

Mechanism for expanding the decoding capacity of transfer RNAs by modification of uridines

Albert Weixlbaumer¹, Frank V Murphy IV¹, Agnieszka Dziergowska², Andrzej Malkiewicz², Franck A P Vendeix³, Paul F Agris³ & V Ramakrishnan¹

One of the most prevalent base modifications involved in decoding is uridine 5-oxyacetic acid at the wobble position of tRNA. It has been known for several decades that this modification enables a single tRNA to decode all four codons in a degenerate codon box. We have determined structures of an anticodon stem-loop of tRNA^{Val} containing the modified uridine with all four valine codons in the decoding site of the 30S ribosomal subunit. An intramolecular hydrogen bond involving the modification helps to prestructure the anticodon loop. We found unusual base pairs with the three noncomplementary codon bases, including a G•U base pair in standard Watson-Crick geometry, which presumably involves an enol form for the uridine. These structures suggest how a modification in the uridine at the wobble position can expand the decoding capability of a tRNA.

It is thought that the primitive genetic code might have been a 'two-letter' code, with the third position in the codon being degenerate¹. However, the present genetic code has evolved so that only 8 of the 16 codon boxes defined by the first two positions are fully degenerate—that is, code for the same amino acid (Fig. 1a). Such degenerate codon boxes are generally decoded by different isoaccepting tRNAs². The four-fold degenerate codon box for valine is decoded by two different *Escherichia coli* tRNAs, tRNA^{Val}_{GAC} and tRNA^{Val}_{cmo⁵UAC} (where the anticodon is specified in the subscript; cmo⁵U denotes uridine 5-oxyacetic acid). Yet valine can be incorporated into MS2 coat protein from all four of its codons using only a single tRNA³. The post-transcriptional modification of uridine to cmo⁵U at the wobble position, nucleotide 34, allows tRNA^{Val}_{cmo⁵UAC} and tRNA^{Pro}_{cmo⁵UGG} to read adenosine, guanosine and uridine efficiently *in vitro* and is required for recognition of adenosine, guanosine, uridine and cytidine *in vivo* in mutant strains lacking other isoacceptors^{4–6}. Given its importance in expanding the tRNA's ability to decode different codons, it is not surprising that the wobble-position U34 is modified to cmo⁵U34 or a derivative (Supplementary Fig. 1 online) in tRNAs recognizing codons in six different codon boxes (Fig. 1a). Thus, this is one of the most widely used tRNA modifications directly involved in decoding.

Previous kinetic and structural work has shown that tRNA selection by the ribosome involves an induced fit to a productive form that allows activation of the accessory GTPase factor EF-Tu^{7,8}. During decoding, the geometry of Watson-Crick base pairing between codon and anticodon at the first two positions is monitored through minor-groove interactions by ribosomal bases A1492, A1493 and G530 (*E. coli* numbering)⁹. This recognition provides the additional binding

energy to induce a transition of the ribosome from an open to a closed form, which is required for tRNA selection¹⁰. Antibiotics or mutations that make the transition more energetically favorable can allow selection of a near-cognate tRNA. However, the formation of Watson-Crick base pairs at the first two positions is not always sufficient for decoding, and previous structural studies have suggested an important role for modifications in a lysine tRNA¹¹.

Here we report the crystal structures of an anticodon stem-loop (ASL^{Val}_{cmo⁵UAC}) bound to all four valine codons in the decoding site of the 30S ribosomal subunit. A recent high-resolution structure of the ribosome¹² shows that the ASL of full-length tRNA makes essentially identical contacts in the entire ribosome to those of an isolated ASL with the 30S subunit⁹, thus validating studies that use crystals of the 30S subunit with ASLs to examine tRNA interactions in the A site. Our structures reveal unusual base-pairing at the wobble position and shed light on how modification of the uridine allows expanded decoding by tRNAs.

RESULTS

In addition to cmo⁵U34, the ASL contained the companion modification N6-methyladenosine at position 37 (m⁶A37; Fig. 1b,c)². The ASL and oligonucleotides corresponding to each of the four valine codons were soaked into crystals of the 30S ribosomal subunit and their structures were determined as described⁹ (see Methods). The four structures were refined to resolutions of between 2.8 and 3.1 Å. Both modifications are clearly visible in all four unbiased difference Fourier maps of ASL^{Val}_{cmo⁵UAC} bound to GU(A/G/U/C). The structures show unusual and unexpected base pairing only at the wobble position (Fig. 2a–d and Supplementary Fig. 2 online).

¹Medical Research Council Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK. ²Institute of Organic Chemistry, Technical University of Lodz, Zeromskiego str. 116, 90923 Lodz, Poland. ³Department of Molecular and Structural Biochemistry, North Carolina State University, Raleigh, North Carolina 27695-7622, USA. Correspondence should be addressed to P.F.A. (paul_agris@ncsu.edu) or V.R. (ramak@mrc-lmb.cam.ac.uk).

Received 9 January; accepted 27 March; published online 13 May 2007; doi:10.1038/nsmb1242

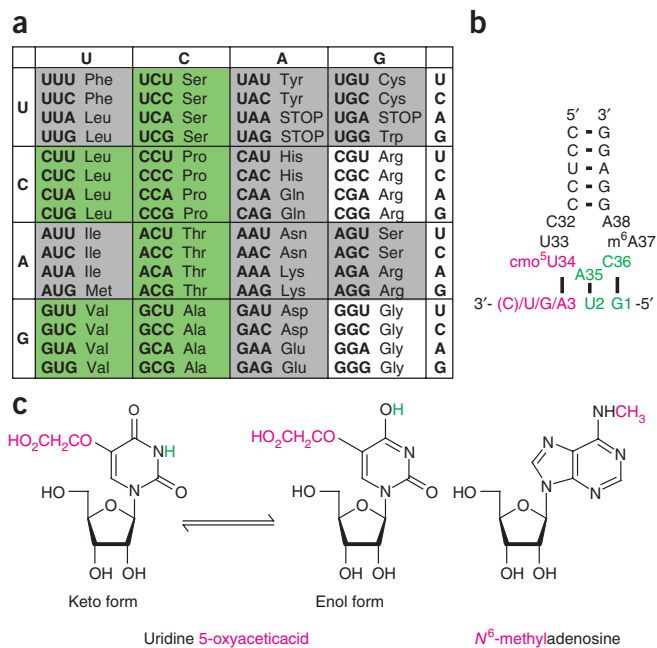


Figure 1 The degeneracy of the genetic code. **(a)** The genetic code. 8 of 16 codon boxes are degenerate (highlighted in green and white). Boxes containing codons read by tRNAs with $\text{cmo}^5\text{U34}$ or derivatives are highlighted in green. **(b)** Pairing between the ASL and the codon on the mRNA. The modification allows the tRNA to read A, G, U and in some cases also C. **(c)** Modified bases in *E. coli* tRNA^{Val} _{cmo^5UAC} . Modifications are shown in magenta. The cmo^5U modification is expected to affect the keto-enol equilibrium.

In all four structures, the global conformation of the 30S subunit and the interactions at the first two codon-anticodon base pairs are very similar to those reported for cognate tRNA^{Phe} (ref. 9). The codon nucleotide at the wobble position is constrained by hydrogen bonds to G530 of the 16S ribosomal RNA and, via a Mg²⁺-mediated bridge, to C518 and ribosomal protein S12 (ref. 9). However, the anticodon of ASL^{Val} _{cmo^5UAC} is more constrained than in the case of an unmodified base, because the ether oxygen (O5) of $\text{cmo}^5\text{U34}$ forms an intramolecular hydrogen bond, in all four base pairs, to the 2'-OH of U33 (Fig. 3a). In this ASL, the 2'-OH of U33 is also within hydrogen-bonding distance from the N7 of A35. The ribose of $\text{cmo}^5\text{U34}$ adopts a C3'-endo conformation in all four structures.

The modification does not seem to have any obvious role in the canonical $\text{cmo}^5\text{U34}\cdot\text{A}$ pair, which adopts the normal Watson-Crick geometry. The $\text{cmo}^5\text{U34}$ is well ordered and adopts a defined orientation in which the carboxyl group points toward the backbone of the opposite strand (Fig. 2a).

The $\text{cmo}^5\text{U34}\cdot\text{G}$ base pair does not adopt wobble geometry

Notably, the $\text{cmo}^5\text{U34}\cdot\text{G}$ base pair adopts Watson-Crick geometry, suggesting that the uridine must be in the enol form. The crystal structure of 5-methoxyuridine showed an increase in the C4-O4 bond length, as well as a decrease in the N3-C4 bond length, indicating a shift in the keto-enol equilibrium of the base¹³ (Fig. 1c). This change in bond lengths was more pronounced than in 5-hydroxyuridine, indicating that the shift in the equilibrium is dependent on the electronic and steric properties of the group attached to C5 (ref. 13). We propose that, in addition,

the pairing of $\text{cmo}^5\text{U34}$ to G induces and stabilizes the enol form. In support of this, theoretical studies have predicted that glycine induces an analogous tautomeric shift in uridine, which can increase the concentration of the enol form by up to five orders of magnitude¹⁴.

Pyrimidine-pyrimidine base pairs

The $\text{cmo}^5\text{U34}\cdot\text{U}$ base pair has a wider minor groove, resulting in only one hydrogen bond (Figs. 2c and 4b), in an A-form helix geometry similar to U•U pairs previously observed^{15,16}. We did not observe an ordered water molecule bridging N3 of $\text{cmo}^5\text{U34}$ with O2 of U3, suggesting that the N3 is at least partially desolvated. G530 has been shown to monitor the second anticodon-codon base pair by forming contacts with the 2'-OH of the anticodon through its N3 and 2'-OH. In the $\text{cmo}^5\text{U34}\cdot\text{U}$ base pair, the ribose of G530 adopts a C2'-endo conformation, thereby positioning its 2'-OH closer to the O2 and 2'-OH of $\text{cmo}^5\text{U34}$ but still maintaining the hydrogen bond between its N3 and the 2'-OH of A35 in the ASL (Fig. 4a). These additional interactions that arise as a result of the modification presumably more than compensate for the otherwise necessary desolvation of the O2 of $\text{cmo}^5\text{U34}$. Notably, the carboxyl group of the modification is within hydrogen-bonding distance of O4 of the uridine in the messenger RNA (Fig. 2c).

The $\text{cmo}^5\text{U34}\cdot\text{C}$ base pair (Fig. 2d) adopts a similar conformation to the $\text{cmo}^5\text{U34}\cdot\text{U}$ and to previously observed U•C mismatches in RNA duplexes, allowing the formation of one hydrogen bond^{17,18}. The arrangement of the bases suggests that both N3 positions have to be desolvated, and we did not find a water molecule bridging them, as observed previously¹⁷. In contrast to $\text{cmo}^5\text{U34}\cdot\text{U}$, the $\text{cmo}^5\text{U34}\cdot\text{C}$ pair does not form the G530 contacts described above, and both bases are less stacked (Fig. 4b,c). The smaller stacking interaction and the lack of contacts to G530, as well as the possible lack of the additional hydrogen bond with the carboxyl group of the modification, should all destabilize tRNA^{Val} _{cmo^5UAC} with a codon ending in cytidine relative to one ending in uridine. This is consistent with biochemical studies for valine and leucine tRNAs^{19,20}.

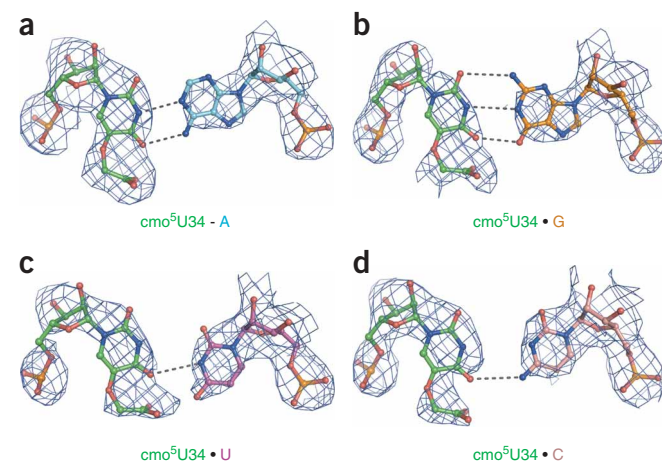


Figure 2 $\text{cmo}^5\text{U34}$ base-pairs with all four bases at the wobble position of the mRNA. Shown are structures of the four base pairs along with unbiased difference Fourier maps. Dotted lines, hydrogen bonds. **(a)** $\text{cmo}^5\text{U34}\cdot\text{A}$. **(b)** $\text{cmo}^5\text{U34}\cdot\text{G}$. **(c)** $\text{cmo}^5\text{U34}\cdot\text{U}$. **(d)** $\text{cmo}^5\text{U34}\cdot\text{C}$. This and the other molecular graphics figures were made with PyMOL.

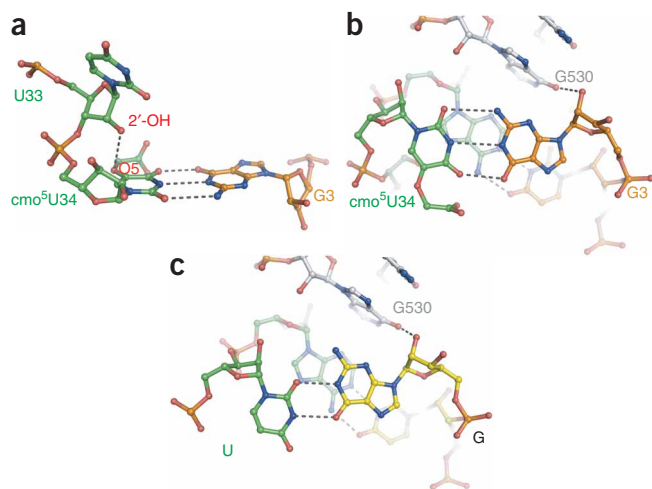


Figure 3 Comparison of $\text{cmo}^5\text{U}\cdot\text{G}$ with standard $\text{U}\cdot\text{G}$ wobble. (a) Hydrogen bond from 2'-OH of U33 to O5 of the modification. This contact inhibits lateral movement of the cmo^5U . (b) Geometry of the $\text{cmo}^5\text{U34}\cdot\text{G}$ base pair. (c) The $\text{U}\cdot\text{G}$ wobble base pair, shown with standard wobble geometry. The U would be completely unstacked.

DISCUSSION

At the beginning of an elongation cycle, the ribosome seems to be delicately balanced, so that the additional binding energy from cognate tRNA, but not near-cognate tRNA, is sufficient to induce a conformational change to an active form rather than increase the affinity of the tRNA^{7,8}. The work here provides a structural rationale for how a modification at the wobble position can provide sufficient additional binding energy so that decoding is facilitated for all four possible base pairs at the wobble position, thus allowing an expanded reading of the code by a single tRNA.

Notably, we found standard Watson-Crick base-pairing geometry not only for the canonical $\text{U34}\cdot\text{A}$ pair but also for the $\text{U34}\cdot\text{G}$ pair. Why is a standard $\text{U}\cdot\text{G}$ wobble geometry not observed in the $\text{U34}\cdot\text{G}$ pair? It has been shown previously that a guanosine in position 34 of tRNA can adopt wobble geometry with a uridine in the codon⁹. However, in all cytoplasmic tRNAs, the reverse pairing of U34 on tRNA with guanosine in the codon occurs only when the U34 is modified (reviewed in ref. 6). In standard $\text{G}\cdot\text{U}$ wobble geometry, the uridine has enhanced stacking with the adjacent 5' base, in contrast with the base on its 3' side^{21,22}. Thus, a $\text{G}\cdot\text{U}$ base pair with a uridine on the 3' rather than the 5' end of a strand is favored. The $\text{U34}\cdot\text{G}$ wobble pair at the end of the codon-anticodon helix in the ribosomal A site, with the uridine on the 5' end of the tRNA, should therefore be energetically less stable and thus presumably not contribute enough binding energy to facilitate the induced conformational change required for efficient decoding. On balance, it seems that the presence of the modification reduces the penalty of stabilizing the enol form enough that a Watson-Crick $\text{cmo}^5\text{U34}\cdot\text{G}$ pair, with its additional hydrogen bond and increased stacking, is the energetically favored conformation. This also explains why $\text{cmo}^5\text{U34}$ is essential. An unmodified U34 would be forced to adopt standard $\text{U34}\cdot\text{G}$ wobble geometry because it could not adopt the enol form efficiently enough and the hydrogen bond to U33 would not lock its position. As the base in the mRNA is held in place by its interactions with G530 as well as with C518 and S12, this would require a movement of U34 toward the major groove and result in it being completely unstacked (Fig. 3b,c). In addition to the modification's role in pairing with guanosine in the

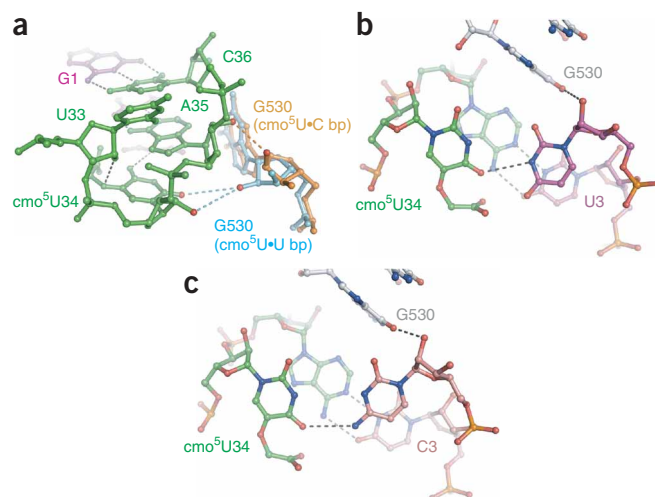


Figure 4 The two pyrimidine-pyrimidine base pairs. (a) Alternative conformations of ribosomal base G530. In the $\text{cmo}^5\text{U}\cdot\text{U}$ base pair, the G530 ribose adopts a C2'-endo conformation, compensating for the desolvation of the O2 and 2'-OH of $\text{cmo}^5\text{U34}$ (cyan). In the $\text{cmo}^5\text{U}\cdot\text{C}$ base pair, it adopts the standard conformation, with a hydrogen bond to the 2'-OH of A35 in the tRNA (beige). (b,c) In the $\text{cmo}^5\text{U}\cdot\text{U}$ base pair, the U stacks better on its 5' base than the C does in the $\text{cmo}^5\text{U34}\cdot\text{C}$ base pair.

codon wobble position, we observe some cross-strand stacking between A35 in the $\text{ASL}^{\text{Val}}_{\text{cmo}^5\text{UAC}}$ and adenosine or guanosine in the wobble codon position. The tRNAs involved in decoding the six degenerate codon boxes mentioned above all contain a purine in position 35, which would provide additional stability through cross-strand stacking for codons ending with adenosine or guanosine. $\text{m}^6\text{A37}$ is involved in a cross-strand stack above the first base pair (Supplementary Fig. 3 online) comparable to that of $\text{t}^6\text{A37}$ in $\text{ASL}^{\text{Lys}}_{\text{UUU}}$ (ref. 11); this stacking also compensates for the otherwise low enthalpy of binding in all four structures.

In the $\text{U34}\cdot\text{U}$ base pair, the carboxyl group of the modification is within hydrogen-bonding distance of the O4 of the codon. However, the pK_a of the carboxyl group in solution is close to 3, suggesting it should be completely deprotonated and thus be unable to form a hydrogen bond with O4 of the codon. Consistent with this, the orbitals of both the carboxyl group and the O4 appear oriented to minimize repulsion. However, a shift of as much as 5 pH units in the pK_a of carboxyl groups has been observed in the interior of enzymes²³. Thus, it remains a formal possibility that a shifted pK_a of the carboxyl group in $\text{cmo}^5\text{U34}$ results in additional hydrogen bonding between the modification on the anticodon and the O4 of the codon. The additional hydrogen bonding would help to explain previous studies showing that the efficiency of $\text{cmo}^5\text{U34}\cdot\text{U}$ decoding is comparable to that of $\text{cmo}^5\text{U34}\cdot\text{G}$ decoding¹⁹.

There are two additional factors that facilitate decoding by tRNAs with $\text{cmo}^5\text{U34}$ of codons ending in uridine or in some cases even cytidine, even though the $\text{cmo}^5\text{U34}\cdot\text{C}$ base pair seems relatively weak. In cases where a single tRNA has been shown to decode all four codons *in vivo*, G-C base pairing occurs at the first two codon positions. Not only are G-C pairs intrinsically more stable than A-U pairs, but the type I A-minor interaction for G-C pairs is also stronger²⁴, providing additional binding energy. This additional energy from having G-C pairs at the first two positions is apparently not sufficient for decoding by itself, because in addition the U34 is required to be modified. We expect that the $\text{cmo}^5\text{U34}$ prestructures

Table 1 Summary of crystallographic data and refinement

	ASL ^{Val} _{cmo5UAC} -GUA	ASL ^{Val} _{cmo5UAC} -GUG	ASL ^{Val} _{cmo5UAC} -GUC	ASL ^{Val} _{cmo5UAC} -GUU
Data collection				
Space group	P4 ₁ 2 ₁ 2			
Cell dimensions				
<i>a</i> , <i>b</i> , <i>c</i> (Å)	402.2, 402.2, 174.9	400.9, 400.9, 174.2	401.9, 401.9, 174.6	400.5, 400.5, 173.4
Resolution (Å)	30–3.0 (3.1–3.0)	50–3.1 (3.2–3.1)	100–2.9 (3.0–2.9)	40–2.8 (2.9–2.8)
<i>R</i> _{sym}	15.9 (63.7)	26.2 (75.0)	17.4 (78.9)	13.4 (79.2)
<i>I</i> / σI	7.4 (2.1)	6.3 (2.0)	8.8 (2.0)	11.5 (2.2)
Completeness (%)	93.7 (87.4)	96.0 (92.9)	97.0 (92.3)	96.8 (91.9)
Redundancy	3.1 (3.1)	4.0 (4.0)	4.2 (4.0)	5.0 (4.9)
Refinement				
Resolution (Å)	30.0–3.1	30.0–3.1	30.0–2.9	30.0–2.8
No. unique reflections	252,303	250,474	308,707	335,672
<i>R</i> _{work} / <i>R</i> _{free}	21.0 / 24.5	22.8 / 26.8	22.3 / 25.4	21.8 / 24.2
No. atoms				
RNA	32,850	32,873	32,870	32,828
Protein	19,238	19,238	19,238	19,238
Ions	242	199	194	255
Paromomycin	42	42	42	42
<i>B</i> -factors				
RNA	66.69	54.29	69.80	62.23
Protein	69.50	59.98	71.67	68.49
Ions	63.44	47.71	55.69	60.59
Paromomycin	55.90	42.19	57.26	51.51
R.m.s. deviations				
Bond lengths (Å)	0.0067	0.0073	0.0067	0.0067
Bond angles (°)	1.18	1.20	1.17	1.17

Structures show ASL^{Val}_{cmo5UAC} bound to each mRNA codon indicated. Values in parentheses are for highest-resolution shell.

the anticodon loop through the hydrogen bond between the 2'-OH of U33 and the O5 of the cmo⁵U34, thereby presumably reducing the entropic cost of binding and making the binding of tRNA to a codon ending in uridine or cytidine sufficiently favorable for decoding in these contexts.

Our structures are also consistent with pre-steady-state kinetics experiments on decoding, which have shown that a tRNA^{Ala} with a noncanonical cmo⁵U34•C pair at the wobble position has a decoding efficiency similar to that of cognate tRNAs, whereas an unmodified A•G pair at the wobble position for tRNA^{Phe} has greatly reduced decoding efficiency²⁵. It would be useful to compare directly the kinetic parameters for the same codon-anticodon pairs with and without the modification. Finally, the structures are also consistent with results that show that the binding affinities of tRNA^{Val} with the modified U34 for ribosomes with codons ending in adenosine, guanosine or uridine are very similar and stronger than that for the codon ending in cytidine (E.M. Gustilo and P.F.A., unpublished data).

In summary, the work described here provides a structural basis for rationalizing longstanding observations on how modifications of uridine at the wobble position facilitate expanded decoding for tRNAs.

METHODS

Materials and crystallization. *Thermus thermophilus* 30S ribosomal subunits were purified, crystallized and cryoprotected as described²⁶ (the final solution being 26% (v/v) MPD, 100 mM potassium-MES (pH 6.5), 200 mM KCl, 75 mM NH₄Cl and 15 mM MgCl₂). ASL^{Val}_{cmo5UAC} was chemically synthesized using a slightly altered phosphoramidite chemistry method²⁷. The standard units and m⁶A were protected with the typical combination of masking

groups (5'-dimethoxytriphenylmethyl chloride (DMTr), 2'-*t*-butyldimethylsilyl (TBDMS), *exo*-amine phenoxacetyl/benzoyl), whereas *p*-nitrophenylethyl was used to block the cmo⁵U carboxyl function. The oligomer was deprotected as described²⁸, and homogeneity was verified by MALDI-TOF analysis, enzymatic digestion to nucleosides and HPLC analysis²⁹. NMR analysis of ASL^{Val}_{cmo5UAC} clearly demonstrated the presence cmo⁵U34 and m⁶A37. The mRNA oligonucleotides were chemically synthesized and gel-purified (Dharmacon) with the sequences 5'-GU(A/C/G/U)AAA-3' (codons underlined). After cryoprotection, the 30S crystals were soaked in cryoprotection buffer containing 80 μM paromomycin, 300 μM ASL and 300 μM of the corresponding mRNA hexanucleotide for at least 48 h as described^{9,10}. Crystals were flash-cooled in liquid nitrogen and stored for data collection.

Data collection and refinement. Crystals were prescreened at European Synchrotron Radiation Facility beamline ID14-2 using short exposures 90° apart. Crystals were then stored in liquid nitrogen before data was collected at European Synchrotron Radiation Facility beamlines ID14-4 and ID29 in a cryostream at 90–100 K. Processing was done using XDS³⁰. The CCP4 package was used for assorted tasks³¹. Coot was used for visualization and building³² and CNS 1.2 was used for refinement³³. Topologies and parameters were used directly or derived using HIC-Up³⁴. A summary of crystallographic data is shown in Table 1. Differences in individual data sets are likely to be the result of slight differences in crystal quality. Figures, as well as alignments between the individual structures, were made using PyMOL (<http://pymol.sourceforge.net>)³⁵. The structure of an RNA A-form helix containing a G•U wobble base pair was aligned to the codon-anticodon helix to compare it with the cmo⁵U•G base pair³⁶. Analysis of stacking interactions and RNA bond angles was done using 3DNA³⁷.

Accession codes. Protein Data Bank: Coordinates for ASL^{Val}_{cmo5UAC} bound to each of four RNA hexamers (5'-GU(A/C/G/U)AAA-3') on the 30S ribosomal subunit have been deposited with accession codes 2UU9 (5'-GUG), 2UUA (5'-GUC), 2UUB (5'-GUU) and 2UUC (5'-GUA).

Note: Supplementary information is available on the Nature Structural & Molecular Biology website.

ACKNOWLEDGMENTS

We thank B. Sproat for help and advice on the polymer chemistry, W.D. Graham for purification and analysis of ASL^{Val}, A. Kelley for purification and crystallization of 30S subunits, C.M. Dunham and S. Petry for help with synchrotron data collection, R. Ravelli, J. McCarthy and G. Leonard for help with data collection on beamline ID14 at the European Synchrotron Radiation Facility, and L. Passmore and M. Schmeing for helpful comments. This work was funded by the UK Medical Research Council (V.R.) and grants from the US National Institutes of Health (P.E.A. and V.R.), the US National Science Foundation (P.E.A.), the Agouron Institute (V.R.), the Austrian Academy of Sciences (A.W.) and the Polish Ministry of Science and Education (A.M.).

COMPETING INTERESTS STATEMENT

The authors declare no competing financial interests.

Published online at <http://www.nature.com/nsmb/>

Reprints and permissions information is available online at <http://npg.nature.com/reprintsandpermissions>

- Crick, F.H.C. The origin of the genetic code. *J. Mol. Biol.* **38**, 367–379 (1968).
- Sprinzi, M., Horn, C., Brown, M., Ioudovitch, A. & Steinberg, S. Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res.* **26**, 148–153 (1998).
- Mitra, S.K., Lustig, F., Akesson, B. & Lagerkvist, U. Codon-anticodon recognition in the valine codon family. *J. Biol. Chem.* **252**, 471–478 (1977).
- Mitra, K. & Frank, J. Ribosome dynamics: insights from atomic structure modeling into cryo-electron microscopy naps. *Annu. Rev. Biophys. Biomol. Struct.* **35**, 299–317 (2006).
- Nasvall, S.J., Chen, P. & Bjork, G.R. The modified wobble nucleoside uridine-5-oxyacetic acid in tRNA^{Phe}(cmo⁵UGG) promotes reading of all four proline codons *in vivo*. *RNA* **10**, 1662–1673 (2004).
- Agris, P.F., Vendeix, F.A. & Graham, W.D. tRNA's wobble decoding of the genome: 40 years of modification. *J. Mol. Biol.* **366**, 1–13 (2007).
- Ogle, J.M. & Ramakrishnan, V. Structural insights into translational fidelity. *Annu. Rev. Biochem.* **74**, 129–177 (2005).
- Rodnina, M.V. & Wintermeyer, W. Fidelity of aminoacyl-tRNA selection on the ribosome: kinetic and structural mechanisms. *Annu. Rev. Biochem.* **70**, 415–435 (2001).
- Ogle, J.M. *et al.* Recognition of cognate transfer RNA by the 30S ribosomal subunit. *Science* **292**, 897–902 (2001).
- Ogle, J.M., Murphy, F.V., Tarry, M.J. & Ramakrishnan, V. Selection of tRNA by the ribosome requires a transition from an open to a closed form. *Cell* **111**, 721–732 (2002).
- Murphy, F.V., IV, Ramakrishnan, V., Malkiewicz, A. & Agris, P.F. The role of modifications in codon discrimination by tRNA(Lys)UUU. *Nat. Struct. Mol. Biol.* **11**, 1186–1191 (2004).
- Selmer, M. *et al.* Structure of the 70S ribosome complexed with mRNA and tRNA. *Science* **313**, 1935–1942 (2006).
- Hillen, W., Egert, E., Lindner, H.J., Gassen, H.G. & Vorbrüggen, H. 5-Methoxyuridine: the influence of 5-substituents on the keto-enol tautomerism of the 4-carbonyl group. *J. Carbohydr. Nucleosides Nucleotides* **5**, 23–32 (1978).
- Dabkowska, I., Gutowski, M. & Rak, J. Interaction with glycine increases stability of a mutagenic tautomer of uracil. A density functional theory study. *J. Am. Chem. Soc.* **127**, 2238–2248 (2005).
- Dirheimer, G., Keith, G., Dumas, P. & Westhof, E. The base pair directory. in *tRNA: Structure, Biosynthesis, and Function* (eds. Söll, D. & RajBhandary, U.L.) 111–112 (American Society for Microbiology, Washington, DC, 1995).
- Tinoco, I.J. The base pair directory. in *The RNA World* (eds. Gesteland, R.F. & Atkins, J.F.) 603–607 (Cold Spring Harbor Laboratory Press, New York, 1993).
- Cruse, W.B. *et al.* Structure of a mispaired RNA double helix at 1.6-Å resolution and implications for the prediction of RNA secondary structure. *Proc. Natl. Acad. Sci. USA* **91**, 4160–4164 (1994).
- Nagaswamy, U. *et al.* NCIR: a database of non-canonical interactions in known RNA structures. *Nucleic Acids Res.* **30**, 395–397 (2002).
- Mitra, S.K. *et al.* Relative efficiency of anticodons in reading the valine codons during protein synthesis *in vitro*. *J. Biol. Chem.* **254**, 6397–6401 (1979).
- Sorensen, M.A. *et al.* Over expression of a tRNA(Leu) isoacceptor changes charging pattern of leucine tRNAs and reveals new codon reading. *J. Mol. Biol.* **354**, 16–24 (2005).
- Masquida, B. & Westhof, E. On the wobble G•U and related pairs. *RNA* **6**, 9–15 (2000).
- Mizuno, H. & Sundaralingam, M. Stacking of Crick Wobble pair and Watson-Crick pair: stability rules of G-U pairs at ends of helical stems in tRNAs and the relation to codon-anticodon Wobble interaction. *Nucleic Acids Res.* **5**, 4451–4461 (1978).
- Fersht, A.R. *Structure and Mechanism in Protein Science* (W.H. Freeman, New York, 1998).
- Doherty, E.A., Batey, R.T., Masquida, B. & Doudna, J.A. A universal mode of helix packing in RNA. *Nat. Struct. Biol.* **8**, 339–343 (2001).
- Kothe, U. & Rodnina, M.V. Codon reading by tRNA^{Ala} with modified uridine in the wobble position. *Mol. Cell* **25**, 167–174 (2007).
- Clemons, W.M., Jr *et al.* Crystal structure of the 30S ribosomal subunit from *Thermus thermophilus*: purification, crystallization and structure determination. *J. Mol. Biol.* **310**, 827–843 (2001).
- Sproat, B.S. RNA synthesis using 2'-O-(tert-butylidimethylsilyl) protection. *Methods Mol. Biol.* **288**, 17–32 (2005).
- Boudou, V. *et al.* Synthesis of the anticodon hairpin tRNA^{Met} containing N-[9-(β-D-ribofuranosyl)-9H-purin-6-yl]carbamoyl-L-threonine (=N6-(((1S,2R)-1-carboxy-2-hydroxypropyl)amino)-carbonyl)adenosine, t6A. *Helv. Chim. Acta* **83**, 152–161 (2000).
- Gehrke, C.W. & Kuo, K.C. Ribonucleoside analysis by reversed-phase high-performance liquid chromatography. *J. Chromatogr.* **471**, 3–36 (1989).
- Kabsch, W. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Cryst.* **26**, 795–800 (1993).
- Collaborative Computational Project Number 4. The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50**, 760–763 (1994).
- Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
- Brünger, A.T. *et al.* Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **54**, 905–921 (1998).
- Kleywegt, G.J. & Jones, T.A. Databases in protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **54**, 1119–1131 (1998).
- DeLano, W.L. *The PyMOL Molecular Graphics System* (Delano Scientific, San Carlos, California, USA, 2006).
- Masquida, B., Sauter, C. & Westhof, E. A sulfate pocket formed by three GoU pairs in the 0.97 Å resolution X-ray structure of a nonameric RNA. *RNA* **5**, 1384–1395 (1999).
- Lu, X.J. & Olson, W.K. 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Res.* **31**, 5108–5121 (2003).