Pergamon

Tetrahedron, Vol. 53, No. 35, pp. 11985-11994, 1997

0040-4020/97 $17.00 + 0.00

PII: S0040-4020(97)00711-4

DNA Aptamer Targets Translational Editing Motif in a tRNA Synthetase

Stephen P. Hale and Paul Schimmel

Department of Biology

Massachusetts Institute of Technology

Cambridge, Massachusetts 02139

Abstract: Potential errors in translation occur when the wrong amino acid is activated by an aminoacyl tRNA synthetase to form a misactivated aminoacyladenylate. The misactivated amino acid can in the next step be attached to a tRNA that has an anticodon different than the ones corresponding to the amino acid. If the misacylated tRNA donates its amino acid to a growing polypeptide chain, then an error of translation occurs. However, certain tRNA synthetases have an editing activity that corrects errors of misactivation and of misacylation. The relationship between these two error-correcting activities in a synthetase has not been clear. We showed recently that an insertion (known as CP1) into the active site of a class I tRNA synthetases has a deacylase activity that hydrolytically removes mischarged amino acids that are attached to tRNAs. In other work, we showed that a specific DNA aptamer, selected from a random pool, could stimulate hydrolytic breakdown of a misactivated aminoacyladenylate bound to a tRNA synthetase. In this work, we photo-crosslinked the DNA aptamer to the tRNA synthetase. A single crosslinked peptide on the synthetase was identified. This peptide is located within the CP1 insertion, adjacent to residues known to affect the amino acid specificity of the tRNA deacylase activity. These results raise the possibility that the CP1 insertion has a role not only in correcting misacylations, but also in the hydrolytic breakdown of misactivated aminoacyladenylates. © 1997 Elsevier Science Ltd.

Introduction

The genetic code is determined at the biochemical level by the specific aminoacylations of transfer RNAs. These reactions are catalyzed by aminoacyl tRNA synthetases[1]. The tRNAs contain the anticodon triplets of the genetic code, so that the fidelity of the code depends on these aminoacylation reactions being highly specific. Because the genetic code is degenerate, more than one anticodon triplet codes for a particular amino acid and these distinct triplets are present in the different tRNA isoacceptors that accept a given amino acid. There is typically one synthetase (in prokaryotes or in the eukaryote cytoplasm) for each amino acid and that synthetase charges all of the tRNA isoacceptors for that amino acid.

The two step aminoacylation reaction catalyzed by tRNA synthetases is:

$$E + AA + ATP \leftrightarrow E(AA\text{-}AMP) + PPi \qquad (1)$$

$$E(AA\text{-}AMP) + tRNA \leftrightarrow AA\text{-}tRNA + AMP + E \qquad (2)$$

In the first step, the synthetase enzyme (E) catalyzes the formation of the tightly bound mixed phosphoanhydride aminoacyladenylate (AA-AMP) intermediate, with the liberation of pyrophosphate (PPi). In the second step, the activated aminoacyl group is transferred to the 3'-acceptor end of the tRNA to form the aminoacyl tRNA (AA-tRNA) together with the release of AMP.

For some tRNA synthetases, distinguishing the correct set of isoacceptors from all other tRNAs in reaction (2) is not as difficult as distinguishing the correct amino acid in reaction (1). The explanation is that the tRNA structure is highly differentiated and, while isoacceptors may have different anticodons, other parts of their structures are held in common and these parts are information rich. For example, the acceptor stem, a region near the amino acid attachment site and far from the anticodon, contains major determinants of aminoacylation efficiency and specificity[2,3]. Even small RNA oligonucleotides that recapitulate just the acceptor stem portion of the tRNA molecule are substrates for specific aminoacylations by many tRNA synthetases[4]. The nucleotides in the stem offer in principle enough variability to distinguish one set of tRNAs from all others.

However, the side chains of certain amino acids are intrinsically difficult to distinguish, at least at the level of binding interactions[5]. Two examples are isoleucine *versus* valine and valine *versus* threonine. Considering isoleucyl-tRNA synthetase, a hydrophobic pocket that holds the isobutyl group of isoleucine would also accommodate the isopropyl group of valine. The latter differs from isoleucine by a single methylene group. The van der Waals interactions contributed by this methylene group would perhaps amount to the order of 2-3 kcal mol^{-1}. As it turns out, isoleucyl-tRNA synthetase misactivates valine with a frequency of about 1/180[6,7]. This misactivation frequency corresponds to a free energy difference of 3.1 kcal mol^{-1}. Similarly, valyl-tRNA synthetase misactivates threonine, presumably because it is isosteric with valine[8]. The misactivation rates in these instances are much higher than seen in the overall fidelity of translation[9,10], so that error-correcting mechanisms are needed to deal with the misactivated complexes.

The editing activities of synthetases correct errors of amino acid activation (reaction (1)) and of aminoacylation (reaction (2)). In the case of isoleucyl-tRNA synthetase (IleRS), hydrolysis of misactivated Val-AMP is stimulated by the addition of $tRNA^{Ile}$ [6]. In addition, an esterase activity removes any valine from Val-$tRNA^{Ile}$ that forms [11]. These editing reactions are:

$$IleRS(Val\text{-}AMP) + tRNA^{Ile} \rightarrow IleRS + Val + tRNA^{Ile} + AMP \quad (3)$$

$$Val\text{-}tRNA^{Ile} + IleRS \rightarrow Val + tRNA^{Ile} + IleRS \quad (4)$$

In reaction (3), nucleotide determinants in $tRNA^{Ile}$ trigger hydrolytic breakdown of the mixed phosphoanhydride intermediate. In reaction (4), an esterase activity specifically distinguishes Val-$tRNA^{Ile}$ from Ile-$tRNA^{Ile}$. Thus, reactions (3) and (4) are examples of tRNA-dependent amino acid discrimination.

Even though three dimensional structures are now known for 11 tRNA synthetases [12-22], these structures have given no insight into the editing activities. The tRNA synthetases are divided into two classes [23], or groups, of ten enzymes each. The division into classes is based on the architecture of the active site domain. IleRS is a class I enzyme that is characterized by a $\beta_6\alpha_4$ nucleotide binding fold of alternating β-strands and alpha-helices. The active-site-containing fold is split into two $\beta_3\alpha_2$ halves by a 276 amino acid insertion known as CP1 (connective polypeptide 1) that joins together the two halves [24]. This insertion was cloned as an isolated protein and shown to have the esterase activity that specifically hydrolyzes valine from Val-$tRNA^{Ile}$ (reaction (4)) [25]. A similar insertion occurs in valyl-tRNA synthetase and this insertion was shown to catalyze deacylation of Thr-$tRNA^{Val}$. Thus, a specific structural motif in these class I enzymes distinguishes misacylated from correctly acylated tRNAs.

To investigate reaction (3) and to determine whether specific nucleotides *per se* and not an amino acid acceptor function were required, we selected for a DNA aptamer that could replace $tRNA^{Ile}$ in the hydrolytic breakdown of enzyme bound Val-AMP [26]. The rationale was that IleRS catalyzes initial attachment of the amino acid to the 2'-hydroxyl at the acceptor end of the tRNA. (After attachment, the aminoacyl group can migrate between 2'-and 3'-positions.) Because DNA lacks the 2'-OH, we reasoned that selection of an aptamer that stimulated hydrolytic breakdown of enzyme-bound Val-AMP

would show that a constellation of nucleotides and not an acceptor function was the important parameter for stimulating the hydrolytic editing reaction. We went through seven cycles of a specialized selection for aptamers that could replace $tRNA^{Ile}$, using a library with a complexity of about 10^{14}. An aptamer was obtained that, when added to IleRS(Val-AMP), induced hydrolysis of the misactivated adenylate, but had no effect on the stability of Ile-AMP bound to IleRS. The activity of the aptamer in this reaction was within about 2-fold of that of $tRNA^{Ile}$ [26].

Considering that we had identified the structural motif in IleRS (and ValRS) responsible for the deacylase activity of reaction (4), and that we had a non-substrate DNA ligand that could replace $tRNA^{Ile}$ in reaction (3), we were motivated to use the DNA aptamer to covalently label the translational editing site on IleRS for reaction (3). The question was whether the region on IleRS responsible for the pre-transfer editing of reaction (3) was distinct from that required for the deacylase activity of reaction (4). Because the aptamer is not a substrate but only a ligand for triggering the editing response, its use as a covalent label offered some clear advantages over that of the natural ligand ($tRNA^{Ile}$) which is also a substrate and would therefore label the catalytic site and not necessarily the site for editing.

Results and Discussion

The 61 nucleotide DNA aptamer used in these studies was selected from a library created by variation of the sequence of a 25 nucleotide segment centered between two fixed 18 nucleotide elements that were used as binding sites for PCR primers. (PCR was used to amplify specific sequences that were selected.) The presence of secondary structure in the selected active aptamer (DNA^{A}) was established by the cooperative melting as followed by uv absorbance with a Tm of 54 °C. The melting profile was independent of concentration of over the range of 0.2 to 2 μM [26]. A computer-predicted secondary structure of the aptamer showed no resemblance to the cloverleaf structure of a tRNA (Figure 1). This aptamer bound to free IleRS with a dissociation constant of 1.5 μM [26].

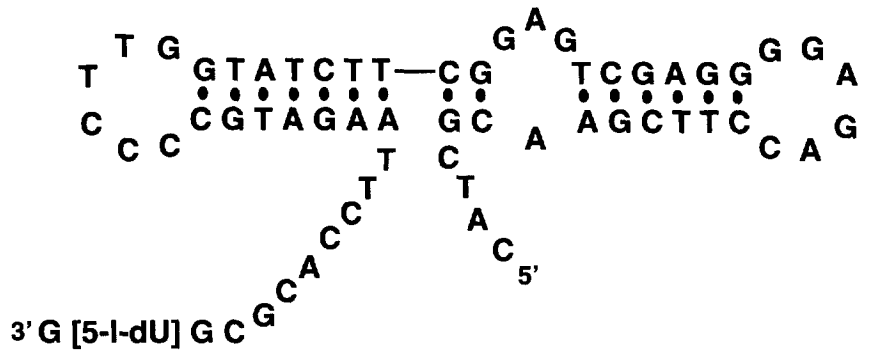

Fig. 1 Nucleotide sequence and predicted secondary structure of 5-I-DNA^A that triggers the specific hydrolysis of Val-AMP bound to IleRS (reaction (3)). The original aptamer DNA^A was modified by substituting the photoactive 5-iodo-U for T at the penultimate position of the 3'-end.

The aptamer was modified to place a photoactive nucleotide at one end (position 60) to give 5-I-DNA^A (Figure 1) [27]. In this location, the photolabel is incorporated into one of the PCR primer binding sites that is fixed in sequence. Separately, we showed that the central 25 nucleotide sequence that was varied during the selection was not, in isolation, able to trigger the editing response (data not shown). Thus, regions that include the PCR primer binding site are required for activity. This requirement for sequences outside of the varied region has also been made in the selection of ribozymes for aminoacylation [28], and is perhaps an inevitable consequence of aptamers being selected in the context of particular flanking nucleotide sequences.

To follow the labeling of IleRS, 5-I-DNA^A was radiolabeled at the 5'-end with ^{32}P. Using a 1:1 ratio of ^{32}P-5-I-DNA^A to IleRS, irradiation was carried out at ambient temperature (~23 ^{o}C) for one hour. The extent of crosslinking was measured by analyzing aliquots of the reaction mixture by polyacrylamide gel electrophoresis under denaturing conditions, where the labeled enzyme-DNA complex could be directly measured by

phosphorimaging. Following one hour of irradiation, 2 % of the enzyme was labeled with ^{32}P-5-DNA^{A}.

The crosslinked complex was subjected to digestion with trypsin under denaturing conditions and subsequently resolved by C18-reverse phase HPLC. A single radiolabeled peak was detected. This peak was subjected to ten cycles of peptide sequencing by the MIT Biopolymers Laboratory. The sequence obtained—^{409}HKTPIIFRAT^{418}....— is located at the C-terminal end of the CP1 insertion (Figure 2). The sequence of the crosslinked peptide is preceded R^{408}, a result which is consistent with the specificity of trypsin for cleavage after arginine and lysine. On the other hand, the peptide contains an internal lysine (K^{410}) and arginine (R^{416}), apparently because of incomplete digestion of the crosslinked complex.

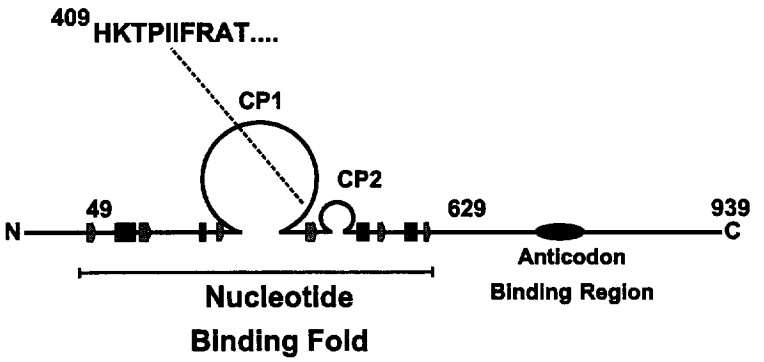

Fig. 2 Schematic illustration of the design of the structure of the 939 amino acid *E. coli* IleRS and the location of the peptide that was covalently labeled by ^{32}P-5-I-DNA^{A}. The active-site-containing nucleotide binding fold consists of alternating β strands (arrows) and α-helices (rectangles). The fold is split by the CP1 insertion. The crosslinked peptide is located at the C-terminal end of the CP1 insertion. The sequence of the first 10 amino acids of the peptide is given.

Most significantly, the crosslinked peptide is located within the CP1 insertion which encodes the deacylase activity of reaction (4) that removes valine from Val-tRNA^{Ile} [25]. Thus, the results of the crosslinking experiment are consistent with CP1 being involved in both editing reactions (3) and (4). Moreover, the crosslinked peptide is adjacent to residues shown by mutagenesis to affect the specificity of reaction (4) [29]. For example, the wild-type enzyme preferentially deacylates Val-tRNA^{Ile} by a factor of about 32-fold over that of Ile-tRNA^{Ile}. Mutation of either H401, Y403, or R408 can greatly enhance (to over 1000-fold) or diminish (to less than 8-fold) the specificity ratio. Thus, the C-terminal side of CP1 may be part of the active site for both editing activities.

Each of these editing activities cleave off the aminoacyl moiety from an acceptor group. In the crystal structure of the class I glutaminyl-tRNA synthetase in complex with tRNA^{Gln}, the CP1 insertion (110 amino acids) extends out from the body of the enzyme and makes contacts with the acceptor helix of tRNA^{Gln}, not far from the amino acid attachment site [13]. Although glutaminyl-tRNA synthetase has no known editing activity, the location of CP1 in the structure and its much larger size in IleRS and ValRS suggest that CP1 should able to interact with the aminoacyl moiety of either the adenylate or the charged tRNA that is bound to the isoleucine or the valine enzyme.

Conclusion

A DNA aptamer was successfully used as a crosslinking agent to implicate a region on a class I tRNA synthetase that is important for hydrolytic breakdown of a misactivated aminoacyladenylate. The region identified is close in the sequence to residues known from other work to be important for a second editing reaction—the hydrolysis of the aminoacyl group from a misacylated tRNA. These results suggest that the same parts of the structure of a synthetase are used for catalysis of hydrolysis of an aminoacyl group from either of two acceptor functions (adenosine monophosphate or the 3'-end of a tRNA). This region lies within an insertion that splits the active site of this class I tRNA synthetase.

Experimental

Isoleucine tRNA Synthetase. Isoleucine tRNA synthetase was isolated as described[30] from *Escherichia coli* cells harboring a multicopy plasmid containing the gene for *E. coli* isoleucyl-tRNA synthetase.

DNA^A Preparation. 5-iodouracil was incorporated into the 3' end of DNA^A during synthesis (Pharmacia LKB Gene Assembler Special, Piscataway, NJ; 5-I-dU phoshoramidite, Glen Research, Sterling, VA) at position T60. All manipulations were done in the dark. 5-I-dU-DNA^A was radiolabeled at the 5' end with ^{32}P-ATP (NEN, Boston, MA) using T4 polynucleotide kinase (New England Biolabs, Beverly, MA). The radiolabeled 5-I-dU-DNA was purified by Sep-Pac C_{18} (Waters, Milford, MA) reverse phase chromatography and concentrated by Speed Vac lyophilization (Savant, Farmingdale, NY). The purified [^{32}P]-5-I-dU-DNA^A was dissolved in 20 mM HEPES (pH 7.5), 150 mM NH_4Cl, 1 mM $MgCl_2$ and denatured at 90 °C for 5 minutes. The sample was refolded by cooling to 25 °C over 10 min.

DNA^A Secondary Structure Prediction. The secondary structure presented for DNA^A was determined using the program RNAdraw (Institute of Medical Phisics, Karolinska Institute, Stockholm, Sweeden). DNA bases were substituted with the corresponding RNA bases and the output structure modified to be consistent with a DNA structure.

5-I-dU-DNA^A to IleRS Crosslinking. IleRS and [^{32}P]-5-I-DNA^A were combined at a molar ratio of 1:1 (10 mM each) in 20 mM HEPES (pH 7.5), 150 mM NH_4Cl, 1 mM $MgCl_2$, 1mM β–mercaptoethanol and incubated at ambient temperature for 10 min. Crosslinking of the [^{32}P]-5-I-DNA^A complex was performed at ambient by photoactivation of the incorporated 5-I-dU of DNA^A. Irradiation was performed for 1 hour in a Rayonet Photochemical Reactor (Hamden, CT) fitted with a 350 nm light source and a <300 nm cutoff filter. A one hour time course analysis of a fraction of the reaction products by polyacrylamide gel electrophoresis[31], and subsequent quantitation using a Molecular Dynamics Phosphorimager SI, determined that the crosslinking reaction was complete (with 2% of IleRS being labeled) at 10 minutes. The remainder of the reaction products were resolved by 10% preparative polyacrylamide gel electrophoresis under

denaturing conditions and the band corresponding to the crosslinked IleRS-DNA^A complex was excised from the gel and eluted passively into water.

IleRS-DNA^A tryptic digestion and peptide purification. The IleRS-DNA^A product was subjected to trypsin (modified sequencing grade trypsin, Boehringer Mannheim, Indianapolis, IN) digestion (1:20 w/w trypsin to IleRS, 3 M urea, 100 mM $CaCl_2$) under denaturing conditions for 4 hours at 37 °C. The digestion products were resolved by C_{18}-reverse phase HPLC (Alltima C_{18}, Alltech, Deerfield, IL) with a linear gradient of acetonitrile in 0.1% TFA. The peptide fraction containing the radiolabel was concentrated by Speed Vac lyophilization and submitted for 10 cycles of sequence analysis (MIT Biopolymers Laboratory, Cambridge, MA) on an Applied Biosystems Model 477A Protein Sequencer with on line Model 120 PTH Amino Acid Analyzer.

References

1. Carter, C. W. Jr. *Ann. Rev. Biochem.* **1993**, *62*, 715-748.
2. Schimmel, P.; Giegé, R.; Moras, D.; Yokoyama, S. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 8763-8768.
3. Giegé, R.; Puglisi, J. D.; Florentz, C. *Prog. Nucleic Acids Res. Mol. Biol.* **1993**, *45*, 129-206.
4. Martinis, S. A.; Schimmel, P. *tRNA: Structure, Biosynthesis, and Function*; Soll, D. ; RajBhandary, U., Eds.; American Society for Microbiology: Washington, D. C., 1995, pp 349-370.
5. Fersht, A. *Enzyme Structure and Mechanism*; 2 ed.; W. H. Freeman and Company: New York, 1985; pp 347-368.
6. Norris, A. T.; Berg, P. *Proc. Natl. Acad. Sci. USA* **1964**, *52*, 330-337.
7. Schmidt, E.; Schimmel, P. *Science* **1994**, *264*, 265-267.
8. Fersht, A. R.; Dingwall, C. *Biochemistry* **1979**, *18*, 2627-2631.
9. Loftfield, R. B. *Biochem. J.* **1963**, *89*, 82-92.
10. Jakubowski, H.; Goldman, E. *Microbiol. Rev.* **1992**, *56*, 412-429.
11. Eldred, E. W.; Schimmel, P. R. *J. Biol. Chem.* **1972**, *247*, 2961-2964.
12. Crick, P.; Bhat, T. N.; Blow, D. M. *J. Mol. Biol.* **1988**, *208*, 83-98.

13. Rould, M. A.; Perona, J. J.; Söll, D.; Steitz, T. A. *Science* **1989**, *246*, 1135-1142.
14. Brunie, S.; Zelwer, C.; Risler, J.-L. *J. Mol. Biol.* **1990**, *215*, 1-14.
15. Cusack, S.; Berthet-Colominas, C.; Hartlein, M.; Nassar, N.; Leberman, R. *Nature* **1990**, *347*, 249-255.
16. Ruff, M.; Krishnaswamy, S.; Boeglin, M.; Poterszman, A.; Mitschler, A.; Podjarny, A.; Rees, B.; Thierry, J. C.; Moras, D. *Science* **1991**, *252*, 1682-1689.
17. Arnez, J. G.; Harris, D. C.; Mitschler, A.; Rees, B.; Francklyn, C. S.; Moras, D. *EMBO J.* **1995**, *14*, 4143-4155.
18. Doublié, S.; Bricogne, G.; Gilmore, C.; Carter, C. W., Jr. *Structure* **1995**, *3*, 17-31.
19. Logan, D. T.; Mazauric, W.-H.; Kern, D.; Moras, D. *EMBO J.* **1995**, *14*, 4156-4167.
20. Mosyak, L.; Reshetnikova, L.; Goldgur, Y.; Delarue, M. *Nat. Struct. Biol.* **1995**, *2*, 537-547.
21. Nureki, O.; Vassylyev, D. G.; Katayanaga, K.; Shimizu, T.; Sekine, S.-i.; Kigawa, T.; Miyazawa, T.; Yokoyama, S.; Morikawa, K. *Science* **1995**, *267*, 1958-1965.
22. Onesti, S.; Miller, A. D.; Brick, P. *Structure* **1995**, *3*, 163-176.
23. Eriani, G.; Delarue, M.; Poch, O.; Gangloff, J.; Moras, D. *Nature* **1990**, *347*, 203-206.
24. Starzyk, R. M.; Webster, T. A.; Schimmel, P. *Science* **1987**, *237*, 1614-1618.
25. Lin, L.; Hale, S. P.; Schimmel, P. *Nature* **1996**, *384*, 33-34.
26. Hale, S. P.; Schimmel, P. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 2755-2758.
27. Willis, M. C.; Hicke, B. J.; Uhlenbeck, O. C.; Cech, T. R.; Koch, T. H. *Science* **1993**, *262*, 1255-1257.
28. Illangasekare, M.; Sanchez, G.; Nickles, T.; Yarus, M. *Science* **1995**, *267*, 643-647.
29. Schmidt, E.; Schimmel, P. *Biochemistry* **1995**, *34*, 11204-11210.
30. Shepard, A.; Shiba, K.; Schimmel, P. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 9964-9968.
31. Maniatis, T.; Fritsch, E. F.; Sambrook, J. *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1982, pp 11.23-11.28.

(*Received* 12 *December* 1996; *accepted* 10 *February* 1997)