Universal Genetic Code and its Natural Variations

Kimitsuna Watanabe, *National Institute of Advanced Industrial Science and Technology, Tokyo, Japan*

Tsutomu Suzuki, *University of Tokyo, Tokyo, Japan*

The universal genetic code is a common language for almost all organisms to translate nucleotide sequences of deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) to amino acid sequences of proteins. However, the genetic code is still evolved. Nonuniversal genetic codes are found in some organisms and organelles. Aminoacyl-transfer RNA (tRNA) synthetases and RNA modifications play a critical role in reassignment of the genetic code.

Advanced article

Article Contents

Online posting date: 14th March 2008

Introduction

The genetic code governs the relationship between the 20 amino acids principally involved in the synthesis of proteins in present-day organisms and all 64 possible arrangements of the four bases, adenine (A), guanine (G), uracil (U) and cytosine (C), in the form of triplets (codons). Almost all extant organisms on earth use a 'universal' genetic code in which 61 triplet sequences correspond to particular amino acids and the remaining three are termination codons. The universal genetic code is believed to appear as a result of a 'frozen accident'. However, a number of organisms as well as mitochondria use 'nonuniversal' codes in which certain codon triplets deviate from those of the universal code. In this article, the molecular bases by which the genetic code and its variants are maintained in organisms are described.

The Universal Genetic Code and its Properties

In the universal genetic code, which was established in 1966 (**Table 1**), each of 61 codons is assigned to a specific amino acid and the remaining three are termination codons. As the number of amino acids is 20, more than one codon can code for the same amino acid, i.e. the code is degenerated. In such codons, which are called 'synonymous', only the third letter of the triplet is changeable; the amino acid phenylalanine (Phe), for example, has two synonymous codons, UUU and UUC. A box in which four codons are synonymous is called a 'family box' or a 'four-codon box', and a set in which a single amino acid has two codons is called a 'two-codon set'. There are eight family boxes for specific amino acids (Ala, Arg, Gly, Leu, Pro, Ser, Thr and Val) and 13 two-codon sets (Asn, Asp, Arg, Cys, Gln, Glu, His, Ile, Leu, Lys, Phe, Ser and Tyr). Irregular cases are seen for Arg, Leu and Ser, each of which (e.g. Arg) has six synonymous codons (AGA, AGG, CGU, CGC, CGA and CGG) consisting of a family box (CGU, CGC, CGA and CGG) and a two-codon set (AGA and AGG). Ile possesses three synonymous codons (AUU, AUC and AUA) – a two-codon set (AUU and AUC) and a single codon (AUA). The other two codons, AUG and UGG, correspond respectively to single amino acids, Met and Trp. The remaining three codons (UAG, UAA and UGA) are stop codons that terminate protein synthesis and do not correspond to any amino acid. They are often referred to by the nicknames, amber for UAG, ochre for UAA and opal for UGA. **See also**: Genetic Code: Introduction; Nucleic Acids as Genetic Material

The genetic code is restricted by the anticodons of transfer ribonucleic acids (tRNAs), the 'adaptor' molecules that attach specific amino acids, bind to specific codons in messenger RNAs (mRNAs) placed on the ribosomal A site. α-amino group of amino acid on the tRNA attacks the carbonyl carbon of peptidyl bond of the P site tRNA, so as to form a new peptide bond. This reaction is catalyzed by peptidyl-transferase centre in 23S ribosomal RNA (rRNA) of the ribosomes. Anticodons (positions 34, 35 and 36 from the 5′ end of the tRNA molecule) and codons (positions 1, 2 and 3 on the mRNA) pair by hydrogen bonding, in which the nucleosides in the second and the third positions (35 and 36) of an anticodon base pair with those in the second and first positions (2 and 1) of a codon through Watson–Crick-type pairing rules, i.e. A pairs with U and G pairs with C (**Figure 1a**). These two Watson–Crick-type base pairs in codon–anticodon interaction are strictly monitored by 16S rRNA of the ribosomes. However, the pairing between the base at

ELS subject area: Genetics and Molecular Biology

How to cite:
Watanabe, Kimitsuna; and, Suzuki, Tsutomu (March 2008) Universal Genetic Code and its Natural Variations. In: Encyclopedia of Life Sciences (ELS). John Wiley & Sons, Ltd: Chichester.
DOI: 10.1002/9780470015902.a0000810.pub2

Table 1 Universal (U) and animal mitochondrial (M) genetic code tables

	U	M		U	M		U	M		U	M
UUU	Phe	Phe	UCU	Ser	Ser	UAU	Tyr	Tyr	UGU	Cys	Cys
UUC			UCC			UAC			UGC		
UUA	Leu	Leu	UCA			UAA	Stop	Stop	UGA	Stop	Trp
UUG			UCG			UAG			UUG	Trp	
CUU	Leu	Leu	CCU	Pro	Pro	CAU	His	His	CGU	Arg	Arg
CUC			CCC			CAC			CGC		
CUA			CCA			CAA	Gln	Gln	CGA		
CUG			CCG			CAG			CGG		
AUU	Ile	Ile	ACU	Thr	Thr	AAU	Asn	Asn	AGU	Ser	Ser
AUC			ACC			AAC			AGC		
AUA		Met[a]	ACA			AAA	Lys	Lys[b]	AGA	Arg	Stop[c]
AUG	Met		ACG			AAG			AGG		
GUU	Val	Val	GCU	Ala	Ala	GAU	Asp	Asp	GGU	Gly	Gly
GUC			GCC			GAC			GGC		
GUA			GCA			GAA	Glu	Glu	GGA		
GUG			GCG			GAG			GGG		

[a]AUA is Ile in echinoderms and platyhelminths.
[b]AAA is Asn in echinoderms and platyhelminths.
[c]AGA/AGG are Ser in most metazoans and Gly in ascidians; AGG is an unassigned codon in *Drosophila*.

the anticodon first position (34) and that at the codon third position (3) does not always conform to the earlier rules; G34 pairs with U3 as well as C3, so that all pairs of codons ending with a pyrimidine, for example, UUU and UUC, are translated to Phe by a single anticodon – GAA. Such irregular pairing is called 'wobble' pairing (**Figure 1b**), and the system by which it occurs is known as the 'wobble rule'. Modified nucleosides (wobble modifications) are often found at the anticodon first position (34), and play an important role in wobble pairing to decode cognate codons (**Figure 1b**). The wobble rules, which were first proposed by Crick (1966), have been examined experimentally both *in vitro* and *in vivo*. The latest version of the codon–anticodon pairing rules, including the wobble rules, is given in **Table 2**, although some of them have been deduced only from analysis of tRNA anticodon and codon usage in certain organisms. **See also**: Transfer RNA; Transfer RNA in Decoding and the Wobble Hypothesis; Watson–Crick Base Pairs

Alternative Codes

A variation in the established genetic code was first found in mitochondria in 1979 (Barrell *et al.*, 1979), and since then a number of other deviations have been reported, not only in mitochondria but also in eubacterial and nuclear systems. Such differences are believed to have arisen during the course of the evolution of living organisms. It has now become clear that almost all genetic code variations occur as a result of changes in the anticodons of tRNAs that decode the relevant codons. The following anticodons are involved in genetic code variations in animal mitochondria (**Table 3**) (reviewed in Yokobori *et al.*, 2001).

1. GNN (N = A, G, U or C) anticodons are used for decoding NNA codons as well as NNU/NNC codons, in which G34 recognizes A3. In mitochondria of platyhelminths and echinoderms, the AAA Lys codon is decoded as Asn by $tRNA^{Asn}$ with the anticodon GUU or GψU. Pseudouridine (ψ), one of the uridine derivatives, is useful for strengthening the interaction between the GψA anticodon and the AAA codon. The original Lys codon, AAG, is read by $tRNA^{Lys}$ with the anticodon CUU. In *Drosophila* mitochondria, the AGA codon is read as Ser by $tRNA^{Ser}$ with the anticodon GCU. In this case the AGG codon is unassigned, its corresponding tRNA being absent. G34 in anticodons seems to have the potential to base pair with A3 in codons so long as no competitor tRNA exists in the system.
2. m^7GNN (m^7G; 7-methylguanosine) anticodons are used for decoding NNA/NNG codons as well as NNU/NNC codons, in which m^7G34 is considered to recognize A3/G3. In mitochondria of molluscs and echinoderms, AGA/AGG codons are read as Ser by $tRNA^{Ser}$ with the anticodon m^7 GCU. m^7G34 is therefore thought to recognize all four bases at the codon third position in mitochondria.
3. $tRNA^{Ser}$ possessing the anticodon UCU recognizes AGA/AGG codons, as well as AGU/AGC, as Ser in platyhelminth and nematode mitochondria. A similar rule for pairing between U34 and N3 is seen in the case of CUN codons decoded by $tRNA^{Leu}_{UAG}$ in most animal mitochondria.
4. A U*CU anticodon is used for decoding AGA/AGG codons as Gly in echinoderm mitochondria, where U* stands for a modified uridine. In this case, U* was found

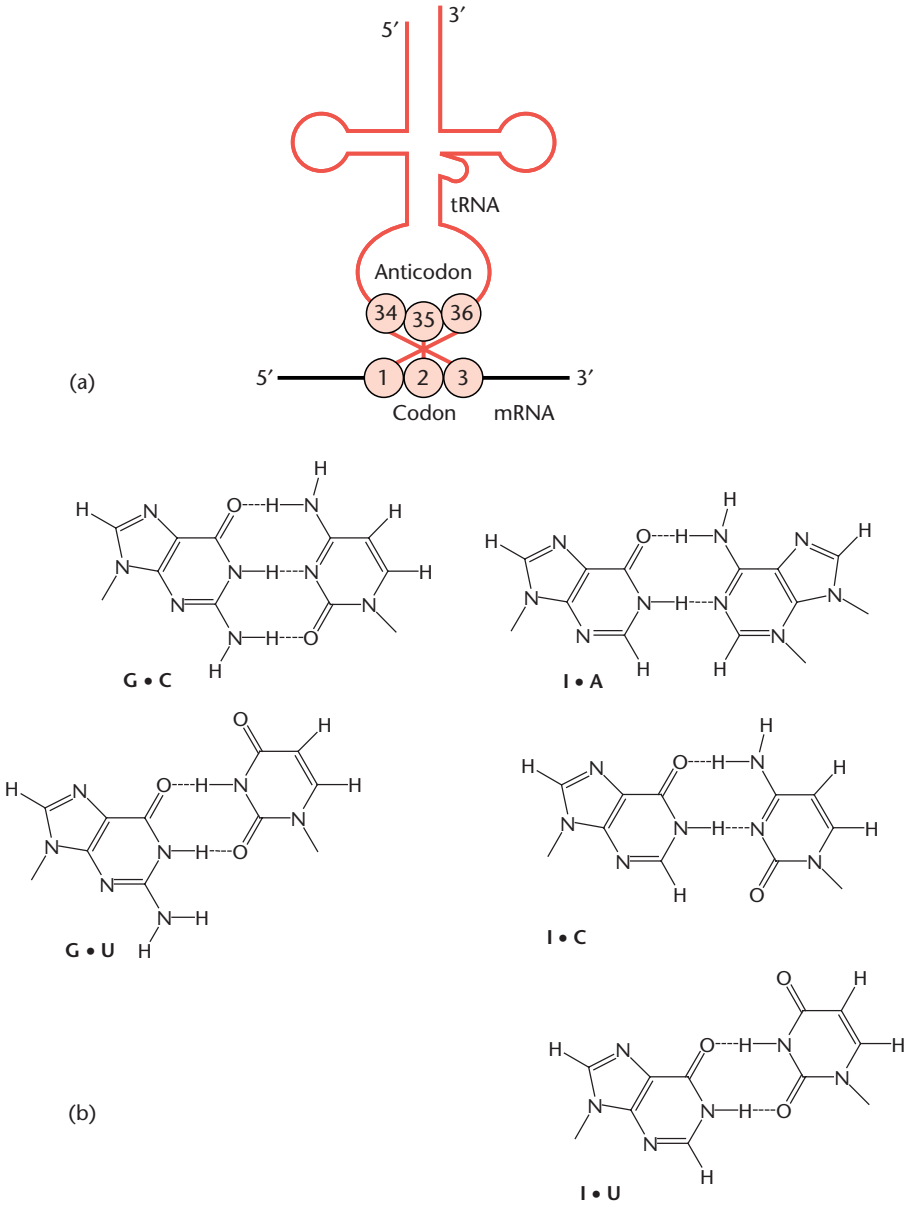

Figure 1 (a) Base pairing between a tRNA anticodon and an mRNA codon. Base pairing between the base at position 34 of the tRNA (wobble base) and that at position 3 of a codon does not always conform to the Watson–Crick base-pairing rule. (b) Wobble base pairings between first position of anticodons and third position of codons. GñU, IñA, IñC and IñU are wobble pairings. I stands for inosine. G–C pair is presented as a typical Watson–Crick base pair for comparison.

to be 5-taurinomethyluridine (τm^5U). Similarly, U*CA anticodon of $tRNA^{Trp}$ decodes the UGA Trp codon, as well as UGG, in yeast and most animal mitochondria. In the case of yeast mitochondria, U* is known to be 5-carboxymethylaminomethyluridine ($cmnm^5U$).

5. The anticodon C*AU of $tRNA^{Met}$, where C* is 5-formylcytidine (f^5C), decodes the AUA codon as well as AUG as Met in most animal mitochondria.
6. Even $tRNA^{Met}$ containing unmodified C34AU appears to decode AUA, only if the nucleotide at position 37, 3′ adjacent to the anticodon is modified to t^6A (N^6 -threonylcarbamoyladenosine) and if no competitor tRNA able to decode the AUA codon exists in the system. **See also**: Evolution of the Genetic Code; Mitochondria: Origin

In **Table 3**, the base-pairing rules for mitochondria are summarized and compared with those for eubacteria and eukaryotic cytoplasm. **See also**: Base Pairing in RNA: Unusual Patterns

Table 2 Wobble modifications found in all domains of life

Symbol	Common name	Third letter of codon	Usage	Source
Cm	2′-*O*-methylcytidine	G	Met AUG, Trp UGG, etc.	A, B, E, AN, P, M, C
ac^4C	N^4-acethylcytidine	G	Met AUG	A, B
L (k^2C)	Lysidine	A	Ile AUA	B
m^5C	5-Methylcytidine	G	Leu UUG	E
f^5C	5-Formylcytidine	G, A	Met AUR	M
f^5Cm	5-Formyl-2′-*O*-methylcytidine	G, (A)	Leu UUR	AN
I	inosine	U, C, A	Family boxes, Ile AUH	B, E, AN, P, C
Gm	2′-*O*-methylguanosine	C	Two-codon sets	B, E, AN, P
m^7G	7-Methylguanosine	(A), (G), (U), (C)	Ser AGN	M
Q	Queuosine	U, C	Two-codon sets	B, AN, P, M
gluQ	Glutamyl-queuosine	U, C	Two-codon sets	B
galQ	Galactosyl-queuosine	U, C	Two-codon sets	AN
manQ	Mannosyl-queuosine	U, C	Two-codon sets	AN
mcm^5U	5-Methoxycarbonylmethyluridine	A, G	Two-codon sets	E
mcm^5s^2U	5-Methoxycarbonylmethyl-2-thiouridine	A, (G)	Two-codon sets	E, AN
mcm^5Um	5-Methoxycarbonylmethyl-2′-*O*-methyluridine	A, (G)	Two-codon sets	M
mnm^5U	5-Methylaminomethyluridine	A, G	Two-codon sets	B
mnm^5s^2U	5-Methylaminomethyl-2-thiouridine	A, (G)	Two-codon sets	B, C
$cmnm^5U$	5-Carboxymethylaminomethyluridine	A, G	Two-codon sets	B, M, C
$cmnm^5s^2U$	5-Carboxymethylaminomethyl-2-thiouridine	A, (G)	Two-codon sets	B, M
$cmnm^5Um$	5-Carboxymethylaminomethyl-2′-*O*-methyluridine	A, (G)	Two-codon sets	B, E, AN
τm^5U	5-Taurinomethyluridine	A, G	Two-codon sets	M
τm^5s^2U	5-Taurinomethyl-2-thiouridine	A, G	Two-codon sets	M
ncm^5U	5-Carbamoylmethyluridine	A, G	Family boxes	E
$mchm^5U$	5-(Carboxyhydroxymethyl)uridine methyl ester	A, G	Two-codon sets	AN
Um	2′-*O*-methyluridine	A, G	Two-codon sets	E, AN, C
s^2U	2-Thiouridine	A, (G)	Two-codon sets	AN
mo^5U	5-Methoxyuridine	A, G, U	Family boxes	A, B
cmo^5U	Uridine 5-oxyacetic acid	A, G, U	Family boxes	B
$mcmo^5U$	Uridine 5-oxyacetic acid methyl ester	A, G, U	Family boxes	B
ψ	Pseudouridine	A	Ile AUA	E

A, archaebacteria; B, eubacteria; E, eukaryotic single cell organisms and fungi; AN, animal cytoplasmic; P, plant cytoplasmic; M, mitochondria and C, chloroplasts.

Missense, Nonsense and Frameshift Mutations and their Consequences

The fidelity of protein synthesis is crucial for the maintenance of genetic information. Early work on basal-level errors in bacteria indicated that missense errors may occur with a frequency ranging between 5×10^{-5} and 5×10^{-3} per codon. The average is estimated to be 5×10^{-4}, which is in the same order as the frequency of transcription errors (Parker, 1989). Although such errors can occur randomly with negative effects, organisms sometimes utilize missense, nonsense and frameshift mutations positively for alternative gene expression. They are mainly caused by the effects of mRNA contexts or tRNA mutations. These nonstandard decoding mechanisms are referred to as 'recoding' (reprogrammed genetic decoding) (Gesteland and Atkins, 1996). **See also**: Gene Expression: Decoding and Accuracy of Translation

Missense alternatives

In both eubacteria and eukaryotes, some cistrons are known to be initiated with non-AUG codons derived from

Table 3 Transfer RNA (tRNA) anticodon decoding of codons in mitochondria, prokaryotes and eukaryotic cytoplasm

	Prokaryotes, eukaryotic cytoplasm		Mitochondria
	Anticodons	Codons	Anticodons
Family box (four-codon box)	5′INN3′ U*NN CNN	5′NNU3′ NNC NNA NNG	5′UNN3′ or m^7GNN[a] or ANN[b]
Two-codon set	5′GNN3′ or QNN	5′NNU3′ NNC	5′GNN3′
	U*NN	NNA NNG	U*NN or f^5CNN[c]
Two-codon set + one codon	5′INN3′ LNN CNN	5′NNU3′ NNC NNA NNG	5′GNN3′ CNN

[a]Found in $tRNA^{Ser}_{GCU}$ of molluscs and echinoderms.
[b]Found in $tRNA^{Thr}_{ACU}$ of *Mycoplasma* spp., and $tRNA^{Arg}_{CGN}$ of yeast and nematode mitochondria.
[c]Found in $tRNA^{Met}$ of bovine, squid, and nematode mitochondria.

a single base substitution in the normal AUG codon, such as GUG-Val, UUG-Leu and AUU-Ile codons in *Escherichia coli*, and the ACG-Thr codon in adeno-associated and Sendai viruses. These codons are recognized by initiator $tRNA^{Met}$ and translated as Met. The strict selection of the initiation codon strongly depends on the context effect of the mRNA sequence surrounding the initiation codon. Alternative usage of the initiation codon is one of the strategies used by cells to regulate the level of protein synthesis. **See also**: Initiator tRNAs in Prokaryotes and Eukaryotes

Missense alternatives result either from the use of misaminoacylated tRNA or ambiguous codon–anticodon interaction on the ribosomes. The aminoacyl-tRNA synthetase recognizes its cognate tRNA with high fidelity through identity elements embedded within the primary and tertiary structure of the tRNA, which are indispensable for it to be discriminated from other noncognate tRNAs. However, some *Candida* spp. seem to utilize misaminoacylation positively, probably for their evolutional advantages: they possess a unique serine tRNA ($tRNA^{Ser}_{CAG}$) that is complementary to the CUG-Leu codon but is charged mostly with Ser and slightly with Leu (3% of serylation *in vivo*). Genetic studies have revealed that both seryl- and leucyl-$tRNAs^{Ser}_{CAG}$ can participate in translation by incorporating Ser or Leu into in-frame CUG codons. CUG codon in *Candida* spp. has been named a 'polysemous codon' by analogy with synonymous codons (Suzuki *et al.*, 1997). It is thought that *Candida* spp. possessing the polysemous codon are in a transition state between the use of CUG-Leu and CUG-Ser during evolution, and that they will eventually bind CUG-Ser. **See also**: Mutations and the Genetic Code

Natural nonsense suppression

It can sometimes happen that a termination codon (UAG, UAA or UGA) is recognized by near-cognate tRNA having natural suppressor activity and translated as any amino acid, resulting in readthrough of the termination codon. In bacteria, the UAG codon seems to be naturally suppressed with glutamine tRNA with a frequency of 7×10^{-3} to 1.1×10^{-4}, while the readthrough frequency of the UAA codon is estimated to be from 9×10^{-4} to less than 1×10^{-5}. The suppression efficiency has been shown to depend on the mRNA context, or a tRNA mutation or mutations in genes encoding ribosomal proteins. **See also**: Nonsense Mutations and Suppression

Programmed readthrough of nonsense codon

There are several instances where nonsense suppression at a defined position plays a significant role in gene expression. Many RNA viruses and retroviruses use nonsense suppression as a common strategy, not only to maximize their genome capacity but also to regulate gene expression. The bacteriophage Qβ produces the elongated coat protein necessary for viral infection by incorporating Trp at the UGA termination codon of a normal coat protein with a frequency of 6.5%. Another example is the readthrough of the UAG codon between the *gag* and *pol* genes in Moloney murine leukaemia virus (MuLV), in which a 5% insertion of glutamine at the UAG codon is important for the viral replication. The pseudoknot structure downstream of the stop codon is known to be a key element responsible for the high level of readthrough. **See also**: Bacteriophages with ssRNA; RNA Structure: Pseudoknots

Nonsense codon for selenocysteine

Selenocysteine (SeCys) is the twenty-first known natural amino acid to be incorporated at a defined in-frame UGA codon. Three and about ten SeCys-containing proteins, including formate dehydrogenase and glutathione peroxydase, have been respectively identified in *E. coli* and mammals. The mechanism for the SeCys insertion at the in-frame UGA codon has been well-studied in the *E. coli* system. A SeCys-specific tRNA with the anticodon UCA is charged with Ser by Seryl-tRNA synthetase, and the resulting Ser-tRNA is converted to SeCys-tRNA by two distinct enzymes, SELA and SELD. A specific EF-Tu (SELB) binds to SeCys-tRNA and transfers to the in-frame UGA codon on the ribosome by interacting with a specific stem–loop structure termed as SECIS (SeCys-insertion sequence) located downstream of the UGA codon in the mRNA. **See also**: Selenocysteine

Nonsense codon for pyrrolysine

In Methanosarcinaceae, UAG codon encodes pyrrolysine (Pyl) in addition to serving as stop signal (Ambrogelly *et al*., 2007). Pyl is the twenty-second natural amino acid which is cotranslationally inserted at the UGA codon. This event requires a new aminoacyl-tRNA synthetase, pyrrolysyl-tRNA synthetase (PylRS).

Frameshift mutations

There are −1 and +1 frameshift mutations. The frameshift site of the −1 shift contains the consensus seven-base sequence, X XXW WWY (before shifting, X and Y are any base and W is A or U) including six bases at the P and A sites in mRNA. Both tRNAs bound to this sequence on the ribosomal P and A sites simultaneously slip by one base in the 5′ direction of the mRNA, where they form new base pairs (**Figure 2**). This is seen in the deoxyribonucleic acid (DNA) polymerase III γ subunit of prokaryotes and the *gal–pol* fusion protein of retroviruses.

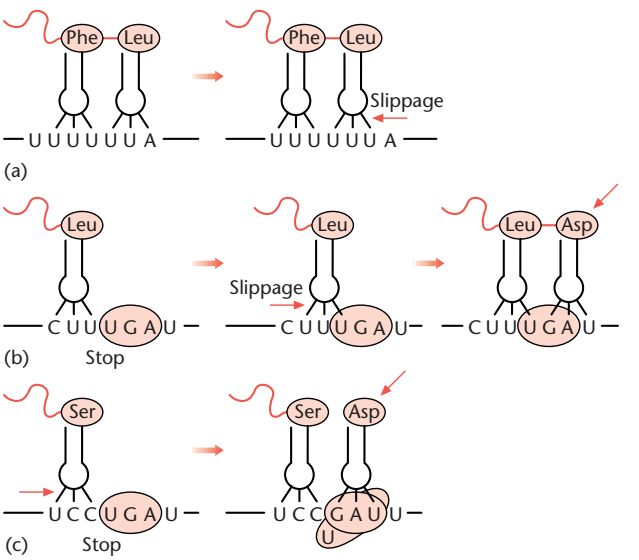

Figure 2 Molecular mechanisms of frameshifts: (a) −1 frameshift, (b) +1 frameshift of *E. coli* RF2 and (c) +1 frameshift of antizyme. Without slippage and rebinding of peptidyl-tRNA, aminoacyl-tRNA binds to the A site, while a U base is skipped.

The +1 shift occurs when the UGA codon or a hungry codon (the amount of tRNA corresponding to this codon is very low) is positioned at the ribosomal A site and translation is temporarily stalled, which is seen in release factor 2 (RF2) of prokaryotes and antizymes of eukaryotes. In both cases, the frameshift site is CUUUGAU/C and the Leu-tRNA bound to the CUU codon at the P site slips by one base in the 3′ direction of mRNA, so as to base pair with UUU codon. The adjacent GAU/C codon enters into the A site, and is recognized by Asp-tRNA, then the translation continues, resulting in full-length polypeptides. **See also**: Bacterial Ribosomes; Gene Expression: Frameshifting

These frameshifts involve various *cis* elements possessing a stem–loop or pseudoknot structure surrounding the frameshift site of the mRNA, which serve to enhance the frameshift efficiency. **See also**: RNA Structure

tRNA hopping (ribosome bypass)

This can be interpreted as an extreme case of the negative direction frameshift. In the case of the T4 phage Gene 60, Gly-tRNA bound to the GGA codon at the P site followed by UGA at the A site hops over 50 bases downstream of the mRNA, lands on the GGA codon, which has newly entered the P site, and the translation continues. It is suggested that a stem–loop structure behind the UGA codon is important for this tRNA hopping.

RNA Editing

In 1986, it was discovered that some RNAs are edited enzymatically after transcription to produce RNAs with different sequences. This phenomenon, named RNA

editing, is another feature giving rise to genetic code variation. RNA editing is classified into two types on the basis of the mechanism employed, insertion/deletion and base modification (Bass, 1993; Nishikura, 2006; Simpson, 1999). **See also**: RNA Editing

Insertion/deletion

This was first found in kinetoplasts of trypanosomes, in which a large number of uridine residues that are not encoded in the DNA are inserted into mRNA, or those encoded in the DNA are deleted during transcription. It is known that guide RNAs with 40–70 nucleotides containing an anchor sequence to bind the 3′ downstream region of the editing site play an important role in this type of editing. Another example is seen in the mitochondria of the slime mould *Physarum polycephalum* and in other members of the Myxomycota phylum, where C insertion mainly takes place, but sometimes insertion of dinucleotides or the substitution of C by U also occur.

Base modification

In chloroplast mRNAs of most green plants, C-to-U editing frequently takes place, in which C base is deaminated by a family of cytidine deaminases. In this editing, defined codons change to their near-cognate codons (Leu, Ser or Pro to Phe, Ser or Pro to Leu, etc.), sometimes creating the initiation codon (ACG-Thr to AUG-Met) or termination codon (CGA-Agr to UGA stop).

An intriguing case is the C-to-U editing that occurs in mitochondrial tRNA of marsupials, in which the anticodon second letter changes from C to U, resulting in a GUC anticodon (finally modified to QUC; Q = queosine) from the GCC anticodon. Thus, the tRNA identity changes from Gly to Asp.

In human and rabbit liver, apolipoprotein B (apoB) mRNA produces apoB100 protein, one of the subunits of low-density lipoprotein. However, in the small intestine, a truncated protein apoB48 is created from the same gene by base modification editing (CAA-Gln to UAA-stop), catalyzed by a complex enzyme including cytidine deaminase APOBEC-1 (ApoB mRNA editing enzyme catalytic polypeptide 1). ApoB48 acts as the structural protein core of chylomicrons.

A-to-I (inosine) editing is observed both in tRNAs and mRNAs. tRNAs frequently utilize I at the first position of anticodons for wobble decoding as described in the next section (see also **Figure 1b** and **Table 2**). In the case of mRNAs, I was first identified in mRNA encoding GluB subunit of AMPA (α-amino-3-hydroxy-5-methyl-lisoxazole-4-propionic acid)-responsive glutamate receptor. Because I is similar to G in its base-pairing nature, the CAG-Gln codon in exon 11 within GluB mRNA is converted to a CIG codon which specifies CGG-Arg codon by this editing (Q/R site editing). A-to-I editing is catalyzed by adenosine deaminases acting on RNA (ADARs). In mammalian, there exist two paralogues of ADARs, ADAR1 and ADAR2. The Q/R site editing of GluB mRNA is catalyzed by ADAR2. The biological importance of the Q/R site editing has been suggested by the finding that ADAR2 knockout mice show increased calcium permeability. It is known that defective editing of GluB mRNA is associated with human diseases, such as sporadic amyotrophic lateral sclerosis (ALS) and malignant human brain tumours. Recently, microRNAs (miRNAs) have been found to be edited by ADARs. I-containing miRNAs alter their specificity to target mRNAs. **See also**: AMPA Receptors

Posttranscriptional Modifications at tRNA Wobble Position

Modified nucleosides are often found in the wobble position (position 34), and these are involved in wobble pairing. Such wobble modifications play critical roles in modulating codon recognition by restricting, expanding or altering the decoding property of the tRNAs. The latest version of the codon–anticodon pairing rules, including the wobble rules, is given in **Table 2**. It should be noted that some of these rules have been deduced on the basis of codon usage and analysis of the tRNA anticodon but still lack biochemical evidence in particular organisms. Detailed information in this section is available from the following reviews: Bjork (1995), Suzuki (2005) and Yokoyama and Nishimura (1995). **See also**: Transfer RNA Modification

A modification

Inosine at tRNA wobble position (I34) expands the ability of the decoding system to decipher three families of codons (NNU, NNC and NNA) (**Figure 1b**) apart from the NNG codon. In bacteria, I is found only in $tRNA^{Arg}$, which has the ICG anticodon. In eukaryotes, I is commonly found in all tRNAs that are responsible for family boxes except for $tRNA^{Gly}$, $tRNA^{Phe}$ and sometimes $tRNA^{Leu}$. As an exception, unmodified A34 is found in $tRNA^{Thr}$ from *Mycoplasma* and mitochondrial tRNAs; this unmodified A34 is assumed to be capable of decoding all four codons. However, there is no experimental evidence to support this issue.

G modification

In bacteria, the G34 of the tRNAs for Tyr, His, Asn and Asp is modified to Q. In eukaryotes, it is further modified to mannosyl-queuosine (manQ) or galactosyl-queuosine (galQ), in which galactose or mannose is attached to the pentadiol ring of the Q base. It is known that Q prefers to base pair with U than with C. However, the exact role the Q modification plays is not fully understood. It has recently been reported that the *E. coli yadB* gene, which is a paralogue of glutamyl-tRNA synthetase, transfers glutamic acid to a cyclopentene moiety of Q to form glutamyl-queuosine (gluQ). Another G modification is 2′-*O*-methyl-guanosine (Gm), which is found in some $tRNA^{Phe}$

molecules isolated from eukaryotes and prokaryotes. Q and Gm may contribute to the prohibition of misreading. In the mitochondrial $tRNA^{Ser}$ of molluscs and echinoderms, G34 is modified to m^7G, which is probably capable of decoding all four codons.

C modification

In prokaryotes and eukaryotes, C34 is sometimes modified to 2′-*O*-methylcytidine (Cm) or 4-acetylcytidine (ac^4C). Cm and ac^4C seem to strengthen the base pairing with G. In the $tRNA^{Met}$ in animal mitochondria, C34 is modified to f^5C. As the AUA codon in the animal mitochondrial genetic code specifies Met instead of Ile, the f^5C modification enables AUA to be deciphered as Met along with AUG. This modification expands the decoding capacity of $tRNA^{Met}$ to assign the nonuniversal genetic code. Lysidine (L, N* or k^2C) is a lysine-containing cytidine derivative that occurs at the wobble position (position 34) of eubacterial and some organellar AUA codon-specific Ile tRNAs ($tRNA^{Ile}$) that have the CAT anticodon. The CAT anticodon also occurs in the Met tRNA that is specific for AUG. The lysidine modification converts the codon specificity of the precursor $tRNA^{Ile}$ with CAU anticodon from AUG to AUA and its amino acid specificity from Met to Ile, and thus prevents the misreading of AUG as Ile and AUA as Met. It is known that an essential gene *tilS* ($tRNA^{Ile}$-lysidine synthetase) is responsible for L synthesis (Soma *et al*., 2003).

U modification

Unmodified uridine is frequently found at the wobble position of tRNAs from *Mycoplasma* spp. and mitochondria that are responsible for family boxes, in which four codons with the same first and second letters but different third letter specify a single amino acid. Although uridine is supposed to only recognize A and G on the 3rd codon letter on the basis of the original wobble rule, U34 can actually base pair with any of the four bases due to its conformational flexibility to enable wobbling with U and C (four-way wobble rule). In bacterial tRNAs, U34 is modified to 5-hydroxyuridine derivatives (xo^5U) such as cmo^5U and mo^5U in tRNAs that are responsible for family boxes. It is known that xo^5U wobble modifications result in the efficient recognition of U and G in addition to A. Thus, the xo^5U modification expands the decoding capacity of U34 in bacterial tRNAs. In contrast, in tRNAs for Lys, Gln and Glu, the U at the wobble position is modified to 5-methyl-2-thiouridine derivatives (xm^5s^2U) such as mnm^5s^2U, $cmnm^5s^2U$, mcm^5s^2U and τm^5s^2U. As these tRNAs are responsible for decoding two-codon sets that end in purine (R) (i.e. NNR), the xm^5s^2U modifications participate in preventing the misreading of the pyrimidine (Y)-ending near-cognate codons (NNY).

References

Ambrogelly A, Palioura S and Soll D (2007) Natural expansion of the genetic code. *Nature Chemical Biology* **3**: 29–35.

Barrell BG, Bankier AT and Drouin J (1979) A different genetic code in human mitochondria. *Nature* **282**: 189–194.

Bass B (1993) *RNA Editing: New Uses for Old Players in the RNA World*. New York: Cold Spring Harbor Laboratory Press.

Bjork GR (1995) *Biosynthesis and Function of Modified Nucleosides*. Washington, DC: ASM Press.

Crick FH (1966) Codon–anticodon pairing: the wobble hypothesis. *Journal of Molecular Biology* **19**: 548–555.

Gesteland RF and Atkins JF (1996) Recoding: dynamic reprogramming of translation. *Annual Review of Biochemistry* **65**: 741–768.

Nishikura K (2006) Editor meets silencer: crosstalk between RNA editing and RNA interference. *Nature Reviews Molecular Cell Biology* **7**: 919–931.

Parker J (1989) Errors and alternatives in reading the universal genetic code. *Microbiology Reviews* **53**: 273–298.

Simpson L (1999) RNA editing: an evolutionary perspective. In: Gesteland RF, Cech TR and Atkins JF (eds) *The RNA World*, 2nd edn, pp. 585–608. New York: Cold Spring Harbor Laboratory Press.

Soma A, Ikeuchi Y, Kanemasa S *et al*. (2003) An RNA-modifying enzyme that governs both the codon and amino acid specificities of isoleucine tRNA. *Molecular Cell* **12**: 689–698.

Suzuki T (2005) *Biosynthesis and Function of tRNA Wobble Modifications*. *In Fine-Tuning of RNA Functions by Modification and Editing*. New York: Springer.

Suzuki T, Ueda T and Watanabe K (1997) The 'polysemous' codon–a codon with multiple amino acid assignment caused by dual specificity of tRNA identity. *EMBO Journal* **16**: 1122–1134.

Yokobori S, Suzuki T and Watanabe K (2001) Genetic code variations in mitochondria: tRNA as a major determinant of genetic code plasticity. *Journal of Molecular Evolution* **53**: 314–326.

Yokoyama S and Nishimura S (1995) *Modified Nucleosides and Codon Recognition*. Washington, DC: ASM Press.

Further Reading

Osawa S (1995) *Evolution of the Genetic Code*. Oxford: Oxford Science Publications.

Parker J (1989) Errors and alternatives in reading the universal genetic code. *Microbiological Reviews* **53**: 273–298.

Watanabe K and Osawa S (1995) tRNA sequences and variations in the genetic code. In: Soll D and RajBhandary UL (eds) *tRNA: Structure, Biosynthesis and Functions*, pp. 225–250. Washington, DC: American Society of Microbiology.