Published in final edited form as:
Nat Struct Mol Biol. 2007 June ; 14(6): 498–502. doi:10.1038/nsmb1242.

Mechanism of expanding the decoding capacity of tRNAs by modification of uridines

Albert Weixlbaumer[1], **Frank V. Murphy IV**[1], **Agnieszka Dziergowska**[2], **Andrzej Malkiewicz**[2], **Franck A. P. Vendeix**[3], **Paul F. Agris**[3,4], and **V. Ramakrishnan**[1,4]

[1]MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK [2]Institute of Organic Chemistry, Technical University of Lodz, Zeromskiego str.116, 90923 Lodz, Poland [3]Department of Molecular and Structural Biochemistry, North Carolina State University, Raleigh, North Carolina 27695-7622, USA

Abstract

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 see unusual base pairs with the three non-complementary codon bases, including a GU 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.

INTRODUCTION

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 are fully degenerate, i.e. code for the same amino acid (Fig. 1a). Such degenerate codon boxes are generally decoded by different isoaccepting tRNAs2. The fourfold degenerate codon box for valine is decoded by two different *E.coli* tRNAs, $tRNA^{Val}_{GAC}$ and $tRNA^{Val}_{cmo5UAC}$ (where the anticodon is specified). 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 uridine 5-oxyacetic acid (cmo^5U) at the wobble position 34 allows $tRNA^{Val}_{cmo5UAC}$ and $tRNA^{Pro}_{cmo5UGG}$ to read A, G and U efficiently *in vitro*, and is required for recognition of A, G, U and C *in vivo* in mutant strains lacking other isoacceptors4-6. Given its importance in expanding the ability to decode different codons, it is not surprising that the wobble position U34 is modified to a cmo^5U34 or its derivatives (Supplementary Fig. 1 online) for tRNAs recognizing codons in six different codon boxes (Fig. 1a), making this 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-Tu7, 8. During decoding, the geometry of Watson-Crick base pairing between

[4]Correspondence should be addressed to Paul F. Agris (paul_agris@ncsu.edu) or V. Ramakrishnan (ramak@mrc-lmb.cam.ac.uk).

Europe PMC Funders Author Manuscripts

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 that is required for tRNA selection10. However, antibiotics or mutations that make the transition more energetically favorable can allow selection of near-cognate tRNA. On the other hand, the formation of Watson-Crick base pairs at the first two positions is not always sufficient for decoding, and previous structural studies have suggested how modifications in a lysine tRNA play an important role11.

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 an isolated ASL does with the 30S subunit 9, thus validating studies using crystals of the 30S subunit with ASLs for studying tRNA interactions in the A site. These 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 cmo^5U34, the ASL contained the companion modification N6-methyladenosine (m^6A37) at position 37 (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 determined as previously 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).

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 $tRNA^{Phe}$ 9. The codon nucleotide at the wobble position is constrained by hydrogen bonds to G530 of 16S rRNA and, via a Mg^{2+}-mediated bridge, to C518 and ribosomal protein S12 9. However, the anticodon of $ASL^{Val}_{cmo5UAC}$ is more constrained than in the case of an unmodified base because the ether oxygen (O5) of cmo^5U34 forms an intra-molecular 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 of the N7 of A35. The ribose of cmo^5U34 adopts a C3′-*endo* conformation in all four structures.

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

The cmo^5U34-G basepair does not adopt wobble geometry

Surprisingly, the cmo^5U34-G base pair adopts Watson-Crick geometry, suggesting that the U 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 base 13(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 13. We propose that, in addition, the pairing of cmo^5U34 to G induces and stabilizes the enol form. In support of this, theoretical studies predicted that glycine induces an analogous tautomeric

shift in U, which can increase the concentration of the enol form by up to 5 orders of magnitude 14.

The pyrimidine-pyrimidine basepairs

The cmo^{5}U34-U base pair has a wider minor groove resulting in only one hydrogen bond (Fig. 2c and 4b), in an A-form helix geometry similar to U-U pairs previously observed 15, 16. We do not observe an ordered water molecule bridging N3 of cmo^{5}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 cmo^{5}U34-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 cmo^{5}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 cmo^{5}U34. Intriguingly, the carboxyl group of the modification is within hydrogen bonding distance from O4 of U on the mRNA (Fig. 2c).

The cmo^{5}U34-C base pair (Fig. 2d) adopts a similar conformation to the cmo^{5}U34-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 do not see a water molecule bridging them as observed previously 17. In contrast to cmo^{5}U34-U, the cmo^{5}U34-C pair does not form the G530 contacts described above and both bases are less stacked (Fig. 4b,c). The smaller stacking interaction, the lack of contacts to G530, as well as the possible lack of the additional hydrogen bonding with the carboxyl group of the modification, should all destabilize $tRNA^{Val}_{cmo5UAC}$ with a codon ending in C relative to one ending in U. This is consistent with biochemical studies for valine and leucine tRNAs 19, 20.

DISCUSSION

At the beginning of an elongation cycle, the ribosome appears to be delicately balanced, such that the additional binding energy from cognate, 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.

Surprisingly, we find standard Watson-Crick base pairing geometry not only for the canonical U34-A pair but also for the U34-G pair. Why is a standard U-G wobble geometry not observed in the U34-G pair? It has been shown previously that a G in position 34 of tRNA can adopt wobble geometry with a U on the codon 9. However, in all cytoplasmic tRNAs, the reverse pairing of U34 on tRNA with G on the codon only occurs when the U34 is modified (reviewed in ref. [6]). In standard G-U wobble geometry, the U has enhanced stacking with the adjacent base on its 5′ side in contrast with the base on its 3′ side 21, 22. Thus, a G-U base pair with a U on the 3′ rather than 5′ end of a strand is favored. The U34-G wobble pair at the end of the codon-anticodon helix in the ribosomal A-site, with the U 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 appears that the presence of the modification reduces the penalty of stabilizing the enol form sufficiently so that a Watson-Crick cmo^{5}U34-G pair, with its additional hydrogen bond and increased stacking, is the energetically favored conformation. This also explains why cmo^{5}U34 is essential. An

unmodified U34 would be forced to adopt standard U34-G wobble geometry because it cannot adopt the enol-form efficiently enough and the hydrogen-bond to U33 would not lock its position. Since the base in the mRNA is held in place by its interaction with G530 as well as with C518 and S12, this would require a movement of U34 towards the major groove and result in it being completely unstacked (Fig. 3b,c). In addition to the role of the modification, for G in the codon wobble position, we observe some cross-strand stacking between A35 in the $ASL^{Val}_{cmo5UAC}$ and A or G 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 A or G. The m^6A37 is involved in a cross-strand stack above the first base pair (Supplementary Fig. 3 online) that is comparable to that of t^6A37 in ASL^{Lys}_{UUU} 11; this stacking also compensates for the otherwise low enthalpy of binding in all four structures.

In the U34-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 as well as the O4 appear oriented to minimize repulsion. On the other hand, a shift of as much as 5 pH units in the pK_a of carboxyl groups has been observed in the interior of enzymes 23. So it remains a formal possibility that a shifted pK_a of the carboxyl group in cmo^5U34 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 cmo^5U34-U decoding is comparable to that of cmo^5U34-G decoding 19.

There are two additional factors which facilitate decoding by tRNAs with cmo^5U34 of codons ending in U or in some cases even C, even though the cmo^5U34-C base pair appears relatively weak. In cases where a single tRNA was shown to decode all four codons *in vivo*, G-C base pairing occurred at the first two codon positions. G-C pairs are not only intrinsically more stable than A-U pairs, but the type I A-minor interaction for G-C pairs is also stronger 24, thus 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 cmo^5U34 prestructures the anticodon loop through the hydrogen bond between the 2′-OH of U33 and the O5 of the cmo^5U34, thereby presumably reducing the entropic cost of binding and thus makes the binding of tRNA to a codon ending in U or C sufficiently favorable for decoding in these contexts.

Our structures are also consistent with presteady state kinetic experiments on decoding which showed that a $tRNA^{Ala}$ with a non-canonical cmo^5U34-C pair at the wobble position had a decoding efficiency similar to that of cognate tRNAs, whereas an unmodified A-G pair at the wobble position for $tRNA^{Phe}$ had drastically reduced decoding efficiency 25. 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 to ribosomes with codons ending in A, G or U are very similar and stronger than that to the codon ending in C (E. M. Gustilo and P.F. Agris, unpublished results).

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 described (the final solution being 26% (v/v) MPD, 100 mM K-MES, pH 6.5, 200 mM KCl, 75 mM NH_4Cl, 15 mM $MgCl_2$)26. The $ASL^{Val}_{cmo5UAC}$ was chemically synthesized using a slightly altered phosphoramidite chemistry 27. The standard units and m^6A were protected with the typical combination of masking groups (5′-DMTr, 2′-TBDMS, exo-amine Pac/Bz), while p-nitrophenyloethyl has been used to block the cmo^5U carboxyl function. The oligomer was deprotected as published 28, and homogeneity was verified by MALDI-TOF analysis, enzymatic digestion to nucleosides and HPLC analysis 29. NMR analysis of $ASL^{Val}_{cmo5UAC}$ clearly demonstrated the presence cmo^5U34 and m^6A37. 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 pre-screened at the ESRF beamline ID14-2 using short exposures 90 degrees apart. Crystals were then stored in liquid nitrogen before data was collected at the ESRF beamlines ID14-4 and ID29 in a cryostream at 90-100 K. Processing was done using XDS 30. The CCP4 package was used for assorted tasks 31. Coot was used for visualization and building 32 and CNS 1.2 was used for refinement 33. Topologies and parameters were used directly or derived using HIC-Up 34. A summary of crystallographic data is shown in Table 1. Differences in individual datasets 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 35. The structure of an RNA A-form helix containing a GU-wobble base pair was aligned to the codon anticodon helix to compare it with the cmo^5U-G base pair 36. Analysis of stacking interactions and RNA bond angles was done using 3DNA 37.

Coordinates

Atomic 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 in the Protein Data Bank (accession codes: 2uu9, 2uua, 2uub and 2uuc)

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank Brian Sproat for his help and advice in the polymer chemistry, W.D. Graham for the purification and analyses of the 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 beam lines ID14 at the European Synchrotron Radiation Facility (Grenoble), and L. Passmore and M. Schmeing for helpful comments. This work was funded by the UK Medical Research Council (VR) and grants from the US National Institutes of Health (PFA and VR), the US National Science Foundation (PFA), the Agouron Institute (VR), the Austrian Academy of Sciences (AW) and the Polish Ministry of Science and Education (AM).

References

1. Crick FHC. The origin of the genetic code. J Mol Biol. 1968; 38:367–79. [PubMed: 4887876]

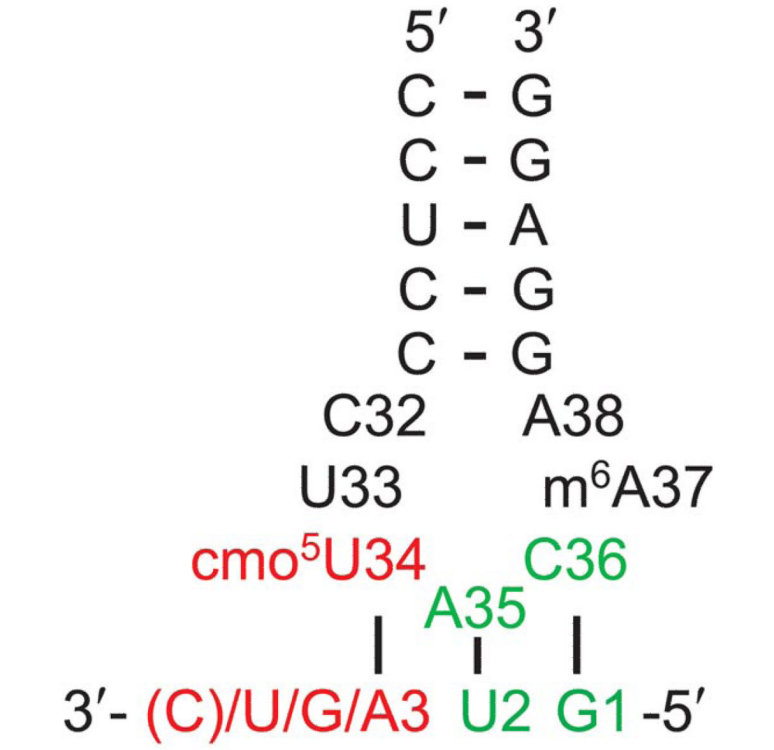
2. Sprinzl M, Horn C, Brown M, Ioudovitch A, Steinberg S. Compilation of tRNA sequences and sequences of tRNA genes. Nucleic Acids Res. 1998; 26:148–53. [PubMed: 9399820]
3. Mitra SK, Lustig F, Akesson B, Lagerkvist U. Codon-anticodon recognition in the valine codon family. J Biol Chem. 1977; 252:471–8. [PubMed: 319094]
4. Mitra K, Frank J. RIBOSOME DYNAMICS: Insights from Atomic Structure Modeling into Cryo-Electron Microscopy Maps. Annu Rev Biophys Biomol Struct. 2006; 35:299–317. [PubMed: 16689638]
5. Nasvall SJ, Chen P, Bjork GR. The modified wobble nucleoside uridine-5-oxyacetic acid in $tRNA^{Pro}(cmo^5UGG)$ promotes reading of all four proline codons in vivo. RNA. 2004; 10:1662–73. [PubMed: 15383682]
6. Agris PF, Vendeix FA, Graham WD. tRNA's wobble decoding of the genome: 40 years of modification. J Mol Biol. 2007; 366:1–13. [PubMed: 17187822]
7. Ogle JM, Ramakrishnan V. Structural Insights into Translational Fidelity. Ann Rev Biochem. 2005; 74:129–177. [PubMed: 15952884]
8. Rodnina MV, Wintermeyer W. Fidelity of aminoacyl-tRNA selection on the ribosome: kinetic and structural mechanisms. Annu Rev Biochem. 2001; 70:415–35. [PubMed: 11395413]
9. Ogle JM, et al. Recognition of cognate transfer RNA by the 30S ribosomal subunit. Science. 2001; 292:897–902. [PubMed: 11340196]
10. 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; 111:721–32. [PubMed: 12464183]
11. Murphy, F. V. t.; Ramakrishnan, V.; Malkiewicz, A.; Agris, PF. The role of modifications in codon discrimination by tRNA(Lys)UUU. Nat Struct Mol Biol. 2004; 11:1186–91. [PubMed: 15558052]
12. Selmer M, et al. Structure of the 70S ribosome complexed with mRNA and tRNA. Science. 2006; 313:1935–42. [PubMed: 16959973]
13. Hillen W,EE, Lindner HJ, Gassen HG, Vorbrüggen H. 5-Methoxyuridine: The influence of 5-substituents on the keto-enol tautomerism of the 4-carbonyl group. J. Carbohydrates-Nucleosides-Nucleotides. 1978; 5:23–32.
14. 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. 2005; 127:2238–48. [PubMed: 15713102]
15. Dirheimer, G.; Keith, G.; Dumas, P.; Westhof, E. tRNA: Structure, Biosynthesis, and Function. Söll, D.; RajBhandary, U., editors. American Society for Microbiology; Washington: 1995. p. 111-112.
16. Tinoco, IJ. The RNA World. Gesteland, RF.; Atkins, JF., editors. Cold Spring Harbor Laboratory Press; 1993. p. 603-607.
17. Cruse WB, et al. Structure of a mispaired RNA double helix at 1.6-A resolution and implications for the prediction of RNA secondary structure. Proc Natl Acad Sci U S A. 1994; 91:4160–4. [PubMed: 7514296]
18. Nagaswamy U, et al. 2003
19. Mitra SK, et al. Relative efficiency of anticodons in reading the valine codons during protein synthesis *in vitro*. J Biol Chem. 1979; 254:6397–401. [PubMed: 376532]
20. Sorensen MA, et al. Over expression of a tRNA(Leu) isoacceptor changes charging pattern of leucine tRNAs and reveals new codon reading. J Mol Biol. 2005; 354:16–24. [PubMed: 16236318]
21. Masquida B, Westhof E. On the wobble G•U and related pairs. RNA. 2000; 6:9–15. [PubMed: 10668794]
22. 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. 1978; 5:4451–61. [PubMed: 724522]
23. Fersht, AR. Structure and mechanism in protein science. W.H. Freeman; New York: 1998.
24. Doherty EA, Batey RT, Masquida B, Doudna JA. A universal mode of helix packing in RNA. Nat Struct Biol. 2001; 8:339–43. [PubMed: 11276255]

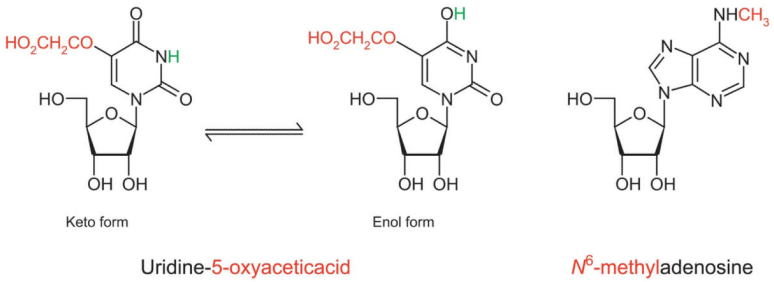
25. Kothe U, Rodnina MV. Codon reading by tRNAAla with modified uridine in the wobble position. Mol Cell. 2007; 25:167–74. [PubMed: 17218280]
26. Clemons WM Jr. et al. Crystal structure of the 30S ribosomal subunit from *Thermus thermophilus*: purification, crystallization and structure determination. J Mol Biol. 2001; 310:827–43. [PubMed: 11453691]
27. Sproat BS. RNA synthesis using 2'-O-(tert-butyldimethylsilyl) protection. Methods Mol Biol. 2005; 288:17–32. [PubMed: 15333895]
28. Boudou V,L,J, Van Aerschot A, Hendrix C, Millar A, Weiss P, Herdewijn P. Synthesis of the anticodon hairpin tRNAfMet containing N-{[9-(b-D-ribofuranosyl)-9H-purin-6-yl]carbamoyl}-L-threonine (=N6-{{[(1S,2R)-1-carboxy-2-hydroxypropyl]amino}-carbonyl}adenosine, t6A). Helvetica Chimica Acta. 2000; 83:152–161.
29. Gehrke CW, Kuo KC. Ribonucleoside analysis by reversed-phase high-performance liquid chromatography. J Chromatogr. 1989; 471:3–36. [PubMed: 2670985]
30. Kabsch W. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. J. Appl. Cryst. 1993; 26:795–200.
31. Collaborative Computational Project Number 4. The CCP4 suite: Programs for protein crystallography. Acta Cryst D. 1994; 50:760–3. [PubMed: 15299374]
32. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr. 2004; 60:2126–32. [PubMed: 15572765]
33. Brünger AT, et al. Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta Crystallogr D Biol Crystallogr. 1998; 54:905–21. [PubMed: 9757107]
34. Kleywegt GJ, Jones TA. Databases in protein crystallography. Acta Crystallogr D Biol Crystallogr. 1998; 54:1119–31. [PubMed: 10089488]
35. DeLano WL. The PyMOL Molecular Graphics System. 2006http://www.pymol.org
36. Masquida B, Sauter C, Westhof EA. sulfate pocket formed by three GoU pairs in the 0.97 A resolution X-ray structure of a nonameric RNA. RNA. 1999; 5:1384–95. [PubMed: 10573129]
37. Lu XJ, Olson WK. 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. Nucleic Acids Res. 2003; 31:5108–21. [PubMed: 12930962]

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

	U	C	A	G	
U	UUU Phe	UCU Ser	UAU Tyr	UGU Cys	U
	UUC Phe	UCC Ser	UAC Tyr	UGC Cys	C
	UUA Leu	UCA Ser	UAA STOP	UGA STOP	A
	UUG Leu	UCG Ser	UAG STOP	UGG Trp	G
C	CUU Leu	CCU Pro	CAU His	CGU Arg	U
	CUC Leu	CCC Pro	CAC His	CGC Arg	C
	CUA Leu	CCA Pro	CAA Gln	CGA Arg	A
	CUG Leu	CCG Pro	CAG Gln	CGG Arg	G
A	AUU Ile	ACU Thr	AAU Asn	AGU Ser	U
	AUC Ile	ACC Thr	AAC Asn	AGC Ser	C
	AUA Ile	ACA Thr	AAA Lys	AGA Arg	A
	AUG Met	ACG Thr	AAG Lys	AGG Arg	G
G	GUU Val	GCU Ala	GAU Asp	GGU Gly	U
	GUC Val	GCC Ala	GAC Asp	GGC Gly	C
	GUA Val	GCA Ala	GAA Glu	GGA Gly	A
	GUG Val	GCG Ala	GAG Glu	GGG Gly	G



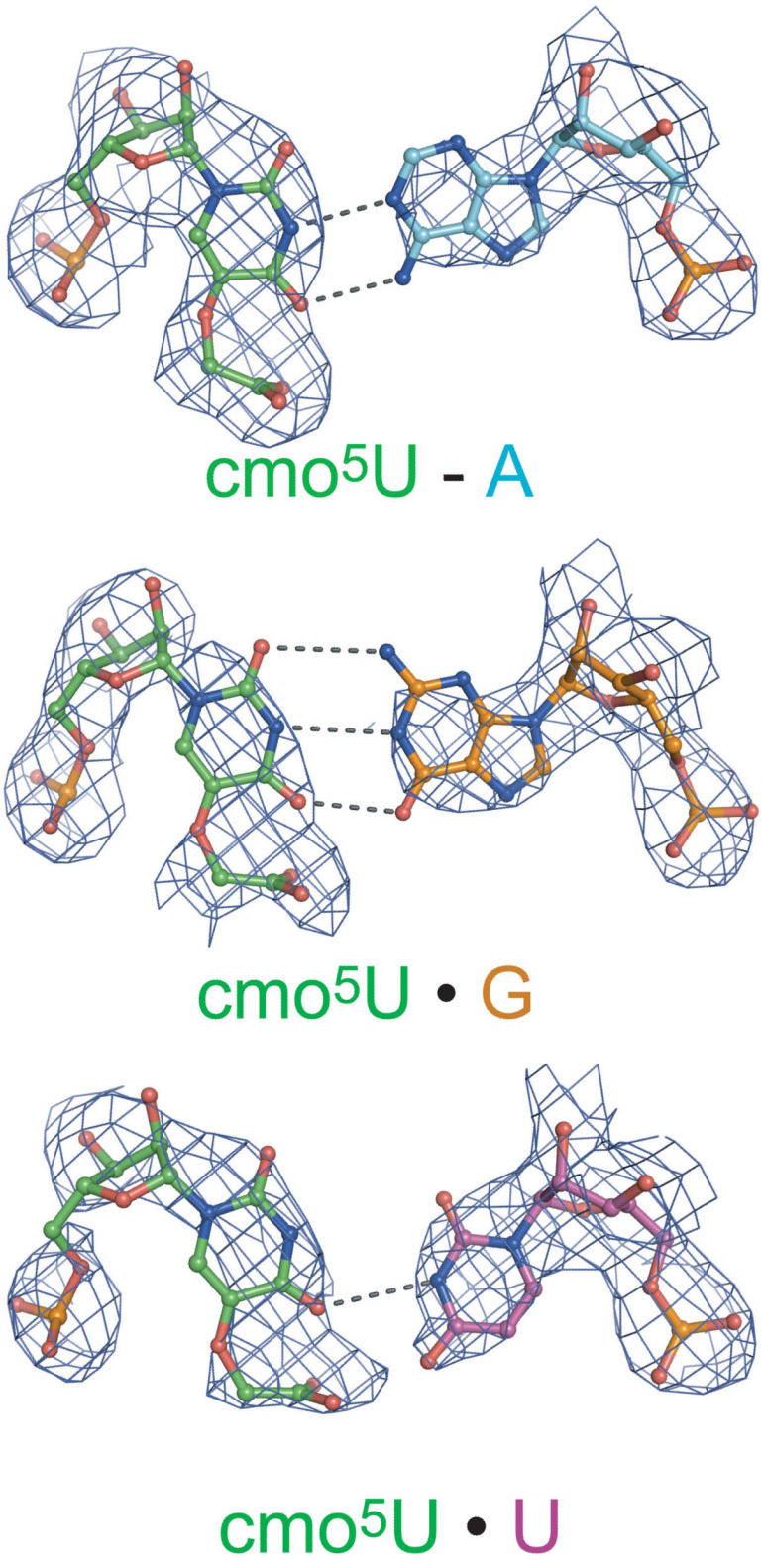


Europe PMC Funders Author Manuscripts

Figure 1.

The degeneracy of the genetic code. **(a)** The genetic code. 8 out of 16 codon boxes are degenerate (highlighted in green and white). Codon boxes where cmo^5U34 or derivatives can be found are highlighted in green. **(b)** The 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)** The modified bases present in *E. coli* $tRNA^{Val(cmo5UAC)}$. The modifications are highlighted in red. For cmo^5U the modification is expected to affect the keto-enol equilibrium.

Europe PMC Funders Author Manuscripts



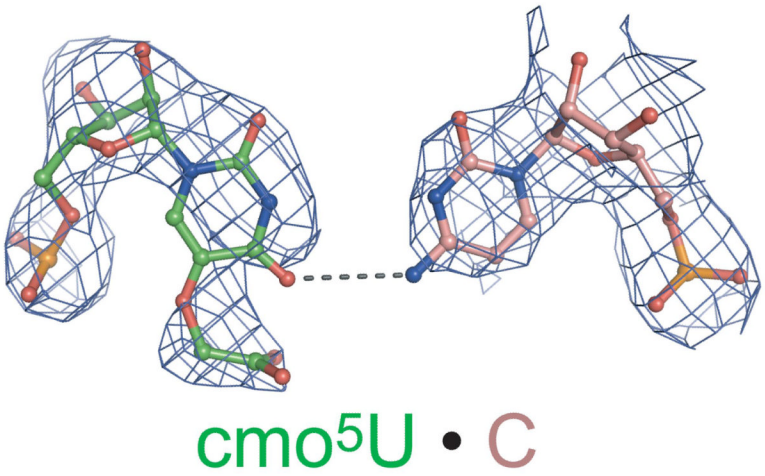
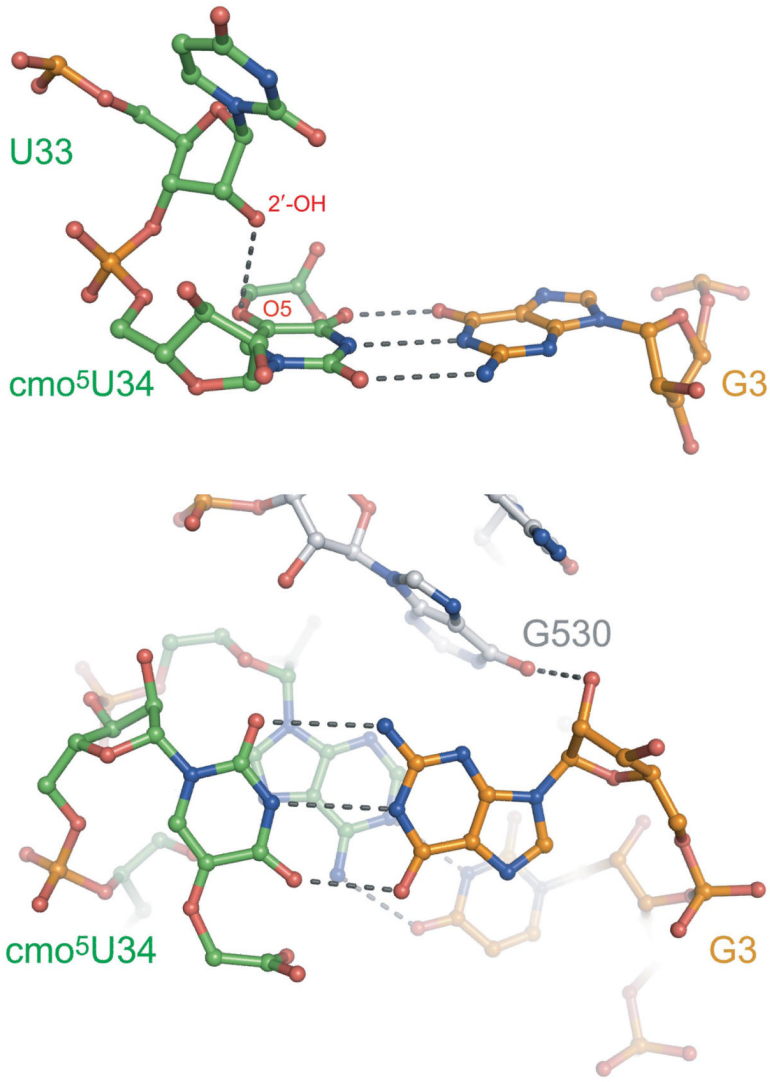


Figure 2.

cmo^{5}U forms base pairs to all four bases at the wobble position. The four base pairs between cmo^{5}U34 and the wobble base in the mRNA observed in the four structures along with unbiased difference Fourier maps. Hydrogen bonds between the bases are indicated. **(a)** cmo^{5}U34-A; **(b)** cmo^{5}U34-G; **(c)** cmo^{5}U34-U; **(d)** cmo^{5}U34-C. This and the other molecular graphics figures were made with PyMOL 35.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts



Europe PMC Funders Author Manuscripts

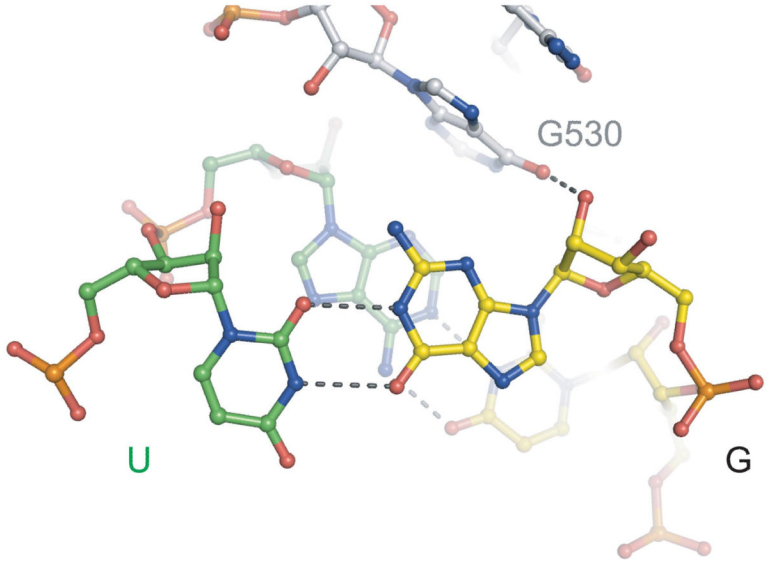
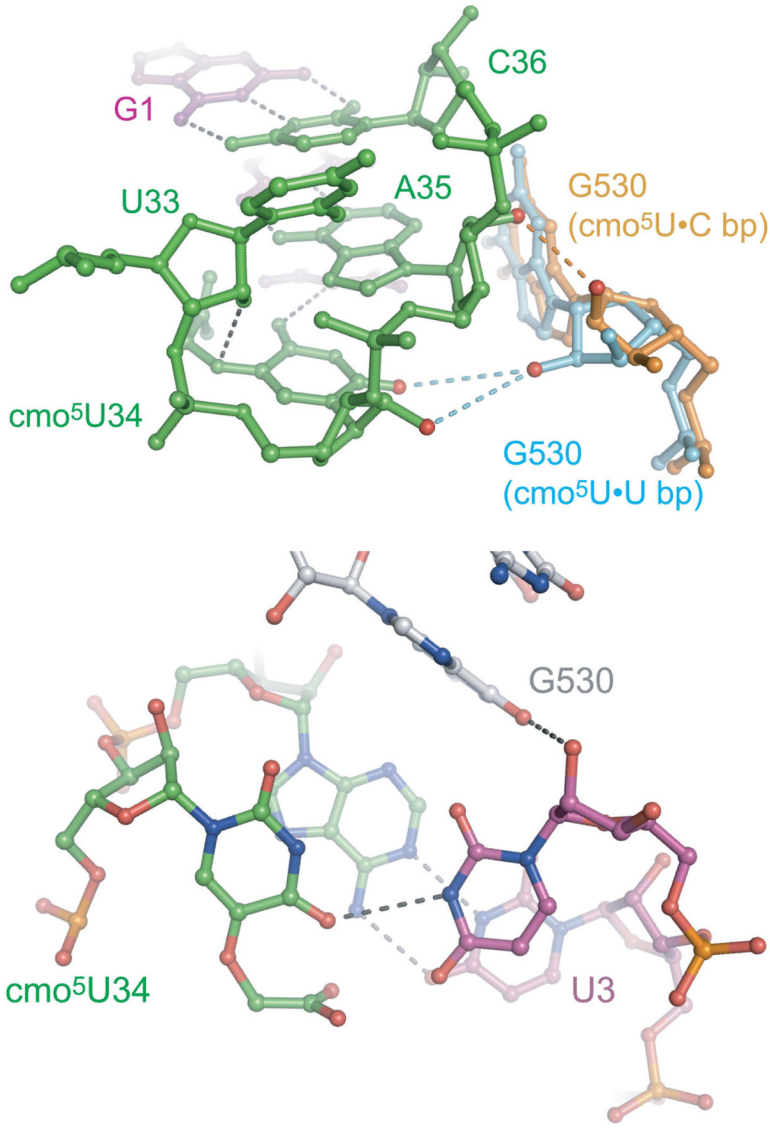


Figure 3.

Comparison of cmo5U-G with standard U-G wobble. **(a)** Hydrogen bond from 2′OH of U33 to O5 of the modification. This contact inhibits lateral movement of the cmo^{5}U. **(b)** Geometry of the cmo^{5}U34-G base pair we observe. **(c)** The U-G wobble base pair shown if it were in standard wobble geometry. Note that the U would be completely unstacked.

Europe PMC Funders Author Manuscripts



Europe PMC Funders Author Manuscripts

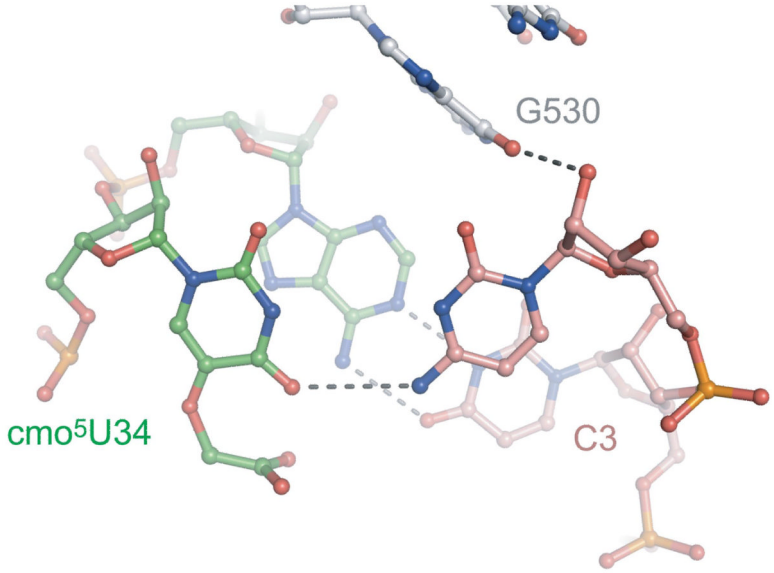


Figure 4.

The two pyrimidine-pyrimidine base pairs. **(a)** Alternative conformations of ribosomal base G530. In the cmo^{5}U-U base pair, the G530 ribose adopts a C2′-*endo* conformation, compensating for the desolvation of O2 and 2′OH of cmo^{5}U34 (cyan). In the cmo^{5}U-C base pair it adopts the standard conformation with a hydrogen bond to the 2′OH of A35 in the tRNA (orange). **(b,c)** In the cmo^{5}U-U base pair the U shows more stacking interaction with its 5′ base than the C does in the cmo^{5}U34-C base pair.

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		$P4_12_12$		
Cell dimensions				
a, *b*, *c* (Å)	*a*=*b*=402.2, *c*=174.9	*a*=*b*=400.9, *c*=174.2	*a*=*b*=401.9, *c*=174.6	*a*=*b*=400.5, *c*=173.4
α, β, γ (°)	α=β=γ=90	α=β=γ=90	α=β=γ=90	α=β=γ=90
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)
R_{sym}	15.9 (63.7)	26.2 (75.0)	17.4 (78.9)	13.4 (79.2)
$I / \sigma I$	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)	30.0-3.0	30.0-3.1	30.0-2.9	30.0-2.8
No. unique reflections	252,303	250,474	308,707	335,672
R_{work} / R_{free}	21.0/24.5	22.8/26.8	22.3/25.4	21.8/24.2
No. atoms				
RNA	32850	32873	32870	32828
Protein	19238	19238	19238	19238
Ions	242	199	194	255
Paromomycin	42	42	42	42
B-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