HHS Public Access

Author manuscript

Nature. Author manuscript; available in PMC 2014 February 01.

Published in final edited form as:

Nature. 2013 August 1; 500(7460): 107–110. doi:10.1038/nature12302.

Unusual base pairing during the decoding of a stop codon by the ribosome

Israel S. Fernández[1], **Chyan Leong Ng**[1], **Ann C. Kelley**[1], **Guowei Wu**[2], **Yi-Tao Yu**[2], and **V. Ramakrishnan**[1]

[1]MRC Laboratory of Molecular Biology, Cambridge CB2 0QH, United Kingdom

[2]Department of Biochemistry and Biophysics, University of Rochester Medical Center, 601 Elmwood Avenue, Rochester, NY 14642, USA

Abstract

During normal translation, binding of a release factor to one of the three stop codons (UGA, UAA or UAG) results in termination of protein synthesis. However, modification of the initial uridine to a pseudouridine (Ψ) allows efficient recognition and read-through of these stop codons by a transfer RNA (tRNA), although it requires formation of two normally forbidden purine-purine base pairs[1]. We have determined the crystal structure at 3.1 Å resolution of the 30S ribosomal subunit in complex with the anticodon stem loop of $tRNA^{Ser}$ bound to the ΨAG stop codon in the A site. The ΨA base pair at the first position is accompanied by the formation of purine-purine base pairs at the second and third positions of the codon, which display an unusual Watson-Crick/Hoogsteen geometry. The structure shows a previously unsuspected ability of the ribosomal decoding center to accommodate non-canonical base pairs.

The genetic code normally requires Watson-Crick base pairing at the first two positions of the codon-anticodon helix, and tolerates certain specific mismatches at the third (wobble) position. The structural basis for this was apparent when three universally conserved bases in the A site of the small (30S) ribosomal subunit were shown to change conformation to monitor Watson-Crick geometry at the first two base pairs in the minor groove of the codon-anticodon helix, while leaving the wobble position relatively unconstrained[2]. These interactions result in additional binding energy that is used to induce global conformational changes that facilitate the hydrolysis of GTP by EF-Tu[3–5], as was predicted by earlier kinetic data[6]. However, recently it was shown that modification of the uridine in stop codons

Correspondence to: Yi-Tao Yu; V. Ramakrishnan.

Author Contributions ISF carried out the crystallographic experiments and analysis and helped write the paper, GW did the in vitro translation assays, CLNG helped with crystallographic data collection, ACK made the 30S subunits, 70S ribosome and $tRNA^{Ser}$, and Yi-TY and VR oversaw the project and helped write the paper.

Author Information The coordinates and structure factors have been deposited in the Protein Data Bank under the accession numbers 4JV5, 4JYA, 4K0K (30S) and 4K0L, 4K0M,4K0P, 4K0Q (70S). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to V.R. (ramak@mrc-lmb.cam.ac.uk) or Y.-T.Y. (YiTao_Yu@URMC.Rochester.edu).

with pseudouridine (Ψ; Fig. 1a) allowed ΨAA and ΨAG to code for serine or threonine and ΨGA for phenylalanine or tyrosine[1].

To understand how the normally forbidden base pairs in such recoding can be accepted by the ribosome, we have determined crystal structures of the anticodon stem loops (ASLs) of $tRNA^{Ser}$ and $tRNA^{Phe}$ bound to the modified stop codons ΨAG and ΨGA, respectively, in the A site of the 30S subunit, as well as the crystal structure of the most common bacterial $tRNA^{Ser}$ bound to ΨAG in the entire 70S ribosome.

We first showed that nonsense suppression by pseudouridine also occurs in bacteria. We constructed three synthetic mRNAs encoding a 6X-histidine tag at the N-terminus and a FLAG tag at the C-terminus (Fig. 1b, top). The three mRNAs contain either an amber stop (UAG), a glutamine sense (CAG). or a pseudouridylated amber (ΨAG) codon inserted just before the FLAG tag. After *in vitro* translation, anti-6X-histidine immunoblotting revealed that the three constructs are translated at the same level (Fig.1b, bottom left). The anti-Flag immunoblot had a signal comparable to background when the normal UAG amber codon was present, showing that normal termination occurred and the downstream Flag sequence was not translated (Fig. 1b, bottom right). However, the presence of the ΨAG amber codon increased the signal from the Flag tag to a level comparable to that of the CAG sense codon, showing that the substitution of U by Ψ results in a strong read-through of the amber stop codon in bacteria, as previously reported in eukaryotes[1]. Given the conservation of the decoding center, it is likely that the recoding also specifies the same amino acids as previously established in eukaryotes.

Since recoding and the decoding center are both conserved across the two kingdoms, we determined the structure of the ASL for the yeast $tRNA^{Ser}$, which has an IGA anticodon (http://gtrnadb.ucsc.edu/Sacc_cere/) (Fig. 1c) bound to a modified ΨAG stop codon in the 30S subunit as done previously for cognate tRNA[2] (Supplementary Methods and Table S1). Clear difference Fourier density was seen for both the codon and the ASL (Fig. 2a), allowing an unambiguous determination of the conformation of the bases involved (Fig. 2b). The overall conformations of the ΨAG codon and ASL in the A site are very similar to those observed previously for phenylalanine codon-ASL pairs in both the 30S subunit[2,3] and the intact 70S ribosome[7]. However the details of each of the codon-anticodon base pairs are strikingly different, and reveal how the ribosome facilitates the decoding of a normally non-cognate tRNA.

At the first position, the Ψ1 of the codon and A36 of the tRNA form the expected Watson-Crick base pair (Fig. 2b, left, and Fig. 3a). The N1 of Ψ is exposed to solvent and does not form an additional interaction with either the ASL or the ribosome. The minor groove of this base pair is recognized by A1493 of the decoding centre (Fig. 3a). However, unlike the case of normal decoding (Fig. 3d), the electron density for A1493 is consistent with a *syn* rather than an *anti* conformation (Fig. S1).

In the second position, the A2 of the codon adopts a *syn* conformation that forms two hydrogen bonds via its Hoogsteen edge[8] with the Watson-Crick edge of G35 of the ASL (Fig. 2b, center and Fig. 3b). Interestingly, the C1′-C1′ distance of 10.9 Å in this non-

canonical base pair is not much larger than the distance of 10.5 Å found in a canonical pyrimidine-purine base pair with Watson-Crick geometry[2]. Presumably this difference is small enough to allow the base pair to be accommodated in the decoding centre of the ribosome. Surprisingly, despite the completely different type of base pair, the conformation and interactions of the ribosomal bases A1492 and G530 are very similar to that seen in normal decoding (Fig. 3b and e).

Yeast $tRNA^{Ser}$ has an inosine in the third (wobble) position of the anticodon[9]. In suppression of a ΨAG codon, the inosine would have to make a previously unknown I-G base pair, which would bring the two normally repulsive O6 groups close together. Similar to the second position, this unusual base pair is formed by a *syn* conformation of G3 making a single hydrogen bond via its Hoogsteen face with the Watson-Crick edge of I34 of the ASL (Fig. 2b, right and Fig. 3c). The potential electrostatic repulsion between the O6 atoms on G3 and I34 is overcome by coordination of both atoms with what is likely a magnesium ion, thus stabilizing the base pair.

To determine the generality of these interactions, we determined two additional structures. The first was that of bacterial $tRNA^{Ser}$ with a CGA anticodon, bound to mRNA in the entire 70S ribosome. Here, only the second base pair involves a purine-purine mismatch, which adopts the same Hoogsteen/Watson-Crick geometry as seen in the 30S (Table S1, Fig. S2a and c). We also determined the structure of the 30S subunit with a ΨGA codon with ASL for $tRNA^{Phe}$ containing a GAA anticodon (Table S1, Fig. S2b), since it was shown that ΨGA could be decoded by phenylalanine[1]. The same type of base pairing is seen, including the *syn* conformation of A1493 at the first position and the Hoogsteen/Watson-Crick interactions at the second and third positions. Since the third position here involves a GA pair instead of an IG, there are two hydrogen bonds and there is no coordinating magnesium ion. These structures show that the unusual Hoogsteen/Watson-Crick base pairs for purine-purine mismatches are seen for both bacterial and eukaryotic tRNAs implicated in recoding of stop codons, and in the context of both 30S subunits and the 70S ribosome.

The structure of $tRNA^{Ser}$ ASL with the unmodified UAG codon in the 30S subunit (data not shown; PDB code: 4K0K) was identical within coordinate error to that with the ΨAG codon. This suggests that any difference in the decoding properties must come from differences in the stability of the final complex due to pseudouridine, rather than differences in structure. This work shows how the normally forbidden codon-anticodon pairs that would be involved in recoding by pseudouridine-containing stop codons are accommodated in the ribosomal decoding center.

At the first codon position, the *syn* conformation of A1493 facilitates an alternate and possibly tighter interaction with the minor groove of the Watson-Crick ΨA pair (Fig. 2b, left, and Fig. 3a). The Watson-Crick/Hoogsteen base pairs at the second and third positions seen here have also been observed in canonical duplex DNA[10], showing that the energetic difference between canonical and Hoogsteen pairs is small enough for the latter to be observed as a transient feature. However, during decoding, the ribosomal bases A1492 and G530 were suggested to specifically recognize Watson-Crick pairing geometry at the second position (Fig. 3e)[2]. Nevertheless, the same interactions can be formed with the Watson-

Crick/Hoogsteen base pair here (Fig. 3b), presumably because its width is similar and the interactions are with the 2′ OH groups of the riboses of the codon-anticodon base pair rather than with the bases themselves. It is likely that these interactions are not as favorable as for a canonical Watson-Crick base pair. However, suppression of stop codons by tRNAs requires only that they are able to out-compete release factors; they do not have to be as efficient as cognate tRNAs on sense codons.

It is not apparent why the pseudouridine modification should facilitate binding of a non-cognate tRNA. The additional hydrogen bond donor N1 of Ψ, which is not present in U (Fig. 1a), is exposed to solvent and does not make additional interactions with the ribosome or ligands. Pseudouridine in RNA has been shown to result in an increased stabilization of helices without noticeable changes in structure[11,12]. Understanding how pseudouridine results in increased stabilization of helices will require both high-resolution data on model systems as well as computational and biochemical studies. Determining the effect of pseudouridine on the rate constants of various steps in decoding[6], or on efficiency of termination, will also help to clarify its role.

In principle, modification of sense codons could also lead to alternate forms of recognition. In this context it is interesting that a pseudouridine in the anticodon of a tRNA leads to altered coding of a sense codon[13]. Finally, the work shows that the decoding center of the ribosome has a previously unsuspected plasticity that under certain circumstances allows alternative non-canonical base pairing, thereby allowing an expansion of the genetic code.

Full Methods

Bacterial *in vitro* translation assay

Reporter mRNAs were synthesized essentially as described previously[1]. Briefly, we generated a DNA template using overlap extension PCR. The template thus generated contained a T7 promoter, followed sequentially by a Shine-Dalgarno (SD) sequence, a His6 tag, a termination codon (TAG), and a Flag tag (see Fig 1), with the sequence:

5′-GGG AGA GGA GGC AAC GAC ATG GCC GCC CAC CAC CAC CAC CAC CAC GCC GGC GCC GAA CAG AAG GAA GAA GAC CGC GAG (T/C/Ψ AG GAC TAC AAG GAC GAC GAC GAC AAG GCC TAG-3′

In vitro transcription with T7 RNA polymerase was used to generate the reporter mRNA transcript containing an authentic termination codon (UAG) immediately upstream of the Flag tag. Using the same strategy, we also generated a control mRNA transcript, where a CAG sense codon replaced the UAG codon. In addition, oligonucleotide-mediated two-piece splint ligation was used to generate another mRNA containing a YAG codon in place of the UAG or CAG codon, as described previously[1]. The 5′ half RNA was in vitro transcribed, ending with GAG at its 3′ end. The 3′ half RNA was chemically synthesized, beginning with ΨAG at its 5′ end. The two RNA fragments were aligned through hybridization with a complementary bridging oligodeoxynucleotide, and ligated by T4 DNA ligase. Thus all three mRNA transcripts have the same sequence, except for the codon immediately upstream of the Flag tag (UAG, CAG or ΨAG).

Translation in bacterial cell lysate and dot blot analysis of proteins

In vitro translation was carried out in a 37-μl reaction using *E. coli* cell lysate (Cosmo Bio Co) according to the manufacturer's instructions. Immediately after the reaction, dot-blot was performed as described[1]. Briefly, small aliquots of the reaction (2.5 μl and 2.5 μl of a 5-fold dilution) were spotted onto the nitrocellulose membrane, and probed with a monoclonal anti-His antibody (H-3; Santa Cruz Biotechnology) or a monoclonal anti-Flag antibody (M2; Sigma-Aldrich). Goat anti-mouse IgG (H+L)-alkaline phosphatase (AP) conjugate (Bio-Rad) was then used as a secondary antibody. His and Flag signals were visualized using 1-Step NBT/BCIP (Pierce).

Crystallization

For 30S experiments, the anticodon stem loops of tRNAs (ASLs) and the hexanucleotide mimics of mRNA with sequence 5′-ΨAGΨAG-3′ (for ASL^{Ser}) or 5′-ΨGAΨGA-3′ (for ASL^{Phe}) were chemically synthesized (Dharmacon). The ASLs were designed with an extra G26-C44 base pair to improve stability. The $tRNA^{Ser}$ was produced and purified as previously reported[16].

Thermus thermophilus 30S ribosomal subunits were purified, crystallized and cryoprotected as previously described[2,7]. In the 30S experiments, crystals were soaked for 48 hr in cryoprotection buffer containing 100 μM of the mRNA, 100 μM of the ASL and then flash frozen in liquid nitrogen. In the case of the ASL for $tRNA^{Phe}$ with a ΨAG codon, the addition of paromomycin to100 μM in the cryoprotection buffer improved diffraction from about 3.7 Å to 3.1 Å, so that structure was solved in the presence of paromomycin. However within the limits of resolution, no difference was seen in the conformation of the decoding center. For the 70S experiment, a complex of the ribosome with mRNA, $tRNA^{fMet}$ in the P site and $tRNA^{Ser}$ in the A site was crystallized as previously described[7]. The chemically synthesized mRNA was identical to that previously used, except for a ΨAG codon in the A site.

Data collection and refinement

Data were collected at Diamond Light Source, Harwell Campues, Oxford, UK beamline I04, integrated and scaled with XDS[17] and refined using REFMAC[18]. Visualization and model building was done with COOT[19]. An initial round of rigid body refinement using the published structure of the *Thermus themophilus* 30S (PDB ID 1FJF) or the 70S ribosome (PDB 2WH1 and 2WH2) was followed by restrained maximum-likelihood refinement. A final "jelly body" refinement in REFMAC[18] further lowered the R-factor and also significantly improved the quality of the maps. The mRNA and ASL or tRNA were clearly seen in difference Fourier maps. The final model including the ligands and conformational changes in the ribosome was refined as before. Density for a solvent atom adjacent to the O6 of G and I at the third base pair was assigned to a Mg^{2+} ion based on its coordination and the fact that a water molecule at this position refined to an abnormally low B factor. All figures were drawn using PyMOL[20].

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank D. Hall and G. Winter for help and advice with data collection at beamline I04, Diamond Light Source, Harwell Campus, Oxford, UK; T. Tomizaki at beamline X06SA for help with data collection at the Swiss Light Source, Villigen, Switzerland; and A. McCarthy at beamline ID14-4, ESRF, Grenoble, France, where screening and initial data collection was done. We thank M. Härtlein for a gift of an overproducing $tRNA^{Ser}$ clone, M. Torrent for advice on yeast tRNA abundance and G. Murshudov for advice and help with data analysis and refinement. V.R. was supported by the UK Medical Research Council (grant U105184332), a Programme Grant and Senior Investigator Award from the Wellcome Trust, the Agouron Institute and the Louis-Jeantet Foundation. Y-T.Y. was supported by a grant from the National Institute of Health (GM104077), and by the University of Rochester CTSA award (UL1TR000042) from the National Center for Advancing Translational Sciences of the National Institute of Health. I.S.F. was supported by a postdoctoral fellowship from the Fundacion Ramon Areces, Spain.

References

1. Karijolich J, Yu YT. Converting nonsense codons into sense codons by targeted pseudouridylation. Nature. 2011; 474:395–398. [PubMed: 21677757]
2. Ogle JM, et al. Recognition of cognate transfer RNA by the 30S ribosomal subunit. Science. 2001; 292:897–902. [PubMed: 11340196]
3. Ogle JM, Murphy FV, Tarry MJ, Ramakrishnan V. Selection of tRNA by the ribosome requires a transition from an open to a closed form. Cell. 2002; 111:721–732. [PubMed: 12464183]
4. Schmeing TM, et al. The crystal structure of the ribosome bound to EF-Tu and aminoacyl-tRNA. Science. 2009; 326:688–694. [PubMed: 19833920]
5. Voorhees RM, Schmeing TM, Kelley AC, Ramakrishnan V. The mechanism for activation of GTP hydrolysis on the ribosome. Science. 2010; 330:835–838. [PubMed: 21051640]
6. Pape T, Wintermeyer W, Rodnina MV. Complete kinetic mechanism of elongation factor Tu-dependent binding of aminoacyl-tRNA to the A site of the *E. coli* ribosome. EMBO J. 1998; 17:7490–7497. [PubMed: 9857203]
7. Selmer M, et al. Structure of the 70S ribosome complexed with mRNA and tRNA. Science. 2006; 313:1935–1942. [PubMed: 16959973]
8. Leontis NB, Westhof E. Geometric nomenclature and classification of RNA base pairs. RNA. 2001; 7:499–512. [PubMed: 11345429]
9. Sprinzl M, Horn C, Brown M, Ioudovitch A, Steinberg S. Compilation of tRNA sequences and sequences of tRNA genes. Nucleic Acids Res. 1998; 26:148–153. [PubMed: 9399820]
10. Nikolova EN, et al. Transient Hoogsteen base pairs in canonical duplex DNA. Nature. 2011; 470:498–502. [PubMed: 21270796]
11. Davis DR, Veltri CA, Nielsen L. An RNA model system for investigation of pseudouridine stabilization of the codon-anticodon interaction in tRNALys, tRNAHis and tRNATyr. J Biomol Struct Dyn. 1998; 15:1121–1132. [PubMed: 9669557]
12. Yarian CS, et al. Structural and functional roles of the N1- and N3-protons of psi at tRNA's position 39. Nucleic Acids Res. 1999; 27:3543–3549. [PubMed: 10446245]
13. Tomita K, Ueda T, Watanabe K. The presence of pseudouridine in the anticodon alters the genetic code: a possible mechanism for assignment of the AAA lysine codon as asparagine in echinoderm mitochondria. Nucleic Acids Res. 1999; 27:1683–1689. [PubMed: 10076000]
14. Huang C, Wu G, Yu YT. Inducing nonsense suppression by targeted pseudouridylation. Nat Protoc. 2012; 7:789–800. [PubMed: 22461068]
15. Winn MD, Murshudov GN, Papiz MZ. Macromolecular TLS refinement in REFMAC at moderate resolutions. Methods Enzymol. 2003; 374:300–321. [PubMed: 14696379]
16. Borel F, Hartlein M, Leberman R. In vivo overexpression and purification of Escherichia coli tRNA(ser). FEBS Lett. 1993; 324:162–166. [PubMed: 7685296]
17. Kabsch W. XDS. Acta Crystallogr D Biol Crystallogr. 2010; 66:125–132. [PubMed: 20124692]

18. Murshudov GN, et al. REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallogr D Biol Crystallogr. 2011; 67:355–367. [PubMed: 21460454]
19. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. Acta Crystallogr D Biol Crystallogr. 2010; 66:486–501. [PubMed: 20383002]
20. DeLano, WL. The PyMOL Molecular Graphics System. 2006. http://www.pymol.org

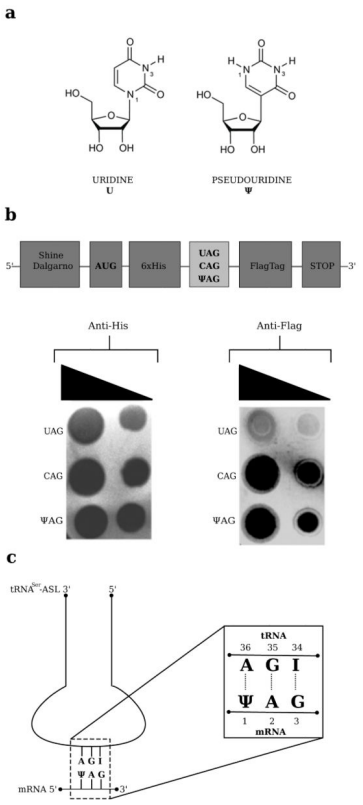

FIGURE 1. Chemical differences between uridine and pseudouride and experimental set-up

(a) Uridine (U, 1-ß-D-ribofuranosyluracil) and pseudouridine (Ψ, 5-ß-D-ribofuranosyluracil). (b) Top, diagram of the three synthetic mRNA constructs designed for the *in vitro* nonsense suppression experiment in bacteria. Bottom, anti-His and anti-Flag immunoblot analysis of the *in-vitro* translation assays in *E. coli*. (c) The $tRNA^{Ser}$ anticodon stem loop (ASL) and mRNA used in this study.

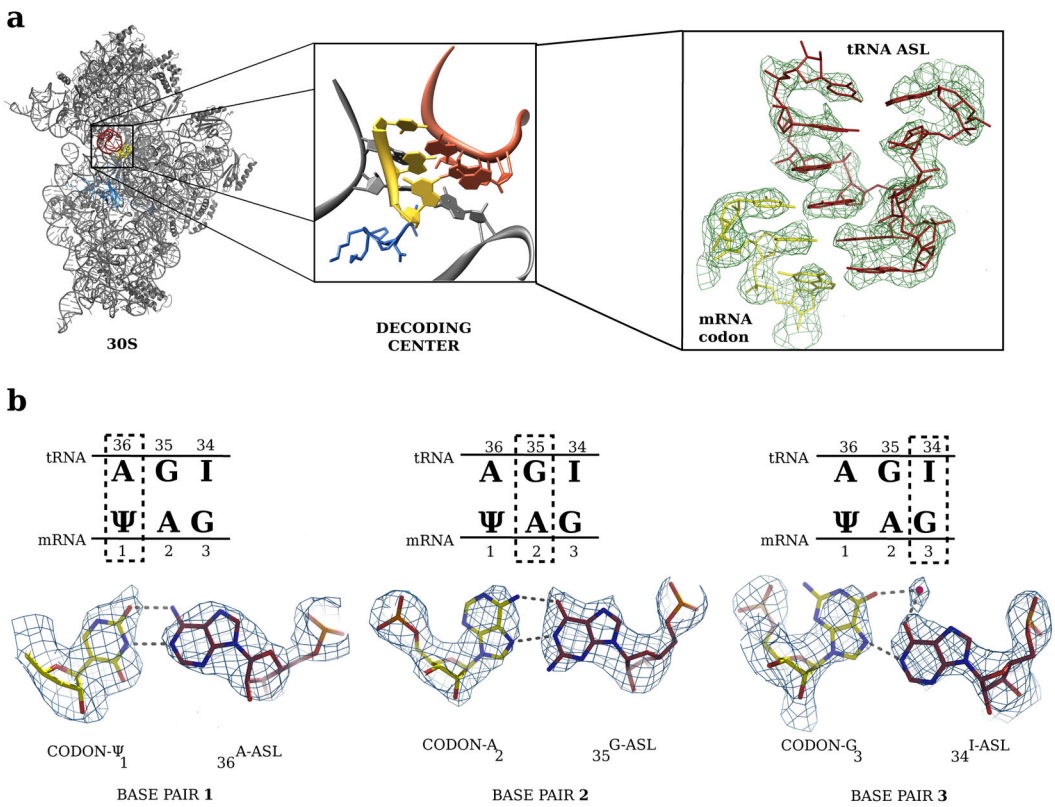

FIGURE 2. Overall and detailed view of the base pairs involved in the codon/anticodon interaction

(a) Overview of the 30S subunit and detail of the 30S decoding site with the ΨAG/ASL-$tRNA^{Ser}$. The ASL-$tRNA^{Ser}$ is depicted in red, the ΨAG codon in yellow, the 16S rRNA in gray and protein S12 in blue. Right, unbiased difference Fourier density for the ASL and codon. (b) The base pairs for the codon-anticodon interaction are shown with boxes over the first (left), second (middle) and third (right) positions; the actual structures of the base pair is shown below in unbiased difference Fourier maps contoured at 1.5σ.

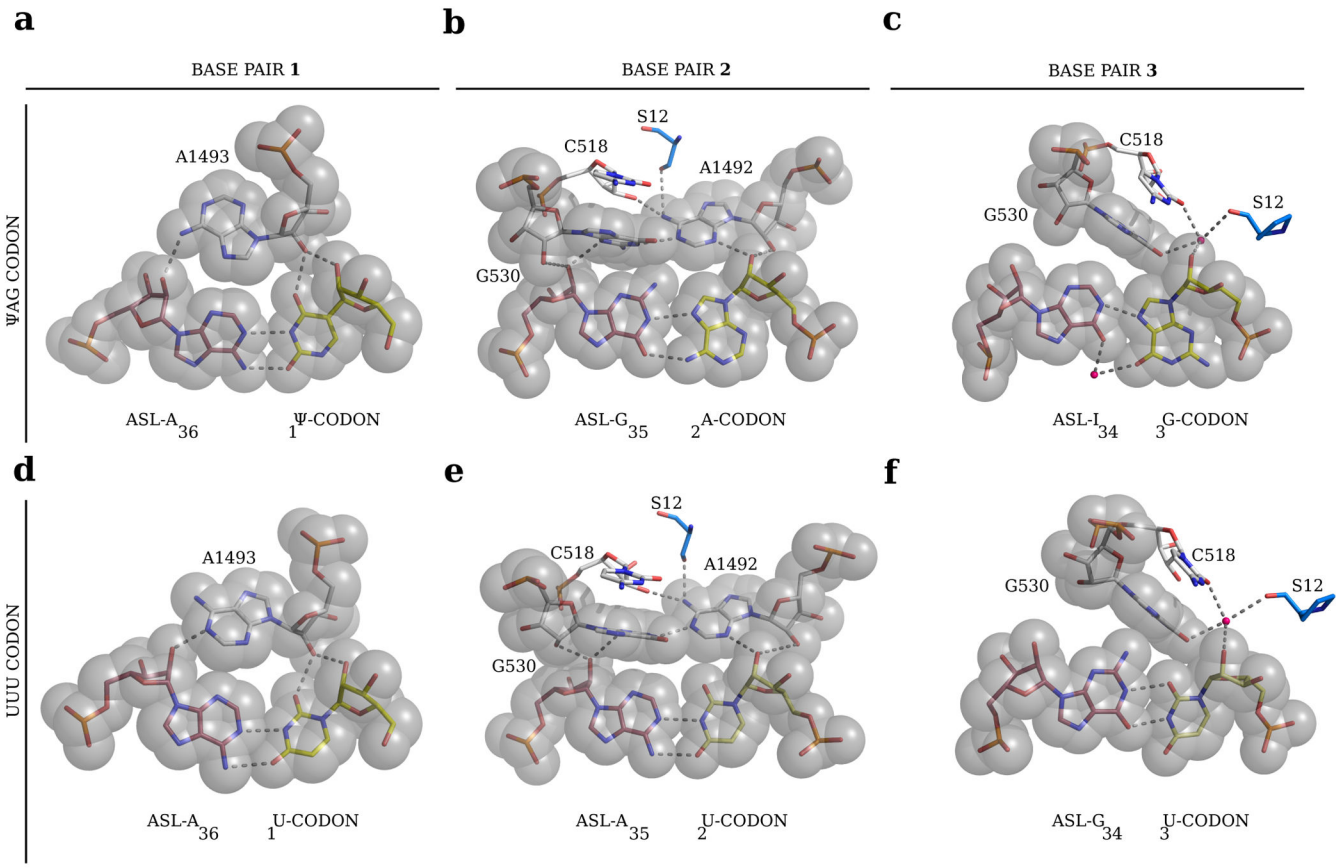

FIGURE 3. Interaction of ribosomal bases with the codon-anticodon base pairs

Details of base pairing for the first, second and third positions of the ΨAG/ASL-tRNA^{Ser} structure with the ribosomal bases involved in minor groove recognition are shown on the left. For comparison, the corresponding positions of the codon-anticodon base pairs of a phenylalanine UUU codon with its cognate ASL anticodon from a previous study[2] are shown on the right.