the exchange protein. Phosphorylation of only approximately 20% of eIF-2 can cause complete inhibition of initiation because there are only 20–25% as many molecules of eIF-2B as eIF-2.

It was also found in reticulocytes that low concentrations of double-stranded RNA could also provoke the phosphorylation of eIF-2, but that a different protein kinase (although of the same substrate-specificity) was activated. This double-stranded RNA-activated protein kinase is also found in nucleated cells, although, depending on the cell type, its endogenous concentration may be quite low. However, when cells are treated with interferon there is an induction of the synthesis of this protein kinase [256]. Attempts have been made to explain why

low concentrations, but not high concentrations, of double-stranded RNA are able to activate this enzyme. It is known that activation of the protein kinase by double-stranded RNA is accompanied by autophosphorylation of the protein kinase, and it has been suggested that this is an intermolecular reaction that is essential for the activation. If this is so, then at high concentrations the ratio of double-stranded RNA to protein kinase would be such that any single molecule of RNA would be unlikely to bind the two molecules of protein kinase required for intermolecular autophosphorylation [264].

The ability of the double-stranded RNA-dependent protein kinase to inhibit viral protein synthesis *in vivo* has been clearly shown for adenovirus [265]. Adenovirus synthesizes a small RNA, VAI RNA, which is required to allow translation of late mRNAs, and this acts by binding to the protein kinase and inhibiting its activity. Other viruses also have mechanisms to inactivate this kinase [265]. In HIV-1 (human immunodeficiency virus type 1), it is possible that the stem–loop structure of the Tat-responsive region RNA acts in a similar way to adenovirus VAI RNA [314]. In contrast, vaccinia virus encodes a truncated homologue of eIF-2α lacking the target serine residue phosphorylated by the double-stranded RNA-dependent protein kinase. This has been shown to abrogate the effect of interferon, presumably by binding to the protein kinase as a pseudo-substrate inhibitor [315]. In the case of influenza virus the precise details of the anti-interferon mechanism are not yet known, but it is interesting that this involves a cellular, rather than viral, protein, which would suggest a role in normal metabolism [316]. Certainly, there is a substantial body of evidence that uninfected cells can regulate their protein synthesis through the double-stranded RNA-dependent eIF-2 protein kinase [266].

(b) *The phosphorylation of eukaryotic EF-2*

Although the dogma of translational regulation insists that control can only be exerted at initiation, there is now clear evidence that the eukaryotic translational translocase, EF-2, is the target of regulatory phosphorylation. Thus, it has been shown that the major (perhaps sole) substrate for Ca^{2+}/calmodulin-dependent protein kinase III among cellular proteins *in vitro* is EF-2 [267], that this phosphorylation inactivates the factor [268], and that there are changes in the extent of phosphorylation of EF-2 in certain situations *in vivo*. It is rather curious, however, that several of these situations are typified by an increase, rather than a decrease, in protein biosynthesis (e.g. stimulation of quiescent cells by insulin [269]); and it must be remarked that there is no obvious rationale for the control of protein synthesis by changes in the cellular concentration of calcium ions.

(c) *The phosphorylation of other components of the protein synthetic apparatus*

Other eukaryotic initiation factors besides eIF-2 have been shown to be phosphorylated *in vivo* [266]. Most strongly implicated in regulation is the phosphorylation of eIF-4E, which in certain circumstances shows a correlation with protein synthetic activity. However, as yet, there has been no demonstration of a direct effect of phosphorylation on the activity of the factor *in vitro*.

Certain eukaryotic ribosomal proteins are also phosphorylated *in vivo* [270, 271]. On the 60S subunit, the acidic proteins related to *E. coli* L7/L12 (section 12.6.1) are phosphorylated, and there is evidence that this is a prerequisite for their incorporation into the subunit, an event that appears to occur in the cytoplasm for these proteins [272, 273].

On the 40S subunit, the basic ribosomal protein, designated (eukaryotic) S6, can accept up to five phosphoryl residues per molecule *in vivo*. The phosphorylation of this protein is greatest in rapidly growing cells and there is much interest in the possible role of this phosphorylation in the transduction of extracellular signals for growth inside the cell. Two ribosomal protein S6 kinases have been characterized [274, 275], one with two different protein kinase domains, but the function of the phosphorylation of ribosomal protein S6 is still uncertain. As it is clear that the phosphorylation is not an obligatory requirement for ribosome function, some more subtle regulatory role is indicated. One intriguing possibility is of an interaction with the system controlling the half-life of those unstable mRNAs that are induced following a mitotic stimulus (section 12.9.3).

Finally it should be mentioned that certain eukaryotic aminoacyl-tRNA synthetases are also phosphorylated, although the regulatory significance of this is as yet unclear [276].

12.9.5 Initiation factor proteolysis [264, 277]

The unorthodox mode of initiation adopted by poliovirus mRNA, which lacks the 5′-'cap' structure, has been described in section 12.5.1. Its lack of requirement for a functional eIF-4F complex is the basis of its inhibition of the translation of capped host mRNAs, which involves proteolytic cleavage of p220, the 220-kDa component of eIF-4F. The poliovirus mRNA, like that of Sinbis virus (Fig. 12.20), is translated into a polyprotein which is cleaved by two virally coded proteases. One of these, protease 2A, has been shown to be necessary for the proteolysis of p220, but is not itself directly responsible for the degradation of the factor. It appears that an, as yet uncharacterized, cellular protease is the enzyme responsible, and this raises the possibility that there exist circumstances where this proteolysis is used as a cellular mechanism for the regulation of protein synthesis.

12.9.6 Suppression and frameshifting [278, 279]

The methods of regulation of protein synthesis discussed so far operate within the framework of rigid adherence to the genetic code. However, in both eukaryotes and prokaryotes there are cases, predominantly although not exclusively involving RNA viruses, where regulation involves the disregard of a termination codon (suppression) or its avoidance by change of reading frame (frameshifting).

(a) *Suppression [278–281]*

Suppression of termination codons is used by certain RNA viruses to effect the synthesis of an alternative extended version of a particular protein, containing additional functional potential. An example of this has already been encountered in the case of the alphavirus, Sinbis virus (Fig. 12.20); and bacteriophage Qβ, mentioned in section 12.9.1, synthesizes minor amounts of a fourth protein, thought to be involved in release of progeny phage particles from the bacterium, as a result of readthrough of the coat protein termination codon. Other examples include several plant viruses, such as tobacco mosaic virus, and certain retroviruses, such as Moloney murine leukaemia virus, where a *gag–pol* 'fusion protein' is produced by the relatively efficient suppression of the *gag* termination codon. In the case of the alphaviruses and retroviruses,

there is proteolytic processing of the fusion protein to generate one or more separate functional polypeptides.

There are two requirements for such specific suppression of termination codons: there must be some distinguishing feature in the 'context' of the codon, and there must be a tRNA (a suppressor tRNA) to insert an amino acid in response to this. As regards a context effect, it has long been known to bacterial geneticists that amber mutations (i.e. artificially induced UAG stop codons) are particularly inclined to be 'leaky' (i.e. to be subject to readthrough suppression) but that the degree of 'leakiness' varied with the position of the mutation. It has been shown in experiments in bacteria that UAG codons followed by a purine, especially A, are well suppressed, whereas those followed by a pyrimidine are weakly suppressed; and it is precisely a UAG codon followed by an A that is suppressed in the readthrough of the Qβ coat protein termination codon. This, incidentally, allows one to rationalize in terms of efficient termination the high frequency with which U follows natural bacterial stop codons, and the low frequency with which UAG occurs (section 12.4.3). The tRNA responsible for the suppression in Qβ has been shown to be a normal *E. coli* $tRNA^{Trp}$ with a 3′-ACC-5′ anticodon, which would predict a C:A mismatch in the middle position of the codon. Although an attempt was made to explain the context effect of the A, 3′ to the suppressed termination codon, in terms of base-pairing with the (universal) U, 5′ to the anticodon, this has now been excluded, and at present the molecular basis of the abnormal codon–anticodon interaction is unknown [282].

In eukaryotic viruses there is no consistent pattern to the suppressed termination codon and the following base: in Sinbis virus it is a UGA codon followed by a C, in Moloney murine leukaemia virus a UAG followed by a G, and in tobacco mosaic virus a UAG followed by a C. Furthermore, in the case of Moloney murine leukaemia virus, mutation of the UAG termination codon to UAA did not abolish its suppression. Thus, it seems likely that a wider context is important, as has been found for frameshifting (below), although there are few other data available at present that bear on this question. Instead, the focus of much work in this area has been identifying the tRNAs responsible for the suppression *in vivo*.

A normal mouse $tRNA^{Gln}$ has been implicated as the suppressor of the UAG termination codon of Moloney murine leukaemia virus, and this has an anticodon 3′-GUC-5′, implying a G:U interaction at the first position of the codon [283]. In the case of tobacco mosaic virus, the UAG suppressor in tobacco leaves has been shown to be a normal $tRNA^{Tyr}$ with an anticodon 3′-AΨG-5′, implying abnormal G:G hydrogen bonding in the 'wobble' position. It is interesting that another natural $tRNA^{Tyr}$ does not act as a suppressor. This has the 5′-anticodon G replaced by queuosine (Q), a modified form of G, which shows the same pattern of base-pairing with C in the wobble position. One imagines that the conformation adopted by the termination codon must be different from that usually adopted in order to allow such abnormal decoding to occur. Other potential natural suppressors identified include two calf liver $tRNA^{Leu}$ species with anticodons 3′-GAC-5′ and 3′-AAC-5′, both of which are able to suppress UAG in tobacco mosaic virus RNA *in vitro*.

It is convenient in this section to describe one case in which there is suppression of certain cellular UGA codons, even though it does not have any regulatory significance. Rather, it is the means by which the genetic code is extended to a twenty first amino acid, *selanocysteine*, an analogue of cysteine in which selenium replaces the sulphur atom

[284, 309]. This amino acid is rare, but is found in *E. coli* formate dehydrogenase and mammalian glutathione peroxidase. In both eukaryotes [285] and prokaryotes [286] a specific tRNA, $tRNA^{[Ser]Sec}$, has been identified as responsible. In each case the primary transcript has an anticodon of sequence 3′-ACU-5′, although the U in the 'wobble' position undergoes a variety of modifications in different organisms. The tRNAs have other unusual structural features that presumably ensure that they only recognize UGA codons in the specific context of the mRNAs for proteins containing selanocysteine. In the case of formate dehydrogenase this context includes 39 nucleotides 3′, and 9 nucleotides 5′, to the codon [287], and a common feature of the context in different cases is a stem–loop structure immediately 3′ to the UGA codon. It has been shown in *E. coli* that the $tRNA^{[Ser]Sec}$ is first charged with serine by the normal seryl-tRNA synthetase, and the serine then undergoes enzymic modification to selanocysteine [288]. A special elongation factor, structurally related to EF-Tu (section 12.4.2), is required to bind this tRNA to the ribosome [289].

(b) *Frameshifting [278, 279]*

The synthesis of extended fusion-proteins by frameshifting is encountered in retroviruses, such as Rous sarcoma virus, mouse mammary tumour virus and human immunodeficiency virus (HIV); in coronaviruses, such as the avian infectious bronchitis virus (IBV); in hepadnaviruses, such as mouse hepatitis virus; and in certain transposons such as Ty (section 7.7.3). However, there are also two examples known in which frameshifting is necessary for the synthesis of normal cellular proteins. Frameshifting in the translation of the mRNA encoded by the *E. coli dnaX* gene is responsible for the out-of-frame termination that generates the 47-kDa γ subunit of DNA polymerase in addition to the 71-kDa τ subunit [290]; and frameshifting is necessary in the normal translation of the *E. coli* mRNA for RF-2 [291]. Although mutant tRNAs have been described that can cause abnormal frameshifting, the natural frameshifting involves normal tRNAs, and the focus of research has been on defining the mRNA context responsible.

The frameshifting in all the above cases is to the '−1' reading frame, with the exception of Ty and RF-2, where it is to the '+1' reading frame. In the case of shifting to the '−1' frame, a common feature of the mRNA in the region immediately preceding the position of shift is a sequence containing tandemly repeated nucleotides. This has been shown to be necessary for the frameshifting, and is thought to facilitate slippage of the (normal) tRNA. The most striking examples of such 'slippery sequences' are A AAAAAC and U UUU UUA in certain of the retroviruses (the codon recognized by the tRNA involved in the frameshift is underlined), although G GGA AAC is also encountered. It is not merely the tandem repetition that is important, as individual nucleotides are also essential in particular cases. However, such slippery sequences, though necessary, are not sufficient for frameshifting. A second feature is required that is thought to cause the ribosome to pause for a sufficient length of time to allow the frameshifting to occur. In some cases this takes the form of 3′-secondary structure: in many retroviruses (although not in HIV) a stem–loop, and in IBV a pseudoknot (Fig. 2.27). In Ty (where the slippery sequence, CUU AGG C, does not conform to the pattern of tandem repeats) the pausing is thought to be caused by a preceding AGG codon, which is decoded by a rare $tRNA^{Arg}$ (cf. section 12.2.4). In the case of RF-2 two features are involved: a 5′-Shine and Dalgarno sequence

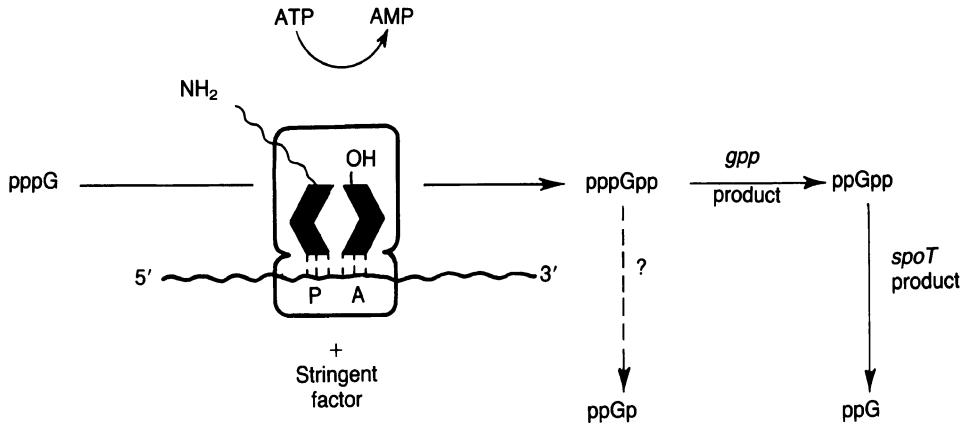

Fig. 12.36 Pathways for the synthesis and degradation of guanine nucleotides involved in the 'stringent response'.

and a UGA termination codon immediately 3′ to the CUU at which the shift occurs. The mechanism by which the Shine and Dalgarno sequence can facilitate frameshifting in this case is obscure, although pausing at the termination codon is easier to envisage. Indeed, the rationale for the frameshifting here is thought to be to regulate the synthesis of RF-2 through control of the ratio of frameshifting to termination. Termination would be expected to vary inversely with the cellular concentration of RF-2, which is required to read the UGA codon. It is curious, however, that there is no comparable regulation in the case of RF-1.

Finally, we should mention some cases where the frameshifting does not involve slippage of the tRNA and ribosome, but 'hopping' over considerable distances. An example of this is gene *60* of bacteriophage T4, where 50 nucleotides are bypassed with almost 100% efficiency. No rationalization of this apparent translational perversity has yet been proposed, but some requirements for its occurrence have been defined, and these include matching sequences at either side of the region bypassed, and a particular region of nascent peptide [292].

12.10 OTHER FUNCTIONS OF THE PROTEIN SYNTHETIC APPARATUS

12.10.1 Ribosomes and the stringent reaction

In addition to their role in protein synthesis, the ribosomes of bacteria such as *E. coli* are able to synthesize guanine nucleotides which act as 'alarmones' to integrate the regulation of RNA and protein synthesis. This phenomenon (which is not found in eukaryotes) occurs when amino acids are limiting, and is known as the 'stringent response' [293, 294]. It results in the selective inhibition of the synthesis of rRNA and tRNA, but not of the bulk of mRNA; hence avoiding wasteful synthesis of more ribosomes and tRNA, and conserving the amino acids provided by protein turnover for the synthesis of essential proteins.

Certain mutations preventing the stringent response map to the *relA* locus. This encodes the 'stringent factor', which is required for the ribosome to respond to the presence of a deacylated tRNA cognate to the codon in the A-site with the synthesis of guanosine 5′-

triphosphate 3′-diphosphate (pppGpp) in an 'idling' reaction (Fig. 12.36). There is further non-ribosomal conversion of pppGpp to yield what is probably the more physiologically important 'alarmone', guanosine 5′-diphosphate 3′-diphosphate (ppGpp); and other guanine nucleotides are generated during the degradation of these so-called 'magic spots'.

The concentration of ppGpp generally correlates well with the inhibition of stable RNA synthesis, but there has been a long controversy regarding the effects of this nucleotide *in vitro* [294, 295]. It is a striking reflection of the shift of scientific interest (and funding) from prokaryotic to eukaryotic systems that, over 20 years after its discovery, the mechanism of action of ppGpp is still unknown.

12.10.2 Other functions of individual ribosomal proteins and elongation factors

In prokaryotes, there are situations in which certain ribosomal proteins and elongation factors are recruited for other roles. It was mentioned in section 12.9.1 that the RNA phages encode a replicase enzyme; however this also requires three host subunits in order to replicate the RNA 'minus' strand. These are EF-Tu, EF-Ts, and ribosomal protein S1. This may merely be a case of the bacteriophage adapting for its own advantage the RNA-binding properties of EF-Tu and ribosomal protein S1. In the case of EF-Tu, it has been suggested that this may involve specific binding to putative tRNA-like features at the 3′-end, although the RNA phages do not show the strong resemblance to tRNA seen in certain plant viruses (section 12.10.3). For ribosomal protein S1 there is the fact that it is able to inhibit the synthesis of the bacteriophage coat protein, and it has been suggested that the presence of this protein in the replicase might either be connected with the temporal lagging of coat protein synthesis behind replicase synthesis, or to play a role in preventing collision between replicase and ribosomes on the RNA [231, 232].

An additional function is suggested in normal cells for the particular eukaryotic ribosomal protein that is termed S27a in mammals and S37 in yeast. In all eukaryotes examined this protein is generated by post-translational processing of a larger species which has *N*-terminal sequences specifying the protein ubiquitin [296]. Ubiquitin is perhaps best known for its role in targetting proteins for degradation [297], but in certain circumstances can apparently have the opposite effect. There is evidence to suggest that the ubiquitin portion remains with the ribosomal protein until it is incorporated into the ribosome [298], and it has been found that a yeast mutant with the ubiquitin tail removed from the ribosomal protein is defective in assembling the 40S ribosomal subunit [299]. It is therefore possible that the ubiquitin sequence acts as a 'chaperonin' for this protein or, a more provoking thought, plays a role in the assembly of the 40S subunit as a whole.

12.10.3 Other functions of aminoacyl-tRNA synthetases

The ribosome is not the only component of the protein synthetic machinery with the capacity to synthesize a nucleotide in an alternative to the reaction normally catalysed. Under abnormal conditions such as heat shock or oxidative stress, the aminoacyl adenylate produced in the first step of the aminoacyl-tRNA synthetase reaction can react with ATP to give diadenosine 5′,5‴-P^1, P^4-tetraphosphate, AppppA [300]:

$$\text{(amino acid-pA)E} + \text{pppA}$$
$$\downarrow$$
$$\text{E} + \text{amino acid} + \text{AppppA}$$

Other related nucleotides have also been reported to be synthesized by this reaction, which can occur in both prokaryotes and eukaryotes. The function of the nucleotide is still in dispute, however. It has been suggested that AppppA may act as an 'alarmone' to induce the heat shock response [300], but this idea is not supported by experiments with *E. coli* mutants in the *apaH* gene, which specifies the hydrolase that degrades the nucleotide [301].

More recently it has been found that specific aminoacyl-tRNA synthetases are components of the intron-splicing machinery of fungal mitochondria [302]. It is possible to produce a mutant Tyr-tRNA synthetase that has lost its ability to charge tRNA with Tyr, but still functions in the splicing reaction. However, this does not exclude a role in RNA recognition of those parts of the enzyme already evolved to recognize features of $tRNA^{Tyr}$, especially as we have already seen that the acceptor stem, with its 5′- and 3′-termini, is one of the features that are recognized by certain tRNA-synthetases (section 12.3.2). The possibility that such recognition is a very ancient feature of life is discussed in the final section of this chapter.

12.10.4 Other functions of tRNAs and tRNA-like structures

In addition to their role on the ribosome, specific tRNAs are important for several other cellular processes. Some of these appear quite natural extensions of the major contemporary role of tRNAs. Thus, certain aminoacyl-tRNAs act as donors of amino acids in the synthesis of the peptide components of the peptidylglycans of bacterial cell walls [303]. A more recently described related function [304] appears to be in the post-translational tagging of the *N*-terminus of proteins with chemically basic amino acids to prepare them for proteolysis by the ubiquitin system [297].

Somewhat more esoteric is the involvement of $tRNA^{Glu}$ in the synthesis of δ-aminolevulinate from glutamate in plants, where it is an important precursor of chlorophyll [305]. For its NADPH-linked reduction to glutamate 1-semialdehyde, glutamate must be in the form of Glu-tRNA. As there seems no obvious chemical advantage in this it has been suggested that the use of the tRNA may be a mechanism for co-ordinating chloroplast protein synthesis and chlorophyll synthesis.

The above examples all involve the amino acid attached to the tRNA. However, there are two virus-related examples involving tRNA structures which, although apparently quite different, may be 'molecular fossils' of a role of this structure that preceded protein biosynthesis. The first example is the use of particular cellular tRNAs by different retroviruses to prime reverse transcription, where a specific sequence on the tRNA hybridizes to the viral 5′-LTR [306], as has already been mentioned in section 6.14. At first sight the second example is quite dissimilar from the first. It is the role of the tRNA-like structures that are found at the 3′-ends of the genomic RNAs of certain plant viruses, such as tobacco mosaic virus [307]. These can be aminoacylated by specific amino acids (His in the virus mentioned), but for many years no plausible function could be suggested for them. An attractive hypothesis has now been proposed by Weiner and Maizels [308], in which it is suggested that the tRNA structures might serve a telomeric function (section 6.16.4). This would be through their recognition by the tRNA nucleotidyl transferase, which

adds CCA to the 3′-end of tRNA transcripts (and the virus RNAs) – a process formally similar to the addition to telomeres of the polymeric C_mA_n sequences by telomerases. The point of convergence between these two examples is Weiner and Maizels' further speculation that tRNAs may have evolved from the 3′-terminal structures of primitive RNA genomes in the 'RNA world' believed to have existed before proteins (and protein synthesis) emerged; and which the proteins that subsequently took over the major part of the replicative function would have had to be able to recognize. In this case the use of extra-genomic tRNA by retroviruses might be more than just fortuitous use of a cellular RNA for its base-complementarity to their genome, but reflect an ancient ability of the replicative enzymes of the virus to recognize the tertiary structure of the tRNA molecule. Indeed, as the original replicative enzymes in an RNA world would have been RNA, a specific base-pairing interaction would also have been a feature of their interaction with the 3′-end of the genome.

REFERENCES

1 Stryer, L. A. (1989) *Biochemistry* (3rd edn), Freeman, New York.
2 Dintzis, H. M. (1961) *Proc. Natl. Acad. Sci. USA*, **47**, 247.
3 Thach, R. E., Cecere, M. A., Sundrararajan, T. A. and Doty, P. (1965) *Proc. Natl. Acad. Sci. USA*, **54**, 1167.
4 Terzaghi, E., Okada, Y., Streisinger, G., Emrich, J., Inouye, M. and Tsugita, A. (1966) *Proc. Natl. Acad. Sci. USA*, **56**, 500.
5 Lodish, H. F. and Jacobsen, M. (1972) *J. Biol. Chem.*, **247**, 3622.
6 Heywood, S. M., Dowben, R. M. and Rich, A. (1967) *Proc. Natl. Acad. Sci. USA*, **57**, 1002.
7 Lacroute, F. and Stent, G. (1968) *J. Mol. Biol.*, **35**, 165.
8 Knopf, P. M. and Lamfrom, H. (1965) *Biochim. Biophys. Acta*, **95**, 398.
9 Kleinkauf, H. and von Döhren, H. (1983) *Trends Biochem. Sci.*, **8**, 281.
10 Crick, F. H. C. (1967) *Proc. R. Soc. London, Ser. B*, **167**, 331.
11 Woese, C. R. (1967) *The Genetic Code. The Molecular Basis for Genetic Expression*, Harper and Row, New York.
12 Jukes, T. H. (1977) in *Comprehensive Biochemistry* (eds M. Florkin and E. H. Stotz), Elsevier, Amsterdam, vol. 24, p. 235.
13 Marshall, R. E., Caskey, C. T. and Nirenberg, M. (1967) *Science*, **155**, 820.
14 Goodman, H. M., Abelson, J., Landy, A., Brenner, S. and Smith, J. D. (1968) *Nature (London)*, **217**, 1019.
15 Marcker, K. A. and Sanger, F. (1964) *J. Mol. Biol.*, **8**, 835.
16 Dube, S. K., Marcker, K. A., Clark, B. F. C. and Cory, S. (1968) *Nature (London)*, **218**, 232.
17 Cory, S., Marcker, K. A., Dube, S. K. and Clark, B. F. C. (1968) *Nature (London)*, **220**, 1039.
18 Marcker, K. A., Clark, B. F. C. and Anderson, J. S. (1966) *Cold Spring Harbor Symp. Quant. Biol.*, **31**, 279.
19 Delk, A. S. and Rabinowitz, J. C. (1974) *Nature (London)*, **252**, 106.
20 Adams, J. M. (1968) *J. Mol. Biol.*, **34**, 571.
21 Miller, C. M., Strauch, K. L., Kukral, A. M., Miller, J. L., Wingfield, P. T. *et al.* (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 2718.
22 Sherman, F., Stewart, J. W. and Tsunasawa, S. (1985) *BioEssays*, **3**, 27.
23 Crick, F. H. C. (1966) *J. Mol. Biol.*, **19**, 548.
24 Borén, T., Elias, P., Samuelsson, T., Claesson, C., Barciszewska, M. *et al.* (1992) *J. Biol. Chem.*
25 Offengand, J. (1982) in *Protein Biosynthesis in Eukaryotes* (ed. R. Pérez-Bercoff), Plenum, New York, p. 1.
26 Lustig, F., Borén, T., Guindy, Y., Elias, P., Samuelsson, T. *et al.* (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 6873.
27 Hillen, W., Egert, E., Lindner, H. J. and Gassen, H. G. (1978) *FEBS Lett.*, **94**, 361.
28 Yokoyama, S., Watanabe, T., Marao, K., Ishikura, H., Yamaizumi, Z. *et al.* (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 4905.
29 Martin, R. P., Sibler, A.-P., Gehrke, C. W., Kuo, K., Edmonds, C. G. *et al.*

(1990) *Biochemistry*, **29**, 956.
30 Sibler, A. P., Dirheimer, G. and Martin, R. P. (1985) *Nucleic Acids Res.*, **13**, 1341.
31 Hunter, W. N., Brown, T. and Kennard, O. (1987) *Nucleic Acids Res.*, **15**, 6589.
32 Grosjean H. J., de Henau S. and Crothers, D. M. (1978) *Proc. Natl. Acad. Sci. USA*, **75**, 610.
33 Lagerkvist, U. (1978) *Proc. Natl. Acad. Sci. USA*, **75**, 1759.
34 Samuelsson, T., Axberg, T., Boren, T. and Lagerkvist, U. (1983) *J. Biol. Chem.*, **258**, 13178.
35 Caron, F. (1990) *Experientia*, **46**, 1106.
36 Osawa, S., Muto, A., Ohama, T., Andachi, Y., Tanaka, R. *et al.* (1990) *Experientia*, **46**, 1097.
37 Jukes, T. H. and Osawa, S. (1990) *Experientia*, **46**, 1117.
38 Grosjean, H. and Fiers, W. (1982) *Gene*, **18**, 199.
39 Konigsberg, W. and Godson, G. N. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 687.
40 Ikemura, T. (1981) *J. Mol. Biol.*, **146**, 1.
41 Bennetzen, J. L. and Hall, B. D. (1982) *J. Biol. Chem.*, **257**, 3026.
42 Garel, J.-P. (1976) *Nature (London)*, **260**, 805.
43 Carter, P. W., Bartkus, J. M. and Calvo, J. M. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 8127.
44 Hoekma, A., Kastelein, R. A., Vasser, M. and de Boer, H. A. (1987) *Mol. Cell. Biol.*, **7**, 2914.
45 Goddard, J. P. (1977) *Prog. Biophys. Mol. Biol.*, **32**, 233.
46 Schevitz, R. W., Podjarny, A. D., Krishnamachari, N., Hughes, J. J. and Sigler, P. B. (1979) *Nature (London)*, **278**, 188.
47 Woo, N. H., Roe, B. A. and Rich, A. (1980) *Nature (London)*, **286**, 346.
48 Westhof, E., Dumas, P. and Moras, D. (1985) *J. Mol. Biol.*, **184**, 119.
49 Quigley, G. J. and Rich, A. (1976) *Science*, **194**, 796.
50 Hoagland, M. B., Stephenson, M. L., Scott, J. F., Hecht, L. I. and Zamecnik, P. C. (1958) *J. Biol. Chem.*, **231**, 241.
51 Lagerkvist, U., Rymo, L. and Waldenström, J. (1966) *J. Biol. Chem.*, **241**, 5391.
52 Julius, D. J., Fraser, T. H. and Rich, A. (1979) *Biochemistry*, **18**, 604.
53 Griffin, B. E., Jarman, M., Reese, C. B., Sulston, J. E. and Trentham, D. R. (1966) *Biochemistry*, **5**, 3638.
54 Zachau, H. G. and Feldman, H. (1965) *Prog. Nucleic Acid Res. Mol. Biol.*, **4**, 217.
55 Yarus, M. (1988) *Cell*, **55**, 739.
56 Schimmel, P. (1991) *Trends Biochem. Sci.*, **16**, 1.
57 Rould, M. A., Perona, J. J., Söll, D. and Steitz, T. A. (1989) *Science*, **246**, 1135.
58 Perona, J. J., Swanson, R. N., Rould, M. A., Steitz, T. A. and Söll, D. (1989) *Science*, **246**, 1152.
59 Chapeville, F., Lipmann, F., von Ehrenstein, G., Weisblum, B., Ray, W. J. *et al.* (1962) *Proc. Natl. Acad. Sci. USA*, **48**, 1086.
60 Fersht, A. R. (1979) in *Transfer RNA: Structure, Properties and Recognition* (eds P. R. Schimmel, D. Söll and J. N. Abelson), Cold Spring Harbor Laboratory Monograph Series, p. 247.
61 Lengyel, P. (1974) in *Ribosomes* (eds M. Nomura, A. Tissières and P. Lengyel), Cold Spring Harbor Laboratory Monograph Series, p. 13.
62 Weissbach, H. (1979) in *Ribosomes: Structure, Function and Genetics* (eds G. Chambliss, G. R. Craven, J. Davies, K. Davis, L. Kahan and M. Nomura), University Park Press, Baltimore, p. 377.
63 Clark, B. F. C. and Petersen, H. U. (1984) *Gene Expression. The Translational Step and its Control*, Munskgaard, Copenhagen.
64 Grunberg-Manago, M. (1980) in *Ribosomes: Structure, Function and Genetics* (eds G. Chambliss, G. R. Craven, J. Davies, K. Davis, L. Kahan and M. Nomura), University Park Press, Baltimore, p. 445.
65 Maitra, U., Stringer, E. A. and Chaudhuri, A. (1982) *Annu. Rev. Biochem.*, **51**, 869.
66 Gualerzi, C. O. and Pon, C. L. (1990) *Biochemistry*, **29**, 5881.
67 Shine, J. and Dalgarno, L. (1974) *Proc. Natl. Acad. Sci. USA*, **71**, 1342.
68 Kozak, M. (1983) Microbiol. Rev., **47**, 1.
69 Seong, B. L. and RajBhandary, U. L. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 8859.
70 Kuechler, E. (1971) *Nature (London) New Biol.*, **234**, 216.
71 Morel-Deville, F., Vachon, G., Sacerdot, C., Cozzone, A., Grunberg-Manago, M.

et al. (1990) *Eur. J. Biochem.*, **188**, 605.
72 Sacerdot, C., Fayat, G., Dessen, P., Springer, M., Plumbridge, J. A. *et al.* (1982) *EMBO J.*, **1**, 311.
73 Reddy, P., Peterkofsky, A. and McKenny, K. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 5656.
74 Gold, L., Stormo, G. and Saunders, R. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 7061.
75 Butler, J. S., Springer, M. and Grunberg-Manago, M. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 4022.
76 Harris, R. J. and Pestka, S. (1977) in *Molecular Mechanisms of Protein Biosynthesis* (eds H. Weissbach and S. Pestka), Academic Press, New York, p. 413.
77 Miller, D. L. and Weissbach, H. (1977) in *Molecular Mechanisms for Protein Biosynthesis* (eds H. Weissbach and S. Pestka), Academic Press, New York, p. 324.
78 Bermek, E. (1978) *Prog. Nucleic Acid Res. Mol. Biol.*, **21**, 63.
79 Clark, B. F. C. (1983) in *DNA Makes RNA Makes Protein* (eds T. Hunt, S. Prentis and J. Tooze), Elsevier Biomedical, Amsterdam, p. 213.
80 Ono, Y., Skoultchi, A., Klein, A. and Lengyel, P. (1968) *Nature (London)*, **220**, 1304.
81 Bourne, H. R., Sanders, D. A. and McCormick, F. (1991) *Nature (London)*, **349**, 117.
82 La Cour, T. F. M., Nyborg, J., Thirup, S. and Clark, B. F. C. (1985) *EMBO J.*, **4**, 2385.
83 Jurnak, F. (1985) *Science*, **230**, 32.
84 Monro, R. E. (1967) *J. Mol. Biol.*, **26**, 147.
85 Lake, J. A. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 1903.
86 Nierhaus, K. H. (1990) *Biochemistry*, **29**, 4997.
87 Moazed, D. and Noller, H. F. (1989) *Nature (London)*, **342**, 142.
88 Caskey, C. T. (1977) in *Molecular Mechanisms of Protein Biosynthesis* (eds H. Weissbach and S. Pestka), Academic Press, New York, p. 443.
89 Tate, W. P. (1984) in *Peptide and Protein Reviews* (ed. M. T. W. Hearn), Marcel Dekker, New York, vol. 2, p. 173.
90 Tate, W. P., Brown, C. P. and Kastner, B. (1990) in *The Ribosome. Structure, Fuction and Evolution* (eds W. E. Hill, A. Dahlberg, R. A. Garrett, P. B. Moore, D. Schlessinger and J. R. Warner), American Society for Microbiology, Washington, p. 393.
91 Capecchi, M. R. and Klein, H. A. (1969) *Cold Spring Harbor Symp. Quant. Biol.*, **34**, 469.
92 Mugola, E. J., Huazi, K. A., Göringer, A. E. and Dahlberg, A. E. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 4162.
93 Kaempfer, R. (1974) in *Ribosomes* (eds M. Nomura, A. Tissières and P. Lengyel), Cold Spring Harbor Monograph Series, p. 679.
94 Naaktgeboren, N., Roobol, K. and Voorma, H. O. (1977) *Eur. J. Biochem.*, **72**, 49.
95 Kohli, J. and Grosjean, H. (1981) *Mol. Gen. Genet.*, **182**, 430.
96 Brown, C. M., Stockwell, P. A., Trotman, C. N. A. and Tate W. P. (1990) *Nucleic Acids Res.*, **18**, 2079.
97 Miller, J. H. and Albertini, A. M. (1983) *J. Mol. Biol.*, **164**, 59.
98 Bossi, L. (1983) *J. Mol. Biol.*, **164**, 73.
99 Pérez-Bercoff, R. (1982) *Protein Biosynthesis in Eukaryotes*, Plenum, New York.
100 Moldave, K. (1985) *Annu. Rev. Biochem.*, **54**, 1109.
101 Pain, V. M. (1986) *Biochem. J.*, **235**, 625.
102 Rhoads, R. E. (1988) *Trends Biochem. Sci.*, **13**, 52.
103 Caskey, C. T., Redfield, B. and Weissbach, H. (1967) *Arch. Biochem. Biophys.*, **120**, 119.
104 Kozak, M. (1978) *Cell*, **15**, 1109.
105 Kozak, M. (1989) *J. Cell Biol.*, **108**, 229.
106 Darnborough, C., Legon, S., Hunt, T. and Jackson, R. (1973) *J. Mol. Biol.*, **76**, 379.
107 Schreier, M. H. and Staehelin, T. (1973) *Nature (London) New Biol.*, **242**, 35.
108 Cigan, A. M., Feng, L. and Donahue, T. F. (1988) *Science*, **242**, 93.
109 Milburn, S. C., Duncan, R. F. and Hershey, J. W. B. (1990) *Arch. Biochem. Biophys.*, **276**, 6.
110 Nielsen, P. J. and Trachsel, H. (1988) *EMBO J.*, **7**, 2097.
111 Lazaris-Karatzas, A., Montine, K. S. and Sonenberg, N. (1990) *Nature (London)*, **345**, 544.
112 Smit-McBride, Z., Schnier, J., Kaufman, R. J. and Hershey, J. W. B. (1989) *J. Biol. Chem.*, **264**, 18527.

113 Strauss, E. G., Rice, C. M. and Strauss, J. H. (1984) *Virology*, **133**, 92.
114 Jackson, R. J., Howell, M. T. and Kaminski, A. (1990) *Trends Biochem. Sci.*, **15**, 477.
115 Hunt, T. (1985) *Nature (London)*, **316**, 580.
116 Hinnebusch, A. G. (1988) *Trends Genet.*, **4**, 169.
117 Hann, S. R., King, M. W., Bentley, D. L., Anderson, C. W. and Eisenman, R. N. (1988) *Cell*, **52**, 185.
118 Florkiewicz, R. Z. and Sommer, A. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 3978.
119 Bernards, A. and de la Monte, S. M. (1990) *EMBO J.*, **9**, 2279.
120 Meerovitch, K., Pelletier, J. and Sonenberg, N. (1989) *Genes Devel.*, **3**, 1026.
121 Macejak, D. G. and Sarnow, P. (1991) *Nature (London)*, **353**, 90.
122 Riis, B., Rattan, S. I. S., Clark, B. F. C. and Merrick, W. C. (1990) *Trends Biochem. Sci.*, **15**, 420.
123 Slobin, L. I. (1980) *Eur. J. Biochem.*, **110**, 555.
124 Nagata, S., Nagashima, K., Tsunetsugu-Yokota, Y., Fujimura, K., Miyazaki, M. *et al.* (1984) *EMBO J.*, **3**, 1825.
125 Collier, R. J. (1975) *Bacteriol. Rev.*, **39**, 54.
126 Van Ness, B. G., Howard, J. B. and Bodley, J. W. (1980) *J. Biol. Chem.*, **255**, 10710.
127 Lee, H. and Iglewski, W. J. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 2703.
128 Neth, R., Monro, R. E., Heller, G., Battaner, E. and Vázquez, D. (1970) *FEBS Lett.*, **6**, 198.
129 Goldstein, J. L., Beaudet, A. L. and Caskey, C. T. (1970) *Proc. Natl. Acad. Sci. USA*, **67**, 99.
130 Craigen, W. J., Cook, R. G., Tate, W. P. and Caskey, C. T. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 3616.
131 Lee, C. C., Craigen, W. J., Muzny, D. M., Harlow, E. and Caskey, C. T. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 3508.
132 Kaempfer, R. (1969) *Nature (London)*, **222**, 950.
133 Wada, K., Aota, S., Tsuchiya, R., Ishibashi, F., Gojobori, T. *et al.* (1990) *Nucleic Acids Res.*, **18**, 2367.
134 Nomura, M., Tissières, A. and Lengyel, P. (1974) *Ribosomes*, Cold Spring Harbor Monograph Series, New York.
135 Chambliss, G., Craven, G. R., Davies, J., Davis, K., Kahan, L. and Nomura, M. (1980) *Ribosomes: Structure, Function and Genetics*, University Park Press, Baltimore.
136 Hardesty, B. and Kramer, G. (1986) *Structure, Fuction and Genetics of Ribosomes*, Springer, New York.
137 Hill, W. E., Dahlberg, A., Garrett, R. A., Moore, P. B., Schlessinger, D. and Warner, J. R. (1990) *The Ribosome. Structure, Function and Evolution*. American Society for Microbiology, Washington.
138 Moore, P. B. (1988) *Nature (London)*, **331**, 223.
139 Bielka, H. (1982) *The Eukaryotic Ribosome*, Springer, Berlin.
140 Wool, I. G. (1990) in *The Ribosome. Structure, Function and Evolution* (eds W. E. Hill, A. Dahlberg, R. A. Garrett, P. B. Moore, D. Schlessinger and J. R. Warner), American Society for Microbiology, Washington, p. 203.
141 Lake, J. (1985) *Annu. Rev. Biochem.*, **54**, 507.
142 Frank, J., Verschoor, A., Wagenknecht, T., Radermacher, M. and Carazo, J.-M. (1988) *Trends Biochem. Sci.*, **13**, 123.
143 Yonath, A. and Wittmann, H. G. (1989) *Trends Biochem. Sci.*, **14**, 329.
144 Milligan, R. A. and Unwin, P. N. T. (1986) *Nature (London)*, **319**, 693.
145 Rich, A. (1974) in *Ribosomes* (eds M. Nomura, A. Tissières and P. Lengyel), Cold Spring Harbor Monograph Series, p. 871.
146 Olsen, H. McK., Grant, P. G., Cooperman, B. S. and Glitz, D. G. (1982) *J. Biol. Chem.*, **257**, 2649.
147 Olsen, H. McK., Lasater, L. S., Cann, P. A. and Glitz, D. G. (1988) *J. Biol. Chem.*, **263**, 15196.
148 Kang, C. and Cantor, C. R. (1985) *J. Mol. Biol.*, **181**, 241.
149 Noller, H. F. (1984) *Annu. Rev. Biochem.*, **53**, 119.
150 Maden, B. E. H., Khan, M. S. N., Hughes, D. G. and Goddard, J. P. (1977) *Biochem. Soc. Symp.*, **42**, 165.
151 Maden, B. E. H. (1980) *Nature (London)*, **288**, 293.
152 Brimacombe, R. (1984) *Trends Biochem. Sci.*, **9**, 273.

153 Walker, T. A. and Pace, N. R. (1983) *Cell*, **33**, 320.
154 Moazed, D., Stern, S. and Noller, H. F. (1986) *J. Mol. Biol.*, **187**, 399.
155 Noller, H. F., Moazed, D., Stern, S., Powers, T., Allen, P. N. *et al.* (1990) in *The Ribosome. Structure, Function and Evolution* (eds W. E. Hill, A. Dahlberg, R. A. Garrett, P. B. Moore, D. Schlessinger and J. R. Warner), American Society for Microbiology, Washington, p. 72.
156 Brimacombe, R., Greuer, B., Mitchell, P., Osswald, M., Rinke-Appel, J. *et al.* (1990) in *The Ribosome. Structure, Function and Evolution* (eds W. E. Hill, A. Dahlberg, R. A. Garrett, P. B. Moore, D. Schlessinger and J. R. Warner), American Society for Microbiology, Washington, p. 93.
157 Westhof, E., Romby, P., Romaniuk, P. J., Ebel, J. P., Ehresmann, C. *et al.* (1989) *J. Mol. Biol.*, **207**, 417.
158 Wittmann-Liebold, B. (1986) in *Structure, Function and Genetics of Ribosomes* (eds B. Hardesty and G. Kramer), Springer, New York, p. 326.
159 Wada, A. and Sako, T. (1987) *J. Biochem. (Tokyo)*, **101**, 817.
160 Wilson, K. S., Appelt, K., Badger, J., Tanaka, I. and White, S. W. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 7251.
161 Leijonmarck, M. and Liljas, A. (1987) *J. Mol. Biol.*, **195**, 555.
162 Lin, A., Wittmann-Liebold, B., McNally, J. and Wool, I. G. (1982) *J. Biol. Chem.*, **257**, 9189.
163 Rich, B. E. and Steitz, J. A. (1987) *Mol. Cell. Biol.*, **7**, 4065.
164 Leader, D. P. (1980) in *Molecular Aspects of Cellular Regulation* (ed. P. Cohen), Elsevier/North-Holland, Amsterdam, vol. 1, p. 203.
165 Traut, R. R., Lambert, J. M., Boileau, G. and Kenny, J. W. (1980) in *Ribosomes. Structure, Function and Genetics* (eds G. Chambliss, G. R. Craven, J. Davies, K. Davis, L. Kahan and M. Nomura), University Park Press, Baltimore, p. 89.
166 Moore, P. B. and Capel, M. (1988) *Annu. Rev. Biophys. Biophys. Chem.*, **17**, 349.
167 Stöffler-Meilicke, M. and Stöffler, G. (1990) in *The Ribosome. Structure, Function and Evolution* (ed. W. E. Hill, A. Dahlberg, R. A. Garrett, P. B. Moore, D. Schlessinger and J. R. Warner), American Society for Microbiology, Washington, p. 123.
168 Oates, M. I., Scheinman, A., Atha, T., Schankweiler, G. and Lake, J. A. (1990) in *The Ribosome. Structure, Function and Evolution* (eds W. E. Hill, A. Dahlberg, R. A. Garrett, P. B. Moore, D. Schlessinger and J. R. Warner), American Society for Microbiology, Washington, p. 180.
169 Garrett, R. (1983) *Trends Biochem. Sci.*, **8**, 189.
170 Dahlberg, A. E. (1989) *Cell*, **57**, 525.
171 Dabbs, E. R., Hasenbank, R., Kastner, B., Rak, K-H., Wartusch, B. *et al.* (1983) *Mol. Gen. Genet.*, **192**, 301.
172 Crick, F. H. C. (1968) *J. Mol. Biol.*, **38**, 367.
173 Darnell, J. E. and Doolittle, W. F. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 1271.
174 Steiner, G., Keuchler, E. and Barta, A. (1988) *EMBO J.*, **7**, 3949.
175 Hall, C. C., Johnson, D. and Cooperman, B. (1988) *Biochemistry*, **27**, 3983.
176 Moazed, D. and Noller, H. F. (1989) *Cell*, **57**, 585.
177 Ehresmann, C., Ehresmann, B., Millon, R., Ebel, J.-P., Nurse, K. and Ofengand, J. (1984) *Biochemistry*, **23**, 429.
178 Li, M., Tzagoloff, A., Underbrink-Lyon, K. and Martin, N. C. (1982) *J. Biol. Chem.*, **257**, 5921.
179 Beauclerk, A. A. D. and Cundliffe, E. (1987) *J. Mol. Biol.*, **193**, 661.
180 Denman, R., Nègre, D., Cunningham, P. R., Nurse, K. *et al.* (1989) *Biochemistry*, **28**, 1012.
181 Helser, T. L., Davies, J. E. and Dahlberg, J. E. (1971) *Nature (London) New Biol.*, **233**, 12.
182 Kurland, C. G., Jörgensen, F., Richter, A., Ehrenberg, M., Bilgin, N. *et al.* (1990) in *The Ribosome. Structure, Function and Evolution* (eds W. E. Hill, A. Dahlberg, R. A. Garrett, P. B. Moore, D. Schlessinger and J. R. Warner), American Society for Microbiology, Washington, p. 513.
183 Gorini, L. (1974) in *Ribosomes* (eds M. Nomura, A. Tissières and P. Lengyel), Cold Spring Harbor Monograph Series, p. 791.
184 Ozaki, M., Mizushima, S. and Nomura, M. (1974) *Nature (London)*, **222**, 333.
185 Hirsch, D. (1971) *J. Mol. Biol.*, **58**, 439.

186 Hopfield, J. J. (1974) *Proc. Natl. Acad. Sci. USA*, **71**, 4135.
187 Ninio, J. (1975) *Biochimie*, **57**, 587.
188 Thompson, R. C. (1988) *Trends Biochem Sci.*, **13**, 91.
189 Woese, C. R. and Wolfe, R. S. (1985) *The Bacteria*, vol. 8, *Archaebacteria*, Academic Press, New York.
190 Kandler, O. and Zillig, W. (1986) *Archaebacteria '85*, Gustav Fischer, New York.
191 Woese, C. R. and Fox, G. E. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5088.
192 Wittmann-Liebold, B., Köpke, A. K. E., Arndt, E., Krömer, W., Hatakeyama, T. *et al.* (1990) in *The Ribosome. Structure, Function and Evolution* (eds W. E. Hill, A. Dahlberg, R. A. Garrett, P. B. Moore, D. Schlessinger and J. R. Warner), American Society for Microbiology, Washington, p. 598.
193 Mathieson, A. T., Auer, J., Ramirez, C. and Böck, A. (1990) in *The Ribosome. Structure, Function and Evolution* (eds W. E. Hill, A. Dahlberg, R. A. Garrett, P. B. Moore, D. Schlessinger and J. R. Warner), American Society for Microbiology, Washington, p. 617.
194 Borst, P., Grivell, L. A. and Groot, G. S. P. (1984) *Trends Biochem. Sci.*, **9**, 48.
195 Attardi, G. (1985) *Int. Rev. Cytol.*, **93**, 93.
196 Chomyn, A. and Attardi, G. (1987) *Curr. Top. Bioeng.*, **15**, 295.
197 Gray, M. W. (1989) *Trends Genet.*, **5**, 294.
198 O'Brien, T. W., Denslow, N. D., Anders, J. C. and Courtney, B. C. (1990) *Biochim. Biophys. Acta*, **1050**, 174.
199 Smith, A. E. and Marcker, K. A. (1968) *J. Mol. Biol.*, **38**, 241.
200 De Bruijn, M. H. L. and Klug, A. (1983) *EMBO J.*, **2**, 1309.
201 Whitfeld, P. R. and Bottomley, W. (1983) *Annu. Rev. Plant Physiol.*, **34**, 279.
202 Sugiura, M. (1989) *Annu. Rev. Cell Biol.*, **5**, 51.
203 Subramanian, A. P., Smooker, P. M. (1990) in *The Ribosome. Structure, Function and Evolution* (eds W. E. Hill, A. Dahlberg, R. A. Garrett, P. B. Moore, D. Schlessinger and J. R. Warner), American Society for Microbiology, Washington, p. 655.
204 Das, R. C. and Robbins, P. W. (1988) *Protein Transfer and Organelle Biogenesis*, Academic Press, San Diego.
205 Pugsley, A. P. (1989) *Protein Targeting*, Academic Press, San Diego.
206 Walter, P., Gilmore, R. and Blobel, G. (1984) *Cell*, **38**, 5.
207 Verner, K. and Schatz, G. (1988) *Science*, **241**, 1307.
208 Rothman, J. E. (1989) *Nature (London)*, **340**, 433.
209 Rapoport, T. A. (1991) *Nature (London)*, **349**, 107.
210 Gierasch, L. M. (1989) *Biochemistry*, **28**, 923.
211 Connolly, T. and Gilmore, R. (1989) *Cell*, **57**, 599.
212 Bieker, K. L. and Silhavy, T. J. (1990) *Trends Genet.*, **6**, 329.
213 Inouye, S., Soberon, X., Franceschini, T., Nakarnura, K., Itakura, K. *et al.* (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 3438.
214 Schelness, G. S. and Blobel, G. (1990) *J. Biol. Chem.*, **265**, 9512.
215 Poritz, M. A., Bernstein, H. D., Straub, K., Zopf, D., Wilhelm, H. and Walter, P. (1990) *Science*, **250**, 1111.
216 Muesch, A., Hartmann, E., Rohde, K., Rubartelli, A., Sitia, R. *et al.* (1990) *Trends Biochem. Sci.*, **15**, 86.
217 Wickner, W. T. and Lodish, H. F. (1985) *Science*, **230**, 400.
218 Rogers, J., Early, P., Carter, C., Calaine, K., Bond, M., Hood, L. *et al.* (1980) *Cell*, **20**, 303.
219 Yost, C. S., Hedgpeth, J. and Lingappa, V. R. (1983) *Cell*, **34**, 759.
220 Singer, S. J., Maher, P. A. and Yaffe, M. P. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 1960.
221 Pelham, H. R. B. (1990) *Trends Biochem. Sci.*, **15**, 483.
222 Pfeffer, S. R. and Rothman, J. E. (1987) *Annu. Rev. Biochem.*, **56**, 829.
223 Douglas, M. G., McCammon, M. T. and Vassarotti, A. (1986) *Microbiol. Rev.*, **50**, 166.
224 Hartl, F.-Z. and Neupert, W. (1990) *Science*, **247**, 930.
225 Baker, K. P. and Schatz, G. (1991) *Nature (London)*, **349**, 205.
226 Keegstra, K. and Bauele, C. (1988) *BioEssays*, **9**, 15.
227 Smeekens, S., Weisbeek, P. and Robinson, C. (1990) *Trends Biochem. Sci.*, **15**, 73.
228 Borst, P. (1989) *Biochim. Biophys. Acta*, **1008**, 1.

229 Dingwall, C. and Laskey, R. A. (1986) *Annu. Rev. Cell Biol.*, **2**, 367.
230 Roberts, B. (1989) *Biochim. Biophys. Acta*, **1008**, 263.
231 Hindley, J. (1973) *Prog. Biophys. Mol. Biol.*, **26**, 269.
232 Zinder, N. D. (1975) *RNA Phages*, Cold Spring Harbor Monograph Series.
233 Min Jou, W., Haegeman, G., Ysebaert, M. and Fiers, W. (1972) *Nature (London)*, **237**, 82.
234 Kozak, M. (1988) *Mol. Cell. Biol.*, **8**, 2737.
235 Kabat, D. and Chappell, M. R. (1977) *J. Biol. Chem.*, **252**, 2684.
236 Lawson, T. G., Cladaras, M. H., Ray, B. K., Lee, K. A., Abrahamson, R. D. *et al.* (1988) *J. Biol. Chem.*, **263**, 7266.
237 Fu, L., Ye, R., Browder, L. W. and Johnston, R. N. (1991) *Science*, **251**, 807.
238 Huang, W.-I., Hansen, L. J., Merrick, W. C. and Jagus, R. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 6359.
239 Romaniuk, P. J., Lowary, P., Wu, H.-N, Stormo, G. and Uhlenbeck, O. C. (1987) *Biochemistry*, **26**, 1563.
240 Nomura, M., Gourse, R. and Baughman, G. (1984) *Annu. Rev. Biochem.*, **53**, 75.
241 Draper, D. E. (1989) *Trends Biochem. Sci.*, **15**, 73.
242 Starzyk, R. M. (1988) *Trends Biochem. Sci.*, **13**, 119.
243 Haile, D. J., Hentze, M. W., Rouault, T. A., Harford, J. B. and Klausner, R. D. (1989) *Mol. Cell. Biol.*, **9**, 5055.
244 Lin, J.-J., Daniels-McQueen, S., Patino, M. M., Gaffield, L., Walden, W. E. *et al.* (1990) *Science*, **247**, 74.
245 Müllner, E. W. and Kühn, L. C. (1988) *Cell*, **53**, 815.
246 Raghow, R. (1987) *Trends Biochem. Sci.*, **12**, 358.
247 Cleveland, D. W. and Yen, T. J. (1989) *New Biologist*, **1**, 121.
248 Brawerman, G. (1981) *Crit. Rev. Biochem.*, **10**, 1.
249 Jackson, R. J. and Standart, N. (1990) *Cell*, **62**, 15.
250 Wickens, M. (1990) *Trends Biochem. Sci.*, **15**, 277.
251 Bernstein, P. and Ross, J. (1989) *Trends Biochem. Sci.*, **14**, 373.
252 Shaw, G. and Kamen, R. (1986) *Cell*, **46**, 659.
253 Malter, J. S. (1989) *Science*, **246**, 664.
254 Baglioni, C. (1979) *Cell*, **17**, 255.
255 Lengyel, P. (1982) in *Protein Biosynthesis in Eukaryotes* (ed. R. Pérez-Bercoff), Plenum, New York, p. 459.
256 Pestka, S., Langer, J. A., Zoon, K. C. and Samuel, C. E. (1987) *Annu. Rev. Biochem.*, **56**, 727.
257 Samuel, C. E. (1991) *Virology*, **183**, 1.
258 Kwong, A. D. and Frenkel, N. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 1926.
259 Kwong, A. D. and Frenkel, N. (1989) *J. Virol.*, **63**, 4834.
260 Cleveland, D. W. (1988) *Trends Biochem. Sci.*, **13**, 339.
261 Farrell, P. J., Balkow, K., Hunt, T. and Jackson, R. J. (1977) *Cell*, **11**, 187.
262 Ochoa, S. (1983) *Arch. Biochem. Biophys.*, **223**, 325.
263 Voorma, H. O., Goumans, H., Amesz, H. and Benne, R. (1983) *Curr. Topics Cell Regul.*, **22**, 51.
264 Schneider, R. J. and Shenk, T. (1987) *Annu. Rev. Biochem.*, **56**, 317.
265 Sonenberg, N. (1990) *New Biologist*, **2**, 402.
266 Hershey, J. W. B. (1989) *J. Biol. Chem.*, **264**, 20823.
267 Nairn, A. C. and Palfrey, H. C. (1987) *J. Biol. Chem.*, **262**, 17299.
268 Ryazanov, A. G., Shestakova, E. A. and Natapov, P. G. (1988) *Nature (London)*, **334**, 170.
269 Levenson, R. M. and Blackshear, P. J. (1989) *J. Biol. Chem.*, **264**, 19984.
270 Leader, D. P. (1980) in *Molecular Aspects of Cellular Regulation*, vol. 1 (ed. P. Cohen), Elsevier/North Holland, Amsterdam, p. 203.
271 Kozma, S. C., Ferrari, S. and Thomas, G. (1989) *Cell Signal.*, **1**, 219.
272 MacConnell, W. P. and Kaplan, N. O. (1982) *J. Biol. Chem.*, **257**, 5359.
273 Vidales, F. J., Robles, M. T. S. and Balleta, J. P. G. (1984) *Biochemistry*, **23**, 390.
274 Jones, S. W., Erikson, E., Blenis, J., Maller, J. L. and Erikson, R. L. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 3377.
275 Kozma, S. C., Ferrari, S., Bassand, P., Siegmann, M., Totty, N. and Thomas, G. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 7365.
276 Clemens, M. (1990) *Trends Biochem. Sci.*, **15**, 172.
277 Sonnenberg, N. (1990) *Curr. Top. Microbiol. Immunol.*, **161**, 23.

278 Valle, R. P. C. and Morch, M.-D. (1988) *FEBS Lett.*, **235**, 1.
279 Atkins, J. F., Weiss, R. B. and Gestland, R. F. (1990) *Cell*, **62**, 413.
280 Murgola, E. F. (1985) *Annu. Rev. Genet.*, **19**, 57.
281 Eggertson, G. and Söll, D. (1988) *Microbiol. Rev.*, **52**, 354.
282 Ayer, D. and Yarus, M. (1986) *Science*, **231**, 393.
283 Kuchino, Y., Beier, H., Akita, N. and Nishimura, S. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 2668.
284 Chambers, I. and Harrison, P. R. (1987) *Trends Biochem. Sci.*, **12**, 255.
285 Lee, B. J., Rajagopalan, M, Kim, Y. S., You, K. H., Jacobson, B. *et al.* (1990) *Mol. Cell. Biol.*, **10**, 1940.
286 Schön, A., Böck, A., Orr, G., Sprinzl, M. and Söll, D. (1989) *Nucleic Acids Res.*, **18**, 1989.
287 Zinoni, F., Heider, J. and Böck, A. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 4660.
288 Leinfelder, W., Forchhammer, K., Veprek, B., Zehelein, E. and Böck, A. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 543.
289 Forchhammer, K., Leinfelder, W. and Böck, A. (1969) *Nature (London)*, **342**, 453.
290 Tsuchihashi, Z. and Kornberg, A. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 2516.
291 Craigen, W. J. and Caskey, T. (1987) *Cell*, **50**, 1.
292 Weiss, R. B., Huang, W. M. and Dunn, D. M. (1990) *Cell*, **62**, 117.
293 Edlin, G. and Broda, P. (1968) *Bacteriol. Rev.*, **32**, 206.
294 Gallant, J. A. (1979) *Annu. Rev. Genet.*, **13**, 393.
295 Lamond, A. I. and Travers, A. A. (1985) *Cell*, **40**, 319.
296 Warner, J. R. (1989) *Nature (London)*, **338**, 379.
297 Finley, D. and Varshavsky, A. (1985) *Trends Biochem. Sci.*, **10**, 343.
298 Müller-Taubenberger, A., Graak, H.-R., Grohmann, L., Schleicher, M. and Gerisch, G. (1989) *J. Biol. Chem.*, **264**, 5319.
299 Finley, D., Bartel, B. and Varshavsky, A. (1990) in *The Ribosome. Structure, Function and Evolution* (eds W. E. Hill, A. Dahlberg, R. A. Garrett, P. B. Moore, D. Schlessinger and J. R. Warner), American Society for Microbiology, Washington, p. 636.
300 Varshavsky, A. (1983) *Cell*, **34**, 711.
301 Farr, S. B., Arnosti, D. N., Chamberlin, M. J. and Ames, B. N. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 5010.
302 Benne, R. (1988) *Trends Genet.*, **4**, 181.
303 Soffer, R. L. (1974) *Adv. Enzymol.*, **40**, 91.
304 Ferber, S. and Ciechanover, A. (1987) *Nature (London)*, **326**, 808.
305 Kannangara, C. G., Gough, S. P., Bruyant, P., Hoober, J. K. *et al.* (1988) *Trends Biochem. Sci.*, **13**, 139.
306 Litvak, S. and Arya, A. (1982) *Trends Biochem. Sci.*, **7**, 361.
307 Haenni, A.-L., Joshi, S. and Chapeville, F. (1982) *Prog. Nucleic Acid Res.*, **27**, 85.
308 Weiner, A. M. and Maizels, N. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 7383.
309 Burk, R. F. (1991) *FASEB J.*, **5**, 2274.
310 Meyer, F., Schmidt, H. J., Plümper, E., Haslik, A., Mersmann, G. *et al.* (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 3758.
311 Lingappa, V. R. (1991) *Cell*, **65**, 527.
312 Herman, P. K., Stark, J. H., DeModena, J. A. and Emr, S. D. (1991) *Cell*, **64**, 425.
313 Silver, P. A. (1991) *Cell*, **64**, 489.
314 Gunnery, S., Rice, A. P., Robertson, H. and Mathews, M. B. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 8687.
315 Beattie, E., Tartaglia, J. and Paoletti, E. (1991) *Virology*, **183**, 419.
316 Lee, T. G., Tomita, J., Hovanessian, A. G. and Katze, M. G. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 6208.
317 Heus, H. A. and Pardi, A. (1991) *Science*, **253**, 191.
318 Mitchell, P., Osswald, M., Schueler, D. and Brimacombe, R. (1990) *Nucleic Acids Res.*, **18**, 4325.
319 Altman, S. (1990) *J. Biol. Chem.*, **265**, 20053.
320 Powers, T. and Noller, H. F. (1991) *EMBO J.*, **10**, 2203.
321 Leberman, R., Härtlein, M. and Cusack, S. (1991) *Biochim. Biophys. Acta*, **1089**, 287.
322 Rould, M. A., Perona, J. J. and Steitz, T. A. (1991) *Nature (London)*, **352**, 213.
323 Ruff, M., Krishnaswamy, S., Boeglin, M., Poterszman, A., Mitschler, A. *et al.* (1991) *Science*, **252**, 1683.
324 Mattheakis, L., Vu, L., Sor, F. and Nomura, M. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 448.