Ribosomal protein L6: structural evidence of gene duplication from a primitive RNA binding protein

Barbara L. Golden, V. Ramakrishnan^{1,2} and Stephen W. White²

Department of Microbiology, Duke University Medical Center, Durham, NC 27710 and ¹Biology Department, Brookhaven National Laboratory, Upton, NY 11973, USA

²Corresponding authors

Communicated by L.N. Johnson

In all cells, protein synthesis is coordinated by the ribosome, a large ribonucleoprotein particle that is composed of >50 distinct protein molecules and several large RNA molecules. Here we present the crystal structure of ribosomal protein L6 from the thermophilic bacterium *Bacillus stearothermophilus* solved at 2.6 Å resolution. L6 contains two domains with almost identical folds, implying that it was created by an ancient gene duplication event. The surface of the molecule displays several likely sites of interaction with other components of the ribosome. The RNA binding sites appear to be localized in the C-terminal domain whereas the N-terminal domain contains the potential sites for protein–protein interactions. The domain structure is homologous with several other ribosomal proteins and to a large family of eukaryotic RNA binding proteins. **Key words:** ribosome/RNA binding protein/translation/X-ray crystallography

Introduction

The process whereby the genetic code is translated into the functional amino acid sequences of proteins is performed by the ribosome. Many of the details involved in this process are poorly understood, and they will remain so until accurate structural information can be obtained. To this end, several groups have undertaken high resolution structural studies on individual ribosomal components. At this time, the structures of several prokaryotic ribosomal proteins have been determined by X-ray crystallography and high-resolution NMR including S5 (Ramakrishnan and White, 1992), S6 (A. Liljas, personal communication) and S17 (Golden *et al.*, 1993) from the small subunit, and L7/L12 (Leijonmarck *et al.*, 1980), L9 (Hoffman *et al.*, 1992) and L30 (Wilson *et al.*, 1986) from the large subunit. Here, we present the crystal structure of ribosomal protein L6 at 2.6 Å resolution.

With a molecular weight of 19 000 Daltons, ribosomal protein L6 is one of the larger proteins of the 50S ribosomal subunit. It appears to be a ubiquitous ribosomal protein, and has been found in prokaryotes, archaeobacteria and eukaryotes with a highly conserved primary structure (Kimura *et al.*, 1981; Ceretti *et al.*, 1983; Suzuki *et al.*, 1990; Scholzen and Arnot, 1991). L6 is associated with the translocation domain of the ribosome and has been localized by immune electron

microscopy near the base of the L7/L12 'stalk' on the large ribosomal subunit at the subunit interface (Hackl and Stöffler-Meilicke, 1988). Crosslinks have been generated between L6 and proteins that are associated with the central protuberance and the stalk of the large subunit (Traut *et al.*, 1985). In *Escherichia coli*, L6 is known to be associated with 23S ribosomal RNA (rRNA) near nucleotides 2473–2481 (Wower *et al.*, 1981) at the base of helix 89 in the tRNA binding domain of the 23S rRNA (Mitchell *et al.*, 1990). Mutations in the protein have been shown to render the ribosome resistant to the antibiotic gentamicin (Kühberger *et al.*, 1979), an antibiotic that reduces the fidelity of the ribosome (Cundliffe, 1990). Since ribosomes that contain these mutant L6 molecules become hyperaccurate, it has been suggested that the protein influences the acyl-tRNA binding site (Kühberger *et al.*, 1979).

L6 purified under non-denaturing conditions from the ribosomes of the thermophilic bacterium *Bacillus stearothermophilus* was crystallized several years ago (Appelt *et al.*, 1983), but no high quality heavy atom derivatives were obtained. The gene that encodes this molecule was recently cloned and overexpressed (Ramakrishnan and Gerchman, 1991) and this provided the quantities of material needed to complete the X-ray analysis. It also enabled us to modify the molecule to create heavy atom derivatives for use in the multiple isomorphous replacement method. Specifically, a valine residue was replaced by a cysteine for binding mercury (site-direct mutant V124C), and two selenium atoms were incorporated by biosynthetically replacing the methionine residues with selenomethionine.

Results

Introducing a cysteine into L6

In order to create a potential mercury binding site in L6, a cysteine residue was introduced into the wild-type *B. stearothermophilus* protein which does not naturally contain this amino acid type. Valine 124 was chosen as the site of mutation because *E. coli* L6 contains a cysteine at the corresponding position (Figure 1). The mutant protein, V124C, crystallized in the same space group and with the same unit cell dimensions as the wild-type protein and could be derivatized with *p*-chloromercuribenzenesulfonate (PCMBs) to create an excellent heavy atom derivative. Since the crystals of the mutant protein are non-isomorphous with wild-type crystals at high resolution, we decided to solve first the structure of the mutant protein by multiple isomorphous replacement. To create a second derivative of V124C, selenium was incorporated as selenomethionine in place of the two methionine residues.

Solving the crystal structure

X-ray diffraction data were collected and processed as described in Materials and methods and the results are

Fig. 1. Amino acid sequence alignments for ribosomal protein L6 (one-letter code). The four sequences include representatives from each of the principal life-forms, two from eubacteria: *B.stearothermophilus* (stea) (Kimura *et al.*, 1981) and *E.coli* (coli) (Ceretti *et al.*, 1983); one from archaeobacteria: *Halobacterium marismortui* (mari) (Scholzen and Arnot, 1991); and one eukaryote: *Rattus norvegicus* [rat (L9)] (Suzuki *et al.*, 1990). Secondary structure elements are named such that the corresponding elements in each domain have similar labels. Apart from the N- and C-termini, all four sequences are highly homologous. The *H.marismortui* and *R.norvegicus* sequences are slightly longer with small insertions in loop regions and at the ends of secondary structure elements. Hydrophobic core residues are boxed and shaded. Conserved residues at the domain interface are bold and italicized. The residues that are proposed to interact with 23S rRNA are boxed and unshaded (see also Figure 2B). Valine 124 in *B.stearothermophilus*, which corresponds to a cysteine in *E.coli*, was chosen as the site to introduce a cysteine point mutation for binding mercury (see text for details).

summarized in Table I. The single PCMBBS site of V124C was located by calculating a difference Fourier map using the 6 Å phases of the wild-type protein that had been determined previously (Appelt *et al.*, 1984). The mercury site was confirmed in a difference Patterson map, and the heavy atom parameters were then refined. Phases for V124C were then calculated using both isomorphous and anomalous differences, and these were used to calculate a difference Fourier map for the selenomethionine derivative. Both selenium sites were easily located in this map and had

corresponding peaks in the difference Patterson maps. Following refinement of all of the heavy atom parameters, a final set of phases was calculated to 2.6 Å using isomorphous data from both derivatives and anomalous data from the PCMBBS derivative. The figure of merit for these phases was 0.619 and this was increased to 0.866 after solvent flattening (Wang, 1985). The final solvent flattened map was of sufficient quality to trace the polypeptide chain unambiguously (Figure 2).

The protein was built outwards from methionine 40,

Table I. Data collection, heavy atom refinement and phasing statistics for the structure determination of ribosomal protein L6 [space group P6₁22 ($a = b = 72.1$ Å, $c = 124.7$ Å)]

Data set	Wild-type L6	V124C		
	Native	Native	Semet	PCMBS
Number of collected reflections	39 928 (6.63%)	36 408 (4.27%)	24 768 (5.33%)	25 720 (5.72%)
R_{sym}^a in parentheses				
Number of unique reflections	7604	6234	6039	5944
Percent complete	93.0	96.4	92.9	91.4
Maximum resolution (Å)	2.40	2.60	2.60	2.60
Number of heavy atom sites	—	—	2	1
R_{cullis}^b (%)	—	—	55.8	62.1
Number of centric reflections	—	—	819	871
Overall phasing power ^c	—	—	1.68	2.31

^a $R_{\text{sym}} = \sum |I_i - I_m| / \sum I_m$ (I_i is the intensity of the measured reflection and I_m is the mean intensity of i observations).

^b $R_{\text{cullis}} = \sum |(F_{\text{PH}} \pm F_{\text{P}}) - F_{\text{H}}| / \sum |F_{\text{PH}} - F_{\text{P}}|$ (F_{P} , F_{PH} and F_{H} are, respectively, the protein, derivative and heavy atom structure factors).

^cPhasing power = $F_{\text{H}}/E_{\text{r.m.s.}}$ (F_{H} is the heavy atom structure factor and $E_{\text{r.m.s.}}$ is the residual lack of closure).

Semet is selenomethionine.

PCMBS is *p*-chloromercuribenzenesulfonate.

Overall figure of merit before solvent flattening = 0.619.

Final overall figure of merit after solvent flattening = 0.866 (solvent content = 0.4).

Fig. 2. Stereoview of electron density maps of ribosomal protein L6. (A) Solvent flattened MIR map of V124C. This region shows adjacent residues phenylalanine 36 and methionine 74 which are located between α -helix N- α -1 (top) and β -strand N- β -4 (bottom) in the hydrophobic core of the N-terminal domain. (B) ($2F_o - F_c$) map of the same region generated from the final model of the wild-type L6 protein at 2.6 Å resolution.

methionine 74 and cysteine 124, using the positions of the heavy atoms as landmarks. The initial model of V124C contained 144 out of the 177 residues in the protein. This structure was refined to an R -factor of 20.5% using all data between 6 and 2.8 Å with $F_o > 2\sigma(F_o)$. At this point, valine 124 was modelled back into the protein and the structure was refined against wild-type data to 2.6 Å. Currently, seven residues are missing from the N-terminus and nine from the C-terminus. The R -factor is 18.3% with a deviation from ideality of 0.016 Å in bond lengths and 3.37° in bond angles.

Structure of the protein

The L6 molecule contains two domains joined at a relative angle of 120° in a head-to-tail fashion to create an elongated, L-shaped structure with approximate dimensions 25 Å × 25 Å × 60 Å (Figure 3). The domains are approximately equal in size; the N-terminal domain spans amino acids 1–79, and the C-terminal domain 80–177. Each domain shares a common basic structure that consists of three distinct secondary structure elements, two three-stranded β -sheets and an α -helix, arranged approximately on the sides of a

Fig. 3. Ribbon diagrams of ribosomal protein L6. The L6 secondary structure elements are colored to show the domain architecture and the similarities between the N- (bottom) and C- (top) terminal domains. Within each domain, β -sheet 1 is red, β -sheet 2 is green, the α -helix is blue and connecting regions are yellow. All ribbon diagrams were drawn using the program 'O'.

triangular prism. The connectivities and topological arrangements of the six β -strands and the α -helix are identical in the two domains. The first β -sheet comprises strands 1, 5 and 6, the second β -sheet comprises strands 2, 3, 4, and the α -helix is found at the very C-terminus immediately following strand 6 (Figure 4). The six β -strand connections conform to the classic 'Greek key' motif (Richardson, 1977). The notation for the secondary structure elements and connecting loops, and their locations within the primary structure are shown in Figure 1.

There are two major differences between the L6 domains. First, the C-terminal domain contains an additional β -strand, C- β -7, which is connected to the end of C- α -1 by an extended loop. C- β -7 associates with the first β -sheet and extends it to a four-stranded structure. Second, N- α -1 contains six turns and is significantly longer than the C- α -1 which contains only 3.5 turns. After omitting these regions, together with the variable loops and turns, 62 α -carbons can be superposed with a root mean squared (r.m.s.) deviation of 1.75 Å (Figure 4). After superposition, the α -helices and β -sheets overlap almost perfectly. It is also evident that C- β -7 and its connecting loop is spatially equivalent to the extra turns at the C-terminus of N- α -1.

The L6 molecule contains many structurally important residues that are highly conserved and which therefore show that its overall fold is invariant between life-forms. These include amino acids that mediate the structure of each domain and others that define the domain interface. Of the

former, the most notable are the hydrophobic core residues and several glycines that occur at tight turns (Figure 1). Although the proline residues are not particularly conserved, those in the *B. stearotherophilus* molecule mostly occur at β turns. The only completely conserved proline in the molecule is proline 11 at the C-terminus of N- β -1 that directs the main chain at this point into the turn connecting sheet 1 with sheet 2. A series of conserved residues makes up the domain interface (Figures 1 and 5). The side chain of leucine 32 makes close contact with the main chain of the C-terminal domain near residue 135 and the aromatic residue at position 82 packs into the interface between the two domains. A series of glycine residues (27, 30, 77, 81 and 134) allows close contact between the domains. In particular, glycine 27 is adjacent to the absolutely conserved valine 78 which preserves the relative orientation of C- β -3 and N- α -1. Finally, the side chains of two absolutely conserved residues, aspartate 136, which serves as the N-cap (Richardson and Richardson, 1988) of C- α -1, and arginine 34 appear to be involved in a charge-charge interaction. Although the distance between these two side chains is too great for this to be a true salt bridge, these residues may be interacting between one or two water molecules.

In addition to these conserved structural residues, each domain contains conserved surface features that are presumably important to their functions in the ribosome. The natures of these sites are very different in the two domains. The C-terminal domain contains two adjacent regions with similar properties. The first (Figure 5) is on the side of the domain and comprises several positively charged residues (lysines 84, 137 and arginine 148), an exposed aromatic residue (tyrosine 163) and a cluster of small residues (glycine 141, alanines 144, 145 and 149) along the face of C- α -1 that borders this region. The side chain of arginine 148 is held in a defined orientation by a well-ordered salt bridge to glutamate 166. The second region (Figure 5) encompasses C- β -7 and the loop segments at the top of the molecule. It contains glycines 90, 92 and 107, lysines 157 and 159, arginines 151 and 162, and tyrosine 156. Also contributing to this region are serine 109 and histidine 110 which are uniquely conserved in eubacterial amino acid sequences. The principal surface feature of the N-terminal domain is a central region that is almost totally devoid of charged residues. None of the amino acids within this area are particularly well conserved although they tend to provide extensive opportunities for hydrogen-bonding interactions.

Discussion

Ribosomal protein structures have provided and will continue to provide invaluable insights into the structural basis of protein-RNA interactions (Wilson *et al.*, 1986; Ramakrishnan and White, 1992). Such interactions are crucial within the ribosome since it is now clear that a fundamental component is a correctly folded RNA moiety (Noller, 1991; Noller *et al.*, 1992). It appears that the 23S rRNA binding site on the L6 molecule is localized to the C-terminal domain. The two adjacent conserved regions on the surface of this domain (see Results and Figure 5) are populated with the types of residues that are present in the proposed RNA binding sites of other proteins (Wilson *et al.*, 1986; Nagai *et al.*, 1990; Hoffman *et al.*, 1991;

Fig. 4. Structural comparisons of the N- and C-terminal domains of ribosomal protein L6. (A) Schematic representation of the C-terminal domain secondary structure elements and their connectivities. The elements comprise a four-stranded β -sheet (sheet 1), a three-stranded β -sheet (sheet 2) and an α -helix. The β -strands are connected in a 'Greek-key' fashion. The N-terminal domain is identical apart from strand 7 (hatched) which is not present. (B) Schematic diagram showing the spatial organization of the secondary structure elements in the C-terminal domain. To facilitate comparison, the elements are shaded according to Figure 4A. The β -sheets and α -helix are arranged on the sides of a triangular prism. Sheet 1 and the α -helix constitute a region that is homologous to the RNA binding protein A-RRM-1 (see text). (C) Stereoview of the least squares superposition of the α -carbons of the two domains as determined by the 'O' program. The N-terminal domain is green, and the C-terminal domain is magenta.

Ramakrishnan and White, 1992). These include positively charged and aromatic side chains for direct interactions with the nucleic acid, and residues with small or no side chains that would allow a large RNA molecule to make a close approach to the protein. Mutations in L6 result in resistance to the antibiotic gentamicin, an aminoglycoside that interacts with 16S rRNA and induces the ribosome to misread the messenger RNA. Such mutations have previously been found in regions that appear to be binding sites for RNA (Ramakrishnan and White, 1992; Golden *et al.*, 1993). Although the sites of mutation have not been identified in L6, preliminary evidence suggests that they, like this proposed RNA binding site, are in the C-terminal domain (S. White, unpublished results).

L6 is known from crosslinking data to have many protein neighbors within the ribosome including L5 (Traut *et al.*, 1985), L7/L12 (Schiebe and Wagner, 1986), L18 (Traut *et al.*, 1985), L19 (Redl *et al.*, 1989) and L31 (Traut *et al.*, 1985). It is also associated with transiently bound elongation factor G (Sköld, 1982). Several surface regions of the protein are good candidates for sites of protein-protein interaction.

These typically contain many polar and non-polar residues and are devoid of charged amino acids. Such a surface is ideal for protein-protein interactions because the polar groups can form geometrically restricted hydrogen bonds that provide a high degree of specificity. Most of these regions on L6 occur in the N-terminal domain and they coalesce to form a band that encircles the middle of the domain. Thus, in the ribosome, the N-terminal domain may mediate many of the protein-protein interactions while the C-terminal half primarily interacts with RNA.

Although the angle between the two domains in the crystal structure appears to be fixed at $\sim 120^\circ$, it is pertinent to ask whether this elbow angle is merely an artifact of the crystal packing. We believe that it represents the actual conformation of the protein in the ribosome for the following reasons. At first sight there might appear to be flexibility between the two halves of the molecule, but the connecting secondary structural elements (N- α -1 and C- β -1) are in fact contiguous and the domains interact through the rather extensive and conserved contact surface described above (see Results and Figure 5). This feature appears to create a rigid

Fig. 5. The conserved domain interface and the putative RNA binding regions of ribosomal protein L6. The backbone in this view is color-ramped red-to-blue, N-terminus to C-terminus to show the path of the polypeptide chain. The side chains of the important residues which are fully described in the text are drawn in magenta, and the α -carbons of glycine residues are drawn as blue CPK atoms. (A) The conserved residues at the domain interface. This contact region appears to form a rigid interface between the two domains. (B) The residues that are proposed to interact with 23S rRNA. This proposed RNA binding region consists of a series of exposed, conserved residues that congregate on one side and across the top of the C-terminal domain to form one contiguous site.

two-domain structure with no flexible tether region. This is supported by proteolytic digestion studies which show that the molecule is sequentially degraded from the C-terminus and is not preferentially cut between the two domains (Littlechild *et al.*, 1987).

The crystal structure of L6 reveals that the protein contains two domains with nearly identical topology. The remarkable similarity of the two domains (Figure 3) suggests that L6 evolved from a single domain protein by a gene duplication event. This almost certainly occurred early in evolution since it is clear from the patterns of conservation in the amino acid sequences that L6 had evolved to its present structure and function before the divergence of the major life-forms. The antiquity of the event is supported by amino acid sequence analyses which fail to detect any significant homologies in the primary structures of the two domains.

Although the L6 domain structure is unique, it is related to a structural motif seen in several other proteins from ribonucleoprotein particles. It has previously been shown that three small proteins, ribosomal proteins L7/L12 and L30 and the RNA binding domain of U1 snRNP A protein (A-RRM-1), have similar structures in which α -helices are packed against one surface of an anti-parallel β -sheet (Leijonmarck *et al.*, 1980, 1988; Wilson *et al.*, 1986; Nagai *et al.*, 1990; Hoffman *et al.*, 1991). It was proposed that

this similarity, which includes identical topologies and connectivities of secondary structural elements, might indicate a common evolutionary origin for these molecules (Wilson *et al.*, 1986; Hoffman *et al.*, 1991; Leijonmarck *et al.*, 1988). This proposal is supported by L6 since its C-terminal domain also conforms to the same basic motif. The comparison with A-RRM-1 is the easiest to visualize and simply involves the replacement of β -sheet 2 with a second α -helix. If this evolutionary scenario is correct, it is clear that the N-terminal domain has undergone considerable changes to arrive at its present structure. For example, it no longer contains a β -strand that is equivalent to C- β -7 and which is an integral part of the basic motif, and it also appears to have lost its RNA binding elements.

We have recently determined the structure of ribosomal protein S5 and found that, in common with L6, it is rather elongated and contains two sub-domains connected by a conserved hydrophobic interface (Ramakrishnan and White, 1992). Although the two S5 domains are not similar either to themselves or to the motif described above, the molecule may also have evolved by the fusion of two smaller molecules (Ramakrishnan and White, 1992). If this turns out to be a general feature of the larger ribosomal proteins, it would support the notion that they evolved from a small number of primitive RNA binding proteins.

Materials and methods

Cells and strains

The gene encoding *B. stearothermophilus* L6 has previously been cloned and overexpressed in *E. coli* (Ramakrishnan and Gerchman, 1991). The L6 gene was then introduced into the plasmid vector pET-13a (S.E. Gerchman and V. Ramakrishnan, unpublished results) which provides kanamycin resistance as a selectable marker (Golden *et al.*, 1993). The *E. coli* host strain for this plasmid was BL21(DE3) (Studier *et al.*, 1990) which has the gene encoding T7 RNA polymerase under control of the *lac* operator in its genome. For labelling with selenomethionine, the plasmid was transformed into the strain B834(DE3) (Wood, 1966; Studier and Moffatt, 1986) which is auxotrophic for methionine.

A cysteine residue was engineered into the L6 protein by using a site-directed mutagenesis procedure that involves three distinct steps of the polymerase chain reaction (PCR) (Nelson and Long, 1989). In the first step, two oligonucleotides were used to make a 'superoligo' that contained both the required mutation within the gene and a unique DNA sequence (duso) in the 3' flanking region. The mixture for this reaction contained 20 ng of template DNA, 50 pmol of each oligonucleotide, 0.2 mM deoxynucleotide triphosphates (dNTPs) and 2.5 U Taq DNA polymerase. The template DNA was the pET-13 plasmid that contains the wild-type L6 gene. Twenty-five cycles of PCR were performed (94°C for 1 min, 45°C for 1 min and 72°C for 2 min) and this was followed by a single round at 45°C for 1 min, and 72°C for 7 min. The second step, which filled in the remaining gene sequence, involved a single cycle of PCR (95°C for 5 min, 45°C for 2 min and 72°C for 15 min) and used a mixture containing 100 ng of the superoligo, 20 ng of template DNA, 0.2 mM dNTPs and 2.5 U Taq DNA polymerase. The conditions for the final PCR step, which specifically amplified the mutant gene via the *duso* sequence, were identical to the first.

Normally, cells were grown at 37°C in an environmental shaker using LB as the growth medium. Cells were grown to an optical density of 0.8 and then induced with 100 mg/l of isopropyl- β -D-thiogalactopyranoside (IPTG). Three hours after induction, the cells were harvested by centrifugation for 10 min at 6000 g. When the protein was metabolically labelled with selenomethionine a defined medium was used. Each liter contained 12 g Na₂HPO₄, 6 g KH₂PO₄, 2 g NH₄Cl, 1 g NaCl, 4 g dextrose, 40 mg kanamycin, 1 mg biotin, 1 mg choline chloride, 1 mg folic acid, 1 mg niacinamide, 1 mg D-pantothenate, 1 mg pyridoxal and 0.1 mg riboflavin, 5 mg thiamine, 2 mM MgSO₄, 1 μ M FeCl₃, 0.2 mM CaCl₂, 40 mg/l each amino acid except methionine and 40 mg/l selenomethionine. Since selenomethionine is somewhat deleterious to cell growth, it was necessary to adapt the cells gradually to this medium. The cells were first inoculated into 2 ml of medium containing 100% methionine and this culture was then sequentially diluted into 2 ml cultures containing 50, 80 and 100% selenomethionine. Finally, the last culture was diluted into a second 2 ml culture containing 100% selenomethionine and this was used to inoculate a 2 l culture that was grown and induced as described above.

Protein purification

The cells were resuspended in a buffer containing 20 mM Tris-Cl pH 8.0, 10 mM MgCl₂, 50 mM NaCl, 0.5% Tween 20, 0.1 mM phenylmethylsulfonyl fluoride (PMSF), 0.05 mM benzamide (BAM), a few crystals of hen egg-white lysozyme and a few crystals of bovine DNase I. After incubation for 30 min at room temperature, the cells were frozen at -20°C for at least 12 h. After thawing, additional PMSF and BAM were added to a final concentration of 0.1 mM and 0.05 mM respectively and cell debris was removed by centrifugation of 15 000 g for 10 min. The supernatant was loaded onto a 40 ml S-Sepharose column equilibrated with buffer A (30 mM HEPES pH 7.5, 50 mM NaCl, 1 mM Na₂S₂O₃, 0.1 mM PMSF and 0.05 mM BAM) and washed with 10 column volumes of buffer A. The protein was eluted with a 10 column volume linear gradient of 0.0–1.0 M NaCl in buffer A. Fractions containing L6 were identified by SDS-PAGE, pooled and dialyzed against buffer B (10 mM Tris-Cl pH 8.0, 20 mM NaCl, 0.1 mM PMSF and 0.05 mM BAM). The protein was then loaded onto a 15 ml DEAE-Sepharose column equilibrated with buffer B. The flow-through, which contained L6, was recovered and concentrated to a final concentration of 10–15 mg/ml using Centricon-10 microconcentrators (Amicon). The purification of V124C was identical except that 10 mM β -mercaptoethanol was present in the resuspension buffer and 3 mM β -mercaptoethanol was present in all other buffers.

Crystallization

L6, V124C and selenomethionine-labelled V124C were crystallized at room temperature by vapor diffusion using the hanging-drop method (Ollis and White, 1987) with conditions that were similar to those previously described (Appelt *et al.*, 1983). In each case, the drops containing 10–15 mg/ml L6,

0.6 M Na/K phosphate pH 7.2–7.8, 3 mM β -mercaptoethanol and 3% (v/v) 1,4-dioxane were sealed over wells containing 1.6–1.8 M Na/K phosphate, 3 mM β -mercaptoethanol and 3% 1,4-dioxane. Crystals began to grow in 3 days and reached maximum size in 7–10 days.

Preparation of a mercury derivative

A mercury derivative of V124C was prepared by soaking crystals for 16 h in a solution containing 2.0 M Na/K phosphate pH 7.6, 0.1 mM sodium *p*-chloromercuribenzenesulfonate (PCMBS) and 3 mM β -mercaptoethanol. Increasing either the concentration of the mercurial or the length of the soak severely damaged the crystal.

X-ray data collection

Data were collected on a RAXIS II image plate system (Molecular Structures Corporation) mounted on a Rigaku RU300 X-ray generator running at 40 kV and 80 mA. Using standard oscillation geometry, 2° oscillation images were collected on each frame. Most of the crystals were mounted with their *c*-axis at an angle of ~10° to the oscillation axis in order to minimize the blind region. The mercury-soaked crystal was aligned so that its *c*-axis was nearly parallel to the oscillation axis. This allowed the Bijvoet pairs to be recorded simultaneously within the same frame to minimize the error in measuring the anomalous differences. The data were integrated, scaled and merged using the RAXIS software package.

Solving the crystal structure

The structure was solved using the standard method of multiple isomorphous replacement. Initial phase calculations were based on isomorphous differences from PCMBS and selenomethionine derivatives, and anomalous differences from the PCMBS derivative. The phases were subsequently improved by solvent flattening using a solvent content of 0.4 (Wang, 1985). All phase and map calculations were performed by the package 'PHASES' (Furey and Swaminathan, 1990). Model building was performed using the program 'O' (Jones *et al.*, 1991) running on an Evans and Sutherland ESV graphics system. The program XPLOR (Brünger *et al.*, 1987) was used to refine the model. Alternate cycles of refinement by simulated annealing followed by inspection of ($2F_o - F_c$) and ($F_o - F_c$) maps and manual rebuilding were used to improve the structure.

Acknowledgements

We thank Stephanie Porter for making the site-directed mutant used in this work; F. William Studier for providing the B834(DE3) cells for incorporating selenomethionine; Ronald Green for providing selenomethionine; Sue Ellen Gerchman and Helen Kycia for technical assistance; and David Hoffman, Christopher Davies and M. Otman for helpful discussions. This work was supported by NIH grant GM44973 (to S.W.W. and V.R.), and by a grant from the Office of Health and Environmental Research of the US Department of Energy (to V.R.).

References

- Appelt, K., Dijk, J., White, S.W. and Wilson, K.S. (1983) *FEBS Lett.*, **160**, 75–77.
- Appelt, K., Tanaka, I., White, S.W. and Wilson, K.S. (1984) *FEBS Lett.*, **165**, 43–45.
- Brünger, A.T., Kuriyan, K. and Karplus, M. (1987) *Science*, **235**, 458–461.
- Ceretti, D.P., Dean, D., Davis, G.R., Bedwell, D.M. and Nomura, M. (1983) *Nucleic Acids Res.*, **11**, 2599–2616.
- Cundliffe, E. (1990) In Hill, W.E., Moore, P.B., Dahlberg, A., Schlessinger, D., Garrett, R.A. and Warners, J.R. (eds), *The Ribosome, Structure, Function and Evolution*. American Society for Microbiology, Washington DC, pp. 479–501.
- Furey, W. and Swaminathan, S. (1990) *Am. Crystallogr. Assn. Abstr.*, Ser. 2, Vol 18, 73.
- Golden, B.L., Hoffman, D.W., Ramakrishnan, V. and White, S.W. (1993) *Biochemistry*, in press.
- Hackl, W. and Stöffler-Meilicke (1988) *Eur. J. Biochem.*, **174**, 431–435.
- Hoffman, D.W., Query, C.C., Golden, B.L., White, S.W. and Keene, J.D. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 2495–2499.
- Hoffman, D.W., Davies, C., Gerchman, S.E., Kycia, J.H., Porter, S.J., White, S.W. and Ramakrishnan, V. (1993) *EMBO J.*, **12**, 0000–0000.
- Jones, T.A., Zou, J.-Y., Cowan, S.W. and Kjeldgaard, M. (1991) *Acta Crystallogr.*, **A47**, 110–119.
- Kimura, M., Rawlings, N. and Appelt, K. (1981) *FEBS Lett.*, **136**, 58–64.
- Kühberger, R., Piepersberg, W., Petzet, A., Buckel, P. and Böck, A. (1979) *Biochemistry*, **18**, 187–193.

- Leijonmarck,M., Ericksson,S. and Liljas,A. (1980) *Nature*, **286**, 824–826.
- Leijonmarck,M., Appelt,K., Badger,J., Liljas,A., Wilson,K.S. and White,S.W. (1988) *Proteins*, **3**, 243–251.
- Littlechild,J., Malcolm,A., Paterakis,K., Ackermann,I. and Dijk,J. (1987) *Biochim. Biophys. Acta*, **913**, 245–255.
- Mitchell,P., Osswald,M., Schueler,D. and Brimacombe,R. (1990) *Nucleic Acids Res.*, **18**, 4325–4333.
- Nagai,K., Oubridge,C., Jessen,T.H., Li,J. and Evans,P.R. (1990) *Nature*, **348**, 515–520.
- Nelson,R.M. and Long,G.L. (1989) *Anal. Biochem.*, **180**, 147–151.
- Noller,H.F. (1991) *Annu. Rev. Biochem.*, **60**, 191–227.
- Noller,H.F., Hoffarth,V. and Zimniak,L. (1992) *Science*, **256**, 1416–1419.
- Ollis,D. and White,S.W. (1987) *Methods Enzymol.*, **182**, 646–659.
- Ramakrishnan,V. and Gerchman,S.E. (1991) *J. Biol. Chem.*, **266**, 880–885.
- Ramakrishnan,V. and White,S.W. (1992) *Nature*, **358**, 768–771.
- Redl,B., Waliczek,J., Stöffler-Meilicke,M. and Stöffler,G. (1989) *Eur. J. Biochem.*, **181**, 351–356.
- Richardson,J.S. (1977) *Nature*, **268**, 495–500.
- Richardson,J.S. and Richardson,D.C. (1988) *Science*, **240**, 1648–1652.
- Schiebe,U. and Wagner,R. (1986) *Biochim. Biophys. Acta*, **869**, 1–7.
- Scholzen,T. and Arnot,E. (1991) *Mol. Gen. Genet.*, **228**, 70–80.
- Sköld,S.E. (1982) *Eur. J. Biochem.*, **127**, 225–229.
- Studier,F.W. and Moffatt,B.A. (1986) *J. Mol. Biol.*, **189**, 113–130.
- Studier,F.W., Rosenberg,A.H., Dunn,J.J. and Dubendorff,J.W. (1990) *Methods Enzymol.*, **185**, 60–89.
- Suzuki,K., Olvera,J. and Wool,I.G. (1990) *Gene*, **93**, 297–300.
- Traut,R.R., Tewari,D.S., Sommer,A., Gavino,G.R., Olson,H.M. and Glitz,D. (1985) In Hardesty,B. and Kramer,G. (eds), *Structure, Function and Genetics of Ribosomes*. Springer-Verlag, New York, pp. 286–308.
- Wang,B.-C. (1985) *Methods Enzymol.*, **115**, 90–112.
- Wilson,K.S., Appelt,K., Bader,J., Tanaka,I. and White,S.W. (1986) *Proc. Natl Acad. Sci. USA*, **83**, 7251–7255.
- Wood,W.B. (1966) *J. Mol. Biol.*, **16**, 118–133.
- Wower,I., Wower,J., Meinke,M. and Brimacombe,R. (1981) *Nucleic Acids Res.*, **9**, 4285–4302.

Received on August 4, 1993