Wobble during decoding: more than third-position promiscuity

Luisa Cochella & Rachel Green

Structural studies by the Ramakrishnan and Agris groups allow us to directly observe how the ribosome's decoding site accommodates non-Watson-Crick base pairs in the third position of the codon-anticodon triplet while maintaining the one-amino-acid-per-codon framework that is central to life.

In the mid-1960s, while the genetic code was being elucidated, evidence began to accumulate that a given tRNA species could recognize more than one codon triplet. A comparison of these codons showed that they differed in the identity of the nucleotide in the third position[1–3]. Francis Crick[4] proposed that there could be limited "wobble pairing" at the third position, allowing some base pairs between the tRNA and the codon with geometries that were merely "close to the standard one," but still preventing wholesale promiscuity. The proposed wobble rules (**Table 1**) were confirmed over and over thereafter, and the list of acceptable third-position base pairs grew as the number of modified bases found in tRNAs' anticodons increased (**Table 1**). Although the wobble rules have long been generally accepted, the conformational deviations in the codon·anticodon region associated with "wobble" and the manner in which these deviations are accommodated in the ribosome's decoding center are completely unknown. Further, although a role for the numerous modified nucleotides in the anticodon of tRNA in permitting and restricting wobble has been documented[5], a detailed molecular explanation for these effects has awaited high-resolution structural information.

In two separate reports on pages 1186 and 1251 of this issue of *Nature Structural & Molecular Biology*, Murphy *et al.*[6,7] describe crystal structures at ~3 Å resolution of different codon-anticodon pairs in the decoding site of the *Thermus thermophilus* 30S subunit. These data illustrate some of the molecular details of the wobble rules. In the first study the authors focus on the interactions made by the purine inosine with two different partners at the third position of the codon (the wobble position)[6]. Inosine was first identified as a component of yeast $tRNA^{Ala}$ by Holley *et al.*[8] and was later found to be more common than adenosine at position 34 of the anticodon[5]. Indeed, inosine had already been found in such a high fraction of the tRNA sequences that were available in 1966 that it became a cornerstone of the original wobble hypothesis. Crick recognized that inosine had the potential for promiscuous interactions and proposed near-standard pairing with U and C and a slightly broader purine·purine pair (I•A) where both bases were found in the *anti* conformation. However, a number of crystal structures have provided examples of purine·purine base pairs that assume *syn*•*anti* arrangements[9–11], suggesting that such a geometry may be more readily accommodated in helical RNA elements. In light of such structures and of the expectation that the decoding center must impose harsh restrictions on the overall width of accepted base pairs, it was speculated that the I•A wobble interaction might also assume a *syn*•*anti* geometry in the decoding center (**Fig. 1a**).

Murphy *et al.*[6] resolve this discrepancy by determining the structures of a $tRNA^{Arg}_{IGC}$ anticodon stem-loop (ASL) bound to 30S ribosomes with two different codons, one creating a rather standard I•C base pair and the other an I•A wobble base pair. The latter reveals how inosine is accommodated to form an unusual wide purine·purine base pair in the context of the ribosome. As originally proposed by Crick[4], the base pair adopts an I_{anti}•A_{anti} arrangement, whereby the increase in base pair width is accommodated by changes in the tRNA anticodon backbone, specifically in the torsional angle β between C5′ and O5′ of the inosine nucleotide (**Fig. 1b**). Such conformational changes were previously predicted by Fuller and Hodgson in 1967 (ref. 12) on the basis of the anticodon structure, and later by Ramakrishnan's group when they observed that in the context of the small subunit, the third nucleotide of the mRNA codon is fixed by both G530 of the 16S rRNA and protein S12 whereas the anticodon base is less constrained[13]. Thus the ability of a tRNA to recognize multiple nucleotides at the wobble position not only requires acceptable hydrogen bonding but in some cases depends on a mildly deformable anticodon loop.

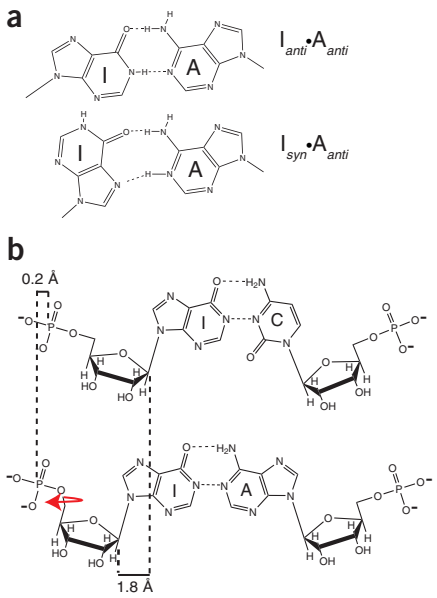


Figure 1 Inosine·adenosine base pair conformations. (**a**) Two alternative arrangements for the I•A base pair. Top, with both bases in the *anti* conformation; bottom, with inosine rotated about the glycosidic bond, in the *syn* conformation. (**b**) Schematic of the conformational change required to accommodate the 'wide' purine·purine base pair I•A in the ribosome's decoding center. Although an increase of 1.8 Å in C1′–C1′ distance is observed as compared to a standard purine·pyrimidine base pair, the difference in P–P distance is only 0.2 Å. Most of the change is absorbed by a substantial rotation (37.8°) of the β torsion angle of the main chain, as indicated by the red arrow.

In the second study we see how post-transcriptional modifications have similarly been used to modulate codon·anticodon interactions so as to maintain fidelity even in the face of degenerate pairing rules[7]. In this analysis, the authors explore the role of post-transcriptional modifications found in the anticodon loop of $tRNA^{Lys}_{UUU}$ in the decoding of both AAA and AAG codons. This particular tRNA species has long been of interest for a number of reasons. First, $tRNA^{Lys}_{UUU}$ carries an N^6-threonyl-

Howard Hughes Medical Institute, Department of Molecular Biology and Genetics, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, USA.
e-mail: ragreen@jhmi.edu

Table 1 Wobble pairing rules

Crick's wobble rules		Modified rules	
Anticodon	Codon	Anticodon	Codon
G	U, C	G	U, C
C	G	C	G
		k^2C	A
A	U	A	U, C, G > A
U	A, G	U	U, A, G > C
		xm^5s^2U, xm^5Um, Um, xm^5U	A > G
		xo^5U	U, A, G
I	U, C, A	I	U, C, A

Pairing rules for third codon position, or wobble position, as first proposed by Crick[4] and further expanded to include modified bases. k^2C, lysidine; xm^5s^2U, 5-methyl-2-thiouridine derivatives; xm^5Um, 5-methyl-2′-O-methyluridine derivatives; Um, 2′-O-methyluridine; xm^5U, 5-methyluridine derivatives; xo^5U, 5-hydroxyuridine derivatives. From Yokoyama and Nishimura[17].

carbamoyladenosine (t^6A) residue at position 37 (**Fig. 2a**), 3′-adjacent to the anticodon (positions 34–36). Bulky purines are often found at this position in other anticodons, and for $tRNA^{Lys}$ this modification is required for binding to the AAA codon. Interestingly, however, this singly modified ASL will not bind to the AAG lysine codon. Only a $tRNA^{Lys}_{UUU}$ carrying both t^6A_{37} and the modification 5-methylaminomethyluridine (mnm^5U) at the wobble position 34 can bind both codons. Finally, although third-position U•U wobble is permitted for a number of tRNAs (**Table 1**), $tRNA^{Lys}_{UUU}$ does not decode the asparagine codon AAU. This specificity has been attributed to the modification at U_{34}, which was speculated to restrict the tRNA's wobble capacity. Again, high-resolution structural data[7] provide a molecular explanation for this collection of biochemical observations.

The bulky purine residue t^6A_{37} apparently serves two important functions. First, the modification on the Watson-Crick face of the nucleotide effectively prevents the formation of an otherwise feasible U_{33}•A_{37} base pair within the tRNA itself that would disrupt the anticodon loop structure. Second, and more interestingly, what emerges here is that the modification provides substantially increased stacking potential through the action of the planar ureido group that increases t^6A_{37}–A_{38} stacking within the ASL, thus preordering the anticodon (**Fig. 2a**). Because of this stabilizing interaction, the overall energetic cost of binding is reduced. Further, the increased stacking potential of t^6A_{37}, as well as steric restrictions imposed by the bulky threonyl moiety, lead to productive interactions with A_1 of the codon, creating a stabilizing cross-strand stack (**Fig. 2a**). Both the preordering of the anticodon and the extra stabilization of binding increase the overall energy of the inherently weak AAA•UUU codon·anticodon interaction.

What, then, is the function of the requisite uridine modification mnm^5U_{34}? It is known that, in general, G•U wobble pairs at the *ends* of helices where G is at the 5′ end show more favorable stacking with the neighboring Watson-Crick pair than do U•G pairs where G is at the 3′ end[14]. Consistent with these biophysical observations, the most commonly accepted wobble interactions (at the end of the codon·anticodon minihelix) include a purine (G or I) at position 34. For $tRNA^{Lys}_{UUU}$, among others; however, the purine is on the 3′ side of the duplex (at position 3 of the codon) and is thus inherently less stabilizing. Moreover, in $tRNA^{Lys}_{UUU}$ the issue of stability is aggravated by the presence of another poor stacker, uridine, in the second anticodon position. It turns out that these tRNA species depend on a modified U at the critical wobble position to compensate for these thermodynamic deficiencies.

In the structure of the doubly modified $tRNA^{Lys}_{UUU}$ ASL with an AAG codon bound to the ribosome[7], we see how mnm^5U_{34} facilitates formation of an otherwise disfavored wobble interaction by stabilizing an unusual conformation of the mnm^5U_{34}•G_3 base pair. This base pair deviates from standard U•G wobble geometry and instead maximizes anticodon stacking (mnm^5U_{34} on U_{35}). Again, accommodation takes place in the anticodon backbone, in this case causing mnm^5U_{34} to shift to a position that is 'halfway' between a Watson-Crick and a common U•G base pair (**Fig. 2b**). Why does mnm^5U_{34} facilitate U•G but not U•U wobble? An explanation for the inability of $tRNA^{Lys}_{UUU}$ to decode AAU codons (with U•U wobble) is that if the U•U pair adopted the normal wobble geometry, anticodon stacking would be disrupted, whereas if it adopted the geometry observed here for mnm^5U_{34}•G_3, stacking would be maintained but there would be no hydrogen bonding interactions with the U in the codon.

Together these examples illustrate a number of important points concerning codon•anticodon interactions. First, the significance of base stacking in defining nucleic acid interactions and structure is showcased by the contributions both from base modifications and from unusual base pair geometries. Stacking has long been known to be a dominant feature in RNA structure—for example, in $tRNA^{Phe}$ 71 of 76 nucleotides form substantial stacking interactions and the three unstackable, nonaromatic dihydrouridines account for the majority of these unstacked nucleotides[15]. Because of this role of stacking, it follows that both the anticodon composition itself and the neighboring nucleotides will have an important impact on the decoding properties of the actual anticodon element. A second principle that emerges is that conformational changes required to accommodate unusual pairing in the ribosome's decoding center are generally absorbed by the anticodon backbone. This is in good agreement with previous studies from the same authors showing that the anticodon wobble position is relatively unconstrained by the ribosome whereas the position of the third codon nucleotide is quite fixed[13].

The changes observed here in the details of the codon·anticodon interaction must ultimately be 'read' by the translation machinery. An important downstream consequence of codon recognition is the crystallographically observable closure of the 30S subunit around cognate aa-tRNA[13]. Each of the codon·anticodon interactions described in these manuscripts triggers domain closure, confirming that the ribosome recognizes them as cognate. It will be interesting to see whether there are effects on the kinetics of tRNA selection for some of these 'less favorable' codon·anticodon pairs or whether the translation machinery has evolved to compensate for these inherent differences as it has in other cases[16].

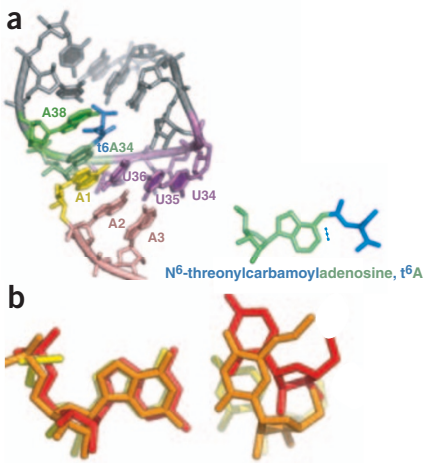
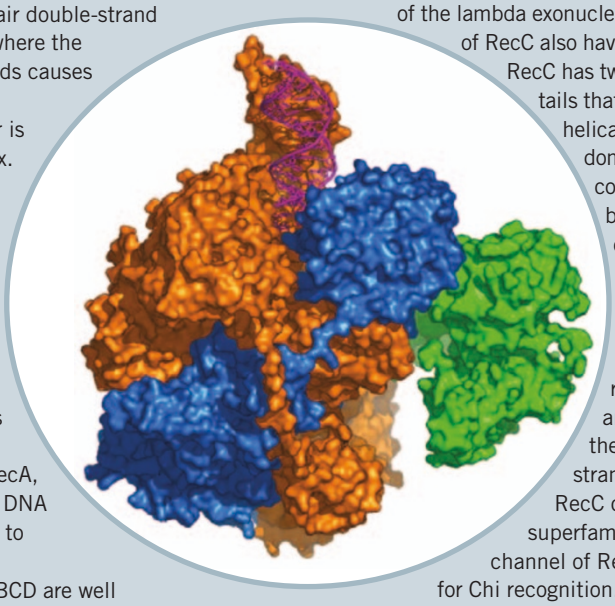


Figure 2 Cartoon structures showing the effects of modified nucleotides during decoding. (**a**) Overview of the codon·anticodon interaction with the anticodon (U_{34}, U_{35}, U_{36}) in purple and the codon in yellow (A_1) and pink (A_2, A_3). This view shows cross-strand stacking of t^6A between A_{38} of the ASL and A_1 of the codon. Inset shows the modified nucleotide t^6A (with the adenosine moiety in green and the modification in blue), with its planar ureido group mimicking a third heterocycle. (**b**) Comparison of the geometries of a Watson-Crick A•U base pair (yellow), a standard G•U wobble base pair (red) and the modified G•mnm^5U base pair (orange), 'halfway' between the other two. Figures made from PDB files from[7](1XMQ and 1XMO), using PyMOL (http://www.pymol.org).

1. Bernfield, M.R. & Nirenberg, M.W. *Science* **147**, 479–484 (1965).
2. Söll, D. *et al. J. Mol. Biol.* **19**, 556–573 (1966).
3. Kellogg, D.A., Doctor, B.P., Loebel, J.E. & Nirenberg, M.W. *Proc. Natl. Acad. Sci. USA* **55**, 912–919 (1966).
4. Crick, F.H. *J. Mol. Biol.* **19**, 548–555 (1966).
5. Agris, P.F. *Nucleic Acids Res.* **32**, 223–238 (2004).
6. Murphy, F.V. & Ramakrishnan, V. *Nat. Struct. Mol. Biol.* **11**, 1251–1252 (2004).
7. Murphy, F.V., Ramakrishnan, V., Malkiewicz, A. & Agris, P.F. *Nat. Struct. Mol. Biol.* **11**, 1186–1191 (2004).
8. Holley, R.W. *JAMA* **194**, 868–871 (1965).
9. Leonard, G.A., Booth, E.D., Hunter, W.N. & Brown, T. *Nucleic Acids Res.* **20**, 4753–4759 (1992).
10. Subramanian, E., Madden, J.J. & Bugg, C.E. *Biochem. Biophys. Res. Commun.* **50**, 691–696 (1973).
11. Carter, R.J., Baeyens, K.J., SantaLucia, J., Turner, D.H. & Holbrook, S.R. *Nucleic Acids Res.* **25**, 4117–4122 (1997).
12. Fuller, W. & Hodgson, A. *Nature* **215**, 817–821 (1967).
13. Ogle, J.M. *et al. Science* **292**, 897–902 (2001).
14. Mizuno, H. & Sundaralingam, M. *Nucleic Acids Res.* **5**, 4451–4461 (1978).
15. Holbrook, S.R., Sussman, J.L., Warrant, R.W. & Kim, S.H. *J. Mol. Biol.* **123**, 631–660 (1978).
16. LaRiviere, F.J., Wolfson, A.D. & Uhlenbeck, O.C. *Science* **294**, 165–168 (2001).
17. Yokoyama, S. & Nishimura, S. *tRNA: Structure, Biosynthesis and Function* (eds. Söll, D. & RajBhandary, U.) 207–223 (American Society for Microbiology, Washington, DC, 1995).

Rec-ing DNA

Double-strand DNA (dsDNA) breaks are caused most frequently by ionizing radiation. The breaks can cause cell death if not repaired, because they prevent the progression of polymerases along the DNA. The cell can repair double-strand breaks through recombination, where the reunion of the broken DNA strands causes a physical exchange of parts. In eubacteria, recombination repair is initiated by the RecBCD complex. RecBCD has a high affinity for blunt DNA ends and so is able to bind the damaged dsDNA. It also possesses helicase and nuclease activities that unwind and chew away dsDNA until the complex encounters a Chi (crossover hotspot instigator) site, a DNA sequence that favors recombination, on the 3′ strand. At the Chi site, RecBCD loads RecA, a recombinase that can catalyze DNA strand exchange, onto the 3′ tail to initiate recombination.



The multiple activities of RecBCD are well studied and can be mapped to its three components: RecB is a nuclease and 3′→ 5′ helicase, RecC recognizes the Chi site, and RecD has 5′→ 3′ helicase activity. To elucidate how RecBCD is able to perform these various tasks, Singleton *et al.* (*Nature* **432**, 187–193; 2004) have solved the structure of *Escherichia coli* RecBCD in complex with a dsDNA substrate, a 19 bp duplex with a 5 bp hairpin on one end and a blunt end on the other.

The structure of the complex shows DNA (purple) bound between RecB (orange) and RecC (blue). RecD (green) is located on the periphery of the complex and interacts with RecC. As had been anticipated, the DNA is bound to RecBCD with its blunt end lying in the complex and the last 4 bp unwound. Both RecB and RecD share structural similarity with members of the superfamily 1 group of helicases. The C-terminal portion of RecB has a fold similar to that of the lambda exonuclease core. Interestingly, two of the domains of RecC also have the same fold as superfamily 1 helicases. RecC has two channels that accommodate the ssDNA tails that run to and from the RecB and RecD helicase subunits. The third and final RecC domain interacts with RecB, makes extensive contacts with the two ssDNA strands, and bears a 'pin' that physically separates the duplex DNA before feeding the 3′ strand to RecB and the 5′ strand to RecD.

The structure of RecBCD reveals how nuclease and motor helicase activities are coupled, allowing for Chi site recognition. The helicase motors of RecB and RecD pull the DNA strands across the pin of RecC, splitting the duplex. Each strand of the DNA is passed through the RecC channels. The ssDNA-binding site in the superfamily 1 helicases corresponds to the 3′ tail channel of RecC, which the authors theorize is a site for Chi recognition. This is supported by previously identified mutations in this region that alter Chi site recognition. The 3′ tail exits from this channel and is passed directly into the nuclease domain of RecB, where it is digested processively. The 5′ tail must first be passed to RecD before reaching the nuclease domain, and is therefore cut less frequently. When a Chi site is recognized by RecC, it is able to bind tightly to the 3′ strand, preventing further digestion of this strand and allowing for more frequent cleavage of the 5′ tail. The challenge is now to understand the subsequent steps of recombination repair (which include recruitment of RecA to the 3′ tail) that allow double strand break repair to begin.

Michelle Montoya