Crystal Structure of tRNA Adenosine Deaminase (TadA) from *Aquifex aeolicus*  

Received for publication, December 27, 2004, and in revised form, January 18, 2005  

Published, JBC Papers in Press, January 26, 2005, DOI 10.1074/jbc.M414541200

Mitsuo Kuratani‡, Ryohie Ishii‡, Yoshitaka Bessho§, Ryuya Fukunaga‡, Toru Sengoku‡, Mikako Shirouzu§, Shun-ichi Sekine‡‡, and Shigeyuki Yokoyama§§

From the ‡Department of Biophysics and Biochemistry, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan, §RIKEN Genomic Sciences Center, 1-7-22 Suehiro-cho, Tsukumi, Yohokama 230-0045, Japan, and †RIKEN Harima Institute at SPring-8, 1-1-1 Kouto, Mikazuki-cho, Sayo-gun, Hyogo 679-5148, Japan

The bacterial tRNA adenosine deaminase (TadA) generates inosine by deaminating the adenosine residue at the wobble position of tRNA^Arg^2. This modification is essential for the decoding system. In this study, we determined the crystal structure of *Aquifex aeolicus* TadA at a 1.8-Å resolution. This is the first structure of a deaminase acting on tRNA. *A. aeolicus* TadA has an α/β/α three-layered fold and forms a homodimer. The *A. aeolicus* TadA dimeric structure is completely different from the tetrameric structure of yeast CDD1, which deaminates mRNA and cytidine, but is similar to the dimeric structure of yeast cytosine deaminase. However, in the *A. aeolicus* TadA structure, the shapes of the C-terminal helix and the regions between the β4 and β5 strands are quite distinct from those of yeast cytosine deaminase and a large cavity is produced. This cavity contains many conserved amino acid residues that are likely to be involved in either catalysis or tRNA binding. We made a docking model of TadA with the tRNA anticodon stem loop.

Inosine (Fig. 1A) was found at the first (“wobble”) position of the tRNA anticodon 40 years ago (1). Crick postulated (2) that inosine is able to pair with U, C, and A. The codons corresponding to each inosine-bearing tRNA are synonymous, which contributes to decreasing the number of isoacceptor tRNAs (3). In eukaryotes, the wobble positions of eight cytoplasmic tRNAs (seven in *Saccharomyces cerevisiae*) bear inosine (4), which is generated by the posttranscriptional hydrolytic deamination of adenosine (5). In most bacteria and plant chloroplasts, only tRNA^Arg^2 (Fig. 1B) has the inosine modification (4).

The enzymes that catalyze inosine generation were cloned recently (6, 7). In eukaryotes, a heterodimeric enzyme comprises one tRNAArg-2 (Fig. 1) has the inosine modification (4). In most bacteria and plant chloroplasts, only tRNA^Arg^2 (Fig. 1B) has the inosine modification (4).

The enzymes that catalyze inosine generation were cloned recently (6, 7). In eukaryotes, a heterodimeric enzyme comprises one tRNAArg-2 (Fig. 1) has the inosine modification (4). In most bacteria and plant chloroplasts, only tRNA^Arg^2 (Fig. 1B) has the inosine modification (4).

EXPERIMENTAL PROCEDURES

**Protein Preparation**—The *A. aeolicus* tadA gene was PCR-amplified from the genomic DNA and subcloned into pET28b (Novagen) between the NdeI and SalI sites. The recombinant protein consists of the 151 amino acid residues from *A. aeolicus* TadA and 20 additional vector-encoded His tag residues (MGSSHHHHHHESGLVPRGSH) at the N terminus for affinity purification. *Escherichia coli* strain BL21(DE3) CodonPlus (Strategene) was transformed with the plasmid. For protein expression, the cells were grown in LB medium at 37 °C to an A_600 of 0.6 and then the expression was induced with 1 mm isopropyl-β-D-thiogalactopyranoside for 3 h. The cells were harvested and sonicated in 20 mM Tris-HCl buffer (pH 8.5) containing 500 mM NaCl, 10 mM imidazole, 1.4 mM 2-mercaptoethanol, and a protease inhibitor mixture, Complete EDTA-free (Roche Applied Science). The insoluble cell debris was removed by centrifugation at 15,000 × g for 10 min at 4 °C. The supernatant was heated to 72 °C for 20 min to denature the E. coli proteins. The heat-treated mixture was centrifuged at 15,000 × g for 15 min at 4 °C. The supernatant was applied to a 10-ml column of nickel-nitrilotriacetic acid Superflow (Qiagen) equilibrated with 20 mM Tris-HCl buffer (pH 8.5) containing 500 mM NaCl, 10 mM imidazole, and 1.4 mM 2-mercaptoethanol. The protein was eluted in one step with 20 mM HEPES-NaOH buffer (pH 7.5) containing 300 mM NaCl, 250 mM imidazole, and 1.4 mM 2-mercaptoethanol. Four volumes of 20 mM HEPES-

*This work was supported by a grant-in-aid for Scientific Research S from the Japan Society for Promotion of Science (JSPS), a grant-in-aid for Scientific Research in Priority Areas from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan, the RIKEN Structural Genomics/Proteomics Initiative (RSGI), and the National Project on Protein Structural and Functional Analyses, MEXT. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (code 1JWW) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

† To whom correspondence should be addressed. Tel.: 81-3-5841-4392; Fax: 81-3-5841-8057; E-mail: yokoyama@biochem.s.u-tokyo.ac.jp.

§ The abbreviations used are: ADAT, adenosine deaminase acting on tRNA; TadA, tRNA adenosine deaminase; CDA, cytidine deaminase; CD, cytosine deaminase; GD, guanine deaminase; MES, 4-morpholinoethanesulfonic acid.
NaOH buffer (pH 7.5) containing 1.4 mM 2-mercaptoethanol were added to the eluate to reduce the NaCl concentration to less than 100 mM, and then this solution was loaded onto a UnoS column (Bio-Rad) using an AKTA system (Amersham Biosciences). The protein was eluted with a linear gradient of 0.1–0.5 M NaCl in 20 mM HEPES-NaOH buffer (pH 7.5) containing 1.4 mM 2-mercaptoethanol. Two peaks appeared in the chromatogram. Each peak fraction was individually diluted and loaded onto the UnoS column again. Both elution patterns were the same as the first elution. Two peaks appeared at the corresponding salt concentrations. The two-peak fractions were mixed and dialyzed against 10 mM HEPES-NaOH buffer (pH 7.5) containing 400 mM NaCl and 5 mM 2-mercaptoethanol. The final purity of the protein was >95% as monitored by FIG. 2.
amino acid residues (Met15-Val108 of yeast CD and Leu7-Ile98 of TadA) share 31% sequence identity (Fig. 2). We carried out structure refinement and building of the model using the program O (23), using the data set up to a 2.5-Å resolution. After several rounds of refinement was performed in the same way using the data set from (5% data set).

RESULTS AND DISCUSSION

Structure Determination—We determined the crystal structure of A. aeolicus TadA at a 1.8-Å resolution by the molecular replacement method. The crystal contains four molecules A, B, C, and D, per asymmetric unit (Fig. 3, A, B, C, and D respectively). The monomer structure (Fig. 3A) consists of a central β-sheet (β1-β5) with two α-helices (α1, α5) on one side of the sheet and three

TABLE I
Data collection and model-refinement statistics

| Data collection | Model-refinement statistics |
|-----------------|-----------------------------|
| Space group | $P_{2_1}$ |
| Unit-cell parameters | $a = 43.2$ Å, $b = 152.0$ Å, $c = 54.1$ Å, $β = 113.4^\circ$ |
| Resolution range (Å) | 50.0–1.8 |
| Unique reflections | 56,975 |
| Redundancy (last shell) | 3.1(1.7) |
| Completeness (%) (last shell) | 96.5(77.7) |
| $I/σ$ (last shell) | 19.1(2.0) |
| $R_{\text{free}}$ (%) (last shell)$^a$ | 7.7(17.6) |
| No. of reflections: working set/test set | 54,092/2,883 |
| No. of protein atoms | 4,972 |
| No. of water molecules | 520 |
| No. of ion atoms | 4 |
| $R_{\text{factor}}$: working set/test set$^b$ | 19.8/24.8 |
| R.m.s. bonds (Å) | 0.007 |
| R.m.s. angles (°) | 1.3 |

$^a$ $R_{\text{free}} = \frac{\sum_j |I_j| - \langle j(hkl)\rangle|\sum_{j} \sum_{hkl} |I_j(hkl)|, \text{where } I_j(hkl) \text{ and } \langle j(hkl)\rangle} {\sum_{j} \sum_{hkl} |I_j(hkl)|}$, respectively.

$^b$ $R_{\text{factor}} = \frac{\sum |F_{o,b} - kF_{c,b}|/\sum F_{o,b}^2}$, where $b$ is a scale factor and $R_{\text{free}}$ is the $R_{\text{factor}}$ for the test set of reflections not used during refinement (5% data set).
α-helices (α2–α4) on the other side. A long loop exists between the β4 and β5 strands and is designated as the β4–β5 loop (residues Lys104–Arg124).

Zinc Ion—In the Fσ – Fc Fourier map contoured at a 9-σ level, each of the monomers contained one strong spherical density, which was assigned as a zinc ion based on the x-ray absorption fine structure data (Fig. 4). Because no zinc ion was added in the buffer during the protein purification and crystallization, A. aeolicus TadA contains endogenous zinc ions. The zinc ion is tetrahedrally coordinated (Fig. 3, A and C) by His52 Nε1 (2.1 Å), Cys82 Sγ (2.3 Å), Cys85 Sγ (2.3 Å), and a water molecule, Water O (2.3 Å). Glu54 Oγ interacts with the zinc-bound water (2.5 Å) (Fig. 3, A and C). These residues are conserved among the CDA superfamily members (Fig. 2). The active-site architectures are also similar within the CDA superfamily (8–12), and the deamination mechanisms may be the same. The zinc ion and Glu54 are proposed to activate the zinc-bound water to form a hydroxide ion. Glu54 may shuttle a proton from the water to the adenosine residue. The zinc-bound hydroxide ion is proposed to attack the C6 atom of the adenosine residue (Fig. 1, A) nucleophilically.

Dimerization State of TadA—In the crystal, molecules A-B and molecules C-D form two apparent dimers along the non-crystallographic 2-fold axes (Fig. 3B). Each dimer is almost spherical with the exception of the C-terminal protrusions and the zinc-containing cavities (Fig. 3D). The dimerization interface is mainly composed of helices α3 and α4 and the β4–β5 loop. The interface is extensive and buries 1300 Å² of the total monomer surface area of 8100 Å². Thirty residues are involved in the dimerization including eight conserved hydrophobic residues (Met55, Ile58, Met84, Ala88, Val112, Phe113, Ile115, and Leu121). In addition, four intersubunit salt bridges are formed between Glu44 O and Lys58 Nε (2.5 Å) and between Asp48 Oγ and Lys59 Nε (2.7 Å). When Asp64 in E. coli TadA, which corresponds to Asp48 in A. aeolicus TadA, was replaced by Glu, the mutant enzyme was fully active in vivo but lost its activity in vitro (7, 25). This residue may contribute to the structural stabilization of the protein, either by itself or through interaction with other protein(s).

The A-B and C-D dimers touch (Fig. 3B) through the A-C, A-D, and B-D contacts of 140, 730, and 150 Å², respectively, of the 8100-Å² monomer surface area. Six residues of the C-terminal helix α5 of molecule A interact with nine residues on the wall of the zinc-containing cavity of molecule D and vice versa. An E. coli TadA mutant lacking the C-terminal 17 res-
TadA and the helix α5 in yeast CD). In addition, the C-terminal motifs of the monomers have similar structures (Fig. 6). First, the β4 and β5 strands run parallel to each other. Second, the N and C termini are located on the same side of the αβ/α β three-layered fold. Third, both of the C-terminal helices (α5 in A. aeolicus TadA and α6 in yeast CD) contact the α1 N-terminal helices.

B. subtilis GD forms an intertwined dimer through C-terminal domain swapping (Fig. 5C). The dimer interface is composed of the helices α3 and α4 and the C-terminal domain. Helices α5 and α6, which are located between the β4 and β5 strands, make the domain swapping possible and have important roles in guanine recognition (11). As a consequence of the swap, the C-terminal motif of one subunit interacts with the other subunit in a manner similar to that of A. aeolicus TadA and yeast CD (Fig. 6C).

Yeast CDD1 and B. subtilis CDA form tetrameric structures. One subunit of yeast CDD1 (Fig. 5D) interacts with the other three subunits via the α2–α6 helices. B. subtilis CDA also has a tetrameric structure similar to that of yeast CDD1. The tetramers are not formed by the dimerization of two TadA-like dimers. Therefore, the quaternary structure of yeast CDD1 (Fig. 5D) completely differs from that of A. aeolicus TadA (Fig. 5A). Also, the C-terminal motifs of B. subtilis CDA (Fig. 6D) and yeast CDD1 (Fig. 6E) share no structural similarities with that of A. aeolicus TadA (Fig. 6A). First, the β4 and β5 strands run antiparallel to each other and only short turns exist between the β4 and β5 strands. Second, the N and C termini are located on the opposite sides of the αβ/α β three-layered fold.

The oligomerization state does not determine whether nucleic acids are accommodated. Based on the structure of yeast CDD1, a model of human activation-induced cytidine deaminase has been made (12) on the assumption that its quaternary structure is similar to a tetramer. On the other hand, an analysis of the human activation-induced cytidine deaminase sequence with the 3D Jury method suggested that the dimeric yeast CD was the best template (27). One problem with using yeast CD as a template was that the intramolecular active site was too small to accommodate large nucleic acid molecules (12). However, the structure of A. aeolicus TadA indicates that activation-induced cytidine deaminase might form a TadA-like dimer for nucleic acid binding.

The RNA-binding Site of A. aeolicus TadA—We next examined the substrate-binding site differences between A. aeolicus TadA and yeast CD to determine which motifs contribute to the specific binding of their substrates, tRNA and cytosine, respectively. The active-site cavity of A. aeolicus TadA (Fig. 5A) is composed of the α2 and α5 helices, and the zinc ion from one subunit, the β4–β5 loops from both subunits, and loop 1 (between α2 and β3) and loop 2 (between α3 and β4) from the other subunit. Both of the β4–β5 loops of the A. aeolicus TadA dimer are extended and cooperate with each other to form the putative RNA-binding site (Fig. 5A). On the other hand, the corresponding regions (the α5 helices) in the yeast CD dimer are involved in the intersubunit interaction but are far from the cytosine-binding site (Fig. 5B). The C-terminal α5 helix of A. aeolicus TadA protrudes outward, whereas the corresponding C-terminal α7 helix of yeast CD bends inward, which makes the cavity narrower and suitable for a small cytosine base (10). In summary, the diversified regions between the β4 and β5 strands and the C-terminal helices have important roles in the specific recognition of the substrates by A. aeolicus TadA and yeast CD.

The Recognition of tRNAArg^2—The anticodon stem loop structure (Fig. 1B) is reportedly sufficient for E. coli TadA to deaminate the adenosine residue at the wobble position (posi-
tRNA\textsubscript{Phe} (Protein Data Bank code 1EHZ). We docked them and the anticodon stem loop moiety (bases 26–43) of yeast tRNA\textsubscript{Phe}. The ribbon model of the anticodon stem-loop is colored yellow. TadA is represented as a surface model. The color scheme in B is the same as in A. The surface of A. aeolicus TadA in C is color-coded according to its electrostatic potential (red, −10 kT/e; blue, +10 kT/e). The molecular surface was produced using the program MSMS.

**Fig. 7.** The putative tRNA\textsubscript{Ado}\textsubscript{2}-binding site of A. aeolicus TadA. A, stereo view of the conserved residues in the cavity of A. aeolicus TadA. One subunit is colored pink and the other is colored cyan. B and C, preliminary docking model of TadA and the anticodon stem-loop moiety (bases 26–43) of yeast tRNA\textsubscript{Phe}. The ribon model of the anticodon stem-loop is colored yellow. TadA is represented as a surface model. The color scheme in B is the same as in A. The surface of A. aeolicus TadA in C is color-coded according to its electrostatic potential (red, −10 kT/e; blue, +10 kT/e). The molecular surface was produced using the program MSMS.
Crystal Structure of tRNA Adenosine Deaminase TadA

7. Wolf, J., Gerber, A. P., and Keller, W. (2002) EMBO J. 21, 3841–3851
8. Betts, L., Xiang, S., Short, S. A., Wofenden, R., and Carter, C. W., Jr. (1994) J. Mol. Biol. 235, 653–656
9. Johansson, E., Mejlhede, N., Neuhard, J., and Larsen, S. (2002) Biochemistry 41, 2563–2570
10. Ko, T. P., Lin, J. J., Hu, C. Y., Hsu, Y. H., Wang, A. H., and Liaw, S. H. (2003) J. Biol. Chem. 278, 19111–19117
11. Liaw, S. H., Chang, Y. J., Lai, C. T., Chang, H. C., and Chang, G. G. (2004) J. Biol. Chem. 279, 35479–35485
12. Xie, K., Sowden, M. P., Dance, G. S., Torelli, A. T., Smith, H. C., and Wedekind, J. E. (2004) Proc. Natl. Acad. Sci. U. S. A. 101, 8114–8119
13. Teng, B., Burant, C. F., and Davidson, N. O. (1993) Science 260, 1816–1819
14. Navaratnam, N., Morrison, J. R., Bhattacharya, S., Patel, D., Funahashi, T., Giannoni, F., Teng, B. B., Davidson, N. O., and Scott, J. (1993) J. Biol. Chem. 268, 20709–20712
15. Muramatsu, M., Sankaranand, V. S., Anant, S., Sugai, M., Kinoshita, K., Davidson, N. O., and Honjo, T. (1999) J. Biol. Chem. 274, 18470–18476
16. Jarmuz, A., Chester, A., Bayliss, J., Gisbourne, J., Dunham, I., Scott, J., and Navaratnam, N. (2002) Genomics 79, 285–296
17. Dance, G. S., Beemiller, P., Yang, Y., Mater, D. V., Mian, I. S., and Smith, H. C. (2001) Nucleic Acids Res. 29, 1772–1780
18. Navaratnam, N., Fujino, T., Bayliss, J., Jarmuz, A., How, A., Richardson, N., Somasekaram, A., Bhattacharya, S., Carter, C., and Scott, J. (1998) J. Mol. Biol. 275, 695–714
19. Otwinowski, Z., and Minor, W. (1997) Methods Enzymol. 276, 307–326
20. Vagin, A., and Teplyakov, A. (1997) J. Appl. Crystallogr. 30, 1022–1025
21. Collaborative Computational Project, Number 4. (1994) Acta Crystallogr. Sect. D Biol. Crystallogr. 50, 760–763
22. Adams, P. D., Pannu, N. S., Read, R. J., and Brunger, A. T. (1997) Proc. Natl. Acad. Sci. U. S. A. 94, 5018–5023
23. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) Acta Crystallogr. Sect. A 47, 110–119
24. Laskowski, R. A., MacArthur, M. W., Moris, A. L., and Thornton, J. M. (1993) J. Appl. Crystallogr. 26, 283–291
25. Poulsen, L. K., Larsen, N. W., Melin, S., and Anderson, P. (1992) Mol. Microbiol. 6, 895–905
26. Holm, L., and Sander, C. (1998) Nucleic Acids Res. 26, 316–319
27. Zaim, J., and Kierzek, A. M. (2003) Nat. Immunol. 4, 1153–1154
28. Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., and Higgins, D. G. (1997) Nucleic Acids Res. 25, 4876–4882
29. Gouet, P., Courcelle, E., Stuart, D. I., and Metoz, F. (1999) Bioinformatics 15, 395–398
30. Gouet, P., Robert, X., and Courcelle, E. (2003) Nucleic Acids Res. 31, 3320–3323