New structural insights into the decoding mechanism: Translation infidelity via a G-U pair with Watson–Crick geometry

Natalia Demeshkina a,b,c, Lasse Jenner a,b,c,1, Eric Westhof d, Marat Yusupov a,b,c, Gulnara Yusupova a,b,c,*

a Département de Biologie et de Génomique Structurales, Institut de Génétique et de Biologie Moléculaire et Cellulaire, Institut de Génétique et de Biologie Moléculaire et Cellulaire, Université de Strasbourg, Illkirch 67400, France
b CNRS, UMR7104, Illkirch 67400, France
c INSERM, U964, Illkirch 67400, France
d Architecture et Réactivité de l’ARN, Université de Strasbourg, Institut de Biologie Moléculaire et Cellulaire, CNRS, Strasbourg 67084, France

Abstract

Pioneer crystallographic studies of the isolated 30S ribosomal subunit provided the first structural insights into the decoding process. Recently, new crystallographic data on full 70S ribosomes with mRNA and tRNAs have shown that the formation of the tight decoding centre is ensured by conformational rearrangement of the 30S subunit (domain closure), which is identical for cognate or near-cognate tRNA. When a G-U forms at the first or second codon–anticodon positions (near-cognate tRNA), the ribosomal decoding centre forces the adoption of Watson–Crick G-C-like geometry rather than that of the expected Watson–Crick wobble pair. Energy expenditure for rarely occurring tautomeric base required for Watson–Crick G-C-like U pair or the repulsion energy due to sterical clash within the mismatched base pair could constitute the only cause for efficient rejection of a near-cognate tRNA. Our data suggest that “geometrical mimicry” can explain how wrong aminoacyltRNAs with G-U pairs in the codon–anticodon helix forming base pairs with Watson–Crick geometry in the decoding center can be incorporated into the polypeptide chain.

© 2013 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.

1. Introduction

The ribosome is the ribonucleoprotein machine responsible for the faithful translation of genetic material into the encoded polypeptide. Ribosomes consist of a large (50S) and a small (30S) subunit, which together compose the 2.5 megadalton (MDa) 70S ribosome. The ribosome maintains the mRNA reading frame during protein biosynthesis by simultaneously fixing tRNAs in their binding sites and grasping onto the messenger RNA (mRNA) chain. During evolution, features of the translation apparatus that are pivotal for mRNA reading frame maintenance in order to achieve efficient and exact translation of genomic information have evolved, but the organization of essential functional sites such as the decoding center was conserved [1]. Ribosomes possess three tRNA binding sites, termed A- (aminoacyl-), P- (peptidyl-) and E- (exit) sites, which are composed of elements from both subunits. The L-shaped tRNA molecules are oriented such that the tRNA anticodon–mRNA codon interactions take place on the 30S subunit, whereas the 3’ CCA termini of tRNA interact with the 50S subunit.

Decoding of the genetic information by codon–anticodon base pairing between the mRNA and the tRNA occurs in the A-site of the small ribosomal subunit. During each cycle of elongation, the ribosome must select the correct aminoacyl-tRNA (aa-tRNA) that matches the codon occupying the A site (i.e. forms Watson–Crick pairs) from a large pool of competing aa-tRNAs. Once the correct codon–anticodon interactions are established in the A-site of the ribosome, a rapid peptide bond formation takes place between the aminoacyl moiety of the A-site tRNA and polypeptide in the P-site. In prokaryotes the aa-tRNA arrives to the ribosome in the form of a ternary complex with elongation factor Tu (EF-Tu) and GTP.

One of the most interesting questions of molecular biology raised in the last decades is what forms the basis for translational fidelity? It was realized that the differences in stability between codon–anticodon interactions of cognate and non-cognate tRNAs were often too small to explain the low overall error frequency (10⁻³ to 10⁻⁵) in the translation process. Thus, the translational machinery must effectively select for the binding of correct aa-tRNA to the ribosome, whereas incorrect (non-cognate or near-cognate) aa-tRNAs are preferentially discarded.

In the 1970s, Hopfield and Ninio proposed the mechanism of “kinetic proofreading” as a way to increase the specificity of aa-tRNAs binding. They suggested a proofreading step driven by the
energy of GTP hydrolysis following the initial selective binding [5,6]. Initial support for the proofreading mechanism came from in vitro data proving that binding of near-cognate tRNAs (that typically carry a single mismatch to the codon in the A-site) resulted in a significant increase in GTP consumption relative to the amount of amino acids incorporated [7]. In these same studies, it was shown that non-cognate tRNAs (normally with more than a single mismatch) did not appear to stimulate the hydrolysis reaction. These observations were consistent with two selection steps separated by GTP hydrolysis. Non-cognate tRNAs are rejected during the initial phase of the selection, but near-cognate ones escape this screening process some of the time and are instead rejected during a second selection phase following GTP hydrolysis. This kinetic proofreading model predicts that slowing down GTP hydrolysis would allow greater time for equilibrium to be reached during the initial selection phase and could thus result in an overall higher fidelity.

 Nevertheless, it was a rather simplified scheme in which the role of the ribosome was only fixing the position of the mRNA codon. Now, it has been recognized that the ribosome plays a crucial and active role for the mechanism of aa-tRNA selection in the decoding site. Based on extensive footprinting, cross-linking, and mutagenesis data, it was demonstrated that several regions in 16S ribosomal RNA (16S rRNA) are implicated in decoding of aa-tRNA recognition (for recent review see [8]). Early biochemical studies positioned the decoding center of the ribosome on the small 30S ribosomal subunit at the interface with the large 50S subunit. This decoding center covers parts of the 16S rRNA that includes nucleotides 1400–1500 of helix 44 (h44), nucleotides 1050–1200 (h43), and 530 loop (h18). Chemical modification protection analysis showed that the bases of the conserved nucleotides G529, G530, A1492 and A1493 are protected by binding of an A-site tRNA [9]. Moreover, the miscoding aminoglycoside paromomycin induced protection of nucleotides 1408 and 1494, just across from 1492 and 1493 in helix 44 [10]. Mutational experiments later demonstrated that these nucleotides are critical for A-site tRNA binding [11,12]. In 1990s the NMR structures [13,14], and latter the X-ray structure [15] of an oligonucleotide corresponding to this region of helix 44 bound to paromomycin revealed that the aminoglycoside stabilized a structure of A1408, A1492, and A1493 that is distinct from that observed in the absence of ligand. It was suggested that the observed conformational changes might mimic those induced by the binding of cognate A-site tRNA to the ribosome. Based on the results of a kinetic analysis in the end of 1990s, an induced fit mechanism of aa-tRNA discrimination on the ribosome that operates in both initial selection and proofreading was proposed [16]. According to this model the cognate codon–anticodon interaction, more efficiently than the near-cognate one, induces the conformation of the decoding center that favors GTPase and accommodation.

 In the early 2000s the breakthroughs in X-ray crystallography of ribosomes and their subunits, allowed for the isolated 30S ribosomal subunit structure to be used as a model to study the selection of cognate and near-cognate anticodon-stem loop fragments and have provided important structural details about this kinetically driven mechanism [17,18]. Together with cryo-EM data [19–23], X-ray data of EF-Tu and tRNA bound to the ribosome [24,25] and kinetic data, an integrated model for decoding was proposed (see for review Refs. in [26–31]).

 The current kinetic model for the mechanisms governing the high level of accuracy was deduced from higher resolution structural approaches, pre-steady state kinetics and single-molecule fluorescence techniques. These studies expanded upon and somewhat altered the initial kinetic model. Unlike earlier analysis of translation using steady-state approaches, pre-steady state kinetics strives to utilize assays that monitor each independent molecular event in isolation using a variety of fluorescent and radioactive probes. The many parameters determined by these approaches, coupled with computational global-fitting techniques, led to a highly detailed picture of the kinetic and thermodynamic framework governing tRNA selection.

2. First structural insight into the decoding mechanism using the 30S ribosomal subunit as a model: A1492, A1493 and G530 of 16S rRNA actively discriminate the geometry of the codon–anticodon helix, and domain closure of the 30S subunit is triggered only by cognate tRNA

About one decade ago, high-resolution structures of the large (50S) and small (30S) subunits of archaeal and bacterial ribosomes were solved [32–34]. At the same time the functional complex of the full Thermus thermophilus 70S ribosome with intact tRNAs and a 36 nucleotide long mRNA was also studied first at 7 Å [35] and later, at 5.5 Å resolution [36]. These studies of 70S provided a general description of tRNA-ribosome and subunit–subunit interactions (inter-subunit bridges) in the context of the whole ribosome and described the tRNA and protein backbones of the bacterial ribosome that had tRNAs bound in the classical A-, P- and E-sites. However, because of the low resolution, these structures lacked details that could reveal the roles of particular atoms in respect to the ribosome function. Therefore, the first definitive localization of the decoding site came from a combination of a 3.1Å structure of T. thermophilus 30S subunits crystals, soaked with an anticodon stem-loop (ASL) of cognate tRNA32 bound to the A-site in the presence of a hexanucleotide mRNA (U6) [17] and a 7 Å structure of the T. thermophilus 70S ribosome containing a complete tRNA bound to the A-site with a long 36-nucleotide mRNA [35].

The current structural view of the mechanism underlying the decoding process was based on X-ray structures of the isolated 30S subunit where crystals were soaked with mRNA mimics U6 and cognate or near-cognate ASLs mimicking tRNA (with a G-U mismatch at either the first or second position) [17,18]. X-ray analysis of these 30S crystals showed that the binding of mRNA and cognate ASL in the A-site induced certain local conformational changes: A1492 and A1493 of 16S rRNA flipped out of the internal loop of h44. This binding also caused the universally conserved base G530 of 16S to switch from the syn conformation present in the vacant 30S subunit (apo-30S structure) [33] to an anti conformation. In their new conformations, A1493 and A1492 interact, respectively, with the minor groove of the first and second base pairs of the codon–anticodon helix, whereas G530 interacts with both the second position of the anticodon and the third position of the codon. A1493 and A1492 adopt type I and type II A-minor motifs, respectively. In addition, the N6 of A1492 is involved in hydrogen bonding to the 2’OH of the universally conserved C518 and to a highly conserved Ser50 of the small subunit protein S12. The third, or ‘wobble’, position of the codon is free to accommodate certain non-canonical base pairs. In those 30S subunit crystal structures, electron density for near-cognate ASLs was observed only when soaking in the presence of the antibiotic paromomycin known to reduce the fidelity of decoding. In the case of the G-U mismatch at the first codon position, this base pair was found in wobble pairing. This geometry results in displacement of the 1st codon position away from the minor groove, thus preventing it from forming the hydrogen bond with A1493. The electron density for the G-U mismatch at the second position did not conform to the expected wobble geometry, but this discrepancy was attributed to spatial averaging of two alternative positions of the G-U pair.

Based on these observations it was proposed that the first two base pairs of the codon–anticodon helix are closely monitored by
the ribosome in their ability to form hydrogen bonds with conserved nucleotides in such a way that the ribosome is able to discriminate between Watson–Crick base pairing and mismatches. In contrast, the environment of the third, or “wobble,” position appears to be suited for accommodating other base-pairing geometries.

Comparison of the apo-30S crystal structure with 30S crystals soaked with cognate ASL and U6 showed that binding of these ligands induced an internal rearrangement in the 30S subunit. This movement, termed “domain closure”, is a tightening of 30S subunit around the A-site by rotations of the head towards the shoulder (S4, G530 loop with surrounding regions of 16S rRNA and S12) and the h44/h27f platform region. In that study it was also shown that binding of paromomycin to the apo-30S flips out A1492 and A1493 of 16S, but does not have an effect on G530, which remains in the syn conformation. With near-cognate ASLs the closed 30S was only observed in the presence of the antibiotic paromomycin. This led to the conclusion that in ribosomal interactions with near-cognate tRNA, deviation from Watson–Crick geometry results in uncompensated desolvation of hydrogen-bonding partners at the codon–anticodon minor groove. As a result, the transition to a closed form of the 30S induced by cognate tRNA is unfavorable for near-cognate tRNA unless paromomycin induces part of the rearrangement.

Based on this hypothesis a mechanistic understanding of ribosomal mutants (for example, ribosomal ambiguity mutation, ram) that affect the fidelity of protein synthesis was suggested. The error-prone ram mutants typically carry altered versions of the small subunit ribosomal proteins S4 and S5 [37–39]. It was observed in the apo-30S structure that these proteins interact closely in the 30S subunit [40]. By comparing the 30S structures in open and closed states it was shown that this interaction is broken during domain closure. It was suggested that the mutations in S4 and S5 lead to disruption of salt bridges that form part of the interaction. By reducing the number of bonds that must be broken for domain closure to take place, these mutants decrease the energy barrier needed for this transition to occur and thereby facilitate the acceptance of tRNAs during the selection process.

3. Recent unanticipated observation from X-ray studies: The structure of full 70S ribosome with near-cognate tRNA at the proofreading state reveals domain closure of the 30S subunit

In a previous X-ray study we observed the expected conformational rearrangements in the 70S ribosome functional complex upon binding of cognate tRNA to the A-site [41,42]. We determined two high-resolution structures of the 70S ribosome complexes in the initiation and elongation states (at 3.35 Å and 3.1 Å resolutions, respectively). All comparisons were made within one crystal form to exclude errors. The 70S ribosome in the elongation state was programmed with a 60 nucleotide-long mRNA containing the SD sequence, AUG codon and tRNA|fMet|,|CAU| in the A/A state when an anticodon stem-loop and acceptor end of tRNA are bound to the A-sites on the 30S and 50S, respectively. The 70S ribosome in the initiation state was programmed with 30 nucleotide-long mRNA containing the SD sequence, AUG codon and tRNA|fMet|,|CAU| in the P- and E-sites. Comparison of these ribosome structures demonstrated that transition from initiation to elongation, i.e. upon binding of cognate tRNA in the decoding site, leads to a ~3-A movement of the shoulder toward the neck of the 30S subunit, whereas the other domains of the subunit remain immobile. These conformational changes lead to a contraction of the downstream mRNA tunnel and, as a result, a network of non-specific interactions between 16S rRNA and mRNA nucleotides is formed. This rearrangement of the 30S subunit in the context of the 70S ribosome is similar to the “domain closure” seen in the isolated 30S subunit but to a slightly lesser extent and without rotation of the head towards the shoulder [18]. It is very probable that the 30S subunit becomes less flexible when it associates with the 50S subunit to create the full 70S ribosome and inter-subunit bridges are formed.

In the study of the initiation complex, the two ribosome molecules per asymmetric unit (named A and B) differed in tRNA content [43,44]. While the A-site of ribosome B was empty, the A-site of the other ribosome (molecule A) contained density of apparent tRNA shape, which we attributed to tRNA|fMet|,|CAU|. The occupancy of this tRNA in the A-site was estimated to be almost equal to the one observed in the elongation complex with cognate A-tRNA. The electron density for tRNA|fMet|,|CAU| showed a well-defined anticodon stem loop, a flexible elbow region and a disordered acceptor end. Superposition of the model built for this tRNA|fMet|,|CAU| on the structure of cognate tRNA|fMet|,|CAU| shows that the position of the tRNA in the A-site is similar to that of cognate tRNA, which is in the classical A/A state. The anticodon CAU of tRNA|fMet|,|CAU| forms a single U-A Watson–Crick base pair with the first nucleotide of the AAA codon in a mini-helix with this codon. Based on these observations, the complex with tRNA|fMet|,|CAU| in the A-site was referred to as near-cognate [45]. However, according to a general definition it should be called non-cognate since tRNA has only one match with the A-codon. A comparison of the ribosome structures in molecule A with this near-cognate tRNA in the A-site and in molecule B with vacant A-site revealed several striking observations. The 30S subunit in the near-cognate complex undergoes the same conformational change (“domain closure”) that was observed upon binding of cognate tRNA to the A-site including the rearrangements of the key nucleotides of decoding center. To investigate these findings further we decided to study models of the ribosome at the proofreading step, where the A-site substrates are classical near-cognate tRNAs.

4. New structural insights about decoding on the ribosome

4.1. X-ray structures of the full 70S ribosome complexes with continuous mRNA and cognate or near-cognate tRNAs at the proofreading step

The most often occurring miscoding events during tRNA selection are G-U mismatches at the first two positions of the codon–anticodon duplexes. Therefore, we dedicated our study to the general problem of how the ribosome responds to and discriminates against G-U mismatches occurring in the decoding center [46]. In parallel with correct (cognate) base pairing we modeled G-U mismatches at all three positions of the codon–anticodon duplexes. Altogether we determined six X-ray structures of the 70S ribosome at 3.1–3.4 Å resolution programmed with 30-nucleotide-long mRNAs containing SD sequence, the AUG codon and tRNA|fMet| in the P-site, and the A-site occupied by either tRNA|Leu|,|GAC|, tRNA|Tyr|,|GLU| and corresponding mRNA codons [46] (Fig. 1A, left panel). In one set of experiments, tRNA|Leu|,|GAC| and tRNA|Tyr|,|GLU| were bound to their respective cognate codons CUC and UAC in the A-site. In the second set of experiments, we modeled near-cognate states of the ribosome. In these complexes the A-site was filled either by tRNA|Lys|,|GAC| and codon UUU with a G-U mismatch in the first position of the codon–anticodon helix or by tRNA|Tyr|,|GLU| and codonUGC with a G-U mismatch in the second position. Furthermore we repeated the second set of experiments with the antibiotic paromomycin present. Analysis of all six crystal structures shows that the A-site of the ribosome was occupied by tRNA at almost equal occupancy and that these tRNAs were found in the A/A state corresponding to the proofreading state of decoding.
A few objections could be raised against the validity of our model system. How is it possible that near-cognate tRNAs are bound with affinities high enough to form crystals containing tRNA with high occupancy if the ribosome rejects near-cognate tRNAs very efficiently? The error frequencies of translation are estimated between $10^{-3}$ and $10^{-5}$, depending on the type of measurement, concentrations and nature of tRNAs that perform misreading, and the mRNA context [4,47,48]. However, other fidelity measurements conducted in buffer systems with tRNA mixtures on heteropolymeric mRNA suggest that in vitro protein synthesis proceeds with even lower fidelity (an error rate of $2 \times 10^{-3}$ to $10 \times 10^{-3}$), thus proposing that further quality control mechanisms exist [3]. These values mean that the ribosome does accept near-cognate tRNAs (i.e. misincorporates) but with a low probability. The only way to model such a low probable but naturally occurring state by X-ray crystallography is to make the ribosome accept near-cognate tRNA by giving high excess of only one type of tRNA (for example, one carrying a mismatch to the first codon position) over the ribosome concentration. Under such non-equilibrium (crystallization) conditions, near-cognate tRNA binds efficiently in the ribosomal A-site. A similar way of obtaining ribosomes with one mismatch in the codon–anticodon helix at high yield in vitro system was recently exploited by Zaher and Green [3,49] allowing the authors to discover the retrospective quality control by the ribosome.

We have strong evidence that in our experimental 70S ribosome model the A-site is specific, but not “omnivorous” (or undiscriminating), and mostly only accept substrates that resemble cognate tRNAs. We determined X-ray structure of the 70S ribosome with a 30 nucleotide-long mRNA with SD sequence, and AUG codon and tRNA$^{\text{Met}}$ in the P-site followed by a polyA tail and as substrate for the A-site tRNA$^{\text{Phe}}_{\text{GCA}}$ was given in high excess (unpublished data). Analysis of this structure showed that 70S ribosome complex had an empty A-site with no electron density in the A-site found. These data demonstrate a natural restriction mode of the decoding center in the 70S ribosome complexes so that non-cognate tRNA cannot serve as a substrate at the accommodation step and most likely for initial selection either.

4.2. The decoding center restricts a G-U mismatch at the first or second codon–anticodon position to form Watson–Crick G-C-like interactions

At the resolution of our X-ray data we could confidently assign the general base pairing. The electron density maps unambiguously demonstrated that U4 and G5 of the A codons UUU and UGC do not show the anticipated wobble interactions with G36 in tRNA$^{\text{Leu}}_{\text{GAG}}$ and U35 in tRNA$^{\text{Val}}_{\text{GUA}}$, respectively [46]. Instead, U4,G36 and G5,U35 at the first and second positions of the codon–anticodon duplexes form base pairs similar to standard Watson–Crick G-C pairs (Fig. 1A, second and third panels). G-C-like G-U or G-T pairs have been shown earlier for RNA and DNA in other X-ray structures [50–52]. When U-G is at the third codon–anticodon position we observe standard U-G wobble pairing (Fig. 1A, last panel) as was shown earlier. However, because hydrogen atoms do not sufficiently scatter X-rays it is therefore very difficult to precisely identify the nature of the hydrogen bond pattern of the mismatches in the near-cognate codon–anticodon helices. Among many mispairing possibilities, either tautomerization or ionization of bases might allow for G-U pairs to obtain Watson–Crick-like geometry.

4.3. Binding of cognate or near-cognate tRNAs to the 70S ribosome leads to the same structural rearrangements in the 30S subunit

Analysis of the decoding center in the crystal structure of the vacant 70S ribosome showed that key nucleotides A1492 and 1493 were well defined and found in the “flipped in” position where they stack with each other and G530 was in the anti conformation.
Paromomycin is a member of the aminoglycoside family of antibiotics that has been known to affect the overall fidelity of translation [57], and as was shown later it accelerates both forward reaction rates (GTPase activation and accommodation) in tRNA selection and reduces the rate of near-cognate tRNA dissociation from A-site [58,59].

As it was mentioned above the first structural insights on the interaction of paromomycin with ribosomal elements were obtained first by NMR analysis [13,14], and latter by X-ray analysis [15], where the antibiotic was bound to an oligonucleotide corresponding to region of decoding center (h44). It was suggested that the observed conformational changes caused by binding of paromomycin might mimic those induced by binding of cognate A-site tRNA to the ribosome.

The structure of the crystals of apo-30S soaked with paromomycin showed that the antibiotic ring I inserts into h44 and helps to flip out bases A1492 and A1493, which were disordered in the structure of the apo-30S [33,40]. Ring I mimics a nucleotide base, stacking against G1491 and hydrogen-bonding with A1408. In addition it forms a tight hydrogen-bond interaction with the phosphate backbone of A1493. In an X-ray study of the Escherichia coli ribosome complex with ribosome recycling factor (RF) [56] it was shown that paromomycin has two binding sites (one in each ribosomal subunit) and a structural explanation for aminoglycoside inhibition of ribosome recycling was proposed, but because of the low resolution of the structure (4.5 Å) a detailed analysis of the interactions was absent. Later, paromomycin was used in X-ray studies of the T. thermophilus 70S ribosome complex to increase binding affinity of tRNAs to the A site [24,25,54,60]. From the high-resolution structures of full 70S complexes with mRNA, P-tRNA and near-cognate A-tRNAs in the presence of paromomycin we showed that upon binding of paromomycin ring I not only forms a tight hydrogen-bond interaction with the phosphate backbone of A1493, which helps lock the flipped-out bases in place, but causes a large positional shift of the A1493 phosphate group [46]. Comparison of these structures with structures in the absence of antibiotic also demonstrates that a movement of H69 accompanied by rearrangements of the intersubunit bridge B2a composed of h44 and H69 occurs in the presence of paromomycin and these distortions are probably initiated by the movement of the A1493 phosphate group. Although this movement only minimally alters the interactions of A1493 with U4-G36 and A1492/G530 with G5-U35, these local changes modulate the B2a bridge. H69 is displaced towards the tRNA, which probably enhances the interaction surface of H69 with the A-tRNA D-stem. In the presence of paromomycin the position of H69 is closer to that observed for the cognate state. Thus, displacement of the A1493 phosphate group relaxes the decoding pocket from the side of the A codon, making the decoding pocket less restrictive, and changes the deformation of the near-cognate codon–anticodon helix. The minor structural rearrangements that we observe with paromomycin are consistent with its moderate effect at the proofreading step [59]. It is also possible that paromomycin increases the lifetime of a tautomeric U or G in a U-G mismatch what potentially simulates stabilization of a near-cognate tRNA in the A-site and as a consequence incorporation of a wrong amino acid into a poly peptide chain. In addition, because the A-site is locked in one state once again that the ribosome is an extremely complex finely-tuned machinery where cross-talk between subunits passes via intersubunit bridges and as a consequence influences the binding of functional ligands.

Another study of the action of aminoglycoside antibiotics that combined both dynamic and structural approaches, showed that particularly paromomycin causes not only miscoding, but also blocks intersubunit rotation, and inhibits translocation [61].
4.5. Comparison of the two experimental models used for structural studies of decoding: The 30S small subunit complex with mimics of functional ligands and full 70S ribosome with mRNA and tRNAs

In light of our findings on complete ribosomes, it appears now that the models initially used to describe the mechanism underlying the decoding process had several shortcomings. First, there were crystal structures of isolated 30S subunits, which implies that inter-subunit contacts are not established and might limit the interpretation [17]. Moreover, the tip of the “spur” (h6) from a symmetry related 30S subunit was mimicking the P-site tRNA and an ASL was used instead of a full tRNA for the A-site (Fig. 2A). The P- and E-site mRNA codons were mimicked by the 3’ end of 16S rRNA so that the hexameric mRNA (position 1–4 visible) that was used was not covalently linked between the P and A site. And finally the near-cognate complexes were solved in the presence of paromomycin, making it difficult to dissect effects of mismatches from those of the antibiotic.

For the study of the mismatch at the first position, the 30S crystals were soaked with an anticodon stem-loop of tRNA\textsubscript{Leu} and U6 mRNA and a classical wobble U\textsubscript{1}/C1 G pair was found[17]. However, the mRNA was not covalently linked between the P- and A-codons and by superimposing the A site GAG anticodons from the near-cognate tRNA\textsubscript{Leu}GAG in the 70S structure [46] and that from the 30S model it was found that the first nucleotide of the A-codon is positioned differently (Fig. 2B and C). Because it does not have the natural restraint coming from being covalently bound to the P codon, this first nucleotide of the A-codon in the 30S structure had the freedom to move so it could form a wobble U-G pair. However, in the 70S ribosome structure the P and A mRNA codons are tightly held which means that the first nucleotide of the A-codon has to form Watson–Crick like interactions with G36 of tRNA\textsubscript{Leu}GAG (Fig. 2B and C). The kink in the mRNA between the A and the P codon [62], which allows simultaneous codon–anticodon pairing and might be important for maintaining the translational reading frame is stabilized by the coordination of a magnesium ion with an mRNA backbone phosphate [54].

In the experiments to study the U-G mismatch in the second codon position (30S crystals soaked with ASL\textsubscript{Ser} and U6) the electron density observed for the U-G base pair was consistent with Watson–Crick rather than wobble geometry (Fig. 2D) [18]. The authors interpreted this as a spatial average of the G/C1 U pair in two alternative positions, each of which allows the 16S decoding center to interact with one half of the minor groove or alternatively by paromomycin favoring tautomeration of the G or U thus allowing pairing with Watson–Crick geometry without any particular implications of this unusual geometry.

After understanding the issues of the 30S model lacking a covalent bond between the A and P mRNA codons on the interpretations of the first codon–anticodon interactions, the experimental electron density observed for the second position showing a Watson–Crick like geometry of the G/C1 U pair corroborates the

![Fig. 2. 30S subunit as a model for the selecton of tRNA. (A) Crystal packing of the 30S subunit in the early studies of decoding [17,18]. Magnification in a square shows arrangement of ligands in the P- and A-sites (see the text). Mimics of the P-site codon (3’ end of 16S rRNA) and the P-site ASL (helix 6 or spur, sp of the neighboring 30S) shown in cyan and grey, respectively; short analogs of the A-site codon (U6) and the A-site ASL soaked to the A site depicted in light magenta and violet, correspondingly. Abbreviations: h, pl, and sh indicate head, platform and shoulder domains of 30S. (B and C) Differences between the positions of the first uridine in the UUU codon base-paired to the GAG anticodon of tRNA\textsubscript{Leu} GAG in the full 70S structure [46] and from the 30S model [18]. (D) Electron density from the near-cognate ASL\textsubscript{Ser}/paromomycin structure around the second codon position, showing A1492 and G530. The appropriate position of a guanine base to pair with the refined position of the codon U is shown in light gray; the expected position of a uridine base forming a typical base pair with the refined position of anticodon G is shown in dark gray. Figure adapted from Ogle, J.M. et al. (2002) Cell 111, 721–732 with permission from Elsevier.](image-url)
re-evaluated principle for decoding based on interpretation of the 70S ribosome structure data.

One of the vulnerable parts of the 30S subunit crystals as an experimental model was the low accessibility of the A-site even to cognate tRNA ASL caused by the way the molecules were arranged in the crystal. Only experimental density for cognate ASL in the presence of paromomycin was published. The occupancy of cognate ASL in A-site in the absence of antibiotic was deduced from a global comparison of conformational changes (movement of 16S rRNA backbone phosphorus atoms after superposition of ASL complex on the native structure) [18]. Such comparisons showed that in the absence of the paromomycin, the changes upon binding of the cognate ASL were relatively smaller (~1 Å) than in the presence of the antibiotic, which probably reflect that the occupancy even for cognate ASL was not so high.

No electron density was observed in the A-site of 30S crystals after disorder and low occupancy of ligands always poses a challenge in X-ray crystallographic studies at low to medium resolution. At 3.6–3.8 Å the authors suggested that disorder of the ASLs was the explanation for the complete absence of tRNA density because they simultaneously observed loss of density for G530 and C1054, however insufficient binding might as well have been the case. From these experiments, a comparison of the overall structures of the 30S subunit with paromomycin and cognate ASL and the 30S subunit with “disordered” near-cognate ASL without paromomycin (or probably with low occupancy of near-cognate ASL) was performed, and the observed domain closure was described and attributed as a feature of cognate tRNA selection. But if the comparison of the movements of 16S RNA backbone phosphorus atoms in the 30S subunit was done between the cognate and near-cognate ASLs without paromomycin, the differences were not more than 1 Å in the shoulder region.

Another example showing that the use of different ribosome complexes (the 30S versus the 70S model system) for the study of decoding can lead to different interpretations is the attempt to rationalize data on ribosomal error-prone ram mutations that affect the fidelity of protein synthesis. Based on the structure of apo-30S the ram mutants, typically carrying altered versions of proteins S4 and S5 [40], were explained by the loss of salt-bridges lowering the energy barrier needed for domain closure to occur and hence increasing the error of decoding as described earlier [18]. However, analysis of 70S ribosome complexes [41] showed that the interactions between S4 and S5 are different from those described for the 30S model system. Moreover, upon binding of A-tRNA and thereby domain closure this interface is almost unchanged [49].

Contemporary kinetic discrimination mechanism operates in two stages: initial selection and proofreading [27,64]. Our structural data demonstrate that the decoding center of the 70S ribosome, which enters the state of initial selection is already “pre-formed” (in contrast with earlier findings where the nucleotides A1492 and A1493 were disordered [40]). In this ribosome the P-site is occupied by tRNA, the A-site is still vacant, and the key functional ligand, the mRNA is bound to the ribosome, with a kink between the A and P codons, a universal feature of the mRNA path on the ribosome. This kink is stabilized by the P-tRNA, 16S rRNA and a magnesium ion (Fig. 3A and B). The base A1493 is already protruding from h44 (“out” conformation) [41,43], while A1492 is held inside the helix 44 by stacking with the conserved base A1913 from H69 from the 50S subunit (Fig. 3C and D). This arrangement of A1492 and 1493 of h44 16S rRNA and A1913 of 23S rRNA is stabilized by neighboring G1494 in 16S rRNA and A1912 in the 23S rRNA through hydrogen bonding. Nucleotide G530 is in syn conformation.

During initial codon-independent interaction with the ribosome, tRNA in the complex with EF-Tu and GTP (ternary complex) binds the ribosome through contacts with the large and small subunits. Proteins L7 and L12 mediate the interaction with the large subunit, while interaction with the small subunit is localized to rRNA, protein, and mRNA in the A site [65–67]. These interactions most likely include tRNA contacts with H69, which cause the conformational change of A1913 from H69: adenine ring rotates about the glycoside bond, while the whole nucleotide rotates additionally. As a result the key nucleotide A1492 becomes free from stacking. According to single molecule FRET experiments during subsequent codon selection, aa-tRNA undergoes fast conformational samplings (fluctuations) within the A-site [68]. During this transition state, the anticodon of aa-tRNA enters the “preformed” 30S decoding site and forms primary interactions with the A-codon. A1492, A1493 and G530 begin to engage the codon–anticodon duplex and the tRNA itself, initiates the conformation change in 30S subunit. As a result upon binding of cognate or near-cognate tRNA to the 70S ribosome, the small subunit undergoes conformational changes in the “shoulder” region around the anticodon loop of the tRNA, where nucleotides G530, A1492 and A1493 form a static part of the decoding center, defining its spatial and stereo chemical properties (Fig. 3E).

Our data showed that the tight decoding center forms a “mould” that constrains the mRNA in such a way that the first two nucleotides of the A codon are limited to form Watson–Crick base pairs (Fig. 3E). When a G-U mismatch is at the first or second codon–anticodon position, the decoding center forces the pair to form Watson–Crick G-C-like interactions and not the expected wobble pairs. Owing to our current data resolution, we cannot precisely identify the hydrogen bond pattern of the mismatches in the near-cognate codon–anticodon helices, but tautomersism is a plausible chemical mechanism. An alternative explanation for the source of tRNA discrimination could be repulsion in the U-G pair. Energy expenditure for formation of tautomers (or repulsion energy) could constitute the sole cause for the very efficient rejection of near-cognate tRNAs by the ribosome (Fig. 3E). Since the keto-tautomer is favored by a factor of about 10$^{-4}$ to 10$^{-6}$, forcing the enol form would amount to a free energy expenditure of ~4 kcal or greater. These values are similar to the ones observed in maintaining translational accuracy. codon–anticodon recognition and the structural rearrangements of the small subunit, which result in forming a tight decoding center, products the tRNA in the A/T state [69] (which perhaps exists in different sub-conformations). Cryo-EM studies of EF-Tu [23,70–72], and recent crystal structures of the 70S ribosome with EF-Tu and aa-tRNA bound to the ribosome brought enormous amount of structural information, but still the

5. Concluding remarks

The new insights into recognition of the base pairs in the cognate and near-cognate models of the 70S ribosome complexes [41,43,46] led us to propose a scheme for the selection mechanism of decoding under the assumption that our data present a snapshot of decoding under the assumption that our data present a snapshot of decoding.
Fig. 3. Decoding center on the 70S ribosomes. (A, B) The P/A-kink positions the A-codon to the interactions with tRNA. The elements involved in the kink formation are shown for the 70S ribosome with the unoccupied A-site (A) and with the A-site bound by cognate tRNA (B). The main contributors to the kink stability (the P-site tRNA, h44 of 16S rRNA and magnesium ions) are illustrated. (C–E) Conformations of the key elements of the decoding center on the 70S ribosome in different functional states and possible scenarios in tRNA selection process: (C), unoccupied center; (D), preformed center with tRNA bound at the P-site; (E), formed or closed state of the center with the cognate or near-cognate tRNA, and possible scenarios in tRNA selection process (see the text). In (C–E) orientations of the center slightly differ from each other to clearly show conformations of G530, A1492 and A1493 of 16S rRNA and A1913 of 23S rRNA in each particular case.
mechanistic details surrounding the long-range signaling responsible for GTPase activation are not fully understood. Recent data demonstrate a number of conformational, probably reversible, events such as proper positioning EF-Tu and GTPase center with respect to the sarcin–ricin loop for the GTPase-activated state, where GTP hydrolysis can occur [68,73]. It was suggested that the conformational changes in the 3OS subunit (shoulder movement upon binding of tRNA) [18] or/and a signal passing via the body of tRNA [74,75] (distortions) can cause these movements.

To understand the high fidelity of protein synthesis on the ribosome it is important to examine the question not only from the point of view of how tRNAs with mismatches are rejected from the ribosome but also how they are incorporated (even at low frequency). One of the possibilities is the mechanism of molecular mimicry of shape, which can allow the mismatch to escape the error discrimination mechanism on the ribosome. Movement of a single proton on one of the mismatched bases alters the hydrogen-bonding pattern such that a base pair forms with an overall shape that it is virtually indistinguishable from a canonical, Watson–Crick base pair in codon–anticodon helix. This idea is similar to the rare tautomer hypothesis of spontaneous mutagenesis during replication, a long-standing concept that has been difficult to demonstrate directly. In their work describing the structure of DNA, Watson and Crick proposed that spontaneous base substitutions could be a consequence of bases spontaneously pairing in rare tautomeric forms [76,77]. They suggested two possible transition mispairs, G–T and A–C, involving the enol form of guanine or thymine and the imino form of adenine or cytosine, respectively. Both mispairs fit well within the dimensions of the DNA double helix to preserve the geometry of a correct Watson–Crick base pair. Thus, the potential importance of mispair geometry to base substitution mutagenesis was implied even before the discovery of DNA polymerases. This notion was elaborated in the rare tautomer hypothesis of spontaneous mutagenesis, which states that mutations arise through the formation of high-energy tautomers at low frequency [78,79].

First structural evidence for replication infidelity via a mismatch with Watson–Crick geometry was provided in the form of a crystal structure of a human DNA polymerase λ variant poised to misinsert dGTP opposite a template T [52]. All atoms needed for catalysis were present at the active site and in positions that overlay with those for Watson–Crick geometry. The mismatch had Watson–Crick geometry consistent with a tautomeric or ionized base pair, with the pH dependence of misinsertion consistent with the latter. These results support the original idea that a base substitution can originate from a mismatch having Watson–Crick geometry.

Another structural evidence for infidelity via a mismatch with Watson–Crick geometry was demonstrated using high-fidelity DNA polymerase (*Bacillus stearothermophilus* DNA polymerase 1 large fragment). In this study, a C-A mismatch was observed within the double helix past the position of incorporation, where it adopts a cognate Watson–Crick base-pair conformation [51]. According to the authors their structures unambiguously demonstrate that tautomeric base pairs can form in the polymerase active site, providing strong support of the rare tautomer hypothesis through direct structural evidence.

Our structural data reveal a mechanism, which explains the principle of discrimination based not on the deviation from Watson–Crick geometry that results in uncompensated desolvation of hydrogen-bonding partners, but on the energy cost of insisting on Watson–Crick geometry by a passive and rigid decoding center (Fig. 3E). Energy cost for formation of tautomers (or repulsion energy due to steric clash within the mismatched base pair) in the case of G-U mismatches could constitute the sole cause for the very efficient rejection of near-cognate tRNAs by the ribosome at each step of decoding process. 

Our data suggest that "geometrical mimicry" can explain how wrong aa-tRNAs with mismatches forming base pairs with Watson–Crick geometry in the decoding center can be incorporated into the poly-peptide chain.

**Acknowledgments**

Our own work summarized in this review was supported by ANR BLAN07-3_190451 (M.Y.), ANR-07-PCVI-0015-01 (G.Y.), FRM (N.D.) and by the European Commission SPINE2.

**References**


Schmeng, T.M. et al. (2009) The crystal structure of the ribosome bound to EF-Tu and aminoacyl-tRNA. Science 326, 688–694.


