Ribosomal Protein S17: Characterization of the Three-Dimensional Structure by ^1H and ^{15}N NMR†

Barbara L. Golden,‡ David W. Hoffman,‡ V. Ramakrishnan,*,§ and Stephen W. White*,‡

Department of Microbiology, Duke University Medical Center, Durham, North Carolina 27710, and Department of Biology, Brookhaven National Laboratory, Upton, New York 11973

Received June 29, 1993; Revised Manuscript Received September 3, 1993®

ABSTRACT: The structure of ribosomal protein S17 from *Bacillus stearothermophilus* was investigated by two-dimensional homonuclear and heteronuclear magnetic resonance spectroscopy. The ^1H and ^{15}N chemical shift assignments are largely complete, and a preliminary structural characterization is presented. The protein consists of five β-strands that form a single antiparallel β-sheet with Greek-key topology. The β-strands are connected by several extended loops, and two of these contain residue types that are frequently seen in the RNA-binding sites of proteins. Additionally, two point mutations that affect antibiotic resistance, translational fidelity, and ribosome assembly are located in these two regions of the protein. Since these potential RNA-binding sites are distributed over a large surface of the protein, it appears that the molecule may interact with several regions of 16S rRNA.

A detailed understanding of ribosome function will be impossible without accurate structural information. To this end, a vast amount of data have been accumulated from this complex ribonucleoprotein particle by such diverse methods as electron microscopy (Frank et al., 1991; Stöffler-Meilicke & Stöffler, 1990), neutron scattering (Capel et al., 1987), crystallography (von Böhlen et al., 1991), chemical cross-linking (Lambert et al., 1983; Brimacombe et al., 1990), and chemical probing (Stern et al., 1989). As a result, the organization of the ribosome is now rather well determined at the current resolution limits of these techniques (Stern et al., 1988b; Brimacombe et al., 1988). In recent years, this information has been supplemented by high-resolution structural studies on individual ribosomal components, and detailed structures have now been determined for several ribosomal proteins and fragments of rRNA[1] (Leijonmarck et al., 1980; Wilson et al., 1986; Cheong et al., 1990; Heus & Pardi, 1991; Ramakrishnan & White, 1992). As more high-resolution information becomes available, it will be possible to integrate it into evolving models of the ribosome and to eventually understand the process of translation at the molecular level. Even with the few ribosomal protein structures currently known, insights have been gained into RNA–protein interactions, mechanisms of antibiotic resistance, and ribosome evolution (Leijonmarck et al., 1988; Ramakrishnan & White, 1992).

Early structural studies on ribosomal proteins were limited to those molecules that could be isolated without denaturation in large amounts from purified ribosomes (Appelt et al., 1981). These tend to be the so-called secondary and tertiary RNA-binding proteins that are located toward the periphery of the ribosome. With the development of cloning techniques for ribosomal protein genes (Ramakrishnan & Gerchman, 1991), it is now possible to isolate the primary RNA-binding proteins that specifically recognize sites on the naked rRNA and which initiate its folding. One such protein is S17, and here we present the features of its secondary and tertiary structure as determined by multidimensional homonuclear and heteronuclear NMR. In common with our previous studies on these proteins (Wilson et al., 1986; Ramakrishnan & White, 1992), the S17 molecule from *Bacillus stearothermophilus* was chosen because its extra thermostability makes it better suited for structural studies.

Ribosomal protein S17 has a molecular mass of 10 000 Da and is located in the small (30S) subunit of the ribosome. It is extremely well conserved in eubacteria (Yaguchi & Wittmann, 1978; Ohkubo et al., 1987; Henkin et al., 1989; Michalowski et al., 1990; Herfurth et al., 1991) and has counterparts in archaebacteria (Kimura & Kimura, 1987; Auer et al., 1989) and eukaryotes (Gantt & Thompson; 1990). Neutron diffraction (Capel et al., 1987, 1988) and immune electron microscopy (Stöffler-Meilicke & Stöffler, 1990) have both clearly shown that S17 is located at the "bottom" of the 30S subunit somewhat remote from the other small subunit proteins. Its closest protein neighbors are S8 and S15, and relatively long cross-linking agents are necessary to bridge the distance to S17 (Expert-Bezançon et al., 1977). Early studies on the assembly of the 30S subunit identified S17 as a primary 16S rRNA binding protein (Held et al., 1974). Subsequent results from protein–RNA cross-linking studies (Kyriatsoulis et al., 1986; Greuer et al., 1987) and chemical footprinting (Stern et al., 1988a, 1989) have shown that S17 is adjacent to 16S rRNA helices 7, 11, and 21 (Brimacombe, 1991). Mutations in S17 indicate that it is crucial to the assembly of the 30S subunit (Herzog et al., 1979) and that it contributes to translational fidelity (Bollen et al., 1975).

† This work was supported by Grant MCB9118369 from the National Science Foundation.

* To whom correspondence should be addressed.

‡ Duke University Medical Center.

§ Brookhaven National Laboratory.

® Abstract published in *Advance ACS Abstracts,* November 1, 1993.

[1] Abbreviations: DDS, sodium 4,4-dimethyl-4-silapentane-1-sulfonate; EDTA, ethylenediaminetetraacetic acid; IPTG, isopropyl-β-D-thiogalactopyranoside; PMSF, α-toluenesulfonyl fluoride; PCR, polymerase chain reaction; DQF-COSY, two-dimensional double quantum filtered *J*-correlated spectroscopy; HMQC-J, heteronuclear multiple quantum coherence *J*-correlated spectroscopy; HMQC, heteronuclear multiple quantum coherence spectroscopy; HSMQC, heteronuclear single- and multiple-quantum coherence spectroscopy; $J_{\alpha NH}$, $^3J_{NH\text{-}C\alpha H}$; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, two-dimensional nuclear Overhauser effect spectroscopy; rRNA, ribosomal RNA; TOCSY, two-dimensional total correlation spectroscopy.

0006-2960/93/0432-12812$04.00/0

MATERIALS AND METHODS

Cloning the Gene for S17 from B. stearothermophilus. The S17 gene was cloned from *B. stearothermophilus* genomic DNA by PCR methods using the general procedure described previously (Ramakrishnan & Gerchman, 1991). The method uses degenerate PCR primers based on the protein primary structure in the absence of a known DNA sequence. The required oligonucleotide degeneracy is considerably reduced by eliminating rare *B. stearothermophilus* codons as possible choices. It was necessary to include an initiator ATG codon at the 5′ end of the gene since the S17 protein sequence lacks an N-terminal methionine (Herfurth et al., 1991). The actual procedure used for S17 is as follows.

PCR amplification using VentR DNA polymerase (NE-Biolabs) was carried out under the following conditions: 40 cycles of melting at 94 °C for 1 min, hybridization at 52 °C for 1.5 min, and primer extension at 72 °C for 3 min. The amplified DNA was cloned into the *Nde*I–*Bam*HI site of the plasmid vector pET-13a, a derivative of pET-11a (Studier et al., 1990). pET-13a was constructed by ligating the 4.2 kb *Alw*nI–*Eco*RI fragment from pET-11a with the 1.2 kb *Alw*nI–*Eco*RI fragment from pET-9a. The former fragment contains the T7 promoter and the *Lac* I gene, and the latter contains the gene that confers kanamycin resistance (Studier et al., 1990).

Four independent clones were sequenced, and those that agreed with the consensus sequence were considered correct. The translated DNA sequence was identical to the published protein sequence apart from amino acid 51. Although this residue had previously been reported to be an arginine (Herfurth et al., 1991), every clone contained a histidine at this position. For protein overexpression, the S17 plasmid was introduced into *E. coli* strain BL21(DE3), which contains an IPTG inducible gene for T7 RNA polymerase.

Cell Growth. It was necessary to adopt several different strategies for growing the inducible cells depending upon the NMR samples to be prepared. For ^{1}H NMR, cells were grown in LB broth. For uniformly labeled ^{15}N samples, the cells were grown on minimal medium using 0.5 g/L $^{15}NH_4Cl$ as the sole nitrogen source. To selectively ^{15}N-label glycine and serine residues, cells were grown in minimal medium supplemented with 25 mg/L of each amino acid except glycine, serine, tryptophan and cysteine. At induction, 50 mg/L [α-^{15}N]glycine was added. Finally, lysine residues were selectively ^{15}N-labeled at the α-N position by growing the cells in minimal medium supplemented with 25 mg/L of each amino acid except lysine. At induction, 50 mg/L [α-^{15}N]-lysine was added to the culture. Cells were grown to an optical density of 0.8 and then induced with 100 mg/L of IPTG. Three hours after induction, the cells were harvested by centrifugation (10 min at 6000*g*).

Protein Purification. Cells were resuspended in buffer A (40 mM phosphate, pH 8.0, 200 mM NaCl, 0.1 mM PMSF, 0.05 mM benzamidine, 0.5 mM EDTA, 1 mM NaN_3, 0.5% w/v Tween-20). To lyse the cells, a few milligrams of chicken egg-white lysozyme was added, and the mixture was frozen at −20 °C overnight. After thawing, polyethyleneimine was added to a final concentration of 1% v/v to precipitate the DNA. Following centrifugation for 30 min at 12000*g* to remove insoluble material, the cell extract was loaded onto an S-Sepharose column equilibrated with buffer A, and the column was washed extensively with buffer A to remove unbound protein. The column was then washed with 10 volumes of buffer B (40 mM phosphate, pH 8.0, 200 mM NaCl, 1 mM NaN_3) to remove the Tween-20 and other small molecules in buffer A that would have eventually contaminated the NMR spectra. The protein was eluted using a linear NaCl gradient of 200–800 mM. Fractions containing pure protein eluted around 500 mM, and these were concentrated and transferred to NMR buffer using Centricon-10 microconcentrators (Amicon). Typical yields were between 20 and 30 mg/L of culture. NMR samples contain 2–3 mM (20–30 mg/mL) S17 in 20 mM phosphate, pH 6.5, and 200 mM NaCl. For spectra collected in H_2O solvent, the sample contained a mixture of 93% H_2O/7% D_2O. For spectra collected in 100% D_2O, the protein was rapidly transferred to the D_2O solution using a 5 cm × 0.25 cm G25 column equilibrated with the D_2O buffer.

NMR Spectroscopy. All spectra were acquired on a Varian Unity 500-MHz spectrometer. For ^{1}H spectra, the sweep width was 6100.4 Hz, and the chemical shifts were referenced relative to DDS. For ^{15}N spectra, the sweep width was 2000 Hz, and the chemical shifts were referenced using an external standard containing 2.9 M $^{15}NH_4Cl$ in 1 M HCl (24.93 ppm). All spectra were acquired in the phase-sensitive mode. Quadrature detection for the homonuclear and HMQC-J experiments was accomplished by the method of States and Haberkorn (States et al., 1982); all other experiments used the tppi method (Marion & Wüthrich, 1983). Low level presaturation of the water signal was used to suppress the solvent signal, and, in the NOESY and DQF-COSY experiments, this was followed by a SCUBA pulse sequence with 28-μs SCUBA delay (Brown et al., 1988). Typically, 1024 complex data points were collected in the F_2 dimension and 400–500 increments were acquired in the F_1 dimension. Spectra were resolution-enhanced by multiplying by a skewed sine bell function prior to Fourier transformation. In the HMQC-J experiment, a higher resolution spectrum consisting of 1024 increments in the ^{15}N dimension was acquired and zero-filled to 4096 points. The temperature was maintained at 25 °C during all the experiments unless otherwise indicated. Data were processed using the program FELIX 1.0 (Hare Research Inc., Woodenville, WA).

Proton assignments were made using ^{1}H NOESY, DQF-COSY (Wüthrich, 1986) and TOCSY (Davis & Bax, 1985) experiments. NOESY and TOCSY experiments were also acquired on samples in D_2O solution to identify slowly exchanging amide protons and to aid in the assignment of aliphatic side chains. Heteronuclear, proton-detected experiments were also carried out to make ^{15}N assignments and to measure the $J_{\alpha NH}$ coupling constants. Both uniformly ^{15}N-labeled and selectively ^{15}N-labeled samples were used in these experiments. HSMQC (Zuiderweg, 1990), HMQC-NOESY, HMQC-TOCSY, and HMQC-J (Gronenborn et al., 1989) experiments were performed on the uniformly ^{15}N-labeled protein sample. Mixing times were 180 ms for the NOESY and HMQC-NOESY experiments and 65 ms for the TOCSY and HMQC-TOCSY experiments. The NOESY pulse sequence included a 180° pulse in the middle of the mixing time. HSMQC experiments were performed on two selectively ^{15}N-labeled samples, one at lysine residues and the other at both glycine and serine.

Structure Determination. Sets of distance and ϕ angle constraints were assembled for a three-dimensional structure calculation of S17. The distance constraints were generated by searching the NOESY spectrum for cross peaks that could be unambiguously identified and which contained no spectral overlap. The distance constraints were classified as large (2.0–3.0 Å), medium (2.0–4.0 Å) and small (2.0–5.0 Å) depending

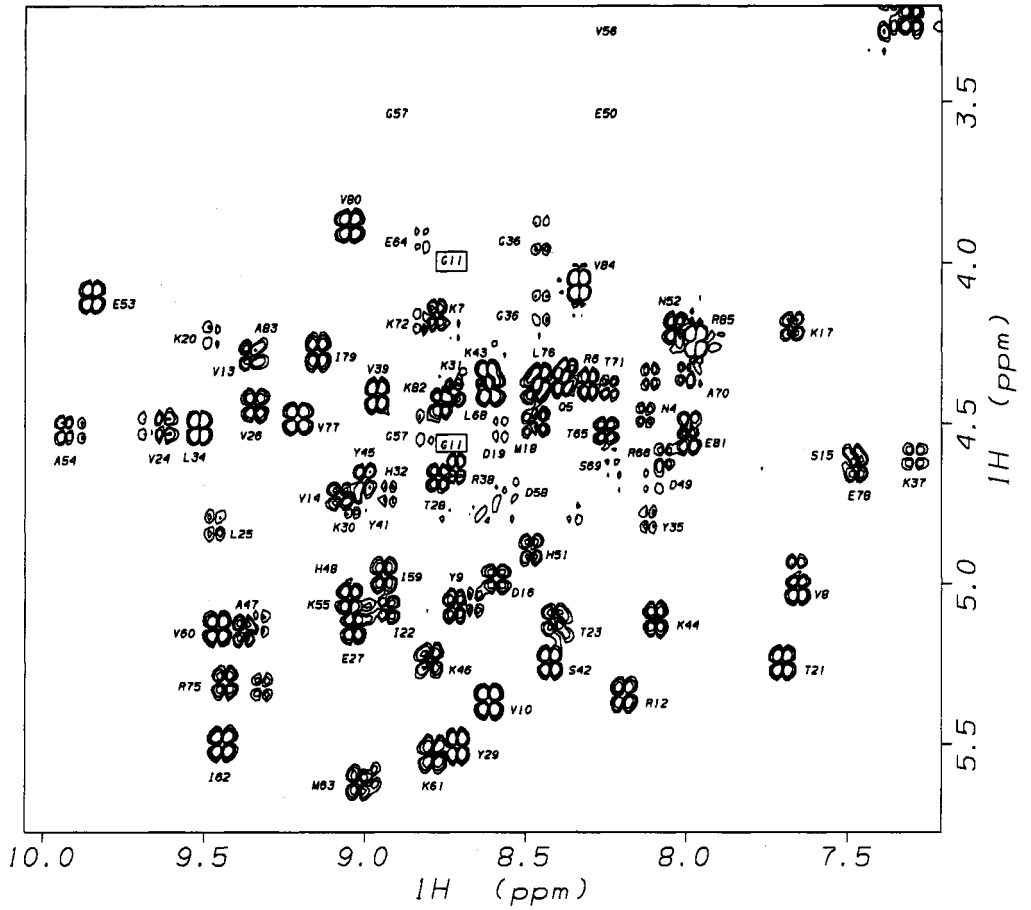

FIGURE 1: Fingerprint region of the DQF-COSY spectrum of ribosomal protein S17. Backbone proton assignments are complete for all residues except three at the N-terminal and arginine 73. Lysine 40 is not labeled due to spectral crowding in that region. Cross peaks for glycine 11, glutamic acid 50, glycine 57, and valine 56 are not visible at this contour level. Note that some of the cross peaks are doubled, for example, those of residues tyrosine 9, valine 24, and methionine 63.

on the intensities of the integrated NOE peaks. In addition, amide protons that were protected from exchange with deuterium solvent were considered to be involved in hydrogen-bonding interactions to carbonyl oxygen atoms within the protein's β-sheet secondary structure. These interactions were therefore constrained to between 1.7 and 2.1 Å. The hydrogen-bonding constraints were not applied to the amide protons of residues 77 and 78 which form a β-bulge. The ϕ angle constraints were calculated from the $J_{\alpha NH}$ coupling constants. These were obtained by measuring the splitting of the amide nitrogen peaks in the F_1 dimension of the HMQC-J spectrum. The ϕ angles were constrained to be $-60° \pm 15°$ (2–4 Hz), $-60° \pm 30°$ (4–4.7 Hz), $-120° \pm 30°$ (8–9 Hz), and $-120° \pm 15°$ (9 Hz or greater).

A low-resolution structure of S17 was determined from the NMR constraints using the program XPLOR (Brünger et al., 1986; Brünger, 1987; Nilges et al., 1988) operating on a Silicon Graphics 4D440S. Initially, a model of the protein was generated by distance geometry. This model was used as the starting structure for dynamical simulated annealing with NMR, geometric, and van der Waals terms included in the XPLOR energy function.

RESULTS

NMR Spectroscopy

(i) General Features of the NMR Spectra. The one- and two-dimensional NMR spectra of S17 demonstrate that the molecule has a well-defined structure. The methyl, α, and aromatic regions of the spectra show a high degree of dispersion of the proton resonances from typical random-coil chemical shifts. Of particular note are the large number of α-proton resonances that appear downfield of the water resonance above 4.7 ppm. This high dispersion is characteristic of proteins with a high content of β-sheet secondary structure (Wishart et al., 1992).

A reproducible feature of all the two-dimensional spectra is a doubling of a subset of the cross peaks (Figure 1 and Table I). This suggests that there exist in solution two structurally distinct populations of the protein. At 25 °C, the major form is present in several-fold excess over the minor form. Cross peaks were assigned to the major or minor form on the basis of NOE spectra which only correlated protons within a species. Such a phenomenon can be caused either by two slightly different chemical species or by two distinct conformations of the protein. An analysis of S17 by mass spectroscopy confirms that there is a single chemical species present in solution with the expected molecular mass (data not shown). This phenomenon does not appear to be the result of *cis–trans* isomerization by one of the two proline residues since neither contributes doubled sets of cross peaks. If two conformations of the protein do contribute to the spectra, they must interchange very slowly since there are no NOE peaks that relate identical protons in the alternate conformations.

(ii) ^1H and ^{15}N Assignments. Sequence-specific ^1H assignments were made by conventional methods, combining information from DQF-COSY/TOCSY spectra and the

Table I: ^1H and ^{15}N Assignments of Ribosomal Protein S17[a]

	amide proton	amide nitrogen	α	β	γ	other
N4	8.13		4.47			δ 6.85, 7.53
Q5	8.38		4.36	2.01, 2.15	2.39	δ 6.90, 7.57
R6	8.31		4.38	1.62, 1.73	1.82, 1.87	δ 3.18
K7	8.78	125.95	4.16	1.74	1.45	
	8.72		4.22			
V8	7.66	122.52	5.02	1.87	0.92	
	7.66		4.96	1.87	0.92	
Y9	8.72	124.65	5.07	2.09, 2.90		H2,H6 6.70
	8.66		5.05	2.03		
V10	8.61	123.40	5.37	1.96	0.90, 0.96	
G11	8.77	113.24	3.96, 4.52			
R12	8.20	121.38	5.34	1.73	1.48	
V13	9.36	128.49	4.29	2.51	0.86, 1.07	
V14	9.08	123.66	4.72	2.45	0.75, 1.00	
S15	7.48	117.08	4.61	3.74, 3.97		
D16	8.59	128.23	4.99	2.29, 3.21		
K17	7.67	121.28	4.20	1.91	1.55	δ 1.76; ε 3.08
M18	8.49	121.12	4.51	2.05	2.55, 2.63	
D19	8.58	125.38	4.52	2.60, 2.73		
K20	9.47	124.70	4.22	2.04	2.34	
	9.44		4.22			
T21	7.71	117.18	5.25	3.75	1.21	
I22	8.94	121.69	5.08	1.85	1.15, 1.50	γME 1.10; δ 0.70
	8.98		5.09			
T23	8.41	119.21	5.11	3.88	1.04	
	8.39		5.13			
V24	9.63	131.70	4.51	1.78	0.51, 0.72	
	9.68		4.51	1.78	0.51, 0.69	
L25	9.47	132.01	4.82	1.31		δ 0.75, 0.83
V26	9.34	133.26	4.45	1.95	0.62, 0.87	
E27	9.04	130.61	5.14	1.97	2.23	
T28	8.77	120.29	4.67	4.04	1.20	
Y29	8.71	123.09	5.51	2.78, 2.90		H2,H6 7.00; H3,H5 6.83
K30	9.04	123.35	4.76	1.79, 1.93	1.41, 1.31	δ 1.53; ε 2.74
K31	8.73	125.32	4.41	1.74	1.33, 1.42	δ 2.10; ε 3.05
H32	8.93	129.99	4.72	3.12, 3.26		H2 7.63; H4 7.29
P33			4.35	2.11, 2.30	1.81, 1.93	δ 2.80, 3.63
L34	9.52	121.85	4.52	1.33	1.43	δ 0.81, 0.92
Y35	8.12	119.36	4.80	2.58, 3.36		H2,H6 6.71
G36	8.45	110.24	3.92, 4.14			
K37	7.29	120.19	4.61	1.67	1.43	
R38	8.72		4.65	1.74, 1.82	1.52	
V39	8.96	126.05	4.41	2.14	0.90, 1.03	
K40	8.47	128.02	4.39			
Y41	8.99	123.46	4.75	2.87, 2.95		H2,H6 7.10; H3,H5 6.77
S42	8.43	119.78	5.25	3.59, 3.67		
K43	8.62	126.52	4.35	1.81	1.69	δ 1.93; ε 3.26
K44	8.10	123.98	5.11	1.58	1.30, 1.40	
Y45	9.00	123.72	4.68	1.97, 2.30		H2,H6 6.50; H3,H5 6.26; H2,H6 6.53; H3,H5 6.22
K46	8.80	126.67	5.24	1.58, 1.70	1.44	δ 2.02; ε 3.00
	8.82		5.25			
A47	9.38	125.48	5.15	1.07		
	9.33		5.13	1.07		
H48	9.07		4.97	3.08, 3.24		H2 7.88; H4 7.17
D49	8.11	120.04	4.68	3.96, 4.08		
E50	8.26	108.07	3.49	1.45, 1.48		
H51	8.49	112.88	4.90	2.97, 3.52		H2 8.36; H4 7.24
N52	8.03	121.54	4.20	2.82, 2.98		δ 7.09, 7.63
E53	9.85	121.38	4.11	1.93	2.19	
A54	9.94	127.24	4.52	1.40		
	9.90		4.52	1.40		
K55	9.05	122.63	5.05	1.77, 1.80	1.45	
V56	8.33	120.50	3.28			
G57	8.81	117.03	3.53, 4.52			
D58	8.55	124.96	4.71	2.56, 2.88		
I59	8.94		4.98	1.93	0.85	
V60	9.46		5.14	2.16	0.62	
K61	8.79	122.01	5.53	1.58, 1.73	1.21, 1.34	
I62	9.45	120.92	5.50	1.71	0.87, 1.45	γME 0.67; δ 0.09
M63	9.02	121.95	5.63	2.03	2.41, 2.70	
	8.98		5.60	2.03	2.41, 2.73	
E64	8.82	130.56	3.92	1.92	2.01	
T65	8.25	119.62	4.52	4.18	1.07	
R66	8.07		4.61	1.84	1.68	δ 3.18
P67			4.62	2.47	1.86, 1.96	δ 3.57
L68	8.61		4.40	1.33, 1.42	1.13	
S69	8.22	116.35	4.69	3.77, 3.95		
A70	8.00		4.35	1.26		
T71	8.24	120.30	4.39	4.15	1.15	
K72	8.82	124.70	4.18			
R73						
F74	*7.81*		*5.40*	2.72, 3.25		H2,H6 7.14; H3,H5 6.88; H4 6.80
	7.68		*5.45*			
R75	9.44		5.31	1.89, 2.04	1.67, 1.83	
	9.33		5.33			
L76	8.46	123.82	4.36	1.68	0.83	
V77	9.21	126.31	4.49	2.02	0.96, 1.03	
E78	7.47	117.13	4.64	2.01	2.10	
I79	9.15	126.93	4.28	1.94	1.12, 1.67	δ 0.82
V80	9.05	132.74	3.88	1.79	0.96	
E81	8.00	126.93	4.55	1.96	2.16, 2.23	
	7.99		4.50	1.96	2.16, 2.23	
K82	8.76	129.73	4.44			
A83	9.34	129.94	4.28	1.39		
V84	8.33	123.20	4.07	2.08	0.98, 1.03	
R85	7.97	132.53	4.25	1.89	1.61, 1.75	δ 3.25

[a] The residues that appear as doublets in the spectra have the chemical shift of the minor form listed below that of the major form. The amide and α chemical shifts for phenylalanine 74 (in italics) were determined from a DQF-COSY spectrum acquired at 45 °C. All other assignments were determined at 25 °C.

NOESY spectrum (Table I). The former were first used to identify protons within each amino acid spin system. Next, neighboring residues in the sequence were determined by the presence of NOE peaks between amide protons and the backbone protons of the preceding residue. When several sequential residue types had been determined, the assignments could usually be uniquely identified within the known primary sequence of the protein. Due to the high degree of dispersion in the spectra, most of the backbone and side chain protons could be specifically assigned.

The NMR spectra are largely unaffected by temperatures as high as 75 °C, and this is consistent with the fact that the molecule is derived from a thermostable bacterium. This proved useful from a practical standpoint since it allowed a DQF-COSY spectrum to be collected at 45 °C. At this higher temperature, the peaks in the spectrum were considerably sharpened. This enabled the severely broadened amide-to-α cross peak of phenylalanine 74 to be identified.

After the ^1H NMR experiments, several ambiguities due to spectral overlap remained, and these were resolved by heteronuclear two-dimensional NMR experiments. Using a protein sample that was uniformly labeled with ^{15}N, HSMQC, HMQC-TOCSY, and HMQC-NOESY spectra were acquired. These experiments were used to assign the

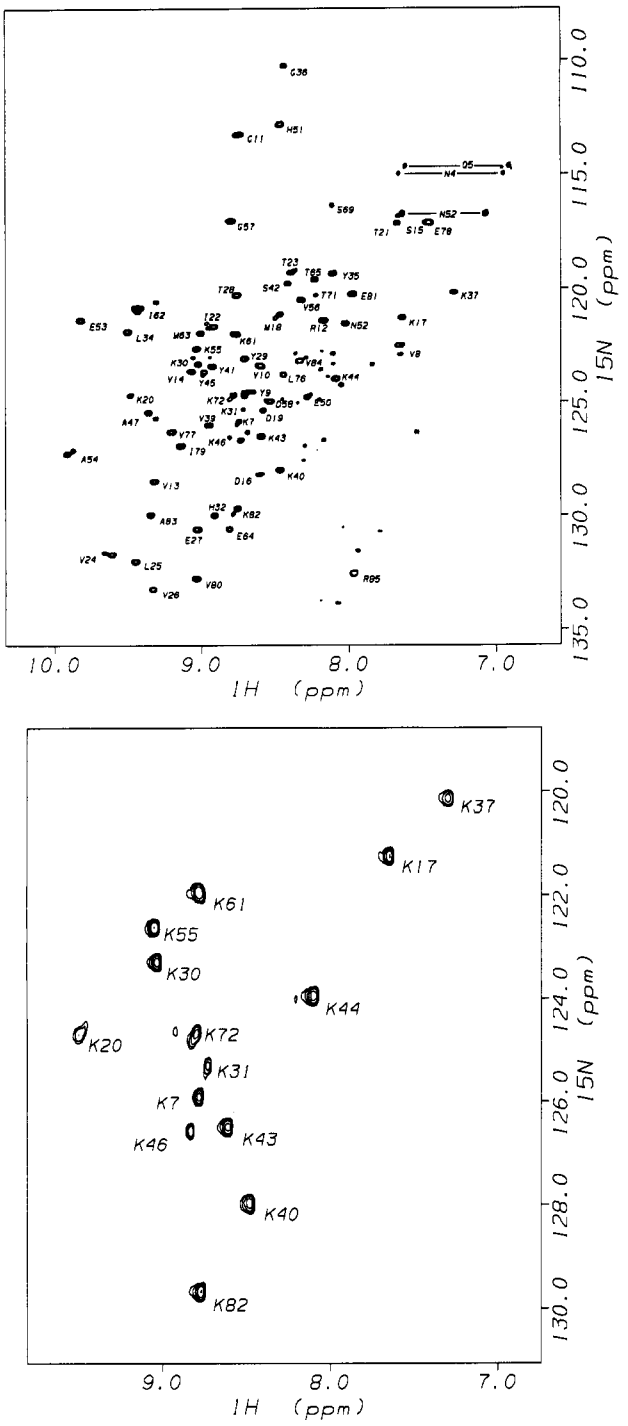

FIGURE 2: (A, top) HSMQC spectrum of uniformly ^{15}N-labeled ribosomal protein S17. Most of the main chain and all of the side chain amide nitrogen chemical shifts have been assigned. (B, bottom) HSMQC spectrum of [α-^{15}N]lysine labeled S17. Only the 14 lysine residues where the protein is selectively labeled can be seen in the spectrum.

amide nitrogen chemical shifts of the protein (Table I). This assignment procedure was considerably simplified by the previously determined proton assignments. These ^{15}N experiments confirmed the proton assignments and enabled the identification of additional residues in the NMR spectra. As a final check of the spectrum assignments, two samples were prepared with ^{15}N specifically incorporated at the amide position of lysine residues or glycine and serine residues. These were chosen because the residues can be easily labeled and because S17, like all ribosomal proteins, contains many (14) lysine residues. HSMQC spectra from these protein samples

Table II: $J_{\alpha NH}$ Coupling Constants of Ribosomal Protein S17[a]

residue	$J_{\alpha NH}$	residue	$J_{\alpha NH}$	residue	$J_{\alpha NH}$
K7	5.1	L34	9.0	K61	9.1
V8	10.4	Y35	6.5	I62	10.6
Y9	8.3	G36	*	M63	9.5
V10	11.1	K37	6.6	E64	3.5
G11	*	R38	*	T65	7.2
R12	7.5	V39	9.2	R66	*
V13	3.6	K40	9.5	P67	*
V14	11.0	Y41	9.5	L68	*
S15	6.1	S42	9.2	S69	5.2
D16	10.0	K43	8.4	A70	*
K17	4.0	K44	8.0	T71	*
M18	9.0	Y45	7.6	K72	*
D19	*	K46	7.2	R73	*
K20	*	A47	8.0	F74	*
T21	10.1	H48	*	R75	*
I22	9.0	D49	*	L76	8.1
T23	9.8	E50	*	V77	12.0
V24	10.6	H51	9.0	E78	4.2
L25	10.0	N52	4.6	I79	8.6
V26	9.6	E53	7.9	V80	10.0
E27	7.5	A54	4.7	E81	7.6
T28	8.9	K55	9.1	K82	5.9
Y29	8.5	V56	6.6	A83	2.0
K30	7.1	G57	*	V84	8.0
K31	*	D58	*	R85	8.8
H32	4.00	I59	*		
P33	*	V60	*		

[a] The coupling constant between the α and amide proton was measured from a HMQC-J spectrum. These values were used to constrain the backbone ϕ angles in the 3D structure calculation as described in the text.

only contain peaks that correspond to the labeled residues (Figure 2). The interpretation of the spectra was therefore straightforward, and it allowed the identification of the previously unassigned residues lysine 31, serine 69, and lysine 72. In total, 376 protons were assigned including the backbone protons for all residues between 4 and 85 except arginine 73.

(iii) Secondary Structure. The NOESY spectrum of S17 contains many strong cross peaks between pairs of α-protons, a characteristic of proteins that contain antiparallel β-sheet structure. These NOE peaks were used as a starting point to identify residues that are cross-strand neighbors within the β-sheet. The β-sheet structure was confirmed and expanded by further assigning interstrand α-to-amide and amide-to-amide NOEs. In this way, a five-stranded antiparallel β-sheet was identified in S17 (Figure 3). The notation for the secondary structure elements within the primary structure of S17 is shown in Figure 4. All 18 protons that are protected from exchange with deuterium solvent during a typical 8-h two-dimensional experiment are involved in hydrogen-bonded interactions between the β-strands.

(iv) Tertiary Structure. It was possible to uniquely assign 267 NOE peaks to specific pairs of protons, 149 involving adjacent residues in the primary structure and 118 representing interactions between nonadjacent amino acids. In addition, 40 ϕ angles were constrained from the $J_{\alpha NH}$ coupling constants in the HMQC-J spectrum (Table II), and 16 nonexchangeable protons (Figure 3) were constrained to be within hydrogen-bonding distance of their target carbonyl atoms. These constraints were first used to generate a structure of S17 using distance geometry. This initial model, which consisted of a fairly flat five-stranded β-sheet, was then used as the starting template for structure determination by simulated annealing. In this manner, a set of reasonable structures with similar features was generated that satisfied the constraints of the NMR data. A ribbon diagram corresponding to one such structure is shown in Figure 5. Although structures generated

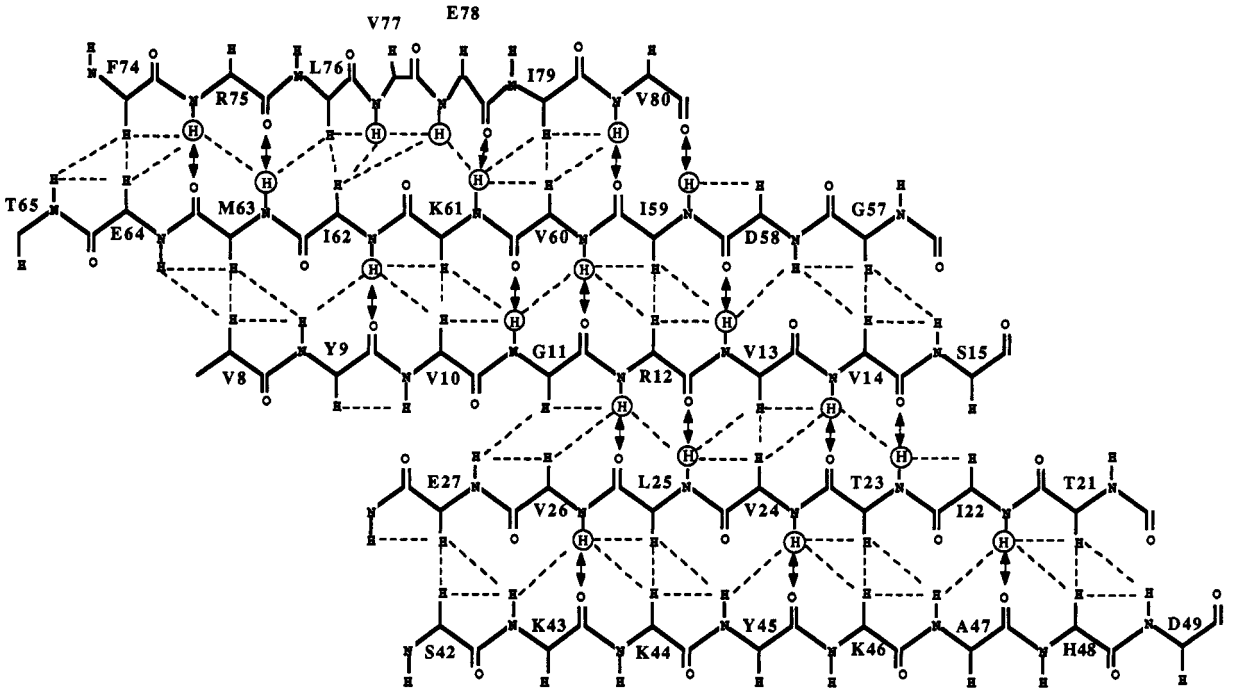

FIGURE 3: The secondary structure of ribosomal protein S17. The molecule contains a five-stranded antiparallel β-sheet, and the observed NOE peaks are indicated by dashed lines. Amide protons that are resistant to deuterium exchange are circled. Atoms considered to be involved in hydrogen-bonded interactions are indicated by double-headed arrows.

by this method are similar in terms of their overall fold and twist in the β-sheet, the side chain positions and backbone conformation in the loop regions are not well defined in the present NMR structure. The average rms difference between the backbone atoms in a set of 10 structures generated in a typical simulated annealing run was 3.0 Å.

All but one of the residues that appear as doublets in the NMR spectra are found in two regions of the secondary structure. The first corresponds to the central residues of β1 and β2. The second is the adjacent ends of β1, β4, and β5 near loop 3. In the three-dimensional model, these two distinct sites are brought together by the left-handed twist in the β-sheet.

Description of the Structure. S17 consists of a five-stranded antiparallel β-pleated sheet in which the strands are arranged with Greek key topology (Richardson, 1977). The twist of the β-sheet is in the usual left-handed sense when viewed across the strands. Apart from a short turn (turn 1) that links strands β1 and β2, the connecting loops are rather extensive. Loop 1 (β2 → β3) is the longest with 15 residues, loop 2 (β3 → β4) contains eight residues and loop 3 (β4 → β5) also has eight residues. Strand β5 contains a one-residue β-bulge in which valine 77 and glutamate 78 occupy the normally single-residue site of the secondary structure element. Since the amide protons of both residues are resistant to exchange with D_2O solvent, both appear to be involved in hydrogen-bonded interactions with cross-strand carbonyl oxygen of lysine 61 (Figure 3).

In a typical β-sheet, alternating side chains along the primary sequence appear on one surface of the sheet. When the available primary sequences of S17 are analyzed, the pattern of conservation in the β-strand regions clearly shows that one surface of the β-sheet is populated with hydrophobic residues (Figure 4). These comprise valines 13, 24, 26, and 60, isoleucines 22, 62, and 79, alanine 47, and leucine 76. In the three-dimensional models, this surface corresponds to the hydrophobic core of the molecule. Glycine 11 appears to be crucial to the integrity of the hydrophobic interior. This is located in the middle of β1 at the absolute center of the β-sheet. It is totally conserved, and the absent side chain may be required for the correct packing of its neighboring hydrophobic core residues.

DISCUSSION

The model of S17 described here shows the overall fold of the molecule in sufficient detail to allow an initial analysis of its structure in terms of its known biochemical properties. However, it is by no means a high-resolution model of the protein and qualifies only as a "first generation" NMR structure (Clore & Gronenborn, 1991). The molecule is comprised of a twisted five-stranded antiparallel β-pleated sheet which gives it the appearance of an incomplete β-barrel. The β-strands are connected by several extended loop regions which appear to be unstructured in this model. However, S17 is known to be very resistant to proteolysis (Littlechild et al., 1987), and the loops probably adopt folded conformations that are not apparent from the current NMR data. Although the protein appears to have two distinct but very similar conformations in solution, the resolution of the model does not allow a detailed description of the alternate conformations at the present time.

It is becoming increasingly clear that an important functional element of the ribosome is the rRNA (Noller, 1991) and that a key role of the ribosomal proteins is to provide a framework for its correctly folded structure. This is especially true for the primary RNA-binding proteins which also direct the folding of the rRNA to its active conformation. Previous structural studies on nucleic acid binding proteins (Ollis & White, 1987) and RNA-binding proteins in particular (Rould et al., 1989; Nagai et al., 1990; Hoffmann et al., 1991; Jessen et al., 1991; Kenan et al., 1991; Ramakrishnan & White, 1992) enable us to predict the expected architecture of the S17 RNA-binding site(s). First, it should contain several highly conserved arginine and lysine residues for interaction with the RNA sugar–phosphate backbone. Second, exposed aromatic residues are often present, presumably to interact

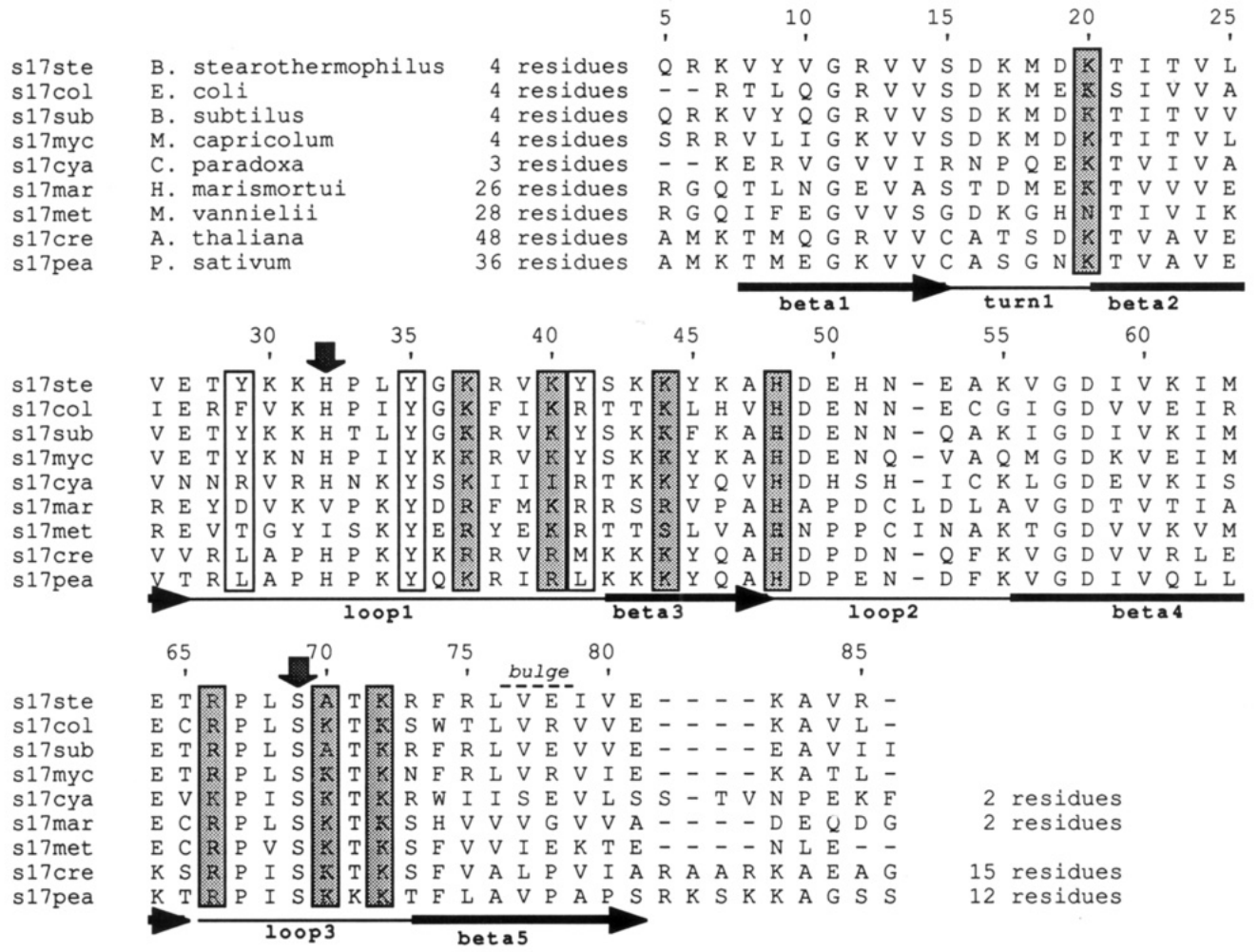

FIGURE 4: Sequence alignment of known S17 primary structures. Of the nine sequences shown, five are eubacterial (*Bacillus stearothermophilus*, *Escherichia coli*, *Bacillus subtilis*, *Mycoplasma capricolum*, *Cyanophora paradox*a), two are archaebacterial (*Halobacterium marismortui*, *Methanococcus vannielii*), and two are chloroplastid (*Arabidopsis thaliana*, *Pisum sativum*). Conserved residues that may interact with rRNA are boxed. Positively charged residues are shaded, and aromatic residues are unshaded. The sites of two phenotypically important point mutations, histidine 32 and serine 69, are indicated by arrows (see text). The locations of the secondary structure elements within the sequences are also indicated. The simultaneous sequence alignment was performed with the program TULLA (Subbiah & Harrison, 1989).

with the RNA bases. Finally, we have demonstrated in ribosomal protein S5 (Ramakrishnan & White, 1992) that point mutations which modulate antibiotic function occur within regions of the molecule that appear to bind RNA. On the basis of these criteria, the likely RNA-binding sites of S17 are the extended loop regions loop 1 and loop 3 (Figure 4). Loop 1 is a particularly good candidate since it contains two highly conserved lysines (37 and 40), three aromatic residues (tyrosines 29, 35, and 41), and a mutation site (histidine 32) that renders the ribosome resistant to neamine. Loop 3 contains the highly conserved residues arginine 66 and lysine 72, and an alanine (70) that is normally a lysine in other species. Loop 3 also contains a mutation site (serine 69) that results in an inability of the small subunit to assemble at 42 °C (Herzog et al., 1979). This mutation (to a phenylalanine) occurs at the very end of the flexible loop and is unlikely to hinder protein folding. More likely, it prevents S17 from correctly binding RNA at restrictive temperatures.

During the past 10–15 years, a large quantity of data have been accumulated on the relative locations of the RNA and protein components within the 30S subunit, and several three-dimensional models of the small subunit have been proposed (Stern et al., 1988b; Brimacombe et al., 1988). Although the models differ to some extent, they agree on the location of S17 at the bottom of the subunit. Early work showed that S17 has a primary RNA-binding site toward the 5′ end of 16S rRNA (Muto et al., 1974; Zimmerman et al., 1974; Mackie & Zimmerman, 1978), and the site has recently been narrowed down to a fragment consisting of nucleotides 1–526 (Weitzmann et al., 1993). Chemical protection experiments confirm these data and reveal that the primary binding site almost certainly comprises the adjacent RNA helices 7 and 11 (Stern et al., 1988a). Thus, the principal function of S17 may be to act as a molecular strut which locks this region of 16S rRNA into its functional conformation. RNA–protein cross-linking studies show that S17 may have a second role. These studies confirm the interaction with helix 11 (Greuer et al., 1987) but also reveal a close association with helix 21 (Kyriatsoulis et al., 1986). Although helix 21 is distant from helix 11 in the secondary structure schemes of 16S rRNA, it folds to the bottom of the subunit to contact S17 in the three-dimensional models. S17 may therefore contribute toward a secondary RNA-binding site in this region of the 30S subunit that helps to lock helix 21 into its correct position.

This RNA-binding scenario might explain some functional aspects of S17 that are not easily rationalized by its location in the 30S models. Two related observations indicate that S17 is associated with the fidelity process in protein synthesis. First, as mentioned above, a mutation in S17 causes the ribosome to become resistant to the antibiotic neamine (Bollen et al., 1975; Yaguchi et al., 1976). Neamine, like the related aminoglycosides neomycin, kanamycin, gentamicin, apramy-

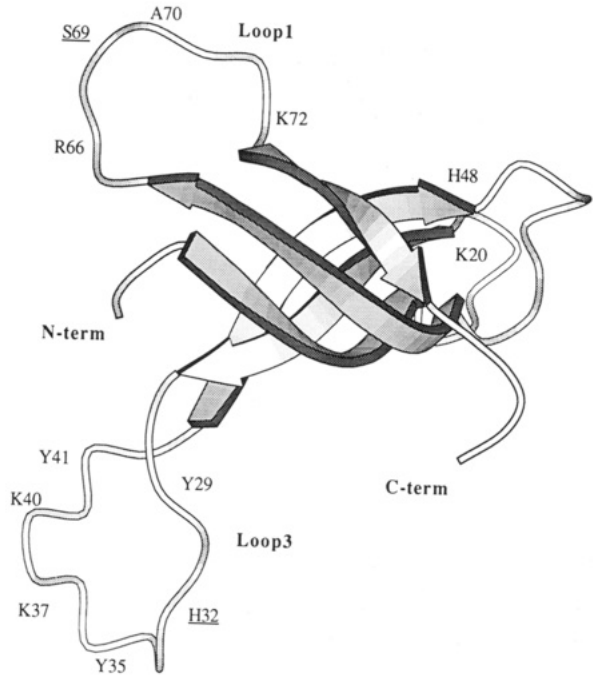

FIGURE 5: Three-dimensional model of ribosomal protein S17. This model shows the left-handed twist in the β-sheet and the relative positions of loop 1 and loop 3. The conserved positively charged residues and the aromatic residues in loop 1 and loop 3 are indicated. The underlined residues are the sites of two-point mutations that may interfere with RNA binding (see text). The ribbon diagram was drawn using MOLSCRIPT (Kraulis, 1991).

cin, and paromomycin, reduces the accuracy of translation, possibly by stabilizing nonspecific interactions of tRNA at the A-site (Woodcock et al., 1991). They have all been shown to interact with 16S rRNA at or close to the A-site around nucleotide 1400 (Moazed & Noller, 1987; Woodcock et al., 1991). Second, in common with the other fidelity-associated proteins S4, S5, and S12, the S17 mutation alters the inherent fidelity of the ribosome. Mutations in S4 and S5 reduce accuracy and those in S12 and S17 increase it (Topisirovič et al., 1977; Matkovič et al., 1980). These observations are not easily explained by the 30S subunit models which show S17 to be distant from both the A-site and the S4/S5/S12 cluster in the so called "proofreading domain". It has been suggested that such long-range effects might represent structural perturbations propagated through the RNA structure (Noller, 1991). Helix 27 of the 16S rRNA is also a component of the proofreading domain (Noller, 1991), and in the 30S models it is positioned directly above helix 11. If, as we suggest, S17 controls the positioning of helix 11, perturbations that are induced by a mutation in S17 could be transmitted to helix 27 and thereby affect translational accuracy.

Of the four ribosomal proteins whose structures have previously been determined, three have been found to contain main chain folds with similar secondary structure topologies. L7/L12 (Leijonmarck et al., 1980), L30 (Wilson et al., 1986), and L6 (unpublished data) all contain an exposed antiparallel β-pleated sheet within an overall structure that is conformationally homologous to a eukaryotic RNA-binding motif (Nagai et al., 1990; Hoffman et al., 1991). It has been postulated that this family of molecules may have evolved from a single, primitive RNA-binding protein (Leijonmarck et al., 1988; Hoffman et al., 1991). It is clear that the structure of S17 is not related to these but instead belongs to a recently recognized family of proteins, named the OB(oligonucleotide/oligosaccharide-binding)-fold (Murin, 1993). This structural motif is found in a group of proteins that share both certain structural features and a common substrate binding site. Included in this group are the major cold shock protein, csp B, from *Bacillus subtilus* (Schindelin et al., 1993; Schnuchel et al., 1993) and the single-stranded DNA-binding protein from bacteriophage f1 (Brayer & McPherson, 1983), both of which look remarkably like S17. With the present model of S17 it is not possible to compare its structure and potential RNA-binding sites in any great detail to this family of proteins. This must await a more detailed model of S17. In order to determine the solution structure of S17 at higher resolution, we plan to label the protein with ^{13}C, and expand the NMR analysis to the three- and four-dimensional experiments that have proven to be so powerful in other systems (Clore & Gronenborn, 1991). We also plan NMR and/or crystallographic investigations of the S17–RNA complex to determine the structural basis of the RNA recognition.

Coordinates have been deposited with the Protein Data Bank, Chemistry Department, Brookhaven National Laboratory, Upton, NY 11973. The PDB filename for the structure described in this paper is 1rip.

ACKNOWLEDGMENT

We thank Sue Ellen Gerchman, Helen Kycia, and Stephanie Porter for technical assistance, and Terrence Oas and M. Otman for helpful suggestions. We also thank Harry Noller for critically reading the manuscript.

SUPPLEMENTARY MATERIAL AVAILABLE

One table listing interresidue NOEs (7 pages). Ordering information is given on any current masthead page.

REFERENCES

Appelt, K., Dijk, J., Reinhardt, R., Sanhuesa, S., White, S. W., Wilson, K. S., & Yonath, A. (1981) *J. Biol. Chem. 256*, 11787–11790.

Auer, J., Spicker, G., & Böck, A. (1989) *J. Mol. Biol. 209*, 21–36.

Bollen, A., Cabezón, T., DeWilde, M., Villarroel, R., & Herzog, A. (1975) *J. Mol. Biol. 99*, 795–806.

Brayer, G. D., & McPherson, A. (1983) *J. Mol. Biol. 169*, 565–596.

Brimacombe, R. (1991) *Biochimie 73*, 927–936.

Brimacombe, R., Atmadja, J., Stiege, W., & Schüler, D. (1988) *J. Mol. Biol. 199*, 115–136.

Brimacombe, R., Greuer, B., Mitchell, P., Osswald, M., Rinke-Appel, J., Schüler, D., & Stade, K. (1990) in *The Ribosome: Structure Function and Evolution* (Hill, W. E., Dahlberg, A., Garrett, R. A., Moore, P. B., Schlessinger, D., & Warner, J. R., Eds.) pp 93–106, American Society for Microbiology, Washington, D.C.

Brown, S. C., Weber, P. L., & Mueller, L. (1988) *J. Magn. Reson. 77*, 166–169.

Brünger, A. T. (1987) XPLOR, Version 3.1 manual, Yale University Press, New Haven, CT.

Brünger, A. T., Clore, G. M., Gronenborn, A. M., & Karplus, M. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 3801–3805.

Capel, M. S., Engleman, D. M., Freeborn, B. R., Kjeldgaard, M., Langer, J. A., Ramakrishnan, V., Schindler, D. G., Schneider, D. K., Schoenborn, B. P., Sillers, I.-Y., Yabuki, S., & Moore, P. B. (1987) *Science 238*, 1403–1406.

Capel, M. S., Kjeldgaard, M., Engelman, D. M., & Moore, P. B. (1988) *J. Mol. Biol. 200*, 65–87.

Cheong, C., Varani, G., & Tinoco, I. (1990) *Nature 346*, 680–682.

Clore, G. M., & Gronenborn, A. M. (1991) *Science 252*, 1390–1399.

Davis, D. G., & Bax, A. (1985) *J. Am. Chem. Soc. 107*, 2821–2822.
Expert-Bezançon, A., Barritault, D., Milet, M., Guérin, M.-F., & Hayes, D. H. (1977) *J. Mol. Biol. 112*, 603–629.
Frank, J., Penczek, P., Grassucci, R., & Srivastava, S. (1991) *J. Cell. Biol. 115*, 597–605.
Gantt, J. S., & Thompson, M. D. (1990) *J. Biol. Chem. 265*, 2763–2767.
Greuer, B., Osswald, M., Brimacombe, R., & Stöffler, G. (1987) *Nucleic Acids Res. 15*, 3241–3255.
Gronenborn, A., Bax, A., Wingfield, P. T., & Clore, M. G. (1989) *FEBS Lett. 243*, 93–98.
Held, W. A., Ballou, B., Mizushima, S., & Nomura, M. (1974) *J. Biol. Chem. 249*, 3103–3111.
Henkin, T. M., Moon, S. H., Mattheakis, L. C., & Nomura, M. (1989) *Nucleic Acids Res. 17*, 7469–7486.
Herfurth, E., Hirano, H., & Wittmann-Liebold, B. (1991) *Biol. Chem. Hoppe-Seyler 372*, 955–961.
Herzog, A., Yaguchi, M., Cabezón, T., Corchuelo, M. C., Petre, J., & Bollen, A. (1979) *Mol. Gen. Genet. 171*, 15–22.
Heus, H. A. & Pardi, A. (1991) *Science 253*, 191–193.
Hoffmann, D. W., Query, C. C., Golden, B. L., White, S. W., & Keene, J. D. (1991) *Proc. Natl. Acad. Sci U.S.A. 88*, 2495–2499.
Jessen, T.-H., Oubridge, C., Teo, C. H., Pritchard, C., & Nagai, K. (1991) *EMBO J. 10*, 3447–3456.
Kenan, D. J., Query, C. C., & Keene, J. D. (1991) *Trends Biochem. Sci. 16*, 214–220.
Kimura, J., & Kimura, M. (1987) *J. Biol. Chem. 262*, 12150–12157.
Kraulis, P. (1991) *J. Appl. Crystallogr. 24*, 946–950.
Kyriatsoulis, A., Maly, P., Greuer, B., Brimacombe, R., Stöffler, G., Frank, R., & Blöcker, H. (1986) *Nucleic Acids Res. 14*, 1171–1186.
Lambert, J. M., Boileau, G., Cover, J. A., & Traut, R. R. (1983) *Biochemistry 22*, 3913–3920.
Leijonmarck, M., Ericksson, S., & Liljas, A. (1980) *Nature 286*, 824–826.
Leijonmarck, M., Appelt, K., Badger, J., Liljas, A., Wilson, K. S., & White, S. W. (1988) *Proteins 3*, 243–251.
Littlechild, J., Malcolm, A., Paterakis, K., Ackermann, I., & Dijk, J. (1987) *Biochim. Biophys. Acta 913*, 245–255.
Mackie, G. A., & Zimmerman, R. A. (1978) *J. Mol. Biol. 121*, 17–39.
Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967–974.
Matkoviç, B., Herzog, A., Bollen, A., & Topisiroviç, L. (1980) *Mol. Gen. Genet. 179*, 135–139.
Michalowski, C. B., Pfanzagl, B., Loeffelhardt, W., & Bohnert, H. J. (1990) *Mol. Gen. Genet. 224*, 222–231.
Moazed, D., & Noller, H. F. (1987) *Nature 327*, 389–394.
Murzin, A. G. (1993) *EMBO J. 12*, 86–867.
Muto, A., Ehresmann, C., Fellner, P., & Zimmerman, R. A. (1974) *J. Mol. Biol. 86*, 411–432.
Nagai, K., Oubridge, C., Jessen, T. H., Li, J., & Evans, P. R. (1990) *Nature 348*, 515–520.
Nilges, M., Clore, G. M., & Gronenborn, A. M. (1988) *FEBS Lett. 239*, 129–136.
Noller, H. F. (1991) *Annu. Rev. Biochem. 60*, 191–227.
Ohkubo, S., Muto, A., Kawauchi, Y., Yamao, F., & Osawa, S. (1987) *Mol. Gen. Genet. 210*, 314–322.
Ollis, D. L., & White, S. W. (1987) *Chem. Rev. 87*, 981–995.
Ramakrishnan, V., & Gerchman, S. E. (1991) *J. Biol. Chem. 266*, 880–885.
Ramakrishnan, V., & White, S. W. (1992) *Nature 358*, 768–771.
Richardson, J. S. (1977) *Nature 268*, 495–500.
Rould, M. A., Perona, J. J., Soll, D., & Steitz, T. A. (1989) *Science 246*, 1135–1142.
Schindelin, H., Marahiel, M. A., & Heinemann, U. (1993) *Nature 364*, 164–168.
Schnuchel, A., Wiltscheck, R., Czisch, M., Herrler, M., Willimsky, G., Gaumann, P., Marahiel, & Holak, T. A. (1993) *Nature 364*, 169–171.
States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286–292.
Stern, S., Changchien, L.-M., Craven, G. R., & Noller, H. F. (1988a) *J. Mol. Biol. 200*, 291–299.
Stern, S., Weiser, B., & Noller, H. F. (1988b) *J. Mol. Biol. 204*, 447–481.
Stern, S., Powers, T., Changchien, L.-M., & Noller, H. F. (1989) *Science 244*, 783–790.
Stöffler-Meilicke, M., & Stöffler, G. (1990) in *The Ribosome: Structure Function and Evolution* (Hill, W. E., Dahlberg, A. D., Garrett, R. A., Moore, P. B., Schlessinger, D., & Warner, J. R., Eds.) pp 123–133, American Society for Microbiology, Washington, D.C.
Studier, F. W., Rosenberg, A. H., Dunn, J. J., & Dubendorff, J. W. (1990) *Methods Enzymol. 185*, 60–89.
Subbiah, S., & Harrison, S. C. (1989) *J. Mol. Biol. 209*, 539–548.
Topisiroviç, L., Villarroel, R., De Wilde, M., Herzog, A., Cabezón, T., & Bollen, A. (1977) *Mol. Gen. Genet. 151*, 89–94.
von Böhlen, K., Makowski, I., Hansen, H., Bartels, Z., Berkovitch-Yellin, H., Zaytzev-Bashan, A., Meyer, S., Paulke, C., Franceschi, F., & Yonath, A. (1991) *J. Mol. Biol. 222*, 11–15.
Weitzmann, C. J., Cunningham, P. R., Nurse, K., & Ofengand, J. (1993) *FASEB J. 7*, 177–180.
Wilson, K. S., Appelt, K., Badger, J., Tanaka, I., & White, S. W. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 7251–7255.
Wishart, D. S., Sykes, B. D., & Richards, F. M. (1992) *Biochemistry 31*, 1647–1651.
Woodcock, J., Moazed, D., Cannon, M., Davies, J., & Noller, H. F. (1991) *EMBO J. 10*, 3099–3103.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.
Yaguchi, M., & Wittmann, H. G. (1978) *FEBS Lett. 87*, 37–40.
Yaguchi, M., Wittmann, H. G., Cabezón, T., DeWilde, M., Villarroel, R., Herzog, A., & Bollen, A. (1976) *J. Mol. Biol. 104*, 617–620.
Zimmerman, R. A., Muto, A., & Mackie, G. A. (1974) *J. Mol. Biol. 86*, 433–450.
Zuiderweg, E. R. P. (1990) *J. Magn. Reson. 86*, 346–357.