Recognizing the D-loop of transfer RNAs

Tamara L. Hendrickson*

Department of Chemistry, Johns Hopkins University, 3400 North Charles Street, Baltimore, MD 21218

Mechanisms that maintain fidelity and repair mistakes are ubiquitous throughout the protein biosynthesis pathway (1). The aminoacyl tRNAs serve as critical turning points in translation, because they link the nucleic acid genetic code with the amino acid building blocks of proteins. Misacylation of tRNAs can have devastating results, affecting the very survivability of an organism. Accuracy in tRNA aminoacylation therefore is paramount to the fidelity of the genetic code. Aminoacyl tRNAs are generated typically through action of the diverse family of aminoacyl-tRNA synthetases (AARSs). One of the hallmarks of these enzymes is the exquisite specificity with which each selects and aminoacylates only its cognate tRNA(s) with only its cognate amino acid (2–5). Herculean efforts spanning nearly 25 years have produced crystal structures for 19 of the 20 AARSs (alanyl-tRNA synthetase is the last holdout). Many of these structures provide detailed molecular insight into the nature of tRNA recognition and discrimination by either providing a structure of the enzyme in complex with its cognate tRNA or suggesting a model that can be tested biochemically. Thus, a clear picture is beginning to emerge, delineating some of the common and not-so-common themes used by the AARSs to discriminate cognate from noncognate tRNAs. [There are several reviews available that describe tRNA recognition by AARSs in detail (4–7).]

Each AARS recognizes specific identity elements within its cognate tRNA(s); these nucleotides are often rigorously conserved and are typically clustered in the anticodon stem loop and the acceptor stem of the tRNA (Fig. 1). [For a comprehensive review on tRNA identity, see Giegé *et al.* (4).] The crystal structures of several different tRNA·AARS complexes demonstrate that most AARSs form specific hydrogen bonding arrays with identity elements in either or both of these tRNA regions (Fig. 1*A*) and occasionally with the variable arm (Fig. 1, VA) as well. In contrast, the rest of a given tRNA is often neglected by its AARS, with the junction of the D-loop and TΨC-loop extending away from the enzyme (for an example, see Fig. 2*B*). On page 13537 of this issue of PNAS, Shimada *et al.* (8) present a detailed structural analysis of ArgRS from *T. thermophilus* and suggest a mechanism by which this enzyme recognizes an unusual D-loop identity element in $tRNA^{Arg}$ (A20).

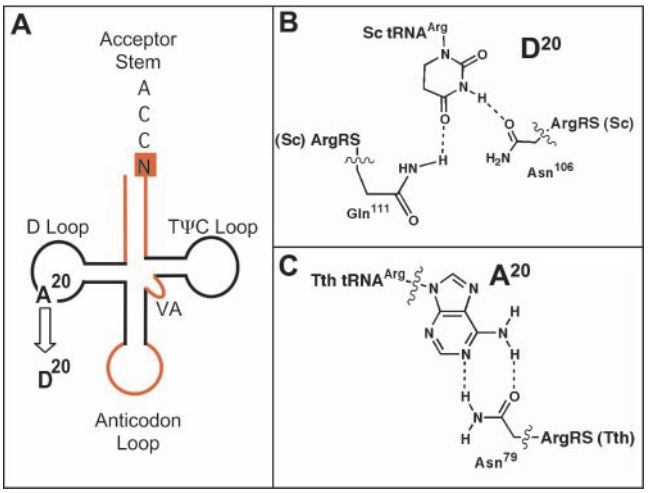

Fig. 1. Recognition of nucleotide 20 by *Saccharomyces cerevisiae* (Sc) and *Thermus thermophilus* (Tth) arginyl-tRNA synthetase (ArgRS). (*A*) The secondary structure of tRNA. Position 20 is an A in *T. thermophilus* $tRNA^{Arg}$ and a D in *S. cerevisiae* $tRNA^{Arg}$. In most tRNAs, nucleotide determinants tend to cluster around the acceptor stem, anticodon loop, discriminator base, and variable arm (VA), all of which are marked in red. (*B*) Recognition of D20 by *S. cerevisiae* ArgRS. (*C*) Recognition of A20 by *T. thermophilus* ArgRS.

Nucleotide A20 is conserved in most known $tRNA^{Arg}$ isoacceptors and has been identified as an identity element both *in vitro* and *in vivo* (9–11). Exceptions include *S. cerevisiae*, *Neurospora crassa*, *Schizosaccharomyces pombe*, and the mitochondria of animals and single cell eukaryotes, where position 20 is less conserved and can be U, C, or D (dihydrouridine; ref. 12). A cocrystal structure of *S. cerevisiae* ArgRS with a $tRNA^{Arg}$ isoacceptor containing D20 has been reported previously (13). In this crystal structure, D20 is recognized by Asn^{106}, Phe^{109}, and Gln^{111} (Fig. 1*B*, Phe^{109} stacks against the D20 aromatic ring and is not shown). On the basis of structural similarities between the *T. thermophilus* ArgRS structure and the published *S. cerevisiae* ArgRS·$tRNA^{Arg}$ complex (13), Shimada *et al.* (8) were able to dock *T. thermophilus* $tRNA^{Arg}$ (containing A20) onto their structure of *T. thermophilus* ArgRS (Fig. 2*A*). In this model, Asn^{106} of *S. cerevisiae* ArgRS is replaced by Val^{74} in *T. thermophilus* ArgRS. The smaller Val^{74} side chain creates a wider cavity to accommodate the larger A20 nucleotide. In contrast, Tyr^{77} and Asn^{79} (the two residues that align with Phe^{109} and Gln^{111} of *S. cerevisiae* ArgRS) are proximal to A20 but are too far removed to form direct contacts. This observation led the authors to propose and biochemically evaluate possible local rearrangements in the *T. thermophilus* ArgRS·$tRNA^{Arg}$ complex. Site-directed mutagenesis of Asn^{79} in ArgRS and/or A20 in $tRNA^{Arg}$, followed by kinetic analyses of each new $tRNA^{Arg}$/ArgRS combination, demonstrated that Asn^{79} indeed is involved directly in recognition of A20. A local structural model was constructed to reflect the proposed conformational reorganization after $tRNA^{Arg}$ binding. In this final model, A20 directly contacts the Asn^{79} carboxyamide side chain via two hydrogen bonds [Figs. 1*C* and 4*C* in the accompanying paper (8)]. The results of these two crystallographic and biochemical analyses of ArgRS reveal a detailed molecular picture of the two different mechanisms used by this enzyme to recognize a D-loop identity element in $tRNA^{Arg}$ (compare Fig. 1 *B* and *C*; refs. 8 and 13).

See companion article on page 13537.

*E-mail: Tamara.Hendrickson@jhu.edu.

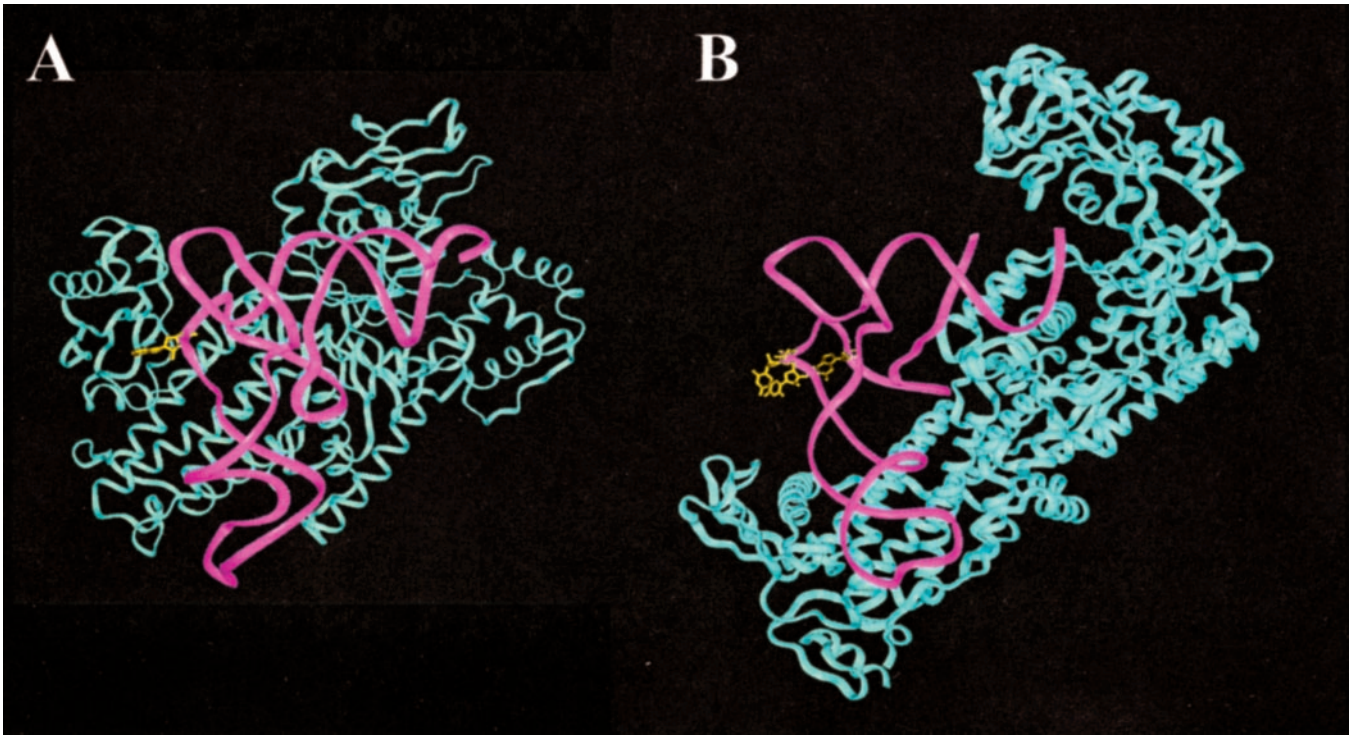

Fig. 2. Structural recognition of D-loop identity determinants. Each AARS is shown in cyan, and each tRNA is shown in magenta. Known D-loop identity elements are shown in stick representation in yellow. (*A*) The crystal structure of *T. thermophilus* ArgRS·tRNA^{Arg}; the N-terminal domain is modeled to contacts A20 (8). (*B*) The crystal structure of *Staphylococcus aureus* isoleucyl-tRNA synthetase IleRS·tRNA^{Ile}, highlighting G16, D20, and D21 (27); IleRS does not form contacts with the D-loop identity elements in tRNA^{Ile}.

Identity elements in the D-loop of tRNAs are rare (4). In addition to tRNA^{Arg}, D-loop determinants have been characterized kinetically in two other tRNAs: tRNA^{Phe} (D20) (14, 15) and tRNA^{Ile} (G16, D20, and D21) (16, 17). Despite cocrystal structures for each corresponding tRNA·AARS complex (8, 13, 18, 19), a key question remains. How is recognition of a D-loop nucleotide communicated to the distal synthetase active site? This question cannot be truly answered without further experimental data. However, the combination of complementary kinetic and crystallographic analyses can offer some insight into the role of each of the known D-loop identity elements.

Remarkably, despite the positional similarities between the different D-loop determinants, the kinetic role of each set of identity elements is unique to its particular tRNA isoacceptor. In tRNA^{Arg}, replacement of A20 with any other nucleotide leads to a dramatic drop in tRNA^{Arg} aminoacylation efficiency *in vivo* and *in vitro*. Furthermore, this decrease has been traced kinetically to a defect in V_{max} (8–11). In the current report from Shimada *et al.*, a similar effect on V_{max} results from mutations in ArgRS; when Asn^{79} is mutated to lysine or arginine, V_{max} drops ≈2 orders of magnitude (8). In contrast, the K_m for tRNA^{Arg} ranges from ≈20% *better* (N79K) to only ≈2-fold worse (N79R) than its K_m with wild-type ArgRS. Thus, maximal catalytic efficiency occurs only when the correct tRNA is recognized, even though binding of this tRNA is unaffected by mutations in the D-loop or in the regions of ArgRS that contact the identity element. These reciprocal sets of data are remarkable, because they suggest that a crucial recognition event occurs at the D-loop of the tRNA, *distal to the enzyme active site*, and that this event is communicated somehow to the catalytic site. One possibility is that the hydrogen bonding array presented by Shimada *et al.* induces a conformational change in ArgRS, tRNA^{Arg}, or both to position tRNA^{Arg} properly in the enzyme active site. Similar conformational changes have been observed in ArgRS for the more common long-range recognition of anticodon determinants (13).

Dihydrouridine 20 (D20) is a key identity element for phenylalanyl-tRNA synthetase (PheRS) both *in vivo* and *in vitro* (14, 15). In contrast to the recognition of tRNA^{Arg} by ArgRS, mutagenesis of D20 in tRNA^{Phe} produces defects in both K_m and k_{cat} (15). This mixed kinetic effect suggests the possibility that a defect in binding (K_m) positions the tRNA incorrectly within the active site, thereby diminishing the catalytic efficacy of the enzyme. In the crystal structure of *T. thermophilus* PheRS·tRNA^{Phe} (18), tRNA-enzyme contacts occur between G19 and a coiled-coil extension in the enzyme's β-subunit, but contacts are not observed with the identity determinant D20. A recent study, however, localized interactions between the β-subunit of PheRS and D20 in tRNA^{Phe} via crosslinking (19), again suggesting that subtle structural perturbations within the tRNA^{Phe}·PheRS complex may take place during aminoacylation. Thus, as suggested for the ArgRS·tRNA^{Arg} complex, conformational flexibility may be integral to the recognition of D-loop identity elements.

Finally, three D-loop identity determinants have been identified in tRNA^{Ile} (16, 17), although the role of nucleotides G16, D20, and D21 is distinct from the D-loop identity elements in tRNA^{Arg} and tRNA^{Phe}. IleRS, the enzyme that aminoacylates tRNA^{Ile}, only poorly discriminates between isoleucine and valine (20, 21). Valine is misactivated occasionally by IleRS to generate valyl adenylate (Val-AMP) and even misacylated tRNA^{Ile} (Val-tRNA^{Ile}). To maintain the fidelity of the genetic code, IleRS catalyzes two hydrolytic proofreading reactions that eliminate both Val-AMP and Val-tRNA^{Ile}, thereby preventing accumulation of Val-tRNA^{Ile} and errors in protein biosynthesis. G16, D20, and D21 in the D-loop of tRNA^{Ile} are recognition elements for these hydrolytic editing reactions only (16, 17) and are quite distinct from the set of

identity elements described for aminoacylation of $tRNA^{Ile}$ (22, 23).

The editing site of IleRS has been characterized by site-directed mutagenesis (24–26) and x-ray crystallography (25, 27) and is ≈25 Å from the enzyme's synthetic active site (where aminoacylation takes place). Thus, for editing to occur the two substrates (Val-AMP and Val-$tRNA^{Ile}$) are translocated from the active site to the remote editing site. This translocation event is $tRNA^{Ile}$-dependent even for Val-AMP (28, 29). In the IleRS·$tRNA^{Ile}$ crystal structure, the acceptor stem of the tRNA has unwound to position the 3′ end in the editing site (27). Therefore, at least in the case of posttransfer editing (hydrolysis of Val-$tRNA^{Ile}$), $tRNA^{Ile}$ undergoes distinct structural rearrangements during translocation, suggesting that the role of the D-loop determinants may be to direct or allow translocation. Remarkably, in the IleRS·$tRNA^{Ile}$ crystal structure, the tRNA D-loop does not form any contacts with IleRS (Fig. 2*B*; ref. 27), and thus an understanding of the mechanism by which these identity elements facilitate editing remains elusive.

Identity elements in the D-loop of tRNAs seem uncommon and are not confined to one type of interaction or mode of recognition by their cognate tRNA synthetase. In fact, the effect of D-loop mutations, as described above, leads to distinctly different consequences in $tRNA^{Arg}$, $tRNA^{Phe}$, and $tRNA^{Ile}$. An apparent common feature of these three tRNA systems is structural flexibility, either within the tRNA or in the corresponding AARS. The conserved L shape of tRNA results in a fixed distance between the active site of an AARS and the cognate tRNA's D-loop; the same tRNA structure facilitates recognition of the anticodon and acceptor stem, traditional hot spots for tRNA identity. The diverse roles of these novel D-loop identity elements, however, argue that the AARS can adapt as necessary to take advantage of additional tRNA recognition motifs and structural flexibility.

I thank Dr. Rebecca Alexander for helpful discussions.

1. Ibba, M. & Soll, D. (1999) *Science* **286,** 1893–1897.
2. Meinnel, T., Mechulam, Y. & Blanquet, S. (1995) in *tRNA: Structure, Biosynthesis and Function*, eds. Soll, D. & RajBhandary, U. (Am. Soc. Microbiol., Washington, DC), pp. 251–292.
3. Sherman, J. M. & Soll, D. (1996) *Biochemistry* **35,** 601–607.
4. Giegé, R., Sissler, M. & Florentz, C. (1998) *Nucleic Acids Res.* **26,** 5017–5035.
5. Ibba, M. & Soll, D. (2000) *Annu. Rev. Biochem.* **69,** 617–650.
6. Cusack, S. (1997) *Curr. Opin. Struct. Biol.* **7,** 881–889.
7. Beuning, P. J. & Musier-Forsyth, K. (1999) *Biopolymers* **52,** 1–28.
8. Shimada, A., Nureki, O., Goto, M., Takahashi, S. & Yokoyama, S. (2001) *Proc. Natl. Acad. Sci. USA* **98,** 13537–13542. (First Published November 6, 2001; 10.1073/pnas.231267998)
9. McClain, W. H. & Foss, K. (1988) *Science* **241,** 1804–1807.
10. Schulman, L. H. & Pelka, H. (1989) *Science* **246,** 1595–1597.
11. Tamura, K., Himeno, H., Asahara, H., Hasegawa, T. & Shimizu, M. (1992) *Nucleic Acids Res.* **20,** 2335–2339.
12. Sprinzl, M., Horn, C., Brown, M., Ioudovitch, A. & Steinberg, S. (1998) *Nucleic Acids Res.* **26,** 148–153.
13. Delagoutte, B., Moras, D. & Cavarelli, J. (2000) *EMBO J.* **19,** 5599–5610.
14. McClain, W. H. & Foss, K. (1988) *J. Mol. Biol.* **202,** 697–709.
15. Peterson, E. T. & Uhlenbeck, O. C. (1992) *Biochemistry* **31,** 10380–10389.
16. Hale, S. P., Auld, D. S., Schmidt, E. & Schimmel, P. (1997) *Science* **276,** 1250–1252.
17. Farrow, M. A., Nordin, B. E. & Schimmel, P. (1999) *Biochemistry* **38,** 16898–16903.
18. Goldgur, Y., Mosyak, L., Reshetnikova, L., Ankilova, V., Lavrik, O., Khodyreva, S. & Safro, M. (1997) *Structure (London)* **5,** 59–68.
19. Moor, N. A., Ankilova, V. N., Lavrik, O. I. & Favre, A. (2001) *Biochim. Biophys. Acta* **1518,** 226–236.
20. Baldwin, A. N. & Berg, P. (1966) *J. Biol. Chem.* **241,** 839–845.
21. Schimmel, P. & Schmidt, E. (1995) *Trends Biochem. Sci.* **20,** 1–2.
22. Nureki, O., Niimi, T., Muramatsu, T., Kanno, H., Kohno, T., Florentz, C., Giegé, R. & Yokoyama, S. f. (1994) *J. Mol. Biol.* **236,** 710–724.
23. Nureki, O., Tateno, M., Niimi, T., Kohno, T., Muramatsu, T., Kanno, H., Muto, Y., Giegé, R. & Yokoyama, S. (1991) *Nucleic Acids Symp. Ser.* **25,** 165–166.
24. Schmidt, E. & Schimmel, P. (1995) *Biochemistry* **34,** 11204–11210.
25. Nureki, O., Vassylyev, D. G., Tateno, M., Shimada, A., Nakama, T., Fukai, S., Konno, M., Hendrickson, T. L., Schimmel, P. & Yokoyama, S. (1998) *Science* **280,** 578–582.
26. Hendrickson, T. L., Nomanbhoy, T. K. & Schimmel, P. (2000) *Biochemistry* **39,** 8180–8186.
27. Silvian, L. F., Wang, J. & Steitz, T. A. (1999) *Science* **285,** 1074–1077.
28. Fersht, A. R. (1977) *Biochemistry* **16,** 1025–1030.
29. Nomanbhoy, T. K., Hendrickson, T. L. & Schimmel, P. (1999) *Mol. Cell* **4,** 519–528.