Published online February 27, 2004

1422–1429 *Nucleic Acids Research, 2004, Vol. 32, No. 4*
DOI: 10.1093/nar/gkh287

Conformational change of pseudouridine 55 synthase upon its association with RNA substrate

Kulwadee Phannachet and Raven H. Huang*

Department of Biochemistry, School of Molecular and Cellular Biology, University of Illinois at Urbana-Champaign, 600 South Mathews Avenue, Urbana, IL 61801, USA

Received November 17, 2003; Revised December 26, 2003; Accepted January 26, 2004

ABSTRACT

Pseudouridine 55 synthase (Ψ55S) catalyzes isomerization of uridine (U) to pseudouridine (Ψ) at position 55 in transfer RNA. The crystal structures of *Thermotoga maritima* Ψ55S, and its complex with RNA, have been determined at 2.9 and 3.0 Å resolutions, respectively. Structural comparisons with other families of pseudouridine synthases (ΨS) indicate that Ψ55S may acquire its ability to recognize a stem–loop RNA substrate by two insertions of polypeptides into the ΨS core. The structure of apo-Ψ55S reveals that these two insertions interact with each other. However, association with RNA substrate induces substantial conformational change in one of the insertions, resulting in disruption of interaction between insertions and association of both insertions with the RNA substrate. Specific interactions between two insertions, as well as between the insertions and the RNA substrate, account for the molecular basis of the conformational change.

INTRODUCTION

Pseudouridine (Ψ), also regarded as 'the fifth nucleotide' because of its abundance (1), is post-transcriptionally converted from uridine through isomerization of the base by a class of enzymes called pseudouridine synthase (ΨS). Ψ is shown to be important for the structural integrity of the transfer RNA (2). The locations of several Ψ in the ribosomal RNA are conserved and are clustered near the peptidyl transfer site, implying their importance for the proper function of the ribosomes (3). Furthermore, it has been shown that the lack of Ψ in the spliceosomal RNA results in loss of splicing function of the spliceosomes (4).

Sequence comparison of pseudouridine synthases groups them into four separate families (5), and structures representing each family are now available (6–9). More recently, a new pseudouridine synthase belonging to the fifth family has been discovered (10). Despite a lack of statistically significant global sequence similarities among the families, superposition of the three-dimensional structures from all four families reveals that they share a core structure (ΨS core) consisting of two domains with βαββαβ topology, as first reported by Stroud and coworkers (6). An aspartic acid was identified to be strictly conserved among the four families of enzymes and was shown to be critical for catalysis, acting as a nucleophile in the isomerization reaction (11–14). This aspartic acid is located in the same structural position in the four families of ΨS, further supporting its essential role in catalysis. While it is generally agreed that the conserved aspartic acid is the nucleophile, opinion differs with regards to the target on the RNA substrate it attacks. The aspartic acid could attack C6 of the base to assist the rotation of the base during reaction (C6 mechanism, Fig. S1A, Supplementary Material) (15,16). Alternatively, the aspartic acid could attack C1′ of the ribose to assist the departure of the base (C1′ mechanism, Fig. S1B, Supplementary Material) (11,16). Santi and coworkers demonstrated the formation of a presumably covalent complex between pseudouridine synthase I (ΨSI) and 5FU-containing tRNA and reported the chemical characterization of the hydrolyzed product of the target nucleotide isolated from the complex (15). More recently, crystal structures of Ψ55S in complex with 5FU-containing RNA have shown the hydrolyzed product of the isomerized 5FU in the active site (7,17). These studies strongly favor the C6 mechanism. On the other hand, the recent crystal structure of 16S rRNA pseudouridine synthase RsuA in complex with a UMP indicates that the C1′ mechanism is also possible (8). With the aim of resolving these two competing mechanisms, we have pursued structural determination of a ΨS-RNA complex trapped in the covalent intermediate state of the reaction. We have chosen Ψ55S over other ΨS as our system because the minimal RNA substrate *in vitro* has been defined, which may increase the chance of obtaining crystals of the ΨS-RNA complex. An additional benefit of this choice is that the study yields insight into molecular recognition of RNA substrate by Ψ55S, as well as evolutionary relationship of Ψ55S compared to other ΨS. We report here our attempts to trap the Ψ55S-RNA covalent intermediate, the structure of apo-Ψ55S, as well as the structure of Ψ55S in complex with 5FU-containing RNA.

MATERIALS AND METHODS

RNA substrates, protein expression and purification

The RNA was from Dharmacon Research (Boulder, CO) and was purified using denaturing polyacrylamide gel electrophoresis (15% acrylamide, 8 M urea). Concentration of RNA was estimated by UV at 260 nm. A plasmid encoding

*To whom correspondence should be addressed. Tel: +1 217 333 3967; Fax: +1 217 244 5858; Email: huang@uiuc.edu

Escherichia coli Ψ55S was a gift from Daniel V. Santi (University of California, San Francisco, CA). The amplification of the gene encoding *Thermotoga maritima* Ψ55S was performed with genomic DNA from *T.maritima* (ATCC 43589D) as the template and two primers, the forward primer 5′-TAGGGCGAATTCAAGGAGATATACATATGAATG-GAATATTT GCTATTGAG-3′ (EcoRI site is underlined), and the reverse primer 5′-TAGGCAAAGCTTTTACCTCG-TGTTGAAGACCTTTCT-3′ (the HindIII site is underlined). The digested PCR product was ligated with the EcoRI/HindIII digested vector pLM-1 (18). The ligated product was then transformed into DH5α cell by heat shock. DNA sequencing indicates that no mutation was introduced during PCR amplification.

The pLM-Ψ55S vector was transformed into BL21(DE3) cells. Cells were grown in LB medium in the presence of 0.1 mg/ml ampicillin at 37°C. When the cell culture was grown to $OD_{600} \approx 0.6$, IPTG was added to 1 mM final concentration. The induced cell was grown for additional 4 h at 37°C and then harvested by centrifuging at 4000 *g* for 20 min at 4°C. Cell pellet was re-suspended in 20 ml of lysis buffer containing 20 mM HEPES (pH 7.0), 50 mM NaCl, 1 mM EDTA and were lysed using a French Press. The cell lysate was centrifuged at 10 000 *g* for 1 h and the supernatant was loaded into a SP Sepharose column equilibrated with SP buffer A (10 mM HEPES, pH 7.0, 10 mM DTT, 10 mM NaCl). The column was eluted with a linear gradient of NaCl (10 mM–1.0 M) with SP buffer B (10 mM HEPES, pH 7.0, 10 mM DTT, 1.0 M NaCl). *Thermotoga maritima* Ψ55S was eluted in fractions between 450 and 550 mM of NaCl. The fractions were combined and 2.5 M $(NH_4)_2SO_4$ stock solution was added to the final concentration of 1 M. The resulting sample was loaded into a HiLoad Sepharose column equilibrated with HiLoad buffer A [10 mM HEPES, pH 7.0, 10 mM DTT, 1 M $(NH_4)_2SO_4$]. The column was eluted with a linear gradient of 1.0–0.0 M of $(NH_4)_2SO_4$. *Thermotoga maritima* Ψ55S was eluted at 450–600 mM of $(NH_4)_2SO_4$. The Ψ55S containing fractions were pooled and dialyzed against a dialysis buffer containing 10 mM HEPES, pH 7.0, 10 mM DTT, 10 mM NaCl. After dialysis, the sample was loaded into a Mono S column. The column was eluted with 20 ml of buffer with linear gradient from 10 mM to 1.0 M of NaCl. Ψ55S was eluted between 350 and 450 mM of NaCl. The protein was then loaded into a Superdex column equilibrated with 10 mM HEPES pH 7.0, 500 mM NaCl, 10 mM DTT. Ψ55S was eluted between 85 and 100 ml. The fractions containing *T.maritima* Ψ55S were combined and concentrated to 5 mg/ml.

SDS gel analysis of formation of stable Ψ55S-RNA complex

For a typical reaction shown in Figure 1B, *E.coli* or *T.maritima* Ψ55S (20 μM) was mixed with equal molar of the stem–loop RNA in 10 μl of reaction buffer containing 10 mM MES (pH 6.5), 50 mM NaCl. After the reaction mixture was incubated at 4°C for 12 h, 5 μl of 2× standard SDS loading buffer containing 4 mM DTT was added. The resulting mixture was further incubated at 25°C for 5 min and 12 μl of each sample was loaded to a 12% SDS gel. The conditions for reaction and gel analysis shown in Figure 1C were similar to those shown in Figure 1B except the concentrations of protein and RNA were 80 μM and, for those samples that required heating, 5 min of incubation at 95°C was used instead of 25°C before gel loading.

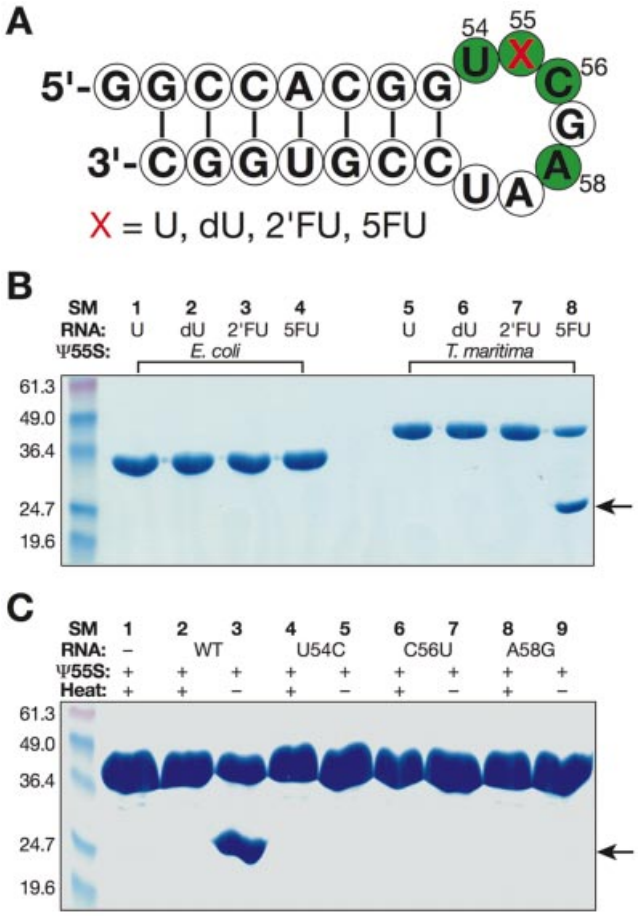

Figure 1. Formation of a stable Ψ55S-RNA complex. (**A**) Nucleotide sequence of a stem–loop RNA employed in this study. The conserved nucleotides required to be substrates of Ψ55S are colored green. U, uridine; dU, 2′-deoxyuridine; 2′FU, 2′-fluorouridine; 5FU, 5-fluorouridine. (**B**) SDS gel analysis of reaction mixtures of Ψ55S and RNA under conditions indicated. (**C**) Effect of mutations of the conserved nucleotides in RNA on the formation of Ψ55S-RNA complex. The band indicated by an arrow is the Ψ55S-RNA complex.

Purification and crystallization of Ψ55S-RNA complex, crystallization of Ψ55S

The *in situ* procedure for the complex formation was carried out on a preparative scale, thereby generating a stable Ψ55S-RNA complex comprising Ψ55S bound to a 22-nucleotide 5FU-containing RNA oligomer (Fig. 1A). This complex was purified on a FPLC system using HiTrap Phenyl Sepharose column, and the sample was concentrated, and the resultant Ψ55S-RNA complex [7 mg/ml, in a buffer containing 10 mM HEPES, pH 7.0, 50 mM $(NH_4)_2SO_4$, 1 mM DTT] was subject to crystallization efforts. Crystals of the complex were obtained using vapor diffusion method by mixing the crystallization sample with an equal volume of a well solution containing 100 mM Tris–HCl, pH 8.0–8.5, 10–12% polyethylene glycol (PEG) 6000, and 100–300 mM NaCl. Crystals grew in diamond shape in 2–3 days at 4°C. Crystals of selenomethionine Ψ55S-RNA complex were obtained under

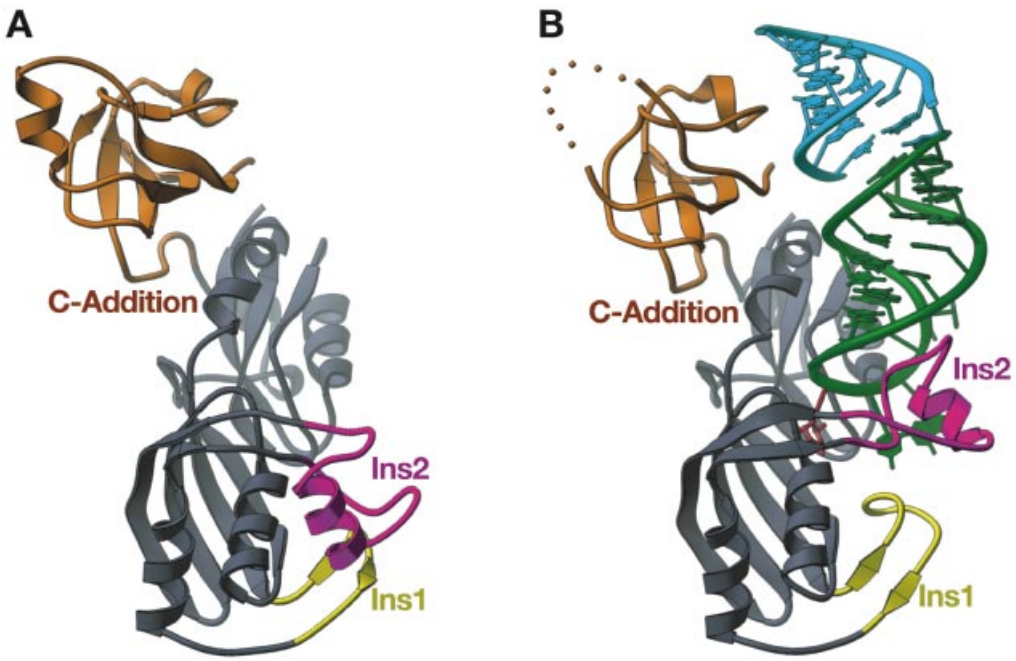

Figure 2. Overall structures of apo-Ψ55S and Ψ55S-RNA complexes. (**A**) RIBBONS representation of Ψ55S. The core structure shared by all four families of ΨS is colored grayish blue. Two insertions (Ins 1 and Ins2) and a C-terminal addition (C-Addition) unique to Ψ55S are yellow, magenta and orange, respectively. (**B**) Ribbon representation of Ψ55S-RNA complex. Ψ55S is colored as in (**A**). The RNA substrate is in green, with the target nucleotide in red. The stem of the RNA from the second complex in the asymmetric unit stacks at the end of the RNA substrate and is colored cyan. This figure and Figures 3–5, were made with RIBBONS (29).

similar conditions. Crystals of Ψ55S were obtained using vapor diffusion method by mixing Ψ55S (5 mg/ml, in a buffer containing 10 mM HEPES, pH 7.0, 100 mM NaCl, 1 mM DTT) with an equal volume of a well solution containing 100 mM HEPES, pH 7.0, 8–10% polyethylene glycol (PEG) 6000, and 0.8–1.0 M NaCl. Crystals grew as thin plate in 3–4 weeks at 4°C.

Data collection, structural determination and refinement

Crystals of Ψ55S-RNA complex were harvested after 2 days of growth at 4°C and stepwise (0.5 h for each step) soaked in a series of cryo-protecting solutions containing all the components of the well solution plus 5, 10, 15, 20 and 25% glycerol. Crystals of Ψ55S were harvested after 3 weeks of growth at 4°C and briefly soaked (3–5 min) in a cryo-protecting solution containing all the components of the well solution plus 20% glycerol. The soaked crystals were then mounted in a nylon loop and flash-frozen in liquid nitrogen. Three-wavelength anomalous dispersion (MAD) data were collected at the 14-BMC beamline at Advanced Photon Source (APS), and the native data for Ψ55S and Ψ55S-RNA complex were collected at the 19-ID beamline at APS. Data were reduced with Denzo and Scalepack (19). For phase determination, the resolution range from 30 to 3.5 Å was chosen. Four of eight expected Se sites (assuming two complexes in the asymmetric unit) were found using program Solve (20). The phases were improved by using program RESOLVE (21). The initial map indicates that there are two complexes in the asymmetric unit, one substantially more ordered than the other. The four Se sites from SOLVE were all located in the better ordered complex, and the complex was manually built with program O (22), followed by refinement with program CNS (23). Phases were improved by several rounds of rebuilding and refinement, allowing small parts of the second complex to be built and refined. The partially built model of the second complex revealed that the two complexes stack head-to-head through base-pairing of the overhanging 5′ guanine residue in RNA (Fig. 1A), resulting in extension of helix from one stem to the other (Fig. 2B). Two-fold non-crystallographic symmetry (NCS) restraints were then introduced in the CNS refinement protocol in order to build more residues of the second complex. Electron density map indicated that the disorder of the second complex was uneven, with the part of molecule further away from dimer interface more disordered. Despite our best efforts, we were not able to build residues 68–107, 150–161, 173–191 and 284–302 of Ψ55S in the second complex because of severe disorder, resulting in a model with R_{free} of 35% and R_{work} of 29%. In order to improve the quality of model, it was necessary to eliminate NCS restraints in the CNS refinement protocol for further rounds of refinements. The final model has R-factor of 26.1% (R_{free} = 31.8%). Attempts to reduce R_{free} to less than 30% were not successful. Final refinement statistics are given in Table 1.

In order to determine the structure of apo-Ψ55S, the refined model of Ψ55S-RNA complex, with the omission of RNA and Ins2 of Ψ55S, was used for the search model for the molecular replacement calculations with program AMORE (24). Four solutions were obtained. After rigid body refinement, density for some of the missing Ins2 was evident and the observable residues were then built with program O, followed by refinement with program CNS. Additional residues were added as refinement proceeded. Phases were improved by many rounds of manual rebuilding and refinement, with 4-fold

Table 1. Statistics of data collection and refinement

Crystals	Inflection	Peak	Hard remote	RNA complex	Apo-Ψ55S
Space group				P41212	P1
Unit cell				135.0, 135.0, 139.5 90.0, 90.0, 90.0	44.7, 64.2, 176.1 87.5, 86.5, 79.35
Resolution (Å)	50–3.5	50–3.5	50–3.5	50–3.0	50–2.9
Wavelength (Å)	0.9800	0.9797	0.9574	0.9793	1.0000
Unique reflections	16 457	16 452	16 402	26 335	40 146
Completeness (%)	96.6	96.6	96.4	99.4	94.5
Average I/s (I)	21.6	18.9	20.1	15.7	14.2
Redundancy	13.2	12.7	12.5	10.3	7.7
R_{sym} (%)	3.1	4.1	3.3	4.6	6.9
Refinement					
Resolution (Å)				50–3.0	50–2.9
Reflections (free)				21 028 (1844)	33 272 (2940)
$R_{crystal}$ (R_{free}) (%)				26.1 (31.8)	23.5 (28.1)
r.m.s.d. bonds (Å)				0.0083	0.0080
r.m.s.d. angles (°)				1.40	1.27

Mean figure-of-merit (FOM) for phasing = 0.52 from SOLVE and 0.66 after RESOLVE.
I/s (I) is the mean reflection intensity/estimated error.
Rsym = $\Sigma|I - \langle I\rangle|/\Sigma I$, where I is the intensity of an individual reflection and $\langle I\rangle$ is the average intensity over symmetry equivalents.
$R_{crystal} = \Sigma| |F_o| - |F_c| |/\Sigma|F_o|$, where F_o and F_c are the observed and calculated structure factor amplitudes.
R_{free} is equivalent to $R_{crystal}$ but calculated for a randomly chosen set of reflections that were omitted from the refinement process.

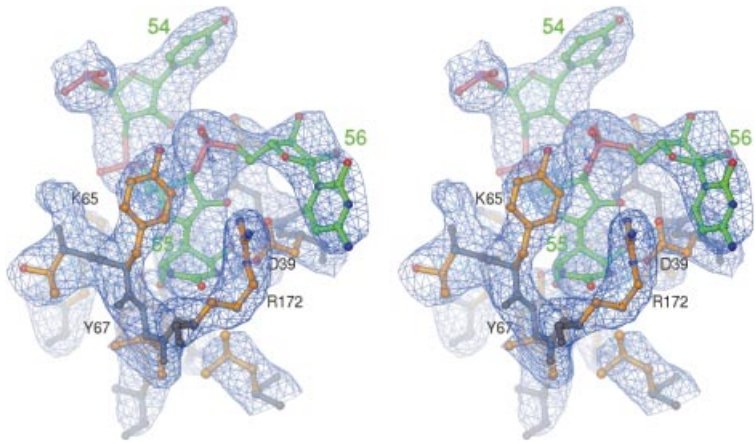

Figure 3. Stereo view of the simulated annealing $F_o - F_c$ omit map showing electron density for FhΨ55 in the structure of Ψ55S-RNA complex. All atoms within 4 Å were omitted from refinement. The main chains of Ψ55S are grayish blue, the side chains are orange and the RNA is green. The hetero-atoms are colored individually, with nitrogen in blue, oxygen in red and phosphate in magenta. The map is contoured at 3.0 σ.

non-crystallographic symmetry (NCS) restraints of Ψ55S (minus Ins2) in all refinement protocols. The final model shows that three of four molecules in the asymmetric unit contain well-ordered structure of Ins2.

RESULTS

Formation of stable Ψ55S-RNA complex

Our initial goal was to trap the covalent intermediate of Ψ55S-catalyzed reaction. Therefore, we have prepared Ψ55S from four different organisms (*E.coli*, *T.maritima*, *Bacillus subtilis* and yeast) and four RNAs that differ from one another with the identity of nucleotide at position 55 (Fig. 1A). Ψ55S was incubated with RNA pair wise and the resulting samples were analyzed by SDS–PAGE without heating. Incubation of the RNA substrate that contains 5FU at position 55 with *T.maritima* Ψ55S resulted in appearance of a new band in an SDS gel, albeit with an increased electrophoretic mobility (Fig. 1B, lane 8). This fast moving band was not observed when the enzyme was incubated with RNA that contain U, dU and 2′FU at position 55 (Fig. 1B, lanes 5–7), or in studies that employ Ψ55S from other organisms (*E.coli*, Fig. 1B, lanes 1–4; *B.subtilis* and yeast, data not shown). Two lines of evidence indicate that this faster-moving band contains a stable protein–RNA complex: (i) treatment of the sample with heat before the SDS gel analysis resulted in disappearance of this band (Fig. 1C, lane 2), indicating that the Ψ55S had not been proteolytically cleaved with the incubation of 5FU-containing RNA; (ii) UV spectrum of the sample extracted from this band

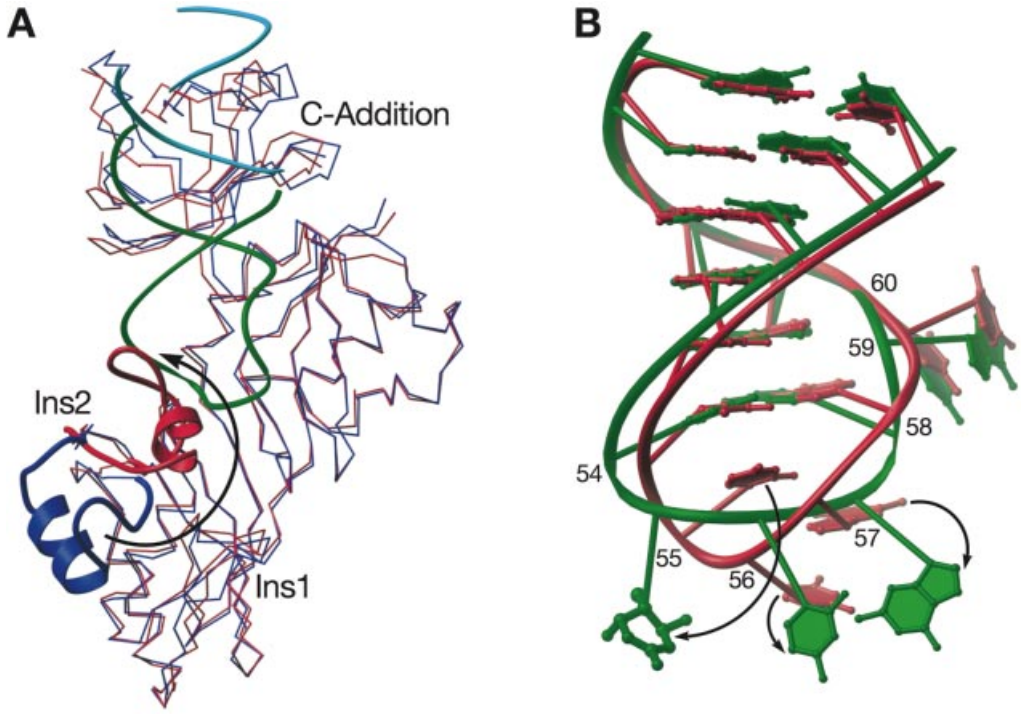

Figure 4. Conformational changes in Ψ55S and RNA. (**A**) Cα superposition of apo-Ψ55S (blue) with the one in Ψ55S-RNA complex (red). The Ins2 of both structures are highlighted with RIBBONS representations. RNAs are colored as in Figure 2B. (**B**) Three-dimensional superposition of the RNA in Ψ55S-RNA complex (green) with the corresponding nucleotides of $tRNA^{Phe}$ (red). Only the C1′ and P atoms of RNA (excluding nucleotides 55–57) were applied for the alignment.

showed the maximum of absorption of 256 nm, indicating that the sample contains RNA (data not shown). Formation of this product requires a 5FU-containing RNA possessing the conserved sequences required to be a substrate of Ψ55S shown in Figure 1A, as incubation of 5FU-containing RNA mutants (U54C, C56U or A58G) with Ψ55S did not produce such a band (Fig. 1C). Furthermore, the conserved aspartic acid (D39 in *T.maritima* Ψ55S) is also required for the formation of this product, as a D39A mutant failed to form such a product (data not shown).

The cause for the increased mobility of the Ψ55S-RNA complex in SDS gel is not clear. It is possible that the combination of the use of a thermostable protein (*T.maritima* is a thermophilic bacterium) and its tight association with RNA caused failure to denature Ψ55S in the complex by SDS loading buffer at 25°C, resulting in a complex with a more compact shape and a corresponding increase in mobility in SDS–PAGE compared to the denatured, unliganded protein. The fact that the same band was not observed with U, dU or 2′FU-containing RNA indicates either that Ψ55S forms a covalent complex with the 5FU-containing RNA, or that Ψ55S binds the 5FU-containing RNA non-covalently but very tightly. In order to understand the nature of the complex, which may have implications for the mechanism of the ΨS-catalyzed reaction, we have purified the complex in large scale and successfully carried out crystallization of the complex.

Structures of Ψ55S and Ψ55S-RNA complex

The structure of Ψ55S-RNA complex was determined at the 3.0 Å resolution using the phase information from the MAD data (Fig. 2B). We have also determined the structure of apo-Ψ55S at the 2.9 Å resolution (Fig. 2A). Electron density maps calculated from the MAD data of the Ψ55S-RNA complex indicate that the Ψ55S is not covalently linked to the RNA, but rather a hydrolyzed product of isomerized 5FU (5-fluoro-6-hydroxy-Ψ, or FhΨ, Fig. 3). Thus, our structure is essentially the same as the one reported earlier by Hoang and Ferre-D'Amare (Fig. S2B, Supplementary Material) (7), as well as the one reported very recently by Stroud and coworkers (Fig. S2A, Supplementary Material) (17). The interactions between Ψ55S and RNA are essentially identical to those from previous reports and will not be presented here. Instead, given the availability of the structure of apo-Ψ55S, we will focus our discussions on structural changes resulting from association of Ψ55S with RNA, the possible evolutionary origin of Ψ55S, and implications of failure to obtain a covalent complex for the possible mechanism of the ΨS-catalyzed reaction.

A different approach of structural description of Ψ55S was employed, which we believe allows us to better understand the role of each part of Ψ55S in substrate recognition from evolutionary point of view. Specifically, instead of sequential assignment of the secondary structure, the structure of Ψ55S is divided into two parts: one is the ΨS core, which is shared by all four families of ΨS (Fig. 2, colored in light blue); the other is a collection of two insertions (Ins1 and Ins2) and a C-terminal addition (C-Addition), which are unique to Ψ55S (Fig. 2, colored in yellow, magenta, and orange, respectively). As shown in Figure 2B, Ins1, Ins2 and the C-Addition all interact with the RNA substrate; their precise roles will be discussed in detail later.

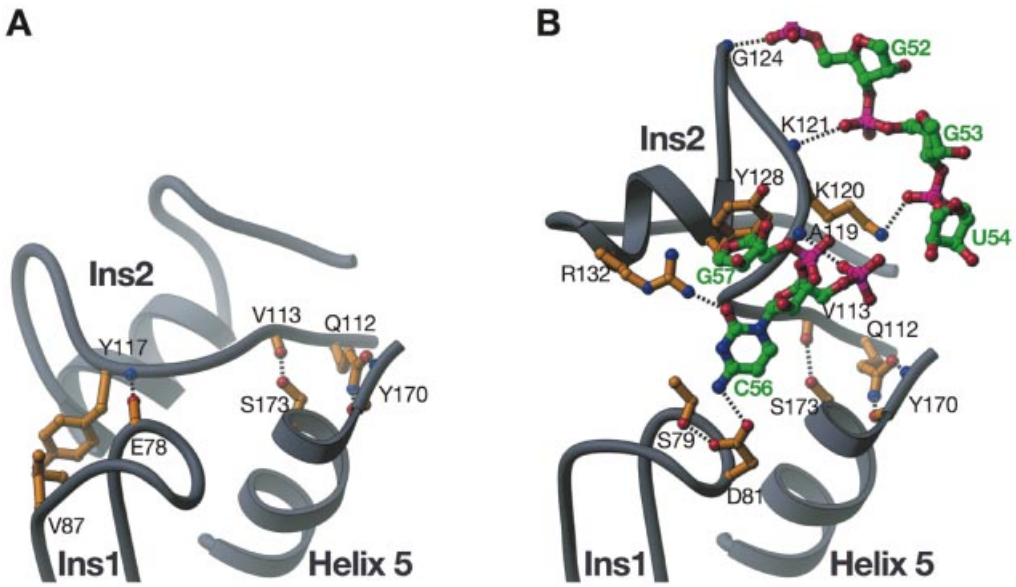

Figure 5. Molecular basis of conformational change in Ins2 upon association of RNA with Ψ55S. (**A**) Interactions between Ins1 and Ins2 in the structure of apo-Ψ55S. (**B**) Interactions of Ins1 and Ins2 with RNA in the structure of Ψ55S-RNA complex. The bases of G52, G53, U54, G57, as well as nucleotide FhΨ55, do not make contact with both Ins1 and Ins2, and are therefore omitted here for clarity. The color schemes are the same as in Figure 3.

Conformational changes and molecular basis of the changes

Three-dimensional superposition of the structure of apo-Ψ55S to that from the Ψ55S-RNA complex reveals that the ΨS core overlaps well (Fig. 4A), indicating that association of RNA does not cause structural changes to the ΨS core. The overlap is not as good for Ins1 and the C-Addition, but the changes are small (slight movement of both Ins1 and C-Addition toward the RNA in the complex). On the other hand, association of RNA with Ψ55S results in a substantial conformational change in Ins2 (Fig. 4A). The change can be depicted as a roughly 180° flip with a twist, resulting in translocation of the central loop in Ins2 by 26 Å.

Conformational changes are also observed in the RNA substrate upon its association with Ψ55S, most notably in the target base, which flips out of the loop (Fig. 4B). The bases at positions 56 and 57 also change, although to a lesser extent than the base at position 55. In addition to the base flipping, there is a slight twist of the RNA backbone in the loop region.

In the absence of RNA, Ψ55S is in a closed form in the insertion region (Figs 2A and 5A). Both Ins1 and Ins2 are sequestered through mutual interactions. The nature of interactions is mainly non-specific, with van der Waals contacts between Ins1 and Ins2, resulting in a buried surface area of 510 Å^2. The modest buried surface area is consistent with the functional requirement that the interaction between these two insertions must be disrupted upon RNA binding (see below). In addition to many non-specific interactions, two specific interactions are also observed: (i) the main chain carbonyl group of E78 in Ins1 accepts a hydrogen bond from the main chain amide of Y117 in Ins2; (ii) the side chain of V87 stacks on the side chain of Y117 (Fig. 5A). V87 and Y117 appear to be conserved among Ψ55S, as sequence alignment of 20 Ψ55S shows V, I or L at position 87, and Y or F at position 117 (data not shown). Since these two amino acids are not involved in binding of RNA substrate, their conservations are likely necessary for the specific interaction in apo-enzyme described above.

Upon RNA binding, the insertion region of Ψ55S opens up as a result of substantial movement of Ins2 (Figs 2B and 5B). Because of this movement, Ins1 and Ins2 no longer interact with each other and instead, they both interact with the RNA substrate. Ins1 only interacts with the base of the conserved nucleotide C56. The side chain of D81, whose position is secured by forming a hydrogen bond with the side chain of S79, accepts a hydrogen bond from the N4 of C56 (Fig. 5B). Both D81 and S79 are strictly conserved amongst Ψ55S from different species.

Ins2 interacts with both sides of the RNA surrounding FhΨ55 but not with FhΨ55 itself: three nucleotides on the 5′ end of FhΨ55 (G52, G53 and U54) and two nucleotides on the 3′ end of FhΨ55 (C56 and G57) (Fig. 5B). The interactions with the nucleotides on the 5′ end of FhΨ55 are not sequence specific. Each of the main chain amides of G124 and K121 and the side chain of K120 donates a hydrogen bond to the phosphate group of G52, G53 and U54, respectively. Three interactions were observed with the nucleotides on the 3′ end of FhΨ55: A hydrogen bond from the main chain amide of A119 to the phosphate backbone of C56; a hydrogen bond from the side chain of R132 to the O2 of C56; and stacking of the side chain of Y128 on the ribose of G57. A119, Y128 and R132 are all strictly conserved in Ψ55S. The interaction of R132 in Ins2 with the O2 of C56, together with the interaction of D81 in Ins1 with the N4 of C56, constitutes specific recognition of the conserved nucleotide C56.

In addition to its interaction with Ins1 (in apo-Ψ55S) or the RNA (in Ψ55S-RNA complex), one end of Ins2 interacts with helix 5 from the ΨS core. Specifically, the side chain of Q112 forms bifurcate hydrogen bonds with the main chain carboxyl group and amide of Y170, and the main chain carbonyl group

of V113 accepts a hydrogen bond from the side chain of S173 (Fig. 5A and B). Both Q112 and S173 are conserved. Since these two interactions are maintained in Ψ55S regardless of whether RNA is bound, they appear to serve as the anchoring point for the conformational change of Ins2 upon association of RNA with Ψ55S.

The active site

As in the case of two previous reported structures of Ψ55S-RNA complexes (7,17), the nucleoside at position 55 is the hydrolyzed product of the isomerized FU, FhΨ55. There are extensive interactions between FhΨ55 and amino acids from the ΨS core (Fig. 3). These extensive interactions, along with the fact that FhΨ55 no longer contains a planar base, may account for tighter binding of the Ψ55S to the 5FU-containing RNA relative to other RNAs (containing U, dU and 2′FU instead of 5FU at position 55), as observed by our SDS gel analysis (Fig. 1B).

DISCUSSION

Insight from the two structures described here, combined with previously reported biochemical and structural studies, allowed us to address two fundamental issues: the molecular recognition of tRNA substrate by Ψ55S and the possible evolutionary origin of Ψ55S. The failure to trap the covalent intermediate may also have implication on the third fundamental issue: the likely mechanism of ΨS-catalyzed reaction.

The minimum requirements for an RNA to serve as a substrate for Ψ55S are a stem–loop structure and four conserved nucleotides, U54, U55, C56 and A58, in which U55 is the target nucleotide for catalysis (25,26). U54 and A58 form a reverse Hoogsteen base pair with each other (Fig. 4B), and therefore play a structural role of the TΨC loop in RNA. This reduces the requirement for the recognition of the RNA substrate to two nucleotides, U55 and C56. The structure of Ψ55S-RNA complex indicates that the amino acids responsible for interacting with FhΨ55 are all from the ΨS core (Fig. 3). Therefore, assuming that substrate recognition and chemical mechanism are the same when a U instead of 5FU is at position 55, the ΨS core is solely responsible for the recognition of U55, as well as for catalysis in U55 to Ψ55 conversion, a conclusion that is consistent with structural comparisons of Ψ55S with other families of ΨS. Recognition of C56 is achieved by its interaction with Ins1 and Ins2, which only exist in Ψ55S. Therefore, it is likely that Ψ55S has evolved from a precursor ΨS, which probably consisted of only the ΨS core, as a consequence of two insertions (Ins1 and Ins2) and a C-terminal addition (C-Addition). Ins1 and Ins2 are responsible for specific recognition of C56. Ins2 has additional interactions with other parts of RNA in the loop region, which may be required for proper binding of RNA substrate with a stem–loop structure (Fig. 5B). The C-Addition was shown to interact with the stem part of RNA from the second complex in an asymmetric unit (Fig. 2B). Therefore, its role is likely to strengthen the binding of the natural substrate, which is an entire tRNA instead of the RNA stem–loop employed in this study. This is consistent with non-specific nature of the interaction as well as low sequence similarities of C-Addition among Ψ55S from different organisms (data not shown).

The ΨS core of the Ψ55S-RNA complex is essentially unchanged when compared to the structure of apo-Ψ55S. In addition, the directions of small conformational changes in Ins1 and C-Addition and substantial change in Ins2 are all toward RNA upon RNA binding (Fig. 4A). This indicates that the RNA-binding cleft in apo-Ψ55S is fully accessible for the RNA substrate. This, together with the fact that the global conformational change in RNA after its association with Ψ55S is not substantial (Fig. 4B), points to the initial event of Ψ55S-RNA association as RNA rigidly docking on the Ψ55S RNA-binding cleft through surface complementarities, followed by slight adjustments in both RNA (such as the backbone in the loop region) and protein (such as slight closing of C-Addition toward the stem of RNA) to better fit each other. Interaction between Ins1 and Ins2 is then disrupted and both Ins1 and Ins2 are now able to interact with their individual RNA targets at the 5′ and 3′ ends of U55 (Fig. 5). These interactions, accompanied by further local structural adjustments of the active site in the ΨS core and the target nucleotide U55 in RNA, result in a final configuration of Ψ55S-RNA complex that is productive for catalysis.

It is unlikely that different crystal packing environments between the apo and the complex crystals are the cause of the conformational changes observed. Three of four molecules in the asymmetric unit of the apo crystal, which are in different crystal packing environment, adopt the same structure shown in Figure 2A, indicating the conformation in the insertion region of apo-Ψ55S is genuine. Furthermore, despite belonging to different space groups, our complex structure overlapped very well with the previously reported similar structures (Fig. S2, Supplementary Material), demonstrating that the structural changes observed are indeed due to RNA binding.

The molecular and energetic bases of dissociation of Ins1 and Ins2 and their association with RNA are unclear. It is possible that the initial binding RNA by Ψ55S weakens the relatively modest interaction between Ins1 and Ins2. The more likely scenario is that the interaction between Ins1 with Ins2 is in equilibrium between association and dissociation states, with the association state predominant when the RNA substrate is not available. Upon RNA binding, the equilibrium shifts to the dissociation state because of interactions of Ins1 and Ins2 with RNA. This is consistent with the observation that Ins2 is disordered in one of four molecules in the asymmetric unit, and also in the structure of apo-Ψ55S from *E.coli* as observed by Stroud and coworkers (17). The modest affinity with which Ins1 and Ins2 associates in apo-Ψ55S may have functional consequences: The association would likely hinder proteolytical degradation in the absence of RNA substrate, while also facilitating release of RNA product when the reaction is complete, which requires dissociation and movement of Ins1 away from RNA.

It is interesting to note that the approach of substrate recognition employed by Ψ55S is different from another RNA modification enzyme, tRNA guanine transglycosylase (TGT), for which, the structures of both the apo-enzyme and the enzyme bound to the RNA substrate are available (27,28). Unlike Ψ55S, the conformation of TGT is essentially unchanged after its association with the RNA substrate. Instead, substantial conformational change occurs in the RNA substrate in order to fit into the substrate-binding cleft in TGT

(28). Hence, our structural studies [Xie *et al.* (28) and this study] of two differing RNA modification enzymes illustrate the plasticity of both protein and RNA substrate in protein–RNA recognition.

We can speculate on the reasons for our inability to trap a covalent enzyme–substrate complex. We failed to observe a covalent intermediate in the structure of Ψ55S-RNA complex, despite our short crystallization time and low incubation temperature (48 h at 4°C). Instead, it is the hydrolyzed product of the isomerized 5FU, FhΨ. The Ψ55S-RNA complex recovered from crystals behaved exactly the same as the purified complex before crystallization in SDS gel, e.g. only the rapidly migrating band was observed in the gel (data not shown), indicating the formation of FhΨ55 might occur before crystallization. The observation only required as little as 10 min incubation of Ψ55S with 5FU-containing DNA (data not shown). These observations make us wonder if FhΨ55 is indeed formed through hydrolysis of the double bond between N1 and C6 in the non-covalent intermediate formed through the C1′ mechanism (Fig. S1B, Supplementary Material), instead of the hydrolysis of the covalent intermediate formed through the C6 mechanism (Fig. S1A, Supplementary Material). Because of our failure to observe the covalent intermediate, the mechanism of ΨS-catalyzed reaction remains uncertain.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at NAR Online.

ACKNOWLEDGEMENTS

We are grateful to Dr D. Santi (University of California, San Francisco) for a plasmid encoding the *E.coli* Ψ55S. We thank the staff of beamlines 14-BMC (K.Brister, T.Teng and R.Pahl) and 19-ID (S.Ginell, A.Joachimiak and Y.Kim) at APS for their support during data collections; Y. Elias, W. Xie and other members of Huang research group; S. Nair, R. Switzer, D. Shapiro and J. Gerlt for many helpful discussions and critical reading of the manuscript. The work was supported by the start-up fund from the University of Illinois and a grant from the National Institute of Health.

REFERENCES

1. Lane,B.G. (1998) Historical perspectives on RNA nucleoside modifications. In Grosjean,H. and Benne,R. (eds), *Modification and Editing of RNA*. ASM Press, Washington, DC, pp. 1–20.
2. Auffinger,P. and Westhof,E. (1998) Effects of pseudouridylation on tRNA hydration and dynamics: a theoretical approach. In Grosjean,H. and Benne,R. (eds), *Modification and Editing of RNA*. ASM Press, Washington, DC, pp. 103–112.
3. Bakin,A., Lane,B.G. and Ofengand,J. (1994) Clustering of pseudouridine residues around the peptidyltransferase center of yeast cytoplasmic and mitochondrial ribosomes. *Biochemistry*, **33**, 13475–13483.
4. Yu,Y.T., Shu,M.D. and Steitz,J.A. (1998) Modifications of U2 snRNA are required for snRNP assembly and pre-mRNA splicing. *EMBO J.*, **17**, 5783–5795.
5. Koonin,E.V. (1996) Pseudouridine synthases: four families of enzymes containing a putative uridine-binding motif also conserved in dUTPases and dCTP deaminases. *Nucleic Acids Res.*, **24**, 2411–2415.
6. Foster,P.G., Huang,L., Santi,D.V. and Stroud,R.M. (2000) The structural basis for tRNA recognition and pseudouridine formation by pseudouridine synthase I. *Nat. Struct. Biol.*, **7**, 23–27.
7. Hoang,C. and Ferre-D'Amare,A.R. (2001) Cocrystal structure of a tRNA Psi55 pseudouridine synthase: nucleotide flipping by an RNA-modifying enzyme. *Cell*, **107**, 929–939.
8. Sivaraman,J., Sauve,V., Larocque,R., Stura,E.A., Schrag,J.D., Cygler,M. and Matte,A. (2002) Structure of the 16S rRNA pseudouridine synthase RsuA bound to uracil and UMP. *Nat. Struct. Biol.*, **9**, 353–358.
9. Sivaraman,J., Iannuzzi,P., Cygler,M. and Matte,A. (2003) Crystal structure of the RluD pseudouridine synthase catalytic module, an enzyme that modifies 23S rRNA and is essential for normal cell growth of *Escherichia coli*. *J. Mol. Biol.*, **335**, 87–101.
10. Kaya,Y. and Ofengand,J. (2003) A novel unanticipated type of pseudouridine synthase with homologs in bacteria, archaea and eukarya. *RNA*, **9**, 711–721.
11. Huang,L., Pookanjanatavip,M., Gu,X. and Santi,D.V. (1998) A conserved aspartate of tRNA pseudouridine synthase is essential for activity and a probable nucleophilic catalyst. *Biochemistry*, **37**, 344–351.
12. Ramamurthy,V., Swann,S.L., Paulson,J.L., Spedaliere,C.J. and Mueller,E.G. (1999) Critical aspartic acid residues in pseudouridine synthases. *J. Biol. Chem.*, **274**, 22225–22230.
13. Conrad,J., Niu,L., Rudd,K., Lane,B.G. and Ofengand,J. (1999) 16S ribosomal RNA pseudouridine synthase RsuA of *Escherichia coli*: deletion, mutation of the conserved Asp102 residue and sequence comparison among all other pseudouridine synthases. *RNA*, **5**, 751–763.
14. Raychaudhuri,S., Niu,L., Conrad,J., Lane,B.G. and Ofengand,J. (1999) Functional effect of deletion and mutation of the *Escherichia coli* ribosomal RNA and tRNA pseudouridine synthase RluA. *J. Biol. Chem.*, **274**, 18880–18886.
15. Gu,X., Liu,Y. and Santi,D.V. (1999) The mechanism of pseudouridine synthase I as deduced from its interaction with 5-fluorouracil-tRNA. *Proc. Natl Acad. Sci. USA*, **96**, 14270–14275.
16. Mueller,E.G. (2002) Chips off the old block. *Nat. Struct. Biol.*, **9**, 320–322.
17. Pan,H., Agarwalla,S., Moustakas,D.T., Finer-Moore,J. and Stroud,R.M. (2003) Structure of tRNA pseudouridine synthase TruB and its RNA complex: RNA recognition through a combination of rigid docking and induced fit. *Proc. Natl Acad. Sci. USA*, **100**, 12648–12653.
18. MacFerrin,K.D., Chen,L., Terranova,M.P., Schreiber,S.L. and Verdine,G.L. (1993) Overproduction of proteins using the expression-cassette polymerase chain reaction. *Methods Enzymol.*, **217**, 79–102.
19. Otwinowski,Z. and Minor,W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods in Enzymology*. Academic Press, San Diego, CA, Vol. 277, pp. 307–326.
20. Terwilliger,T.C. and Berendzen,J. (1999) Automated MAD and MIR structure solution. *Acta Crystallogr. D Biol. Crystallogr.*, **55**, 849–861.
21. Terwilliger,T.C. (1999) Reciprocal-space solvent flattening. *Acta Crystallogr. D Biol. Crystallogr.*, **55**, 1863–1871.
22. Jones,T.A., Zou,J.-Y., Cowan,S.W. and Kjeldgaard,M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr.*, **A47**, 110–119.
23. Brünger,A.T., Adams,P.D., Clore,G.M., DeLano,W.L., Gros,P., Grosse-Kunstleve,R.W., Jiang,J.-S., Kuszewski,J., Nilges,M., Pannu,N.S. *et al.* (1998) Crystallography & NMR System: A New Software Suite for Macromolecular Structure Determination. *Acta Crystallogr.*, **D54**, 905–921.
24. Navaza,J. (1994) AMoRe: an automated package for molecular replacement. *Acta Crystallogr.*, **A50**, 157–163.
25. Gu,X., Yu,M., Ivanetich,K.M. and Santi,D.V. (1998) Molecular recognition of tRNA by tRNA pseudouridine 55 synthase. *Biochemistry*, **37**, 339–343.
26. Becker,H.F., Motorin,Y., Sissler,M., Florentz,C. and Grosjean,H. (1997) Major identity determinants for enzymatic formation of ribothymidine and pseudouridine in the T psi-loop of yeast tRNAs. *J. Mol. Biol.*, **274**, 505–518.
27. Romier,C., Reuter,K., Suck,D. and Ficner,R. (1996) Crystal structure of tRNA-guanine transglycosylase: RNA modification by base exchange. *EMBO J.*, **15**, 2850–2857.
28. Xie,W., Liu,X. and Huang,R.H. (2003) Chemical trapping and crystal structure of a catalytic tRNA guanine transglycosylase covalent intermediate. *Nat. Struct. Biol.*, **10**, 781–788.
29. Carson,M. (1987) Ribbon models of macromolecules. *J. Mol. Graph.*, **5**, 103–106.