Translating the genetic code from nucleotide information to amino acid sequence is the function of aminoacyl-tRNA synthetases (ARSs). By specifically recognizing tRNAs and catalyzing the incorporation of their cognate amino acids to the 3′ terminal ribose of the tRNA, ARSs ensure that each amino acid incorporated by the ribosome into a growing polypeptide matches the information contained in the codon-anticodon pairing.

The tRNA aminoacylation reaction by ARSs proceeds in two enzymatic steps that take place within a single active-site domain of the enzymes:

1) ARS + aa + ATP → ARS-aaAMP + Ppi
2) ARS-aaAMP + tRNA → ARS + AMP + aa-tRNA

In the first part of the reaction, the enzymes activate the amino acid with ATP to form aminoacyl-adenylate, with release of pyrophosphate. Next, the amino acid is transferred, via the formation of an ester bond, to a hydroxyl group of the ribose of the terminal adenosine at the 3′-end of the tRNA, thus generating aminoacyl-tRNA and AMP. Each amino acid is recognized by its own specific ARS, which is universally distributed [4].

The typical tRNA is made up of 76 nucleotides that fold into a cloverleaf structure consisting of four stems and three loops. The 3′-end terminates in the universal CCA3′-OH with the terminal A3′-OH being the amino acid attachment site (Fig. 1). The seven base pairs adjacent to the amino acid acceptor site constitute the acceptor stem. The anticodon triplet is in a loop at the opposite end of the cloverleaf.
Recognition of tRNAs by ARSs depends mostly on molecular interactions with the acceptor stem and the anticodon loop of the tRNA. The active-site domains of the enzymes bind the acceptor arm of the tRNA molecule, where the amino acid is attached. Recognition of the anticodon is achieved through domains attached to the active sites. These domains are not universally conserved, and vary from enzyme to enzyme and from species to species [5, 6]. The ARSs are evenly divided into two classes of ten enzymes each [7, 8]. All enzymes within a class appear to have evolved from a single-domain ATP-binding protein. Insertions into and variations on this domain established a framework for binding the tRNA acceptor stem. Over the course of evolution additional domains were added to this core structure [9].

**Evolution of aminoacyl-tRNA synthetases**

At the biochemical level, the genetic code is established by the action of tRNA synthetases. For that reason, ARSs are at the center of research on, and theories of, the origin of life [9-12], and their evolution is thought to be closely connected with the early development of the genetic code.

Despite the close functional relationship between the aminoacylation activities of tRNA synthetases and the genetic code, attempts to relate the evolution of these enzymes to the development of the code have been inconclusive [12]. However, recent work on the structures of synthetase–tRNA complexes and on the evolutionary relationships within this family of enzymes has provided a framework within which specific features of the code and the synthetases can be examined. In particular, synthetases may have developed as pairs of tRNA-binding proteins that mirrored the increase in complexity of the tRNA molecules [8]. Thus, the extant families of synthetases are a product of genetic code development, and their structural relationships could reflect the nature of intermediate steps in the establishment of codon–amino acid interactions.

The 20 ARSs comprise two distinct families of enzymes (Fig. 2) [7, 8, 13], each of which originated from an ancient, distinct
single-domain protein [9]. This domain contains the active site for adenylate synthesis (condensation of an amino acid with ATP to give the aminoacyl adenylate) and for attachment of the activated amino acid to the 3′-end of tRNA. However, the structures of the ancestral domains for the two families are unrelated and, in spite of extensive research, no evidence exists for a common ancestor.

The two families of enzymes are known as class I and class II, and, with one exception, the class to which a synthetase is assigned is constant throughout evolution [13-18]. The 11 class I enzymes are characterized by an active site domain that has a Rossman nucleotide-binding fold composed of alternating β-strands and α-helices [19, 20]. From statistical analyses of the several hundred sequences of class I enzymes available from prokaryotes, archaea, and eukaryotes, the synthetases can be further divided into three subgroups (Fig. 2) [18]. As expected, these subgroup groupings are also consistent with structural comparisons between the enzymes within the class.

The three subclasses are designated Ia, Ib, and Ic. Each subclass is thought to have its own common ancestor that arose after the progenitor of the entire class. The enzymes in each subclass show a tendency to recognize amino acids that are chemically related. For example, members of subclass Ia recognize hydrophobic amino acids, such as the branched aliphatics (Ile, Leu, and Val), and amino acids with sulfur-containing residues (Met and Cys). Arginyl-tRNA synthetase is also assigned to this subgroup. Subclass Ib enzymes recognize charged amino acids (Glu and Lys) and the derivative Gln. Subclass Ic enzymes recognize the aromatics Tyr and Trp.

By contrast, the active sites of the ten class II enzymes are made up of a seven-stranded β-sheet with flanking α-helices [21, 22]. Statistical analyses of sequences of class II enzymes show that, similar to their class I counterparts, they can be divided into three subgroups (Fig. 2). Subclass IIa enzymes recognize groupings of chemically similar side-chains, such as aliphatics (Ala and Pro) and polar (Ser, Thr, Pro, His) side chains as well as Gly. The charged side-chains Asp and Lys, and the derivative Asn are recognized by enzymes in subclass IIb, whereas subclass IIc synthetase recognizes the remaining aromatic, Phe.

Thus, the two classes have a certain ‘symmetry’, with enzymes for chemically (or sterically) similar amino acids placed across from each other in subclasses with similar numbers of enzymes. This symmetry between the two classes is also seen in other ways. For example, members of class I approach the acceptor stem of tRNA from the minor-groove side, whereas those of class II approach from the major-groove side [23]. Also, class I enzymes attach the amino acid to the 2′-hydroxyl whereas those of class II charge the 3′-hydroxyl [24, 25]. Based on all of these considerations, the division of the 20 enzymes into two classes of almost identical size, with specific and matching subclasses, does not seem coincidental. This organization is consistent with the development of paired synthetases, ultimately giving rise to two classes with equal numbers of enzymes [8].

An ancestral complex of one tRNA and two ARSs
Since the enzymes in the two classes bind to opposite sides of the tRNA acceptor stem, the possibility that a single acceptor stem can simultaneously bind a synthetase from each class was investigated by molecular modeling. Indeed, subclass-specific pairings can be made without any steric clashes (Fig. 3) [8]. However, not all combinations of class I and class II active-site domains on the acceptor stem can be paired together. Several combinations result in large steric clashes between the core regions of the proteins. Moreover, although the details of how each synthetase binds its cognate tRNA are idiosyncratic, with both translational and rotational changes occurring in the precise fit on the acceptor stem, these changes are coordinated so that the result is subclass-specific pairings [8].

Most striking is the way that tyrosyl-tRNA synthetase (TyrRS) and phenyl-tRNA synthetase (PheRS) are mutually accommodating. The binding of each of these enzymes involves large rotational and translational changes relative to other members in its respective class. And yet, these large changes are exactly compensated so that the active site domains of both enzymes can still bind simultaneously onto a tRNA (Fig. 3).

Thus, the remarkable symmetry between the two extant classes of synthetases can be viewed as a consequence of the interaction of specific synthetase pairs in complex with tRNA. These pairs might have evolved to cover and protect the acceptor stem in environments (such as high temperature) where the structure of tRNA was especially vulnerable [8]. At this stage, the charging reactions might have been catalyzed by other molecules, possibly ribozymes, that were later replaced by the synthetases [26–29].

A correlation between ARS pairs and the structure of the genetic code
The structural pairings of tRNA synthetases link specific class I and II subclasses and suggest that an interaction existed between the two ancestral proteins of the paired subclasses [30]. The evolution of extant enzymes from these ancestral pairs can explain the similarities found among the two synthetase classes. This physical link between subclass ancestors has important implications with regard to the ancestral nature of the translation machinery and the genetic code.

This model implies the existence of a primitive genetic code that translated proteins using a reduced number of tRNAs or mini-helix-like structures that might have been tRNA precursors. The association of ancestral synthetases in pairs suggests a smaller number of total tRNA entities and a reduced number of amino acids. If synthetases had simply replaced a set of 20 pre-existing aminoacylating ribozymes that recognized a set of 20 different tRNAs, there would be no reason to expect the symmetrical nature of the two extant synthetase classes.

The primitive genetic code was probably capable of encoding single-domain proteins with the complexity of synthetase active sites that specifically recognized RNA structures. This translation machinery must have had a degree of coding accuracy. Ambiguous codons might have been used for families of related side-chains, and the actual distribution of incorpo-
rated side-chains might have depended on the relative concentrations of these similar amino acids. Indeed, recent experiments demonstrated that even contemporary organisms, such as *Escherichia coli*, can tolerate high levels of coding ambiguity [31].

Expansion and establishment of the final genetic code were probably achieved through duplication and mutation of the genes encoding tRNAs [32–34]. Equivalent duplications and changes of the associated pairs of synthetases would have followed the expansion of the tRNA set. This implies that a relationship should exist between extant codons for amino acids that are recognized by synthetases whose ancestors were paired.

As shown in Fig. 4, specific patterns of codon distribution can be seen when these are organized according to the proposed pairings of synthetase subclasses. This depiction shows that each of the three pairings of subclasses is related to the code in a way that is peculiar to the specific pairing. Thus, the codon assignments associated with 16 amino acids appear to have been constrained by selective pressures that operated through the subclass pairings of the associated synthetases. As expected, these constraints are pair-specific. Moreover, such constraints would only be expected for code assignments achieved during the period in which aminoacyl-tRNA synthetase pairs existed. Early and late amino acid codon assignments that took place before the advent of synthetase pairing, or after the separation of these complexes, would not necessarily follow these correlations.

In a putative RNA world, aminoacylation is thought to have originated with ribozymes that catalyzed attachment of specific amino acids to tRNA precursors [26-29]. Subsequently, ribozymes would be replaced with ribozyme-like RNA–peptide complexes and, eventually, with ribonucleoproteins (RNPs). RNPs with aminoacylation activity could be the precursors of modern tRNA synthetases.

Coded peptide synthesis could have emerged during this evolutionary process. With a simple code, a limited set of amino acids would initially have been used to generate proteins [32]. Early synthetases could have originated as crude proteins that covered and protected the tRNA acceptor stem and participated either directly or indirectly (in conjunction with a ribozyme-like catalyst) in aminoacylation. At this point, the emerging class I and class II tRNA synthetases might have been under strong selective pressure to develop amino acid or ATP affinities to increase the efficiency of the aminoacylation reaction. This resulted in the formation of the ATP-binding pockets that are conserved among all members within a given ARS class, irrespective of their subclass. During this phase of evolution, the development of synthetases and the emerging tRNA structure were undoubtedly coupled and, ultimately, one of the two synthetase domains replaced the aminoacylation activity of the pre-existing ribozyme.

Figure 3. Pairs of sterically compatible aminoacyl-tRNA synthetases (ARSs) active sites around a single tRNA.
Duplication and divergence of this RNP complex generated at least three different species: the ancestors of the three subclass-pairings suggested by extant synthetase structures (Fig. 3) [8]. As the genetic code became more complex, the assignment of codons to new amino acids required the duplication and divergence of existing tRNAs and their associated synthetase domains. Specific recognition of the newly incorporated amino acid was achieved by evolving an active site with new specificity in the other member of the synthetase pair (Fig. 4). This established the symmetry in side-chain specificity seen among the paired synthetase subclasses (Fig. 2) [30].

From this point onwards, each complex evolved to optimize its amino acid and tRNA specificity. Better recognition of the tRNA molecule, particularly through new interactions with the anticodon, was achieved through the incorporation of idiosyncratic domains into each individual complex, and the original active site pairs split apart as a consequence of steric clashes or other factors [30].

New functions of aminoacyl-tRNA synthetases

ARSs have traditionally been considered housekeeping enzymes, solely dedicated to the aminoacylation of their specific tRNAs. However, as our understanding of biological processes increases and the amount of sequence information grows, it becomes increasingly evident that these enzymes are involved in many other biological processes. In bacteria and yeast, these new functions tend to include transcription or translation regulation, whereas in higher eukaryotes ARSs seem to have incorporated sophisticated roles in cell cycle control and signaling that are unrelated to the protein synthesis apparatus.

Gene transcription and translation control by ARSs

In bacteria, ARSs are involved in the regulation of gene expression. This phenomena was first described in Bacillus subtilis [35], but the regulation mechanism controlled by Escherichia coli threonyl-tRNA synthetase (ThrRS) is the best-documented and understood [36].

E. coli ThrRS negatively regulates the expression of its own gene (thrs) at the translational level [37]. The enzyme binds to a site, the operator, positioned in the leader region of its own mRNA and inhibits initiation of translation by competing with the 30S ribosomal subunit [38]. The operator is composed of four structural domains. Domain 1 is single-stranded and carries the Shine-Dalgarno (SD) sequence and the initiation codon. Domain 3 is also single-stranded and links two stem–loop structures (domains 2 and 4) that carry sequence analogies with the anticodon loop of tRNAThr (Fig. 5).

The binding sites for ThrRS and the ribosome are overlapping in the leader RNA [39]. While the ribosome recognizes domain 1 and domain 3, ThrRS binds specifically and symmetrically to the two stem-loops in domains 2 and 4 in a way that mimics tRNA anticodon recognition. This leads to competition between ThrRS and the 30S ribosomal subunit [40]. Interestingly, the two essential regulatory events are controlled by dif-
different domains of the protein. Operon recognition involves the catalytic and C-terminal domains of the synthetase, and ribosome competition is achieved by the N-terminal domain of ThrRS (Fig. 5).

tRNA<sup>Thr</sup> also participates directly in this mechanism by being an efficient competitor of the mRNA operator region [41]. Thus, tRNA<sup>Thr</sup> and the operator compete for ThrRS binding, and the ribosome and the enzyme compete for mRNA binding. When the growth rate increases, the cellular concentrations of tRNA<sup>Thr</sup> and 30S increase, causing ThrRS to mainly recognize tRNA<sup>Thr</sup> and, as a result, full translation of thrS. By contrast, when the concentrations of tRNA<sup>Thr</sup> or 30S decrease, ThrRS binds to its own mRNA, causing repression of expression and mRNA degradation.

Regulation of expression of Saccharomyces cerevisiae aspartyl-tRNA synthetase (AspRS) was the first example of a eukaryotic ARS regulated via a feedback mechanism similar to the one described above. Frugier and co-workers demonstrated that yeast AspRS binds tightly and specifically to its own mRNA both in vitro and in vivo [42]. Expression of AspRS is regulated by a mechanism that necessitates recognition of the 5′ extremity of AspRS mRNA by its translation product, and depends on the cellular tRNA<sup>Asp</sup> concentration. This regulation leads to synchronized expression of AspRS and tRNA<sup>Asp</sup> [43, 44].

ARSs are also implicated in the quality control of protein synthesis. Ribosomes can stall on an mRNA if the message has no stop codon or if there is no cognate tRNA available for a particular codon [45]. To rescue stalled ribosomes and eliminate partially completed polypeptides from the cell, bacteria use a quality-control mechanism mediated by an RNA known as tmRNA, which is aminoacylated with alanine by alanyl-tRNA synthetase (AlaRS).

TmRNA has the unique capacity to act as both a tRNA and mRNA. Its cellular function is to release stalled ribosomes and to induce the C-terminal tagging of prematurely truncated proteins with a protease targeting sequence [45, 46].

Cell regulation functions of ARSs

In eukaryotic cells, the existence of a link between protein synthesis and signal transduction was first suggested by the observation that cells treated with interferon (IFN-γ) strongly over-expressed tryptophanyl-tRNA synthetase (TrpRS) [47, 48]. A direct relationship between the two processes was later demonstrated in human cells by Wakasugi and co-workers, who showed that a human tyrosyl-tRNA synthetase (TyRS) has cytokine functions in addition to its role in protein synthesis [49]. The intracellular form of this enzyme contains an N-terminal catalytic domain and a C-terminal structure, which is homologous to human endothelial monocyte-activating polypeptide II (EMAPII).

Under apoptotic conditions in culture, full-length TyrRS is secreted from cells and digested by leukocyte elastase (an extracellular protease) producing two distinct cytokines: one formed by the active site domain of the enzyme (mini-TyrRS), and a second one that corresponds to the EMAPII-like domain of the native TyrRS [49].

The NH<sub>2</sub>-terminal catalytic domain of TyrRS contains a conserved Glu-Leu-Arg (ELR) motif within its sequence that is identical to that found in CXC chemokines, such as the angiogenic factors interleukin-8, Gro-α, Gro-β, and Gro-γ. Indeed, mini-TyrRS induces angiogenesis and functions as a chemottractant for polymorphonuclear leukocytes (PMNs) [49]. The extra COOH domain of human TyrRS has cytokine activities like those of mature human endothelial monocyte-activating polypeptide II, and becomes an immune-cell stimulant for migration and production of tumor necrosis factor (TNF), tissue factor, and myeloperoxidase.

A second form of cellular signaling function was later discovered for tryptophanyl-tRNA synthetase (TrpRS), an enzyme related in structure to TyrRS. In human cells TrpRS exists in two forms, the major form corresponds to the full-length protein, and a truncated form (mini-TrpRS) results from alternative splicing and lacks most of the NH<sub>2</sub>-terminal extension. Although both human full-length TrpRS and mini-TrpRS are enzymatically active in aminoacylation, only the shorter form is active as an inhibitor of vascular endothelial growth factor (VEGF)-induced angiogenesis [50], thus displaying angiostatic activity. As mentioned before, production of this NH<sub>2</sub>-terminally truncated variant is stimulated in vitro by IFN-γ in a variety of cells [51].

Recently, Tzima and co-workers reported that a truncated TrpRS binds at intercellular junctions of endothelial cells. Using genetic knock-outs, binding was established to depend on vascular endothelial (VE)-cadherin, a calcium-dependent adhesion molecule that is essential for normal vascular development. Binding of the truncated TrpRS inhibited activation of vascular endothelial growth factor (VEGF)-induced extracellular receptor kinase (ERK) activation and cell migration [52].

ARS multi-synthetase complex

In eukaryotic cells, ARSs display a higher level of organization in the form of multi-ARS complexes. At least nine different enzymes, glutamyl-tRNA synthetase (GlutRS), prolyl-tRNA synthetase (ProRS), isoleucyl-tRNA synthetase (IleRS), leucyl-tRNA synthetase (LeuRS), methionyl-tRNA synthetase (MetRS), glutaminyl-tRNA synthetase (GlnRS), arginyl-tRNA synthetase (ArgRS), lysyl-tRNA synthetase (LysRS), and aspartyl-tRNA synthetase (AspRS), have been identified in the mammalian forms of these complexes [53, 54]. In addition, these complexes also contain non-enzymatic factors, known as p43, p38, and p18 (Fig. 6). tRNA channeling has been suggested as a possible function for these mammalian multi-ARS complexes, as an efficient way to control substrate distribution during sequential reactions [55].

Many of the ARSs that form the multi-ARS complex also have non-canonical functions. In the case of LysRS, it has been shown that this enzyme is involved in the regulation of microphthalmia transcription factor (MITF) transcriptional activity [56]. MITF is a basic helix-loop-helix leucine zipper DNA-binding protein and its activity is inhibited by its interactions with a tumor-suppressor protein, Hint. In quiescent leukemia cells, LysRS forms a trimeric complex with MITF and Hint but, upon immunologic activation, LysRS-synthesized Ap4A binds to
zymatic factors are labeled p18, p38, and p43.

A second example of non-canonical regulatory function comes from human MetRS, which constitutes one of the uncommon cases of nuclear localization of ARSs [58]. MetRS is translocated to the nucleus under proliferative conditions to augment rRNA synthesis in nucleioli. The presence of MetRS in the nucleoli depends on the integrity of rRNA and the activity of RNA polymerase, suggesting that MetRS plays a role in rRNA synthesis [58] and act as a sensor protein coupling translation and rRNA biosynthesis [59].

Human GinRS has also been implicated in the control of cell proliferation and in the regulation of cell death through an antagonistic interaction with ASK1, a protein kinase that has a critical role in apoptosis [60]. The interaction of the two proteins is stimulated by glutamine, which can suppress cell death [60]. Like ARSs, ARS-associated factors in the multi-ARS complex play diverse roles in processes other than protein synthesis; p43 is secreted as an active cytokine [61-64], inducing the synthesis of various pro-inflammatory cytokines and chemokines, such as TNF-α, interleukin-8, monocyte chemotactic protein-1, and interleukin-1β, from monocytes [60], as well as intracellular adhesion-molecule-1 [65]. Seemingly, p43 also plays a complex role in angiogenesis. Although it induces migration of endothelial cells at low concentrations, at high concentrations it can suppress vascular growth by blocking the proliferation and triggering apoptosis of endothelial cells [65].

The sequence of p43 contains a caspase-cleavage site, which releases the C-terminal domain of p43 from the multi-ARS complex. This process was thought to trigger the secretion of the cytokine component from p43, causing the disintegration of the multi-ARS complex to block protein synthesis [66]. However, p43 processing does not appear to affect the function of the complex, and it turns out that the uncleaved form of p43 is the active cytokine [64]. The role of proteolytic cleavage of p43 in apoptosis is thus unclear at this point.

Another component of the multi-ARS complex, p38, also has an unexpected additional role. This protein can bind to FUSE-binding protein (FBP), a transcriptional activator of the myc gene, promoting its ubiquitylation and proteasome-dependent degradation [67]. When the expression of endogenous p38 is abolished, myc is overexpressed owing to the lack of p38-mediated suppression, which causes hyperproliferation of lung cells. The consequent malfunction of the lung causes p38-/- mice to die at birth, although they survive development through the prenatal stage [67]. It is not known yet whether the smallest cofactor, p18, is also multi-functional. It shares limited sequence similarity with elongation-factor subunits [68]. Park and colleagues reported that p18 directly interacts with serine/threonine kinases to activate p53, and works as a potent tumor suppressor [69].

Aminoacyl tRNA synthetase-like proteins

Large-scale sequencing efforts have revealed that many genomes contain ARS-like proteins that are evolutionarily related to functional domains from canonical ARSs [70]. The functions of these proteins, however, are largely unknown, and possibly completely unrelated to those of their homologous ARSs.

A subset of these ARS-like proteins is associated with the synthesis of amino acids and cofactors. Sisler et al. described a class of ARS-like proteins (HisZ) in Lactococcus lactis that are homologous to the catalytic domain of histidylation-tRNA synthetase (HisRS). These proteins lack aminoacylation activity, but are instead essential components of the first enzyme in the pathway for histidine biosynthesis: ATP phosphoribosyltransferase (His G) [71].

The observation of a protein linking amino acid synthesis and protein synthesis implies an early connection between the biosynthesis of amino acids and proteins [70]. Indeed, genetic studies in yeast have identified a protein kinase, GCN2, that acts as a primary sensor of amino acid starvation [72]. This protein displays four functional domains: an N-terminal domain that binds other activators, a protein kinase domain, a domain highly homologous to HisRS, and a C-terminal domain conferring binding properties to GNC2. Amino acid starvation causes the accumulation of uncharged tRNAs that bind to the HisRS-related domain of GNC2 and activate the adjacent protein kinase, which, in turn, phosphorylates the translation factor eIF2α, inhibiting protein synthesis [73]. The biological meaning of this reaction remains unclear.

Some ARS-like proteins are involved in RNA modification. Such is the case of YadB, a recently described GluRS-like protein that chemically modifies tRNA [74]. YadB has the ability to
activate glutamate [75], although without the need for tRNA binding that characterizes GluRS [76]. YadB does not attach activated glutamate to tRNA^Glu or tRNA^Glu^k but instead to the anticodon region of tRNA^Glu^k [77]. Salazar et al. have shown that, rather than transferring glutamate to the 3’ end of the tRNA as GluRS would do, YadB attaches glutamate to the hypermodified nucleoside queuosine at the first anticodon position of tRNA^Glu^k, leading to the formation of glutamyl-queuosine [74]. The biological meaning of this reaction is also unclear.

Other ARS-like proteins are involved in the control of translational fidelity. For instance, a homologue of a ProRS domain, the ProX protein from Clostridium sticklandii, efficiently and specifically hydrolyzes Ala-tRNA^Pro^k [78], thus preventing the misincorporation of alanine instead of proline into proteins by misacylated tRNA^Pro^k. Similarly, autonomous AlaRS domain homologues (AlaX proteins) from Methanosarcina barkeri and Sulfolobus solfataricus hydrolyze Ser-tRNA^Ser^k and Gly-tRNA^Gly^k substrates [78]. Finally, YbaK, a protein from Haemophilus influenzae with high sequence identity to the prokaryotic ProRS editing domain, has been shown to be capable of deacylating Ala-tRNA^Pro^k [79].

Finally, a few other ARS-like proteins are involved in cellular transport processes. For instance, MetRS-like proteins, such as Trbp111 in Aquifex aeolicus [80] and Arc1P in Saccharomyces cerevisiae (homologues of mammalian p43 discussed above), have been shown to bind to tRNA [81, 82] and are involved in the nuclear transport of tRNA. In addition, Arc1P forms a complex with MetRS and GluRS [83], and could act as an organizer of tRNA synthetases in yeast [55].

### Aminoacyl-tRNA synthetases in human disease

Ever since initial reports in the 1980s that linked ARSs to human autoimmune diseases [84], the links between ARSs and health disorders have grown steadily. So far, these connections appear to be mainly related to the canonical aminoacylation function of ARSs, but suggestions of pathologic effects related to the non-canonical functions discussed above are starting to emerge.

The better-characterized pathologies related to tRNA and ARS are muscular disorders caused by deficiencies in mitochondrial tRNA aminoacylation. Aminoacylation of mitochondrial tRNAs is carried out by ARSs encoded in the nuclear genome. Some of these enzymes are specific for this organelle, while others function also in the cytoplasm. Incorrect translation of the mitochondrial genome causes severe depletion of the respiratory pathways and results in several disorders that often affect muscular tissue.

Mutations affecting the mitochondrial translation apparatus generate diseases such as mitochondrial myopathy, Leigh syndrome, Leber’s hereditary optic neuropathy, chronic progressive external ophthalmoplegia, Kearns Sayre syndrome, maternal myopathy and cardiomyopathy, progressive encephalopathy, diabetes mellitus, and deafness. More than 100 mutations involving rRNA and tRNA have been found in mtDNA (see http://www.mitomap.org).

Mitochondrial encephalopathy with lactic acidosis and stroke-like episodes (MELAS) is a clear example of a disease caused by mutations in mitochondrial tRNAs that decrease their aminoacylation efficiency. Several tRNA mutations are linked to the MELAS phenotype, most of them affecting the mitochondrial tRNA^Leu^k (UUR). The most frequent are A3243G [85], A3252G [86], G3244A, T3291C [87], C3256T [88], and T3271C [89]. Some of these mutations decrease leucyl-tRNA synthetase efficiency compared to wild-type tRNA^Leu^k due to folding alterations of the tRNA molecule (Fig. 7).

Myoclonic epilepsy with ragged red fibers (MERRF) is a second example of mitochondrial disease related to tRNA mutations. Symptoms of the disease include myoclonus, variable seizures, ataxia, dementia, corticospinal tract degeneration, peripheral neuropathy, optic atrophy, deafness, central hypoventilation syndrome with respiratory failure, and myoclonic epilepsy [90]. Several mitochondrial tRNA^lys^ mutations have been reported to be the cause of MERRF; for example: A8344G [91, 92], G8361A [93], T8356C [94], and G8361 (which includes other phenotypes like Leigh’s syndrome, myoclonus or myopathy with trigeminal, and proximal myopathy). The A8344G mutation causes a decrease in tRNA^lys^ aminoclaylation by tlycyl-tRNA synthetase, and the lack of aminoclaylated tRNA^lys^ might be the cause of premature termination of translation close to tyrosine codons [95].

Thus, most mitochondrial diseases linked to tRNA aminoacylation seem to be caused by changes in the tRNA structure. However, mutations in genes coding for ARSs have also been linked to human disorders. For instance, it has been reported that a mutation in the human mitochondrial leucyl-tRNA synthetase gene may represent a novel type-2 diabetes susceptibility gene [96].

In addition, mutations in glycyI-tRNA synthetase (GlyRS) have been found in Charcot-Marie-Tooth (CMT) disease patients. CMT is the most commonly inherited neurological disorder, found world-wide in all races and ethnic groups. Type 2D CMT could be caused by two mutations in the GlyRS gene: Gly-240→Arg or Gly-71→Gly [97]. Similarly, distal spinal muscular atrophy type V is also associated with mutations in GlyRS (Leu-129→Pro, Gly-526→Arg) [97]. Finally, a connection between certain tryptophanyl-tRNA synthetase polymorphisms and the pathophysiology of vascular angiogenesis and homeostasis has been proposed, although, as yet, there is no evidence for a direct link between these polymorphisms and susceptibility to myocardial infarction [98].

Synthetases also act as antigens in human inflammatory myopathies [99]. Antibodies directed against histidyl-tRNA synthetase (anti-Jo-1) are the most commonly produced by patients with polymyositis, a disease characterized by weakness and wasting of muscle. Anti-Jo-1 also causes interstitial lung disease (ILD) and arthritis [84].

Other anti-ARS autoantibodies have also been found in patients with polymyositis and interstitial lung disease; anti-PL-12 antibodies are directed at AlaRS and cause similar symptoms as anti-Jo-1 myositis [100]. Anti-PL-7 autoantibodies react against ThrRS and constitute an uncommon myositis-associ-
Aminoacyl-tRNA synthetases: a complex system beyond protein synthesis

ated antibody type [101] that also causes pulmonary fibrosis and dermatomyositis with erosive arthropathy [102]. Anti-EJ antibodies directed against GlyRS are associated with myositis and interstitial lung disease [103]. Anti-OJ antibodies react against IleRS and are responsible for myositis and severe interstitial lung disease. Anti-OJ antibodies are the only anti-ARS antibodies that are directed against a member of the multi-enzyme complex [104]. Anti-KS are autoantibodies against AsnRS that are present in patients with interstitial lung disease and inflammatory arthritis, but without evidence of myositis [105].

Aminoacyl-tRNA synthetases as anti-infective targets

ARSs represent ideal targets for antibiotic development because they are essential enzymes of universal distribution whose ancestral nature allows for the selection of specific inhibitors [106]. In addition, they are soluble, stable, easy to express and purify in large amounts, and are straightforward to assay by one or more methods. X-ray structures are available for all synthetases, and much is known about the mechanism of the aminoacylation reaction.

Although there are several natural inhibitors directed against aminoacyl-tRNA synthetases, i.e. borrelidin (ThrRS), furanomycin (IleRS), granaticin (LeuRS), indolmycin (TrpRS), ochratoxin A (PheRS), cispentacin (ProRS), etc., most suffer from lack of inhibitory activity, poor specificity, or poor bioavailability. As a result, only one of these molecules has so far been developed into a commercial antibiotic (Table 1). Pseudomonic acid A (mupirocin, marketed as Bactroban) is the best known natural inhibitor of a synthetase [107] (Fig. 8). It is synthesized by Pseudomonas fluorescens, inhibits isoleucyl-tRNA synthetase from several bacterial pathogens, and shows about 8,000-fold selectivity for pathogenic IleRS over mammalian IleRS [108]. Unfortunately, its low systemic bioavailability restricts its application to the treatment of topical infections. In the search for better ARS inhibitors, several mupirocin analogues showing broad anti-infective activities, including anti-
Table 1. Examples of known ARS inhibitors.

<table>
<thead>
<tr>
<th>ARS</th>
<th>Compound</th>
<th>Organisms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Methionyl-tRNA synthetase</td>
<td>REP8839 [127, 128]</td>
<td>Staphylococcus aureus (even resistant strains), S. epidermidis, S. pyogenes and other Gram positive bacteria</td>
</tr>
<tr>
<td>Valyl-tRNA synthetase</td>
<td>SB– 203207 analogues [129]</td>
<td>bacteria</td>
</tr>
<tr>
<td>Leucyl-tRNA synthetase</td>
<td>Agrocin 84 [130]</td>
<td>Agrobacterium tumefaciens</td>
</tr>
<tr>
<td></td>
<td>SB– 203207 analogues [129]</td>
<td>bacteria</td>
</tr>
<tr>
<td>Arginyl-tRNA synthetase</td>
<td>Spartane sulfate [136]</td>
<td>Lupinus spp.</td>
</tr>
<tr>
<td></td>
<td>Aminoalkyl adenylates [137]</td>
<td>S. aureus</td>
</tr>
<tr>
<td>Glutamyl-tRNA synthetase</td>
<td>Glutamylsulfanoyladenosine and Pyroglutamylsulfamoyladenosin [138]</td>
<td>E. coli</td>
</tr>
<tr>
<td>Glutaminyl-tRNA synthetase</td>
<td>Glutaminol adenylate and Methyl phosphate ester [138]</td>
<td>E. coli</td>
</tr>
<tr>
<td>Tyrosyl-tRNA synthetase</td>
<td>SB-219383 [140] [141]</td>
<td>Staphylococcus spp.</td>
</tr>
<tr>
<td>Tryptophanyl-tRNA synthetase</td>
<td>Chuangxinmycin [142]</td>
<td>Bacteria, E. coli and Shigella dysenteriae</td>
</tr>
<tr>
<td></td>
<td>Indomycin [143] [144]</td>
<td>E. coli and Staphylococcus spp.</td>
</tr>
<tr>
<td>Seryl-tRNA synthetase</td>
<td>SB-217452 [145]</td>
<td>S. aureus and rat</td>
</tr>
<tr>
<td>Threonyl-tRNA synthetase</td>
<td>Borrelidin [146]</td>
<td>E. coli</td>
</tr>
<tr>
<td>Prolyl-tRNA synthetase</td>
<td>Quinoline inhibitors [147]</td>
<td>C. albicans</td>
</tr>
<tr>
<td>Aspartyl-tRNA synthetase</td>
<td>nonhydrolyzable aspartyl adenylate analogs [148]</td>
<td>E. coli</td>
</tr>
</tbody>
</table>

Figure 8. Examples of two ARS inhibitors: agrocin 84 vs. leucyl adenylate (left), and pseudomonic acid A vs. isoleucyl adenylate (right).
gal, antimiycoplasmal, and herbicidal activities, have been reported [109–117].

More recently, a set of methionyl-tRNA synthetase inhibitors that are active against staphylococcal and enterococcal antibiotic-resistant strains have been described [118]. The same group has developed ethanamine inhibitors that block phenylalanyl-tRNA synthetase from Staphylococcus aureus. Some of the latter have a selectivity that is 5,000-fold higher against bacterial enzymes than against mammalian PheRS [119]. Thus, interest in ARSs continues to grow but a solid pipeline of potential lead compounds is still lacking.

In addition to their potential application as compounds active against cellular pathogens, recent reports on the role of some aminocyl-tRNA synthetases and their cognate tRNAs in the packaging of retroviruses have awakened interest in the possibility of targeting these proteins in anti-viral therapies. The best known case of involvement of an ARS in viral packaging is that of tRNA<sub>Lys</sub> and LysRS packaging into human immunodeficiency virus type 1 (HIV-1); however, other retroviruses use tRNA<sub>Lys</sub> as primer for reverse transcriptase (RT) [120], and tRNA<sub>Lys</sub> is used as primer for Moloney murine leukemia viruses (MuLV) [121].

During HIV-1 virion packaging, three human tRNA<sub>Lys</sub> isoacceptors are associated with the viral genomic RNA. The tRNA<sub>Lys</sub> isoacceptor anneals near the 5′ end of the viral RNA genome to a 18-nucleotide sequence (primer-binding site), where it is used as primer by the viral reverse transcriptase for the transcription of the viral genome into DNA [122]. The role of tRNA<sub>Lys</sub> isoacceptors that are also packed in the virion is still unclear [122].

Incorporation of tRNA<sub>Lys</sub> into the virion also requires the packaging of Gag-Pol protein and lysyl-tRNA synthetase. While the interaction between tRNA<sub>Lys</sub> and LysRS is essential for packaging [123], aminoacylation of tRNA<sub>Lys</sub> seems to be dispensable [124]. In fact, the HIV-1 virion contains deacylated tRNA<sub>Lys</sub>, whereas the host cell contains mostly acylated tRNA<sub>Lys</sub> [123]. The presence of a deacylating system, which might be necessary for the RT to extend the 3′ adenosine tRNA<sub>Lys</sub>, has been proposed [125].

Lysyl-tRNA synthetase plays a second role in the packaging of HIV-1. During early assembly of HIV-1, an assembly complex is formed by genomic RNA, Gag, GagPol, tRNA<sub>Lys</sub>, and LysRS. Changes in cellular expression of LysRS result in corresponding changes in viral infectivity and in the amounts of LysRS, tRNA<sub>Lys</sub>, and viral RT packed in the virions [126]. It has been proposed that the altered viral contents of RT resulting from alterations in cellular LysRS concentrations result from the ability of LysRS to inhibit premature activation of Gag-Pol viral protease within the complex. Increases and decreases in cellular LysRS expression are accompanied by five- to eight-fold increases and five-fold decreases, respectively, in the cytoplasmic proteolysis of Gag and GagPol. Accordingly, overexpression of LysRS in the cell reduces viral protease activity [126]. All these considerations suggest that specific ARS inhibitors may, in the future, find additional indications in the treatment of retroviral infections.

Acknowledgements

This work was supported by grant BIO2003-02611 from the Spanish Ministry of Science and Education. We thank all of the members of the Gene Translation Laboratory at the Barcelona Biomedical Research Institute for support and discussions.

References

[3] Stent, G.S. (1968) That was the molecular biology that was. Science. 160, 390-395
quence similarities among the family of aminoaeryl-IRNA synthetases. Biochimie, 68, 1071-1078


[24] Fraser, T.H., Rich, A. (1975) Amino acids are not all initially attached to the same position on transfer RNA molecules. Proc Natl Acad Sci USA, 72, 3044-3048

[25] Sprinzl, M., Cramer, M. (1975) Site of aminoacylation of tRNAs from Escherichia coli with respect to the 2’– 3’- hydroxyl group of the terminal adenosine. Proc Natl Acad Sci USA, 72, 3049-3053


Aminoacyl-tRNA synthetases: a complex system beyond protein synthesis


Kobayashi, Y., Momoi, M.Y., Tominaga, K., Momoi, T.,


[113] PLC, Smithkline Beecham. (1996a) (Hetero)-aryl ketones derivatives with antibacterial properties. *US patent* 5,536,745,

[114] PLC, Smithkline Beecham. (1996b) (Hetero)-aryl ketones derivatives with antibacterial properties. *JP patent* 9,157,269,


About the authors

The laboratory of Gene Translation of the Institute for Biomedical Research of Barcelona was founded in 2003. Its main scientific focus lies on the study of protein synthesis in human pathogens. In particular, the laboratory explores the many functional connections that the protein synthesis apparatus has established with other cellular pathways during evolution. Thus, we investigate the evolution of central components of the translation machinery, we dissect the additional roles that these components play in the biology of species of biomedical interest, and we seek biomedical applications based on these findings.

Lluís Ribas de Pouplana is the director and founder of the Gene Translation Laboratory. He holds a Ph.D. from the University of Edinburgh, was a postdoctoral scientist at the Department of Biology of the Massachusetts Institute of Technology, and a senior scientist and assistant professor of Molecular Biology at the Scripps Research Institute of San Diego.

Teresa Bori Sanz received a Ph.D. from the University of Birmingham in 2005 for her studies in structure-function relationships of the collagen receptor glycoprotein VI. She is currently carrying out postdoctoral research at the Gene Translation Laboratory.

Tanit Guitart graduated in Biology in 2005 and she is currently a Ph.D. student at the Gene Translation Laboratory.