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

Israel S. Fernández1, Chyan Leong Ng1, Ann C. Kelley1, Guowei Wu2, Yi-Tao Yu2 & V. Ramakrishnan1

During normal translation, the binding of a release factor to one of the three stop codons (UGA, UAA or UAG) results in the 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 the formation of two normally forbidden purine–purine base pairs1. Here we determined the crystal structure at 3.1 Å resolution of the 30S ribosomal subunit in complex with the anticodon stem loop of tRNA⁡Ser⁢ψ 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 show an unusual Watson–Crick/Hoogsteen geometry. The structure shows a previously unsuspected ability of the ribosomal decoding centre 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 unconstrained2. These interactions result in additional binding energy that is used to induce global conformational changes that facilitate the hydrolysis of GTP by the elongation factor EF-Tu3–5, as was predicted by earlier kinetic data6. However, recently it was shown that modification of the uridine in stop codons to Ψ (Fig. 1a) allowed ΨAA and ΨAG to code for serine or threonine and ΨGA to code for phenylalanine or tyrosine1.

To understand how the normally forbidden base pairs in such recording can be accepted by the ribosome, we determined the crystal structures of the anticodon stem loops (ASLs) of tRNA⁡Ser and tRNA⁡Phe encoding a 6x-histidine tag at the amino terminus and a Flag tag at the carboxy 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 eukaryotes7. Given the conservation of the

Figure 1 | Chemical differences between uridine and pseudouridine, and experimental set-up. a. Uridine (U; 1β-D-ribofuranosyluracil) and pseudouridine (Ψ; 5β-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 Escherichia coli. c. The tRNA⁡Ser⁢ψ ASL and mRNA used in this study.

1MRC Laboratory of Molecular Biology, Cambridge CB2 0QH, UK. 2Department of Biochemistry and Biophysics, University of Rochester Medical Center, 601 Elmwood Avenue, Rochester, New York 14642, USA. †Present address: Institute of Systems Biology, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia.
decoding centre, it is likely that the recoding also specifies the same amino acids as previously established in eukaryotes.

As recoding and the decoding centre are both conserved across the two kingdoms, we determined the structure of the ASL for the yeast tRNA\textsuperscript{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 previously described for cognate tRNA\textsuperscript{Ser} (Supplementary Methods and Supplementary Table 1).

The codon and the ASL could clearly be seen in difference Fourier maps (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\textsuperscript{23} and the intact 70S ribosome.\textsuperscript{2} However, the details of each of the codon–anticodon base pairs are markedly different, and reveal how the ribosomefacilitates 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 in normal decoding (Fig. 3d), the electron density for A1493 is consistent with a syn rather than an anti conformation (Supplementary Fig. 1).

In the second position, the A2 of the codon adopts a syn conformation that forms two hydrogen bonds via its Hoogsteen edge\textsuperscript{9} with the Watson–Crick edge of G35 of the ASL (Fig. 2b, centre, and Fig. 3b). Interestingly, the C1′–C1′ distance of 10.9 Å in this non-canonical base pair is not much greater than the distance of 10.5 Å found in a canonical pyrimidine–purine base pair with Watson–Crick geometry.\textsuperscript{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 those seen in normal decoding (Fig. 3b, c).

Yeast tRNA\textsuperscript{Ser} has an inosine in the third (wobble) position of the anticodon.\textsuperscript{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. As in 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 probably 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\textsuperscript{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 ribosome (Supplementary Table 1 and Supplementary Fig. 2a, c). We also determined the structure of the 30S subunit with a ΨGA codon bound to the ASL of tRNA\textsuperscript{Phc} containing a GAA anticodon (Supplementary Table 1 and Supplementary Fig. 2b), as it was shown that ΨGA could be decoded by phenylalanine.\textsuperscript{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. Because 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 the recoding of stop codons, and in the context of both 30S subunits and the 70S ribosome.

The structure of tRNA\textsuperscript{Ser} ASL bound to the unmodified UAG codon in the 30S subunit (data not shown; Protein Data Bank (PDB) accession

---

**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\textsuperscript{Ser}. The ASL of tRNA\textsuperscript{Ser} is depicted in red, the ΨAG codon in yellow, the 16S rRNA in grey and the S12 protein 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 are shown below in unbiased difference Fourier maps contoured at 1.5σ. Subscripts indicate residue numbers.
In RNA, Watson–Crick pairing at the second position 2 (Fig. 3e). Nevertheless, the same interactions can be formed with the Watson–Crick/Hoogsteen base pair described 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 favourable as for a canonical Watson–Crick base pair. However, suppression of stop codons by tRNAs requires only that they are able to outcompete release factors; they do not have to be as efficient as cognate tRNAs on sense codons.

It is not apparent why the Ψ 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. Ψ in RNA has been shown to result in an increased stabilization of helices without noticeable changes in structure11,12. Understanding how Ψ 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 Ψ on the rate constants of various steps in decoding6, or on efficiency of termination, will also help to clarify its role.

In principle, modification of sense codons could also lead to alternative forms of recognition. In this context it is interesting that a Ψ in the positions of the codon–anticodon base pairs of a phenylalanine UUU codon with its cognate ASL anticodon from a previous study7 are shown on the bottom (d–f). Subscripts indicate residue numbers.

**METHODS SUMMARY**

A combination of in vitro transcribed and chemically synthesized RNAs were used to generate the three mRNA versions used in the in vitro translation assays in E. coli. Aliquots of the in vitro translation reactions were spotted onto nitrocellulose membranes and probed with monoclonal anti-His or anti-Flag antibodies14. Crystallization of 30S ribosomal subunits and their complex formation with tRNA ASLs and hexanucleotide mRNAs were done as previously reported2. An initial round of refinement with REFMAC15 using only the 30S subunit or the empty 70S ribosome as starting models to avoid any model bias, clearly showed additional differences in density that could be attributable to mRNA and tRNA, or to ASL, as shown in Fig. 2.

Received 19 October 2012; accepted 15 May 2013.
Published online 30 June 2013.

1. Karijolich, J. & Yu, Y.-T. Converting nonsense codons into sense codons by targeted pseudouridylation. Nature **474**, 395–398 (2011).
2. Ogle, J. M. et al. Recognition of cognate transfer RNA by the 30S ribosomal subunit. Science **292**, 897–902 (2001).
3. Ogle, J. M., Murphy, F. V., Tarry, M. J. & Ramakrishnan, V. Selection of RNA by the ribosome requires a transition from an open to a closed form. Cell **111**, 721–732 (2002).
4. Schmeing, T. M. et al. The crystal structure of the ribosome bound to EF-Tu and aminoacyl-tRNA. Science **326**, 688–694 (2009).
5. Voorhees, R. M., Schmeing, T. M., Kelley, A. C. & Ramakrishnan, V. The mechanism for activation of GTP hydrolysis on the ribosome. Science **330**, 835–838 (2010).
6. Pape, T., Wintermeyer, W. & Rodnina, M. V. Complete kinetic mechanism of elongation factor Tu-dependent binding of aminoacyl-tRNA to the A site of the E. coli ribosome. EMBO J. **17**, 7490–7497 (1998).
7. Selmer, M. et al. Structure of the 70S ribosome complexed with mRNA and tRNA. Science **313**, 1935–1942 (2006).
8. Leonis, N. B. & Westhof, E. Geometric nomenclature and classification of RNA base pairs. RNA **7**, 499–512 (2001).
9. Sprinzl, M., Horn, C., Brown, M., Loidl, V. & Steinberg, S. Compilation of tRNA sequences and sequences of tRNA genes. Nucleic Acids Res. **26**, 148–153 (1998).
10. Nikolova, E. N. et al. Transient Hoogsteen base pairs in canonical duplex DNA. *Nature* **470**, 498–502 (2011).

11. Davis, D. R., Veltri, C. A. & Nielsen, L. An RNA model system for investigation of pseudouridine stabilization of the codon-anticodon interaction in tRNA<sup>val</sup>, tRNA<sup>his</sup> and tRNA<sup>tyr</sup>. *J. Biomol. Struct. Dyn.* **15**, 1121–1132 (1998).

12. Yarian, C. S. et al. Structural and functional roles of the N1- and N3-protons of Y at tRNA's position 39. *Nucleic Acids Res.* **27**, 3543–3549 (1999).

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.* **27**, 1683–1689 (1999).

14. Huang, C., Wu, G. & Yu, Y.-T. Inducing nonsense suppression by targeted pseudouridylation. *Nature Protocols* **7**, 789–800 (2012).

15. Winn, M. D., Murshudov, G. N. & Papiz, M. Z. Macromolecular TLS refinement in REFMAC at moderate resolutions. *Methods Enzymol.* **374**, 300–321 (2003).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank D. Hall and G. Winter for help and advice with data collection at beamline I04, Diamond Light Source; T. Tomizaki at beamline X06SA for help with data collection at the Swiss Light Source; and A. McCarthy at beamline ID14-4, ESRF, where screening and initial data collection were done. We thank M. Härlein for the gift of an overproducing tRNA<sup>val</sup> 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-Jeanet Foundation. Y.-T.Y. was supported by a grant from the National Institutes of Health (GM104077), and by the University of Rochester CTSA award (UL1TR000042) from the National Center for Advancing Translational Sciences of the National Institutes of Health. I.S.F. was supported by a postdoctoral fellowship from the Fundacion Ramon Areces.

Author Contributions I.S.F. carried out the crystallographic experiments and analysis and helped write the paper, G.W. did the in vitro translation assays, C.L.N. helped with crystallographic data collection, A.C.K. made the 30S subunits, 70S ribosome and tRNA<sup>val</sup>, and Y.-T.Y. and V.R. oversaw the project and helped write the paper.

Author Information The coordinates and structure factors have been deposited in the PDB under accessions 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).
METHODS

**Bacterial in vitro translation assay.** Reporter mRNAs were synthesized essentially as described previously1. 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’-GGGGAGAGGCGACGACATGGCGCCCGGCACCCACCAACCCACGACGACGACGACGACGACAAGGCCTAG-3’

GCCGGCGCCCGACAGGAAAGAAGACCCCGAG(T/C/P)AGGACCTACAAGGACGACGACGACGGCCTAG-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 transcript, where a CAG sense codon replaced the UAG codon. In addition, the silent bridging oligodeoxynucleotide, and ligated by T4 DNA ligase.

**Translation in bacterial cell lysate and dot-blot analysis of proteins.** In *vitro* translation was carried out in a 37°C reaction using *E. coli* cell lysate (Cosmo Bio Co) according to the manufacturer’s instructions. Immediately after the reaction, dot-blot analysis was performed as described1. Briefly, small aliquots of the reaction (2.5 μl and 2.5 μl of a fivefold 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 conjugate (Bio-Rad) was then used as a secondary antibody. His and Flag signals were visualized using 1-Step NBT/BCIP (Pierce).

**Crystalization.** For 30S experiments, the ASLs of tRNAs and the hexanucleotide mimics of mRNA with the sequence 5’-ΨAGΨAG-3’ (for ASL1wt) or 5’-ΨGAΨGA-3’ (for ASL1wt) were chemically synthesized (Dharmacon). The ASLs were designed with an extra G26–C44 base pair to improve stability. The tRNASer was produced and purified as previously reported5. *Thermus thermophilus* 30S ribosomal subunits were purified, crystallized and cryoprotected as previously described12. In the 30S experiments, crystals were soaked for 48 h in cryoprotection buffer containing 100 μM of the mRNA and 100 μM of the ASL, and then flash frozen in liquid nitrogen. In the case of the ASL for tRNAPro with a ΨAG codon, the addition of paromomycin to 100 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 paromycin. However, within the limits of resolution, no difference was seen in the conformation of the decoding centre. For the 70S experiment, a complex of the ribosome with mRNA, tRNAAsp10 in the P site and tRNASer in the A site was crystallized as previously described1. 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 the Diamond Light Source, beamline I04, integrated and scaled with XDS17 and refined using REFMAC18. Visualization and model building was done with COOT19. An initial round of rigid body refinement using the published structure of the *T. thermophilus* 30S (PDB accession 1FJF) or the 70S ribosome (PDB accessions 2WH1 and 2WH2) was followed by restrained maximum-likelihood refinement. A final ‘jelly body’ refinement in REFMAC18 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 earlier. The density for a solvent atom adjacent to the O6 of G and I at the third base pair was assigned to a Mg2+ 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 PyMOL20.

16. Borel, F., Hartlein, M. & Leberman, R. *In vivo* overexpression and purification of *Escherichia coli* tRNA18. *FEBS Lett.* **324**, 162–166 (1993).
17. Kapush, W., XDS. Acta Crystallogr. D **66**, 125–132 (2010).
18. Murshudov, G. N. et al. REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallogr. D **67**, 355–367 (2011).
19. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. Acta Crystallogr. D **66**, 486–501 (2010).
20. DeLano, W. L. The PyMOL Molecular Graphics System. http://www.pymol.org (2006).