Structure of the S15,S6,S18-rRNA Complex: Assembly of the 30*S* Ribosome Central Domain

Sultan C. Agalarov,[1,2]* G. Sridhar Prasad,[3]* Peter M. Funke,[1,4]* C. David Stout,[3]† James R. Williamson[1]†

The crystal structure of a 70-kilodalton ribonucleoprotein complex from the central domain of the *Thermus thermophilus* 30*S* ribosomal subunit was solved at 2.6 angstrom resolution. The complex consists of a 104-nucleotide RNA fragment composed of two three-helix junctions that lie at the end of a central helix, and the ribosomal proteins S15, S6, and S18. S15 binds the ribosomal RNA early in the assembly of the 30*S* ribosomal subunit, stabilizing a conformational reorganization of the two three-helix junctions that creates the RNA fold necessary for subsequent binding of S6 and S18. The structure of the complex demonstrates the central role of S15-induced reorganization of central domain RNA for the subsequent steps of ribosome assembly.

The recent explosion of structural information on the bacterial ribosome has set the stage for detailed models explaining both the function and the assembly of this large ribonucleoprotein (RNP) that connects genotype to phenotype through mRNA-templated polypeptide synthesis. Stunning low-resolution electron density maps of the 30*S* and 50*S* subunits and the 70*S* ribosome have recently appeared (*1–4*), in anticipation of atomic-resolution details of tRNA binding, mRNA translocation, and peptidyl transferase activity. Additionally, we are poised to address, at the molecular level, unresolved questions about the process of ribosome assembly, whereby ~50 ribosomal proteins and two large and one small ribosomal RNAs spontaneously assemble into a functional RNP.

The bacterial 30*S* ribosomal subunit is a large RNP with perhaps the greatest wealth of available biochemical and structural information. Composed of ~21 small-subunit ribosomal proteins, designated S1, S2, . . . S21, and the 1542-nucleotide 16*S* ribosomal RNA, the 30*S* subunit can be reconstituted from purified components in vitro, and the ordered nature of the assembly was revealed by the elegant work of Nomura (*5*) (Fig. 1A). Six proteins bind independently to 16*S* ribosomal RNA (rRNA), including S4, S7, S8, S15, S17, and S20. After assembly of these primary binding proteins, a second set of proteins binds the growing RNP, including S5, S6, S9, S12, S13, S16, S18, and S19. In turn, the secondary binding proteins potentiate binding of the remaining proteins, including S2, S3, S10, S11, S14, and S21.

The 30*S* subunit consists of the 5′, central, and 3′ domains, each of which can be assembled into an independently folding RNP complex (*6–8*). These functional domains correspond to the body, platform, and head, respectively, of the 30*S* particle. The central domain is nucleated by protein S15, after which proteins S6 and S18 bind cooperatively, followed by protein S11, and finally S21 (*5*) (Fig. 1A). Protein S8 is a primary binding protein that also binds to the central domain; however, it is not required for assembly of any of the other central domain proteins. The minimal binding site for S15 is localized near a three-helix junction in the central domain (*9*, *10*), and binding of S15 to this RNA is accompanied by a large conformational change in the junction region (*11*, *12*).

Recently, we identified by deletion analysis a core central domain RNA capable of binding proteins S15, S6, S18, and S11, and a smaller RNA fragment (*Tth* T4 RNA) capable of binding proteins S15, S6, and S18 (Fig. 2) (*13*). The *Tth* T4 RNA consists of helices 22 and 23a and portions of helices 20, 21, and 23b from 16*S* rRNA and contains both three-helix junctions that form the core of the central domain (*13*). This result was foreshadowed by earlier findings that fragments from the central domain of 16*S* rRNA were protected from ribonuclease by proteins S6, S8, S15, and S18 (*14*); however, it was somewhat surprising that half of the central domain RNA was dispensable for formation of the protein core structure. Proteins S8, S11, S6, and S18 each have hydroxyl-radical footprints in the core subdomain (Fig. 1B) (*15*) and in addition have secondary footprints to the accessory subdomain composed of helices 19, 24, 25, 26, 26a, and 27. Here we describe the structure of the *Tth* T4 RNP, the first atomic-resolution multiprotein complex from the ribosome, along with the insights gained into RNA-protein recognition and the ordered assembly of the 30*S* subunit.

[1]Department of Molecular Biology and the Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, CA 92037, USA. [2]Institute for Protein Research, Pushchino, Russia. [3]Department of Molecular Biology, The Scripps Research Institute, La Jolla, CA 92037, USA. [4]Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: dave@scripps.edu, jrwill@scripps.edu

Overview of the *Tth* T4 RNP Structure

The x-ray crystal structure of the *Tth* T4 RNP was determined by multiple isomorphous replacement (MIR) methods with seven heavy-atom derivatives and alternate rounds of model building and refinement (Table 1 and Fig. 3). The 70-kD *Tth* T4 RNP forms a noncrystallographic symmetry (NCS)–related dimer in the asymmetric unit and has many intermolecular RNA-protein and RNA-RNA contacts. Electron density was not observed for several terminal bases in helix 20, for 17 bases in helix 23b, or for the first 35 NH_2-terminal residues of protein S18. Except for minor differences, the two copies of the *Tth* T4 RNP in the asymmetric unit have similar structures.

Several features in the *Tth* T4 RNA are important for RNA tertiary structure and protein recognition (Fig. 2B). The lower three-helix junction is formed by coaxial stacking of helix 21 and helix 22, with helix 20 at an acute angle to helix 22. The upper three-helix junction is formed by coaxial stacking of helix 23b on helix 22, with short helix 23a folded onto helix 22. The continuous, coaxially stacked portions of helices 21, 22, and 23b form an extended structure that is roughly 75 Å long. The bulged nucleotide C748 and the purine-rich internal loop of helix 22 result in a gradual 40° bend, orienting helices 20 and 23a toward each other on one face of helix 22. The *Tth* T4 RNP structure is extremely similar to the conformation reported in the 5.5 Å structure of the 30*S* ribosomal subunit and is consistent with neutron-scattering studies (*16*).

The lower three-helix junction is stabilized by non–Watson-Crick base pairs between phylogenetically conserved nucleotides among eubacterial 16*S* rRNAs (Fig. 2B) (*17*). The bases U652 and A753 form a reverse Hoogsteen base pair that stacks on helix 21 (Fig. 2D). In addition, the U652 O4 group, which is not directly involved in this base-pairing interaction, extends directly across the junction to form a hydrogen bond with the G752 O2′ on the opposite strand. Above this A:U base pair is a triple-base interaction between junction nucleotides G654 and G752 and residue C754 in helix 20 (Figs. 2, B and D, and 3). A sharp bend in the RNA backbone between G752 and C754, characterized by C_2′-endo ribose conformations, positions C754 above the U652:A753 pair and places helix 20 at an acute angle relative to helix 22. The base of C754 adopts the

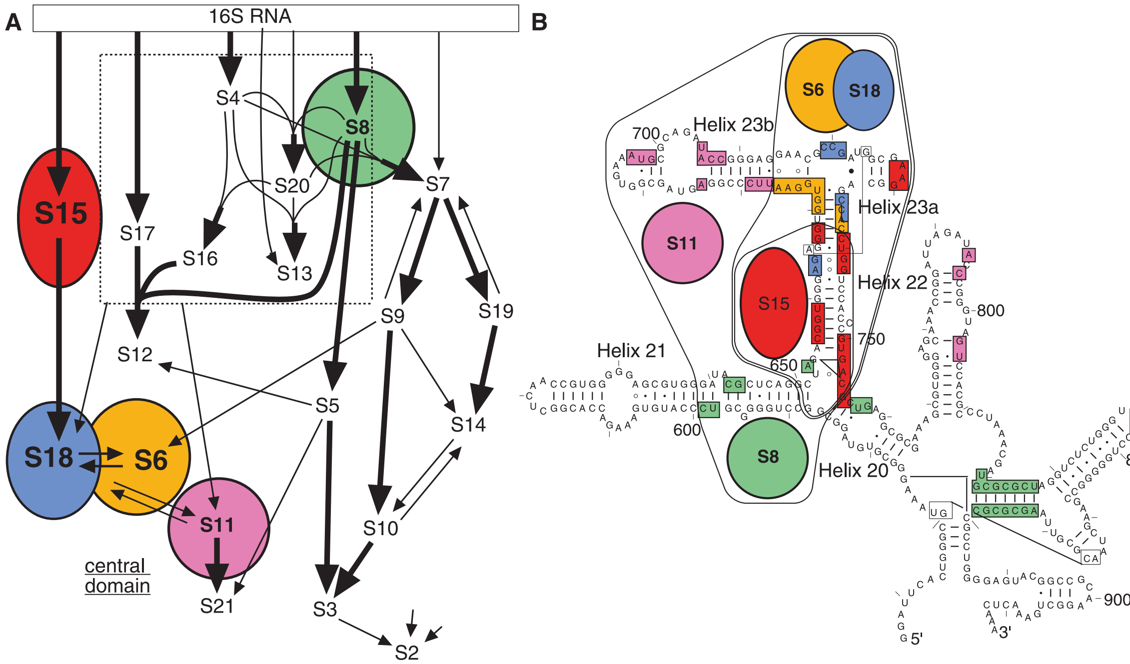

Fig. 1. Assembly of the 16*S* RNA central domain. (**A**) The Nomura 30*S* assembly map (*5*). S15 potentiates the binding of S6 and S18, then S11, and finally S21, which together constitute the central domain. There is no S21 protein in *T. thermophilus*. (**B**) Central domain protein contacts. The S15 minimal RNP, *Tth* T4 RNP, and central domain core RNP are outlined. Protein contacts from the *Tth* T4 RNP structure and from hydroxyl-radical footprinting experiments done in *Escherichia coli* (*15*) are combined and mapped onto the *T. thermophilus* central domain rRNA. Footprints are color-coded for S8 (green), S15 (red), S6 (orange), S18 (cyan), and S11 (pink).

Table 1. Crystallographic analysis. *Tth* T4 RNPs were prepared by reconstitution of *Tth* T4 RNA with a mixture of core proteins from the *T. thermophilus* 30*S* ribosomal subunit (S6, S8, S11, S15, S17, S18) as described (*13*). Crystals of the *Tth* T4 RNP were grown at room temperature to 0.5 mm by 0.2 mm by sitting-droplet vapor diffusion methods from 1.8 M $(NH_4)_2SO_4$, 20 mM $MgCl_2$, and 50 mM K^+cacodylate (pH 6.5). The crystals, belonging to the space group $P6_5$ with unit cell parameters a = 169.5 Å, b = 169.5 Å, and c = 113.8 Å, contain an NCS-related *Tth* T4 dimer in the asymmetric unit and have 75% solvent by volume. Heavy-atom soaks were carried out at 1 to 10 mM added metal ion in mother liquor for 12 to 48 hours. Crystals were transferred to a cryoprotectant containing 20% glycerol in mother liquor before direct freezing in liquid nitrogen (77 K). Data were collected at 100 K at Stanford Synchrotron Radiation Laboratory beam lines 9-1 (λ 0.98 Å) and 7-1 (λ 1.08 Å) and on a Rigaku FR rotating copper anode (λ 1.54 Å) with Mar345 image plates. The data were processed and scaled with Mosflm and CCP4 programs (*30*). Heavy-atom derivative sites were confirmed by Patterson and difference Fourier methods with XtalView (*31*). Experimental phases were obtained by MIR to 3.5 Å with Mlphare (*32*) using seven heavy-atom derivative data sets. These phases were extended to 2.8 Å by NCS averaging, solvent flattening, histogram matching, and SigmaA phase combination from model building with DM (*33*). The NCS averaging was used only in the initial phase improvements because of small differences in the two copies of the *Tth* T4 RNP. All model building was done with Xfit (*31*). R_{free} calculations were performed on 2.3% of the data, and the model was refined by iterative rounds of positional and *B*-factor refinement to 2.6 Å with CNS (*34*). Residues 35 to 40 of S18 are modeled as polyalanine owing to weak electron density for the side chains.

	Native	$UO_2(NO_3)_2$	$UO_2(NO_3)_2$	$HgCl_2$	EMP*	K_2HgI_4	cis-$Pt(NH_3)_2Cl_2$	trans-$Pt(NH_3)_2Cl_2$
				Data				
λ (Å)	0.98	1.08	1.54	0.98	0.98	0.98	0.98	0.98
Resolution (Å)	2.60	2.98	3.44	4.00	4.50	4.20	3.90	5.47
Total reflections	264,392	169,248	98,705	16,943	40,192	48,325	61,976	13,672
Unique reflections	56,665	37,918	24,626	10,246	11,317	13,591	16,941	4,011
Completeness† (%)	99.9	99.7 (100)	99.9 (99.7)	64.4 (70.5)	99.9 (100)	99.9 (99.3)	99.9 (100)	81.7 (83.7)
$I/\sigma(I)$	9.9 (1.7)	8.0 (2.4)	3.6 (1.9)	2.8 (2.0)	2.6 (1.9)	2.4 (2.0)	3.7 (1.9)	3.4 (2.0)
R_{symm} (I)‡ (%)	5.1 (39.9)	6.7 (30.8)	14.7 (38.3)	18.2 (36.9)	23.1 (37.5)	25.1 (35.8)	17.5 (38.9)	13.5 (38.2)
				Phasing				
No. of sites		2	2	2	2	6	6	7
R_{iso}§ (%)		20.1	17.2	16.8	19.1	18.5	14.7	31.8
R_{cullis}‖		0.72 (0.83)	0.74 (0.73)	0.88 (0.97)	0.88 (0.95)	0.79 (0.89)	0.82 (0.96)	0.85 (0.91)
Phasing Power¶		1.57 (1.32)	1.44 (1.67)	0.80 (0.61)	0.78 (0.67)	1.24 (1.01)	1.11 (0.77)	1.03 (0.86)
Figure of merit	0.550 (0.321) for 23,473 reflections to 3.5 Å							

Refinement

Resolution range (Å)	50.0–2.60	rms deviations		Average *B* factor (Å^2) (copy A/copy B)	
Reflections	55,857	Bonds (Å)	0.024	T4 RNA (1874/1834 atoms)	50.2/55.9
R_{cryst}# (%)	26.6	Angles (°)	2.22	S6 (825/825 atoms)	55.6/59.7
R_{free} (%)	29.8	Dihedrals (°)	19.5	S15 (737/737 atoms)	41.5/49.9
		Impropers (°)	1.8	S18 (376/376 atoms)	79.1/75.5

*EMP, ethyl mercuric phosphate. †Value for highest resolution shell in parentheses. ‡$R_{symm}(I) = \Sigma_{hkl}\Sigma_j |I_{hkl,j} - \langle I_{hkl}\rangle| / \Sigma_{hkl}\Sigma_j |I_{hkl,j}|$, where $\langle I_{hkl}\rangle$ is the mean intensity of the multiple $I_{hkl,j}$ observations for symmetry-related reflections. §$R_{iso} = \Sigma\, \| F_{ph}| - |F_p\| / \Sigma |F_p|$, where $|F_{ph}|$ and $|F_p|$ are the derivative and native structure factor amplitudes, respectively. ‖$R_{cullis} = \Sigma\, |(|F_{ph}| - |F_p|) - |f_{calc}|| / \Sigma |F_{ph}|$ where $|f_{calc}|$ is the calculated heavy-atom structure factor amplitude for centric reflections. ¶The phasing power is the ratio of the mean calculated heavy-atom structure factor amplitude to the mean lack of closure error. #$R_{cryst} = \Sigma\, \|F_o| - |F_c\| / \Sigma |F_o|$.

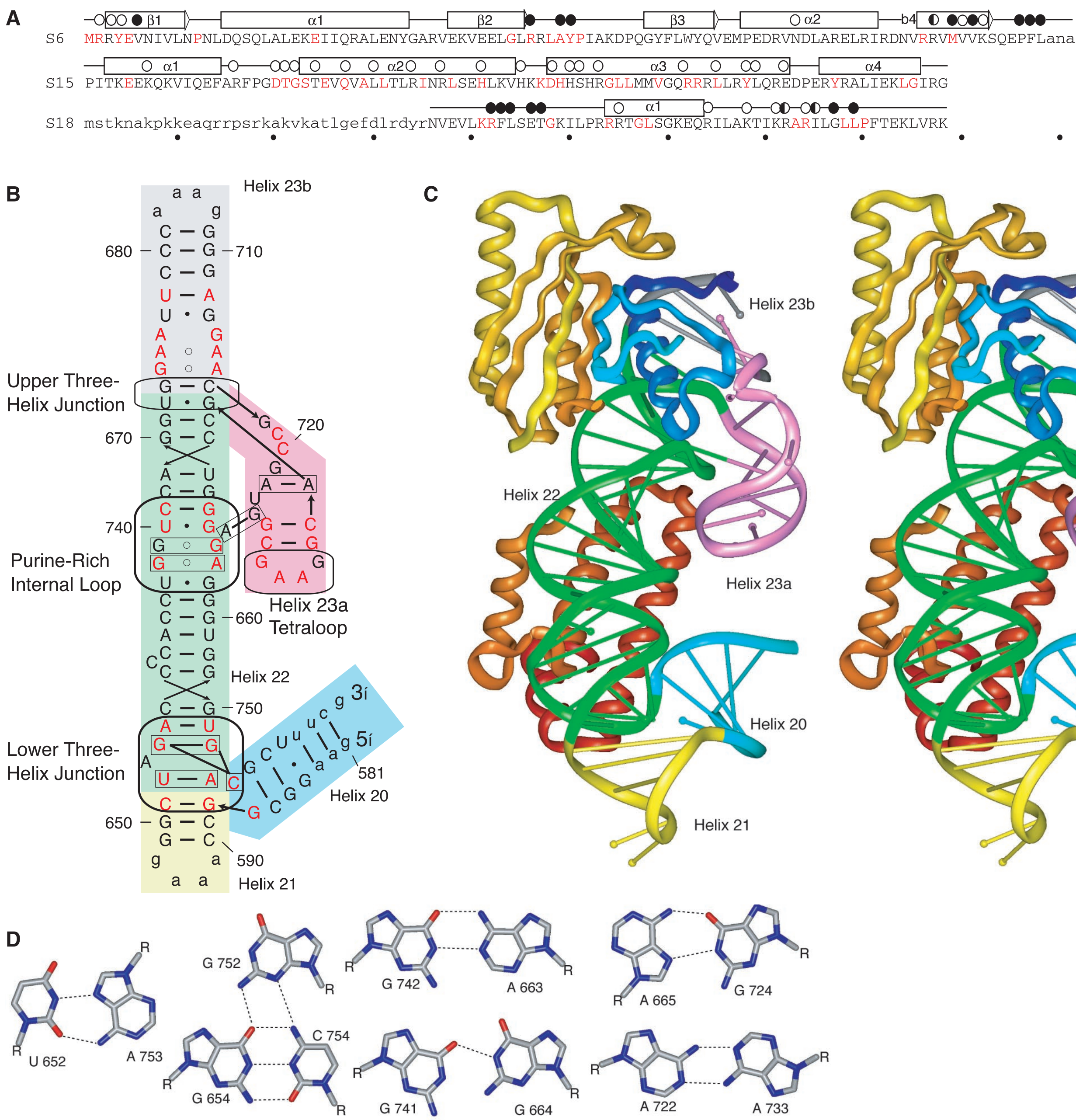

Fig. 2. *Tth* T4 RNP overview. (**A**) Sequence of proteins S6, S15, and S18 from *T. thermophilus* (*19*, *24*, *35*). Residues in lowercase are not observed in the electron density, and residues 35 to 40 in S18 are modeled as polyalanine. Colored residues are conserved >80% across six prokaryotes (*E. coli*, *T. thermophilus*, *Bacillus subtilis*, *Mycobacterium tuberculosis*, *Haemophilus influenzae*, *Helicobacter pylori*). Open circles indicate residues that make close contacts to the RNA (<3.5 Å), filled circles indicate residues involved in S6:S18 protein-protein contacts, and secondary structure elements are indicated above each sequence. Abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr. (**B**) Sequence of the *Tth* T4 RNA from the central domain of *T. thermophilus* 16*S* rRNA. The secondary structure and the general topology of the tertiary structure are shown. Helical domains are color-coded as follows: helix 20 (blue), helix 21 (yellow), helix 22 (green), helix 23a (pink), and helix 23b (gray). Bases in red are conserved >95% across all known eubacterial sequences (*17*), and the residue numbering is consistent with the *E. coli*. sequence. Bases in lowercase were added to close truncated helix 21 and helix 23b and to stabilize helix 20. The five RNA helices are connected by two separate three-helix junctions at the ends of helix 22. In the lower junction, helix 21 stacks coaxially under helix 22, and helix 20 makes an acute angle with helix 22. In the upper junction, helix 23a folds down parallel to helix 22, and helix 23b coaxially stacks on helix 22. Noncanonical base pairs are indicated by rectangular boxes. (**C**) Stereo ribbon diagram of the *Tth* T4 RNP. Nucleotides 676 to 716 are not observed in the electron density, nor are S18 residues 1 to 34. The RNA helices are colored as in (B), S15 is red, S6 is orange, and S18 is blue. Figure created with InsightII. (**D**) Noncanonical base pairs. Bases are rendered as sticks, ribose moieties are labeled R, hydrogen-bonding interactions are indicated by dashed lines, and atoms are color-coded as follows: carbon (gray), nitrogen (blue), and oxygen (red). The U652:A753 reverse Hoogsteen pair and the G654:G752:C754 base-triple are found in the lower three-helix junction, as shown in (B). The A663:G742 and G664:G741 base pairs are found in the purine-rich loop in helix 22, the interhelical A665:G724 base pair forms between helix 22 and helix 23a, and the symmetric A722:A733 pair closes helix 23a.

syn conformation and forms a Watson-Crick pair with G654, and both of these bases hydrogen bond with G752 to form the triple base. Although residue G587 is without a formal base-pairing partner in helix 20, it stacks on the end of helix 20, with the guanine N1 and N2 forming hydrogen bonds across the junction to A753 O2′ and C754 phosphate, respectively.

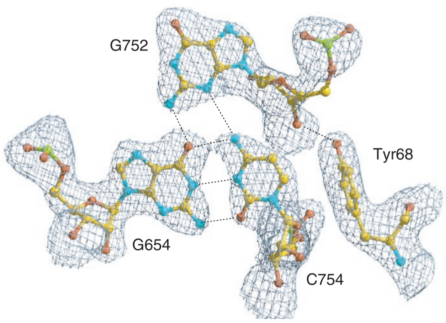

Fig. 3. Electron density for the *Tth* T4 RNP triple-base–S15 interaction at 2.6 Å resolution. The map is calculated with all data in the resolution range 37.7 to 2.6 Å with σ_A weighted coefficients $2m|F_o| - D|F_c|$, contoured at 1.2σ. Nucleotides G654, G752, and C754 in the lower three-helix junction are shown interacting with Tyr[68] from S15. Figure created with Xfit.

Two noncanonical base-pairing interactions are found in the highly conserved purine-rich internal loop of helix 22, including the G742:A663 base pair and the G741:G664 base pair (Fig. 2, B and D). The internal loop also contains an unanticipated tertiary interaction, in which A665 is flipped out of helix 22 and inserted into helix 23a, forming a base pair with G724 and stacking within helix 23a. This interhelical base pair fixes the orientation of helix 23a with respect to helix 22, thereby stabilizing the global conformation of the nearby upper three-helix junction. Immediately above the A665:G724 pair in helix 23a is the symmetric A722:A733 base pair, consistent with the observed covariation of these two positions as either A:A or G:G (*17*).

The S15 protein, a highly basic four–α-helix bundle, binds to the *Tth* T4 RNA along helix 22 by making contacts to the lower three-helix junction, to the minor groove of helix 22 above the purine-rich internal loop, and to the GAAG tetraloop in helix 23a (Figs. 2C and 4). The S15 contacts to the lower three-helix junction stabilize its tertiary fold, while the S15 contacts above the internal loop of helix 22 and to the tetraloop of helix 23a stabilize the tertiary fold of the upper three-helix junction. The tertiary structure of the upper three-helix junction forms

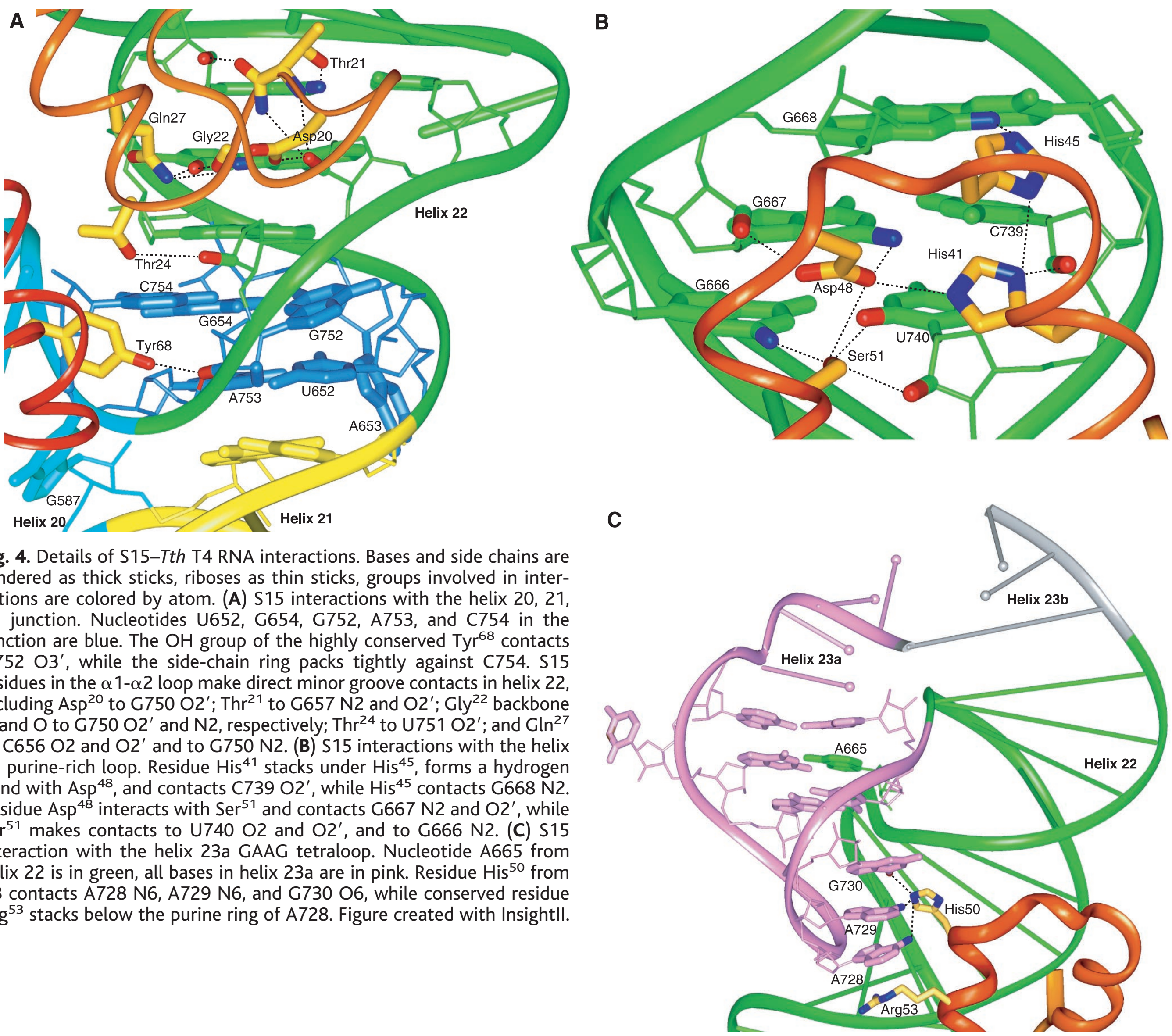

Fig. 4. Details of S15–*Tth* T4 RNA interactions. Bases and side chains are rendered as thick sticks, riboses as thin sticks, groups involved in interactions are colored by atom. (**A**) S15 interactions with the helix 20, 21, 22 junction. Nucleotides U652, G654, G752, A753, and C754 in the junction are blue. The OH group of the highly conserved Tyr[68] contacts G752 O3′, while the side-chain ring packs tightly against C754. S15 residues in the α1-α2 loop make direct minor groove contacts in helix 22, including Asp[20] to G750 O2′; Thr[21] to G657 N2 and O2′; Gly[22] backbone N and O to G750 O2′ and N2, respectively; Thr[24] to U751 O2′; and Gln[27] to C656 O2 and O2′ and to G750 N2. (**B**) S15 interactions with the helix 22 purine-rich loop. Residue His[41] stacks under His[45], forms a hydrogen bond with Asp[48], and contacts C739 O2′, while His[45] contacts G668 N2. Residue Asp[48] interacts with Ser[51] and contacts G667 N2 and O2′, while Ser[51] makes contacts to U740 O2 and O2′, and to G666 N2. (**C**) S15 interaction with the helix 23a GAAG tetraloop. Nucleotide A665 from helix 22 is in green, all bases in helix 23a are in pink. Residue His[50] from α3 contacts A728 N6, A729 N6, and G730 O6, while conserved residue Arg[53] stacks below the purine ring of A728. Figure created with InsightII.

the binding site for the proteins S6 and S18, which bind cooperatively as a heterodimer (*18*). The S6 protein, a mildly acidic four-stranded antiparallel β sheet flanked by two α helices, makes RNA contacts along the minor groove at the junction of helix 22 and helix 23b (Figs. 2C and 5). The S18 protein, which consists of an α helix surrounded by an ordered polypeptide coil, binds to S6 along its β sheet and makes contacts to the RNA backbone in the upper three-helix junction and to single-stranded bases in helix 23a.

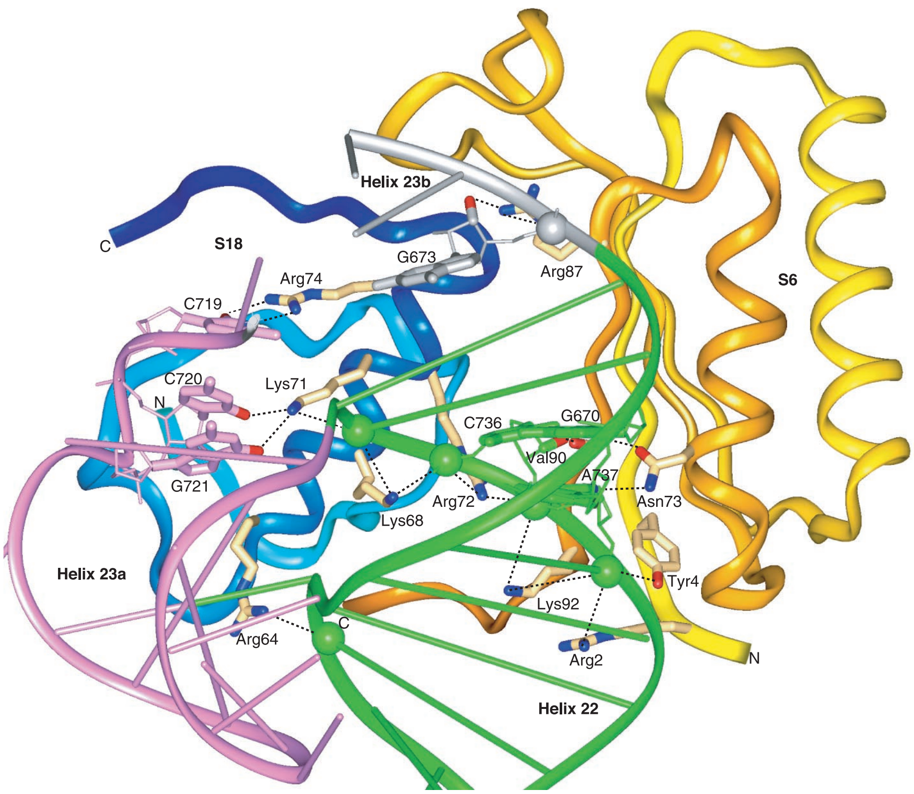

Fig. 5. Details of S6:S18–*Tth* T4 RNA interaction. Molecules are rendered and colored as in Fig. 4, with phosphate groups shown as spheres. S6 residues located near the NH_2-terminus, in α2, and in β4 make electrostatic and hydrogen-bonding contacts to the *Tth* T4 RNA in the minor groove between helix 22 and helix 23b. These contacts include Arg^{2}, Tyr^{4}, and Lys^{92} to A737 and C738 phosphates, Arg^{87} to G673 phosphate and O3′, Val^{90} carbonyl oxygen to C736 O2′, and Asn^{73} to G670 N2 and A737 N3. The charged S18 residues Lys^{68}, Lys^{71}, and Arg^{72}, from the COOH-terminal end of the α helix, contact the phosphate groups of C735, C736, and A737 in helix 22 near the upper three-helix junction. Residue Arg^{64}, which is located near the other end of the S18 α helix, contacts the G664 phosphate located across the narrowed major groove of helix 22 near the interhelical A665:G724 base pair. Residues Lys^{71} and Arg^{74} also make four base-specific contacts to the single-stranded nucleotides C719, C720, and G721 in helix 23a. Figure created with InsightII.

The Protein-rRNA Interfaces

The S15 helices α1, α2, and α3 form a planar, slightly twisted RNA binding face, with the α-helical axes aligned roughly parallel to helix 22 (Fig. 2C). In the *Tth* T4 RNP, S15 α1 packs tightly with the other helices, similar to the nuclear magnetic resonance structure of the free protein (*19*) but unlike the crystal structure of the free protein (*20*), in which α1 lies distal to the core (*21*). There are three principal regions of S15 that make specific contacts to the RNA. Residues located both in the loop region between helices α1 and α2 and in the COOH-terminal end of helix α3 interact with the RNA backbone of the lower three-helix junction and with adjacent nucleotides in the minor groove of helix 22 (Fig. 4A). At the opposite end of the S15 protein, residues in the α2-α3 loop interact with the minor groove of helix 22 above the purine-rich internal loop, one helical turn away from the lower three-helix junction (Fig. 4B). Residues in and near the α2-α3 loop also make direct contact with the GAAG tetraloop in helix 23a (Fig. 4C). There are no protein-protein contacts between S15 and either S6 or S18, consistent with conclusions based on neutron-scattering experiments (*22*). The solvent-accessible surface of S15 makes no contacts in the small subunit, but forms an intersubunit bridge with the 715 loop in 23*S* rRNA (*23*).

Proteins S6 and S18 bind across the upper three-helix junction, making contacts to the minor groove of helix 22 and helix 23b, to single-stranded nucleotides in helix 23a, and to the folded RNA backbone (Figs. 2C and 5). The S6 binding site for S18 is a concave surface made of one strand of the β sheet, the loop between β2 and β3, and the extended COOH-terminal coil. Residues from S6, located on the edge of the protein formed by α2, β4, and the NH_2-terminus, contact the minor groove of helix 22 and helix 23b in the upper three-helix junction. The structure of S6 in the *Tth* T4 RNP is similar to the crystal structure of the free protein (*24*), with most of the differences located in the loop regions and at the termini.

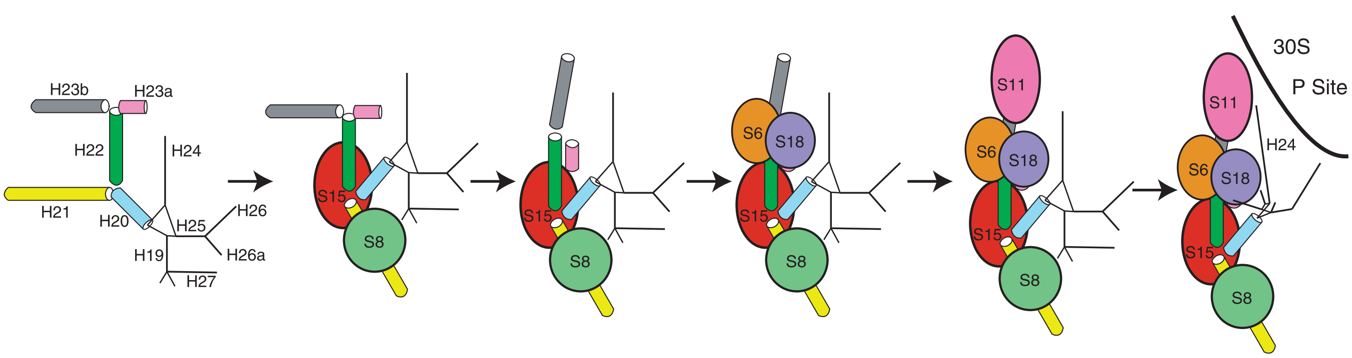

Fig. 6. Assembly mechanism for the central domain. The primary binding proteins S15 and S8 bind independently to the central domain of 16*S* rRNA early in the assembly process. S15 binding is coupled to a conformational change in the lower three-helix junction, in which helices 21 and 22 coaxially stack and helix 20 forms an acute angle with helix 22. Subsequently, the upper three-helix junction undergoes an S15-induced conformational change, thus creating the binding site for the heterodimer of proteins S6 and S18. Once these two proteins have bound the growing RNP, protein S11 binds to complete the "core" of the central domain (*13*). Finally, the remainder of the central domain rRNA assembles onto the core, forming the functional elements of the 30*S* ribosomal P-site.

The S18 protein, unlike S15 and S6, contains a single small element of regular secondary structure, yet it forms a compact structure tightly packed against S6 and the RNA (Figs. 2C and 5). Residues along one face of the α helix and residues 42 to 47 from the coil region form the protein-protein interface with S6, which is characterized by van der Waals contacts and salt-bridge interactions. The S18 α helix lies across the upper three-helix junction and contacts phosphates in helix 22 and single-stranded nucleotides in helix 23a.

Central Domain Assembly

Based on the array of biochemical data and the insights gained from the *Tth* T4 RNP structure, we propose a model for the assembly mechanism of the central domain of the 30*S* ribosomal subunit (Fig. 6). Biochemical and biophysical characterization of the lower three-helix junction indicates that the angle between these helices is ~120° in the absence of either protein S15 or Mg^{2+} ions (*12*). Binding of S15 is accompanied by a conformational change in the RNA whereby helix 20 forms an acute angle with helix 22, and helices 21 and 22 are coaxially stacked. The existence of these two conformations of the lower junction is supported by gel-mobility and transient electric birefringence studies, and the conformation of the bound junction is clearly seen in the structure of the *Tth* T4 RNP (Fig. 2C) and in the structure of the 30*S* subunit (*1*).

Biochemical studies of the S15-rRNA interaction indicated that the upper junction and helix 23b can be deleted with no detectable decrease in the binding affinity of S15 (*9*, *10*). Therefore, we propose that stabilization of the tertiary structure near the upper three-helix junction, which is the binding site for proteins S6 and S18, occurs subsequent to S15 binding. Nucleotides in the upper three-helix junction show enhanced sensitivity to chemical probes upon S15 binding and subsequent protection from these probes upon S6:S18 binding (*18*). These data are consistent with a conformational change in the upper three-helix junction upon S15 binding. In fact, protections in the GAAG tetraloop of helix 23a led to the proposal that S15 was a tetraloop binding protein (*25*). Although helix 23a and its GAAG tetraloop are dispensable for S15 binding to rRNA, S15 does make contacts to the GAAG tetraloop in the *Tth* T4 RNP complex.

Furthermore, the internal loop of helix 22 is not important for S15 binding because it can be replaced by Watson-Crick base pairs in a triple-mutant RNA that has a continuous helix 22 and shows wild-type affinity for S15 (*26*). To test our assembly hypothesis, we created this mutant (G663C, G664C, A665Δ) in the *Tth* T4 RNA. The internal loop of helix 22 was replaced by G:C pairs, and the interhelical A665:G724 base pair, which stabilizes the upper three-helical junction, was disrupted. Reconstitution of this mutant *Tth* T4 RNA with central domain proteins gave an RNP that bound protein S15 normally, showed weak (~10%) binding to protein S6, and exhibited no binding to protein S18 (*27*). This result strongly supports the role of S15 in the stabilization of the RNA tertiary structure in the upper junction that is required for S6:S18 binding. Binding of proteins S6 and S18 has long been known to be cooperative (*5*), but the thermodynamic details of their association are not yet known. Because the structure of S18 is quite irregular, it is unlikely that S18 is folded alone. It is more likely that S18 folds upon binding to S6 to make an RNA-binding heterodimer or that S6 weakly associates with the S15-RNA complex that serves as a scaffold for cooperative folding and assembly of S18.

The subsequent steps in central domain assembly, consistent with the available biophysical information, are also shown in Fig. 6. The protein S8 binds independently of the other central domain proteins and is depicted in the model binding to helix 21 early in assembly, in parallel with S15. After binding of S6:S18, protein S11 can bind to complete the core RNP structure. Once the core is formed, the secondary subdomain of helices 19, 24, 25, 26, 26a, and 27 can assemble onto the core RNP scaffold (*13*).

Interestingly, highly conserved regions of this secondary subdomain that are implicated in ribosome function are not part of the structural core of the central domain. The 690 loop of helix 23b and the 790 loop in helix 24 have both been implicated in P-site tRNA binding (*28*). Helix 27, which lies at the interface between the 5′, central, and 3′ domains, has been implicated as a functional switch in translation (*29*). Apparently, the functionally important and potentially flexible regions of the central domain RNA are not involved in directing assembly of the domain but rather are displayed on the surface of a preassembled RNP core. Hence, formation of the core of the central domain is a prerequisite for organization of subsequent structures essential for ribosome function.

Our studies indicate that the sequential assembly of the central domain is characterized by alternating rounds of RNA conformational change and protein binding. The primary binding protein S15 stabilizes a specific rRNA tertiary structure in the upper three-helix junction necessary for subsequent protein binding and stabilizes a tertiary structure in the lower three-helix junction necessary for further assembly of other RNA helices onto this core structure. These events may reflect general principles of the assembly of large RNPs.

References and Notes

1. W. M. Clemons Jr. *et al.*, *Nature* **400**, 833 (1999).
2. A. Tocilj *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 14252 (1999).
3. N. Ban *et al.*, *Nature* **400**, 841 (1999).
4. J. H. Cate, M. M. Yusupov, G. Z. Yusupova, T. N. Earnest, H. F. Noller, *Science* **285**, 2095 (1999).
5. W. A. Held, B. Ballou, S. Mizushima, M. Nomura, *J. Biol. Chem.* **249**, 3103 (1974).
6. C. J. Weitzmann, P. R. Cunningham, K. Nurse, J. Ofengand, *FASEB J.* **7**, 177 (1993).
7. S. C. Agalarov *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 999 (1998).
8. R. R. Samaha, B. O'Brien, T. W. O'Brien, H. F. Noller, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 7884 (1994).
9. R. T. Batey and J. R. Williamson, *J. Mol. Biol.* **261**, 536 (1996).
10. A. A. Serganov *et al.*, *RNA* **2**, 1124 (1996).
11. R. T. Batey and J. R. Williamson, *RNA* **4**, 984 (1998).
12. J. W. Orr, P. J. Hagerman, J. R. Williamson, *J. Mol. Biol.* **275**, 453 (1998).
13. S. C. Agalarov and J. R. Williamson, *RNA* **6**, 402 (2000).
14. R. J. Gregory *et al.*, *J. Mol. Biol.* **178**, 287 (1984).
15. T. Powers and H. F. Noller, *RNA* **1**, 194 (1995).
16. The root means square deviation (rmsd) for 182 common CA atoms and 71 common P atoms from the 5.5 Å 30*S* structure [PDB entry 1GD7 (*1*)] and the *Tth* T4 RNP is 2.7 Å. The rmsd for the phosphates in the two copies of the T4 RNA in the asymmetric unit was 1.33 Å. Weak electron density was observed above helix 22 in the *Tth* T4 RNA that is consistent with the coaxially stacked arrangement of this helix in the 30*S* model, although we were not able to build an atomic model for this region.
17. R. R. Gutell, *Nucleic Acids Res.* **22**, 3502 (1994).
18. P. Svensson, L. Changchien, G. R. Craven, H. F. Noller, *J. Mol. Biol.* **200**, 301 (1988).
19. H. Berglund, A. Rak, A. Serganov, M. Garber, T. Hard, *Nature Struct. Biol.* **4**, 20 (1997).
20. W. M. Clemons Jr., C. Davies, S. W. White, V. Ramakrishnan, *Structure* **6**, 429 (1998).
21. rmsd values for the Cα atoms (Å): S15a versus S15b = 1.09, S15a versus PDB entry 1A32 (*20*) = 0.93, S15a versus PDB entry 1AB3 (*19*) = 3.26, where a and b refer to the two copies of S15 in the asymmetric unit.
22. M. S. Capel, M. Kjeldgaard, D. M. Engelman, P. B. Moore, *J. Mol. Biol.* **200**, 65 (1988). Distances between protein centers of mass in the *Tth* T4 RNP and in the 30*S* subunit as measured by neutron scattering, in parentheses: S6a-S15a = 42 Å (70 ± 5), S6a ± S18a = 22 Å (33 ± 3), S15a-S18a = 45 Å (70 ± 5). The center of mass of the *Tth* S18 protein was calculated without 34 COOH-terminal residues that are disordered in the structure.
23. G. M. Culver, J. H. Cate, G. Z. Yusupova, M. M. Yusupov, H. F. Noller, *Science* **285**, 2133 (1999).
24. M. Lindahl *et al.*, *EMBO J.* **13**, 1249 (1994) (PDB entry 1RIS); D. E. Otzen, O. Kristensen, M. Proctor, M. Oliveberg, *Biochemistry* **38**, 6499 (1999) (PDB entry 1LOU). rmsd values for the Cα atoms (Å): S6a-S6b = 0.46, S6a-IRIS = 0.89, S6a-1LOU = 1.89, S18a-S18b = 1.13, where a and b refer to the two copies of the asymmetric unit.
25. C. Zwieb, *Nucleic Acids Res.* **20**, 4397 (1992).
26. R. T. Batey and J. R. Williamson, *J. Mol. Biol.* **261**, 550 (1996).
27. A. C. Agalarov and J. R. Williamson, unpublished results.
28. D. Moazed and H. F. Noller, *J. Mol. Biol.* **211**, 135 (1990).
29. J. S. Lodmell and A. E. Dahlberg, *Science* **277**, 1262 (1997).
30. A. G. W. Leslie, *Acta Crystallogr.* **D50**, 760 (1994).
31. D. E. McRee, *J. Struct. Biol.* **125**, 156 (1999).
32. Z. Otwinowski, *Acta Crystallogr.* **D50**, 760 (1994).
33. K. Cowtan and P. Main, *Acta Crystallogr. D* **54**, 487 (1998).
34. A. T. Brunger *et al.*, *Acta Crystallogr. D* **54**, 905 (1998).

35. Supported by NIH grant GM53757 (J.R.W.). The coordinates of the *Tth* T4 RNP complex have been deposited in the Protein Data Bank (accession number 1EKC). We thank D. Shcherbakov for providing the *T. thermophilus* S18 sequence in advance of publication. We thank G. Joyce, R. T. Batey, J. D. Puglisi, and J. Dinsmore for critical review of the manuscript, and the staff at Stanford Synchrotron Radiation Laboratory for their assistance.

19 January 2000; accepted 9 March 2000