Crystal Structure of TruD, a Novel Pseudouridine Synthase with a New Protein Fold

Received for publication, February 16, 2004, and in revised form, March 1, 2004
Published, JBC Papers in Press, March 3, 2004, DOI 10.1074/jbc.C400072200

Yusuf Kaya‡, Mark Del Campo‡, James Ofengand, and Arun Malhotra§

From the Department of Biochemistry and Molecular Biology, University of Miami School of Medicine, Miami, Florida 33101-6129

TruD, a recently discovered novel pseudouridine synthase in *Escherichia coli*, is responsible for modifying uridine13 in tRNA13 to pseudouridine. It has little sequence homology with the other 10 pseudouridine synthases in *E. coli* which themselves have been grouped into four related protein families. Crystal structure determination of TruD revealed a two domain structure consisting of a catalytic domain that differs in sequence but is structurally very similar to the catalytic domain of other pseudouridine synthases and a second large domain (149 amino acids, 43% of total) with a novel α/β fold that up to now has not been found in any other protein.

The most abundant modification seen in structured RNAs (transfer, ribosomal and splicing RNAs) is the isomerization of uridine (U) to pseudouridine (5-ribofuransyl, Ψ, Fig. 1A). Ψ is made by a set of enzymes called Ψ synthases, which select specific U residues in a polynucleotide chain for isomerization to Ψ. Ψ synthases are ubiquitous as putative synthase genes have been found in all genomes so far sequenced (1). In *Escherichia coli*, the Ψ formation activity of 10 putative gene products has been experimentally demonstrated (see references cited in Table II of Ref. 2). These Ψ synthases are related by a set of amino acid sequence motifs and can be grouped into four families based on the degree of amino acid sequence homology (3–5). A fifth family has recently been described whose members have no apparent sequence homology with the others. It was discovered by identifying the *E. coli* member of this family as an enzyme catalyzing the formation of Ψ13 in tRNA13 (Fig. 1B) and was named TruD (6). Fifty-eight homologs of TruD, distributed among the genomes of Eubacteria, Archaea, and Eukarya were identified (6). The function of the homologs is not known except for Pus7p (YOR243c) of *Saccharomyces cerevisiae*, which makes Ψ13 in tRNA as well as Ψ5 in U2 snRNA (7). Because all of the organisms known to have Ψ13 in their tRNAs also have a TruD homolog, it is reasonable to infer that TruD homologs in those organisms with tRNA Ψ13 are the responsible synthases. This need not, however, rule out synthase specificity for other sites in homologs which lack tRNA Ψ13.

Crystal structures of a member from each of the original four families of Ψ synthases have revealed a striking correspondence of both overall structure and of certain conserved amino acids in the catalytic module, including an aspartate residue at the active site (8–13). This raises the question of how the catalytic center of TruD is organized, since there is virtually no amino acid sequence homology, despite the presence of an essential aspartate residue (6). In this work, we report the crystal structure of TruD to a resolution of 2.2 angstroms. TruD folds into a V-shaped molecule with two distinct modules. One (57% of the total amino acids) closely resembles the catalytic module of the previously determined Ψ synthases despite the lack of amino acid sequence similarity. The other module (43%) lacks any similarity in either sequence or structure to known protein domains but is conserved in all TruD homologs. This domain, which we term the TRUD domain, appears to be a novel RNA binding domain.

EXPERIMENTAL PROCEDURES

The *truD* gene was cloned into the NdeI and BamHI sites of the pET28a vector (Novagen Inc.) as described previously (6). Native and selenomethionine (SeMet)-labeled TruD protein were expressed and purified essentially as described for RluD (14). After purification, TruD was concentrated to 20 mg ml−1 by Amicon Ultra centrifugation (Milipore Inc.). Diffraction quality crystals of TruD were grown at room temperature by the hanging drop vapor diffusion method by mixing equal volumes of protein (from 5 to 10 mg ml−1 in 20 mM HEPES, pH 8.0, 250 mM NaCl, 2 mM EDTA, 2 mM dithiothreitol, 25% (v/v) ethylene glycol) and a reservoir solution containing 0.1 M MES, pH 6.0, 12–14% (v/v) polyethylene glycol 8000, and 25% (v/v) ethylene glycol. The crystals were tetragonal (space group P41) with a single molecule of TruD in the asymmetric unit. X-ray diffraction data were collected on a single flash-frozen SeMet TruD crystal and a single flash-frozen native TruD crystal at the X12-1 beamline, National Synchrotron Light Source, Brookhaven National Laboratory. Details of data collection and refinement are presented in Table I (in Supplementary Materials).

Diffraction data were processed using HKL (15), and initial phases were determined for a SeMet derivative by single wavelength anomalous dispersion (SAD) in SOLVE (16). SOLVE located three of the five expected selenium sites using SAD data to 2.2 Å and generated an initial electron density map with an overall figure of merit of 0.29. This map was improved to a figure of merit of 0.61 using density modification and solvent flattening in RESOLVE (17). The automated model building function of RESOLVE placed ~65% of the TruD model into the electron density, and the model was extended by several rounds of model building and refinement in O (18) and CNS (19), respectively. Using this model, the structure of native TruD was solved by molecular replacement in CNS followed by subsequent rounds of model building and refinement as above. The N-terminal polyhistidine tag and nine amino acids at the C terminus were not observed in the electron density profile.

**EXPERIMENTAL PROCEDURES**

The *truD* gene was cloned into the NdeI and BamHI sites of the pET28a vector (Novagen Inc.) as described previously (6). Native and selenomethionine (SeMet)-labeled TruD protein were expressed and purified essentially as described for RluD (14). After purification, TruD was concentrated to 20 mg ml−1 by Amicon Ultra centrifugation (Milipore Inc.). Diffraction quality crystals of TruD were grown at room temperature by the hanging drop vapor diffusion method by mixing equal volumes of protein (from 5 to 10 mg ml−1 in 20 mM HEPES, pH 8.0, 250 mM NaCl, 2 mM EDTA, 2 mM dithiothreitol, 25% (v/v) ethylene glycol) and a reservoir solution containing 0.1 M MES, pH 6.0, 12–14% (v/v) polyethylene glycol 8000, and 25% (v/v) ethylene glycol. The crystals were tetragonal (space group P41) with a single molecule of TruD in the asymmetric unit. X-ray diffraction data were collected on a single flash-frozen SeMet TruD crystal and a single flash-frozen native TruD crystal at the X12-1 beamline, National Synchrotron Light Source, Brookhaven National Laboratory. Details of data collection and refinement are presented in Table I (in Supplementary Materials).

Diffraction data were processed using HKL (15), and initial phases were determined for a SeMet derivative by single wavelength anomalous dispersion (SAD) in SOLVE (16). SOLVE located three of the five expected selenium sites using SAD data to 2.2 Å and generated an initial electron density map with an overall figure of merit of 0.29. This map was improved to a figure of merit of 0.61 using density modification and solvent flattening in RESOLVE (17). The automated model building function of RESOLVE placed ~65% of the TruD model into the electron density, and the model was extended by several rounds of model building and refinement in O (18) and CNS (19), respectively. Using this model, the structure of native TruD was solved by molecular replacement in CNS followed by subsequent rounds of model building and refinement as above. The N-terminal polyhistidine tag and nine amino acids at the C terminus were not observed in the electron density profile.
RESULTS AND DISCUSSION

TruD folds into a V-shaped molecule with two domains: a catalytic domain and a TRUD domain linked together by three extended loops (Fig. 2, A and B). The catalytic domain adopts a mixed α/β fold characteristic of ψ synthases from the four other families. The TRUD domain (residues 155–303) forms a compact fold that is tilted away from the catalytic domain to form a deep cleft in TruD which is lined with basic residues from each domain (Fig. 2C). Several conserved basic residues extend down and out from this cleft to provide an elongated positively charged surface that is similar in length to the longer anticodon stem-loop. The N-terminal end of TruD (residues 1–25) forms a long extended loop (loop L1) that wraps around both the TRUD and catalytic domains (Fig. 2A), before plunging into the middle of a seven strand β sheet consisting of two anti-parallel sheets (β10-β11-β5-β1; β3-β2-β4). Another extended loop (L3; residues 154–161) leads from the catalytic domain into the TRUD domain. A third loop (L19; 300–306) from the TRUD domain extends back into the catalytic domain and forms one end (β10-β11) of the seven-strand β sheet.

Despite the extensive topological connections between the catalytic and TRUD domains, the interface between them is not very extensive and covers only about 1100 Å² of solvent accessible surface area (~11% of the total solvent accessible surface of TruD). The interaction surface has several gaps and is only about half hydrophobic, which is characteristic of transient non-obligate protein-protein interactions (22). The interface between the catalytic and TRUD domains also has several buried well ordered water molecules that can be seen in the electron density. These are indications of a weak interface between the two TruD domains. The fact that the cleft between the catalytic domain and the TRUD domain is too narrow to accommodate the width of a tRNA leads us to propose that the angle between the two domains can vary and that the V shaped TruD undergoes a hinge-like motion. This hypothesis is also supported by the fact that there are several highly conserved glycines in the extended loops that connect these two domains. Movement upon binding RNA has been demonstrated for TruB. Comparison of apo and RNA-bound structures of TruB showed that a 29-residue loop reorients upon binding a stem-loop RNA (9, 23). Our structure cannot distinguish between the opening of the cleft as a result of tRNA binding to TruD and triggering this change or the binding of tRNA serving to trap the open form preexisting in equilibrium with the closed form seen in the crystal.

TruD may also gain access to U13 by reorganizing the RNA binding structure upon binding. Recent co-crystal structures of tRNA guanine transglycosylases (TGTs) with RNA have
shown that substantial reorganization of RNA structure can occur without concomitant changes in protein structure. Prokaryotic TGT rearranges the end of a stem-loop RNA to fit its active site, with little change of the enzyme (24). Archaeosine TGT (an archaeal enzyme) completely unfolds the D arm of tRNA before converting the classic i-form tRNA to a \( \lambda \)-form to gain access to G15, but the enzyme structure changes little from the tRNA-free form (25). Docking of the \( \lambda \)-form tRNA into the cleft of TruD was attempted but too many clashes were found (data not shown). Attempts at docking tRNA from \( T.\ thermophilus \) (26) produced fewer clashes; however, significant unfolding of the D-arm would be required to bring U13 into the catalytic pocket of TruD. Determining the actual mechanism TruD uses to gain access to U13 by co-crystallization with tRNA is under way.

How does TruD choose tRNA from all the other tRNAs in the cell? In yeast, the TruD homolog, Pus7p, has been suggested to recognize the sequence motif Pu9(G/C)UN, and that U11 is particularly important (7). In E. coli, only tRNA \( ^{\text{Glu}} \) has U13 and only tRNA \( ^{\text{Gln}} \) (NCC) has U13. The corresponding sequences are, respectively, \( ^{\text{CGUCyAG}} ^{\text{CGUAA}} \) and \( ^{\text{CGUAUA}} ^{\text{CGUAA}} \), both of which have U11 and both differ from the yeast sequence at position 9. Discounting the variation at position 9, since tRNA \( ^{\text{Glu}} \) is a substrate, the failure of tRNA \( ^{\text{Gln}} \) (NCC) to be a substrate must be due to some more subtle feature distinguishing the two tRNAs.

A DALI (27) search revealed that the closest structural neighbors of TruD in the Protein Data Bank are other \( \Psi \) synthases. TruA \( (Z \text{ score} = 7.8, \text{r.m.s.d.} 3.1 \AA \text{ over 156 Ca atoms}) \) and TruB \( (Z \text{ score} = 7.3, \text{r.m.s.d.} 3.2 \AA \text{ over 144 Ca atoms}) \) are the most similar followed by RsuA \( (Z \text{ score} = 5.9, \text{r.m.s.d.} 2.9 \AA \text{ over 119 Ca atoms}) \) and RluD \( (Z \text{ score} = 3.8, \text{r.m.s.d.} 3.1 \AA \text{ over 105 Ca atoms}) \). An alignment is shown in Fig. 3. The resulting superposition of the backbones of TruA and TruB, two \( \Psi \) synthases that also bind tRNAs, with TruD shows that the structural overlap is extensive except for the PUA and the TRUD domains (Fig. 4A). The overlap in the catalytic fold of TruA, TruB, RsuA, and RluD was characterized previously (11, 12) and can be simply summarized as the sharing of ten elements, six \( \beta \) strands, two \( \alpha \) helices, and two loops, in a linear order along the amino acid sequence: \( \beta_1, \beta_2, a_1, L_4, \beta_3, \beta_5, \beta_10, \beta_11, L_{21}, o(8) \) (Fig. 2A). The catalytic aspartate always appears in the loop L1. This arrangement is also shared in the TruD family except that the first \( \beta \) strand in TruD corresponds to the last \( \beta \) strand in the other four families (Fig. 3). It is quite interesting that a way has been found to rearrange the linear order of these elements in TruD while keeping the same fold.

Five conserved motifs in the TruA, TruB, RsuA, and RluA families of \( \Psi \) synthases have been characterized from comparison of the TruA, TruB, RsuA, and RluD structures (12). The corresponding structural alignment of TruD shows that it also contains these five motifs (Fig. 3). The major difference is that motif IIIa appears first in TruD. Residues from four motifs come together in the largely hydrophobic active sites of the TruA, TruB, RsuA, and RluD \( \Psi \) synthases (12). The essential catalytic Asp (motif II), a basic residue (usually Arg or Lys) that forms a salt bridge to the Asp (motif III), a tyrosine that provides stacking interactions with the uracil of the substrate uridine (motif IIa) (9, 23), and two hydrophobic residues (an Ile or Val from motif III and a Leu from motif IIIa). In TruD, the Asp\( ^{\text{Glu}} \) from motif II has been conserved, but its salt bridge partner is no longer a conserved basic residue from motif III because Thr\( ^{\text{Glu}} \) occupies this location (Fig. 4B). Instead, Lys\( ^{\text{Glu}} \) from loop 1 is the salt bridge partner of Asp\( ^{\text{Glu}} \). Phe\( ^{\text{Glu}} \) replaces Tyr from motif IIa, but a Phe could serve the same stacking function as the Tyr. Leu from motif IIIa has been replaced by the bulkier hydrophobic side chain Phe\( ^{\text{Glu}} \), and a smaller hydrophobic residue, Ala\( ^{\text{Glu}} \), has replaced the hydrophobic Ile or Val from motif III.

The TRUD domain folds into a novel three-dimensional protein fold that is not found in any other protein. A DALI (27) search with either the entire TruD structure or only the TRUD domain alone did not find any structural homologs of the TRUD domain in the Protein Data Bank. Moreover, a search of the non-redundant sequence databases with PSI-BLAST (28) found that the protein sequence of the TRUD domain is always associated with a TruD-type catalytic domain and that it is not found on its own or attached to another type of protein as a separate module. Furthermore, there are no TruD-type catalytic domains that lack the TRUD domain insert. Although we found no known putative sequence in any other protein significantly similar to the TRUD domain, it is always possible that one will be found in a new genome or that this fold will be present in a different (as of yet unsolved) protein structure without sequence conservation.

The TRUD domain is characterized by two conserved sequence motifs (motifs IV and V in Ref. 6). The first of these...
motifs (IV; residues 155–169 in Fig. 3) forms a long extended loop starting at the bottom of the inter-domain interface and extends up to form part of the cleft that likely binds tRNA. This loop starting at the bottom of the inter-domain interface and 169 in Fig. 3) forms a long extended motifs (IV; residues 155–160) that form a part of the hydrophobic core of the TRUD domain. Arg164 and Phe165 in this motif are solvent exposed side chains that extend out of the TruD cleft and would likely be involved in tRNA binding. Gln163 is also solvent exposed but forms a weak salt bridge with the putative catalytic Asp80 in our structure.

In summary, TruD folds into a V-shaped molecule with a secondary sequence motif (V, residues 196–211 in Fig. 3) lies along a long helix a4 that extends all along the length of the TRUD domain. Many of the conserved residues in this motif contribute to the hydrophobic core of the TRUD domain. Arg210 in this motif forms a salt bridge to the highly conserved Gln163 is also solvent exposed but forms a weak salt bridge with the putative catalytic Asp80 in our structure.

Acknowledgments—Data were collected at beamline X12C of the National Synchrotron Light Source, Brookhaven National Laboratory. National Synchrotron Light Source financial support comes principally from the Offices of Biological and Environmental Research and of Basic Energy Sciences of the United States Department of Energy and from the National Center for Research Resources of the National Institutes of Health.

REFERENCES

1. Ofengand, J. (2002) FEBS Lett. 514, 17–25
2. Del Campo, M., Kaya, Y., and Ofengand, J. (2001) RNA (N. Y.) 7, 1603–1615
3. Koonin, E. V. (1996) Nucleic Acids Res. 24, 2411–2415
4. Gustafsson, C., Reid, R., Greene, P. J., and Santi, D. V. (1996) Nucleic Acids Res. 24, 3756–3762
5. Ofengand, J., and Budd, K. E. (2000) in The Ribosome: Structure, Function, Antibiotics, and Cellular Interactions (Garrett, R. A., Douthwaite, S. R., Liljas, A., Matheson, A. T., Moore, P. B., and Noller, H. F., eds.) pp. 175–189, ASM Press, Washington, D. C.
6. Kaya, Y., and Ofengand, J. (2003) RNA (N. Y.) 9, 711–721
7. Behn-Ansamt, I., Urban, A., Ma, X., Yu, Y.-T., Motorn, Y., and Branlant, C. (2003) RNA (N. Y.) 9, 1371–1382
8. Foster, P. G., Huang, L., Santi, D. V., and Strood, R. M. (2000) Nat. Struct. Biol. 7, 23–27
9. Hoang, C., and Ferré-D’Amare, A. R. (2001) Cell 107, 929–939
10. Sivaraman, J., Sauvé, V., Larroque, R., Stura, E. A., Schrag, J. D., Cygler, M., and Matte, A. (2002) Nat. Struct. Biol. 9, 353–358
11. Sivaraman, J., Iannuzzi, P., Cygler, M., and Matte, A. (2004) J. Mol. Biol. 335, 87–101
12. Del Campo, M., Ofengand, J., and Malhotra, A. (2004) RNA (N. Y.) 10, 215–229
13. Mueller, E. G. (2002) J. Mol. Biol. 273, 558–566
14. Del Campo, M., Ofengand, J., and Malhotra, A. (2003) Acta Crystallogr. Sect. D Biol. Crystallogr. 59, 1871–1873
15. Otwinowski, Z., and Minor, W. (1997) Methods Enzymol., 276, 307–326
16. Terwilliger, T. C., and Berendsen, J. (1999) Acta Crystallogr. Sect. D Biol. Crystallogr. 55, 849–861
17. Terwilliger, T. C. (2000) Acta Crystallogr. Sect. D Biol. Crystallogr. 56, 965–972
18. Jones, T. A., Zou, J.-Y., Cowan, S. W., and Kjeldgaard, M. (1991) Acta Crystallogr. Sect. A 47, 110–119
19. Brumberg, A. K., Adams, P. D., Clore, G. M., DeLano, W. L., G ros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Ku szewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Acta Crystallogr. Sect. D Biol. Crystallogr. 54, 905–921
20. Laskowski, R. A., MacArthur, M. V., Moss, D. S., and Thornton, J. M. (1993) J. Appl. Crystallogr. 26, 283–291
21. Lovell, S. C., Davis, I. W., Arendall, W. B., III, de Bakker, P. I. W., Word, J. M., Frantais, M. G., Richardson, J. S., and Richardson, D. C. (2003) Proteins Struct. Func t. Genet. 55, 437–450
22. Noren, M. A., and Thornton, J. M. (2003) EMBO J. 22, 3486–3492
23. Pan, H., Agarwala, S., Moustakas, D. T., Finer-Moore, J., and Strood, R. M. (2003) Proc. Natl. Acad. Sci. U. S. A. 100, 12648–12653
24. Xie, W., Liu, X., and Huang, R. H. (2003) Nat. Struct. Biol. 10, 781–788
25. Ishitani, R., Nureki, O., Nami ki, N., Okada, N., Nishimura, S., and Yokoyama, S. (2003) Cell 113, 383–394
26. Sekine, S., Nureki, O., Dubois, D. Y., Bernier, S., Chénevert, R., Lapointe, J., Vassylyev, D. G., and Yokoyama, S. (2003) EMBO J. 22, 676–688
27. Holm, L., and Sander, C. (1990) J. Mol. Biol. 233, 123–138
28. Altschul, S. F., Madden, T. L., Zehlen, A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Nucleic Acids Res. 25, 3389–3402
29. Nicholls, A., Sharp, K. A., and Lipman, D. J. (1997) Nucleic Acids Res. 25, 3389–3402
30. Delano, W. L. (2002) The PyMOL Molecular Graphics System at pymol.org
Crystal Structure of TruD, a Novel Pseudouridine Synthase with a New Protein Fold
Yusuf Kaya, Mark Del Campo, James Ofengand and Arun Malhotra

J. Biol. Chem. 2004, 279:18107-18110.
doi: 10.1074/jbc.C400072200 originally published online March 3, 2004

Access the most updated version of this article at doi: 10.1074/jbc.C400072200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

Supplemental material:
http://www.jbc.org/content/suppl/2004/03/22/C400072200.DC1

This article cites 31 references, 3 of which can be accessed free at
http://www.jbc.org/content/279/18/18107.full.html#ref-list-1