Quaternary Structure and Catalytic Activity of the *Escherichia coli* Ribonuclease E Amino-Terminal Catalytic Domain†

Anastasia J. Callaghan,‡ J. Günter Grossmann,§ Yulia U. Redko,|| Leopold L. Ilag,⊥ Martin C. Moncrieffe,‡ Martyn F. Symmons,‡ Carol V. Robinson,⊥ Kenneth J. McDowall,|| and Ben F. Luisi*,‡

Department of Biochemistry, University of Cambridge, 80 Tennis Court Road, Cambridge CB2 1GA, U.K., Molecular Biophysics Group, CCLRC Daresbury Laboratory, Daresbury, Warrington WA4 4AD, U.K., Astbury Centre for Structural Molecular Biology, Faculty of Biological Sciences, University of Leeds, Leeds LS2 9JT, U.K., and Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, U.K.

Received June 27, 2003; Revised Manuscript Received September 22, 2003

ABSTRACT: RNase E is an essential endoribonuclease that plays a central role in the processing and degradation of RNA in *Escherichia coli* and other bacteria. Most endoribonucleases have been shown to act distributively; however, Feng *et al.* [(2002) *Proc. Natl. Acad. Sci. U.S.A. 99*, 14746−14751] have recently found that RNase E acts via a scanning mechanism. A structural explanation for the processivity of RNase E is provided here, with our finding that the conserved catalytic domain of *E. coli* RNase E forms a homotetramer. Nondissociating nanoflow-electrospray mass spectrometry suggests that the tetramer binds up to four molecules of a specific substrate RNA analogue. The tetrameric assembly of the N-terminal domain of RNase E is consistent with crystallographic analyses, which indicate that the tetramer possesses approximate D_2 dihedral symmetry. Using X-ray solution scattering data and symmetry restraints, a solution shape is calculated for the tetramer. This shape, together with limited proteolysis data, suggests that the S1-RNA binding domains of RNase E lie on the periphery of the tetramer. These observations have implications for the structure and function of the RNase E/RNase G ribonuclease family and for the assembly of the *E. coli* RNA degradosome, in which RNase E is the central component.

The degradation of messenger RNA in *Escherichia coli* is initiated principally through the hydrolytic activity of the essential endoribonuclease, RNase E (*1*, *2*). But in addition to its role in RNA degradation, RNase E also processes specific RNA molecules to create functionally active forms, including precursors of ribosomal RNA (*3*−*5*), transfer RNA (*6*), the RNA component of RNase P (M1 RNA; *7*), and 10Sa RNA (*8*), which is the tmRNA that rescues those ribosomes stalled by aberrant termination.

With limited sequence specificity, RNase E tends to cut transcripts within single-stranded segments (*9*−*11*) that are rich in A/U nucleotides (*12*). One well-studied substrate of RNase E is RNAI (*13*), which is the antisense regulator of the replication of ColE1-type plasmids. Like many mRNAs, RNAI is cleaved by RNase E close to the 5′ end (*14*, *15*). Certain sites preferred by RNase E, e.g., near the 5′ end of RNAI, can be cleaved as short oligonucleotides, whereas others, such as the b site in the 9S precursor of 5S rRNA, appear to require the presence of additional, but as yet uncharacterized, structural elements in the RNA substrate (*16*). In many cases, the cleavage by RNase E is "5′ end-dependent", in that the enzyme will efficiently cut substrates having a monophosphate group at the 5′ end, but not transcripts that have a 5′-triphosphate group or lack an accessible 5′ end (*17*, *18*).

The RNase E gene encodes one of the largest polypeptides in *E. coli* (1061 residues). The N-terminal domain of RNase E (residues 1−529) possesses the site of catalytic activity (*19*), and mutants that express this domain alone are viable. Such mutants can process ribosomal RNA satisfactorily, although they degrade bulk RNA more slowly and are outgrown by cells expressing the full-length wild-type RNase E (*20*−*22*), suggesting that the C-terminal domain (residues 530−1061) confers a selective advantage.

The C-terminal domain of *E. coli* RNase E functions as a scaffold in the assembly of an ATP-dependent RNA helicase (RhlB), a phosphate-dependent 3′-exonuclease [polynucleotide phosphorylase (PNPase)[1]], and a glycolytic enzyme (enolase); the association of these (and other) proteins forms a complex known as the RNA degradosome (*23*−*26*). It is likely that the selective advantage provided by the full-length protein, in comparison with the shorter N-terminal domain, is related to the coordination of RNA degrading and processing activities within the degradosome (*25*−*27*).

Molecular genetic studies have identified the segments of the C-terminal domain of *E. coli* RNase E that function in self-association and the binding of enolase, RhlB, and PNPase (*28*) and have indicated that arginine-rich segments

† This work was supported by the Wellcome Trust. K.J.M. is the recipient of a fellowship from the Royal Society.

* To whom correspondence should be addressed: Department of Biochemistry, University of Cambridge, 80 Tennis Court Road, Cambridge CB2 1GA, U.K. Telephone: +44 (0)1223 766019. Fax: +44 (0)1223 766002. E-mail: ben@cryst.bioc.cam.ac.uk.

‡ Department of Biochemistry, University of Cambridge.

§ CCLRC Daresbury Laboratory.

|| University of Leeds.

⊥ Department of Chemistry, University of Cambridge.

[1] Abbreviation: PNPase, polynucleotide phosphorylase.

10.1021/bi0351099 CCC: $25.00
Published on Web 11/04/2003

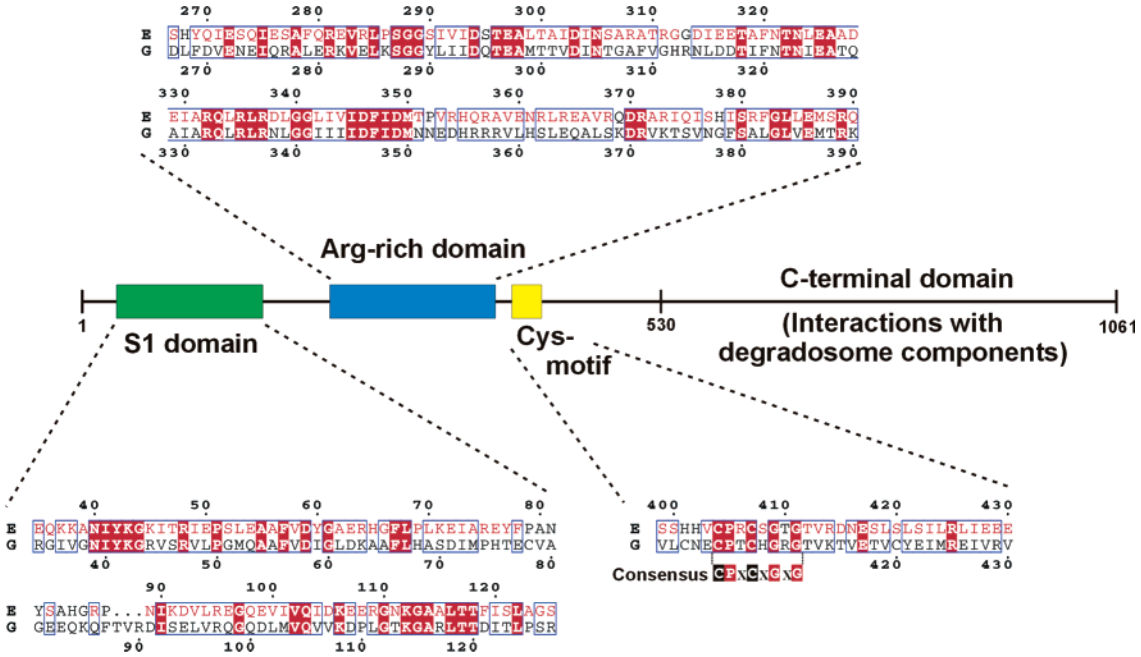

FIGURE 1: Schematic of the *E. coli* RNase E sequence showing subdomains. The conserved regions in the N-terminal domain are shown as colored boxes with residue numbers indicated for the *E. coli* RNase E RNA binding S1 domain (green), a central Arg-rich domain (blue), and a conserved Cys motif (yellow). Isolated sequence motifs also occur in the C-terminal region but are not shown. The multiple-sequence alignments are from ClustalW (*56*) alignments of bacterial RNases E with *E. coli* RNase E and RNase G, but for simplicity, only the last two sequences are shown (top and bottom sequences labeled E and G, respectively, and numbered appropriately). The sequences are formatted with ESPript (*57*) and grouped into RNases E and G separately. Blue boxes represent regions of global similarity between the E and G representatives [Riesler substitution matrix (*57*) with a similarity score of better than 60%], while red highlighted residues are identical among all the representatives. Residues shown with red text are considered similar throughout the RNase E examples alone (Riesler matrix but 70% threshold). Conserved Cys positions, likely to be Zn ligands, are highlighted in black within the consensus sequence of the Cys motif. Sequences in alignments but not shown here were from *Salmonella typhi*, *Haemophilus influenzae*, *Neisseria meningitidis*, *Pseudomonas aeruginosa*, *Xylella fastidiosa*, *Xanthomonas axonopodis*, *Yersinia pestis*, *Pasteurella multocida*, *Vibrio cholerae*, and *Buchinera aphidicola*.

function in RNA binding (*16*, *19*). This C-terminal domain is generally found in RNase E homologues from the enterobacteriaceae and related Gram-negative bacteria, although it varies in length, composition, and the arrangement of defined sequence motifs (Pfam release 7.8; *29*). The C-terminal domain is punctuated with regions that have low levels of sequence complexity, which suggests it has little secondary structure, and some of the conserved motifs are highly Pro/Gly-rich, which indicates they may have an extended conformation (J. Aurikko, personal communication). For most bacterial species, it is not yet known if these motifs recruit components to form an assembly analogous to the *E. coli* degradosome. In the RNase E from *Streptomyces coelicolor*, the domain organization is permuted so that the segment corresponding to the C-terminus of *E. coli* precedes, rather than follows, the catalytic domain (*30*). Like the *E. coli* C-terminal domain, the permuted domain of *Streptomyces* also binds PNPase.

Homologues of the *E. coli* RNase E N-terminal catalytic domain have been identified in other bacterial species and in archae, plastid genomes, and the nuclear genomes of several higher plants (*31*−*34*). *E. coli* itself contains a paralogue of RNase E, known as RNase G (the product of the *cafA* gene) which is closely homologous to the RNase E N-terminal domain (Figure 1). RNase G appears to overlap functionally somewhat with RNase E (*35*−*37*). Sequencing has identified RNase G-type enzymes in almost all known bacteria, with some organisms additionally possessing an RNase E-type enzyme (*32*, *38*, *39*).

Within the conserved N-terminal domain of RNase E and/or RNase G, there is an S1-like subdomain, which is a highly distributed and ancient single-stranded RNA-binding motif [*E. coli* RNase E residues 35−125; Figure 1 (*40*)]. The S1 domain is required for RNA cleavage by RNase E (*41*). Also present is an arginine-rich region (residues 267−390; Figure 1) that may contact the RNA (*31*). A conserved cysteine motif [residues 404−411 (CPxCxGxG); Figure 1] is located near the arginine-rich region (*32*), and these motifs are observed to coordinate zinc in the chaperone DnaJ (PDB entry 1EXK). We have found that RNase E binds zinc (A. J. Callaghan, J. G. Grossmann, L. Murphy, and E. Garman, unpublished observations), and it seems likely that the cysteine motif could form a metal-sharing dimerization interface or an RNA recognition surface.

Here, we show using a variety of biophysical and functional techniques that the N-terminal domain of *E. coli* RNase E forms a tetramer, and that this is catalytically active. We propose a model for the quaternary structure based on the findings of crystallography and X-ray solution scattering. The quaternary structure may occur throughout the wider family of the bacterial, archeal, and plastid RNase E and RNase G proteins. Finally, we discuss the implications of the RNase E quaternary structure for the assembly and activity of the RNA degradosome.

MATERIALS AND METHODS

Expression Vectors. Plasmid pRne529-N for the production of the RNase E N-terminal domain (residues 1−529) with

an oligohistidine tag at the N-terminus was provided by V. Kaberdin (Vienna Biocenter, Vienna, Austria). The construct was based on pET16b (Novagen). The primers that were used were 5′-GGAATTCCATA*TG*AAAAGAATGTTAATCAACG (start codon in italics and *Nde*I site underlined) and 5′-CGGGATCC**TTA**CAGCGCAGGTTGTTCCG (antistop codon in bold and *Bam*HI site underlined). The integrity of the clone was confirmed by DNA sequencing.

Expression and Purification. The N-terminal domain protein was purified from cultures of *E. coli* BL21(DE3) cells harboring the pRne529-N plasmid, using modifications of the immobilized metal affinity chromatography system used previously (*42*). Cell pellets were resuspended in binding buffer [20 mM Tris-HCl (pH 7.9), 5 mM imidazole, 0.5 M NaCl, and Complete EDTA-free protease inhibitors (Roche)]. Protein fractions were buffer exchanged into binding buffer containing 10 mM DTT, 10 mM $MgCl_2$, 0.5 mM EDTA, and 5% (v/v) glycerol, and this sample was further purified by gel filtration chromatography (Superdex 200 column, Amersham Biosciences). The purified protein was homogeneous as determined by SDS−PAGE and amino-terminal sequencing and had no detectable nucleic acid by UV absorption.

Oligonucleotide Synthesis. The RNA decamer 5′P-RNA10-FAM [ACAGU↑AUUUG sequence (cleavage site indicated with an arrow), 5′-monophosphate, 3′-fluorescein (FAM)] was based on the cleavage site at the 5′ end of pBR322 RNAI (*19*, *42*, *43*) and was synthesized by C. Adams (University of Leeds). RNA decamer 5′P-2′*O*-RNA10-FAM (5′-monophosphate, 3′-FAM) was synthesized with protective 2′-*O*-methyl groups at all positions by C. Hill (University of Cambridge).

Enzyme Activity Assay. Reactions were initiated by mixing the N-terminal domain with 5′P-RNA10-FAM in 25 mM Bis-Tris-Propane-HCl (pH 8.3), 100 mM NaCl, 15 mM $MgCl_2$, 0.1% (v/v) Triton X-100, and 1 mM DTT. Final substrate and enzyme concentrations were 100 and 1 nM, respectively. Aliquots of reaction mixtures (incubated at 37 °C) were removed at time intervals and the reactions quenched by mixing the aliquots 1:1 with 20 mM EDTA, and the reaction products were then analyzed by ion-pair reverse phase chromatography using a Dionex HPLC system with an in-line fluorescence detector and automated sample injector (*44*). The column was calibrated using a one-nucleotide ladder generated by incubating the 10-mer substrate with S1 nuclease as described previously (*44*).

Chemical Cross-Linking and Limited Proteolysis. The N-terminal domain was cross-linked with dimethyl suberimidate (Pierce), and the reaction products were analyzed by SDS−PAGE. Limited proteolysis was carried out with 0.1% (w/w) chymotrypsin (Pierce), 50 mM Tris-HCl (pH 8.0), and 1 mM $CaCl_2$. Reactions (25 °C) were quenched [aliquots boiled in SDS loading buffer (NuPage)], and the mixtures were analyzed by SDS−PAGE (NuPage, Invitrogen), electroblotting onto PVDF membranes (Bio-Rad) and amino-terminal sequencing (PNAC Facility, University of Cambridge).

Analytical Size Exclusion Chromatography. Samples of the purified N-terminal domain (typically 1 mg/mL) were run on an analytical size exclusion column (Superdex 200 HR 10/30, Amerhsam, Pharmacia). The column was equilibrated with running buffer [20 mM Tris-HCl (pH 7.9) and 500 mM NaCl] and calibrated using a Molecular Weight Marker Kit for gel filtration chromatography (Sigma).

Circular Dichroism. Spectra of the N-terminal domain [1 mg/mL; 50 mM sodium potassium phosphate buffer (pH 7.5)] were collected with an Aviv CD 215 spectrometer (1 mm path length cuvette, 25 °C, 260−195 nm).

Nanoflow-Electrospray Mass Spectrometry. The N-terminal domain and N-terminal domain−RNA complexes were buffer exchanged into 1 M ammonium acetate (Micro Bio-Spin chromatography columns, Bio-Rad). Mass spectra of protein (3 μg/μL) and the protein−RNA complex (3 μg/μL, 2-fold stoichiometric excess of 5′P-2′*O*-RNA10-FAM) were collected on a modified Q-ToF2 mass spectrometer (*45*) equipped with a nanoflow Z-spray source (Micromass, Manchester, U.K.). Nano-ES capillaries were prepared in-house from borosilicate glass tubes [1 mm outside diameter and 0.78 mm inside diameter (Harvard Apparatus, Holliston, MA)] and coated with gold [SEM sputter coater (Polaron, Newhaven, U.K.)]. Capillary tips had inner diameters of 1−5 μm. Pressures and accelerating potentials in the mass spectrometer were adjusted to preserve noncovalent interactions (positive ion mode), in addition to the following settings: capillary voltage of 1.7 kV, cone gas of 100 L/h, sample cone of 90−200 V, extractor cone of 0−10 V, hexapole ion-guide stage pressure of 6.0×10^{-2} mbar, and ToF analyzer pressure of 4.0×10^{-7} mbar. Stripping of nonspecific adducts was achieved by accelerating ions through the collision cell with the collision energy set at 100−200 V. External calibration was with cesium iodide, and calibration, acquisition, and processing were carried out using MassLynx software (Micromass).

Analytical Ultracentrifugation. Sedimentation velocity data of the N-terminal domain (3 mg/mL) were obtained using a Beckman Optima XL-I analytical ultracentrifuge (Beckman Coulter) with an interference optical system. Epon double-sector centerpieces containing 400 μL of protein solution and 410 μL of sample buffer were centrifuged at 52 000 rpm and 20 °C and data acquired at 5 min intervals with a radial increment of 0.002 cm. Data analysis was performed with Sedfit (*46*). Buffer viscosity (0.01059 cp), density (1.02 g/cm^3), protein partial specific volumes (0.736 g/cm), and the degree of hydration (0.43 g/g) were estimated using the program Sednterp (*47*).

X-ray Solution Scattering. Data were collected at Station 2.1, Synchrotron Radiation Source, Daresbury Laboratory, using a two-dimensional multiwire proportional counter at sample-to-detector distances of 1.25 and 4.25 m. Data from matching buffers were collected and subtracted from the protein profiles. The radius of gyration R_g and intraparticle distance distribution function [$p(r)$] were calculated from the scattering data using the indirect Fourier transform method program GNOM (*48*). This also allowed estimation of the maximum particle dimension D_{max}. Molecular shapes were reconstructed from the scattering profile (after merging high- and low-resolution data) using the *ab initio* procedure of simulating annealing with clustered spheres representing amino acid residues (*49*) (gasbor versions 1.7, 1.8, and 2.0). This results in a three-dimensional (3D) arrangement of scattering centers (spheres 3.5 Å in diameter) that reproduce the one-dimensional (1D) scattering curve in accordance with the experimental data. These are displayed using the space fill option to produce the overall molecular shape.

Table 1: Statistics of Crystallographic Data

space group[a]	*P*1
unit cell dimensions	$a = 73.24$ Å
	$b = 75.57$ Å
	$c = 109.39$ Å
	$\alpha = 94.95°$
	$\beta = 102.03°$
	$\gamma = 91.76°$
no. of reflections (unique)	358 196 (29 332)[b]
resolution range (Å)	50−3.4
completeness (%) (3.52−3.40 Å shell)	97.6 (94.5)
R_{merge}[c] (%) (3.52−3.40 Å shell)	8.8 (32.9)
I/σ (3.52−3.40 Å shell)	12.5 (2.8)
Matthew's coefficient (Da/Å^2) (four molecules assumed)	2.38
solvent content (%) (four molecules assumed)	47.9

[a] Data at a wavelength of 1.28 Å (100 K) collected on station 5.0.2, Advanced Light Source, Lawrence Berkeley National Laboratory (Berkeley, CA). [b] Data processed with the HKL2000 suite (*55*). [c] $R_{merge} = \sum_{hkl}\sum_j(|I_j - \langle I_{hkl}\rangle|)/\langle I_{hkl}\rangle$, where I_j is the individual measurement of the intensity of a reflection with a mean intensity $\langle I_{hkl}\rangle$ and the sum $\sum_{hkl}$ is over all *h*, *k*, and *l* Miller indices.

Crystallization and Data Collection. Crystallization (vapor diffusion) was carried out by mixing protein (7 mg/mL) 1:1 with 0.18 M Li_2SO_4, 0.09 M Tris-HCl (pH 8.5), 27% (w/v) polyethylene glycol 4000, and 10% (v/v) glycerol. A single crystal (300 μm × 100 μm × 50 μm) with added 30% (v/v) polyethylene glycol 200 was immediately frozen at 100 K. Statistics of collected data are given in Table 1. A self-rotation function was calculated with POLARRFN, AMORE, and other programs from the CCP4 suite (*50*). The crystal contained the intact N-terminal domain as determined by N-terminal sequencing and mass spectrometry analysis (L. Packman, University of Cambridge).

RESULTS

The recombinant N-terminal domain of *E. coli* RNase E (residues 1−529) was purified to homogeneity. A circular dichroism spectrum of the purified sample indicated a mixed α/β composition (data not shown), which is consistent with a folded conformation and with predictions of the secondary structure based on the sequences of the wider RNase E/RNase G family.

Earlier studies of RNase E using yeast two-hybrid analyses indicated a propensity for self-association (*28*). In agreement with these *in vivo* data, we observe that the N-terminal domain elutes at the expected size for a tetramer from size exclusion chromatography (Figure 2A). Treatment with the cross-linking reagent dimethyl suberimidate gave a pattern on SDS−PAGE gels consistent with a tetrameric species (Figure 2B). The purified domain cleaves 5′P-RNA10-FAM, a 3′-fluorescein-labeled version of a decanucleotide substrate derived from the major cleavage site in pBR322 RNAI (Figure 2C,D). These observations, made with the N-terminally tagged protein, are consistent with the activity and oligomerization state of the purified domain containing a C-terminal histidine tag (*44*), indicating that the presence and position of the tag do not affect the protein function.

Evaluation of the N-terminal domain by sedimentation velocity analytical ultracentrifugation revealed that the protein is predominantly tetrameric, having an apparent sedimentation coefficient of 7.88 S and an approximate molecular mass of 300 kDa (Figure 3). Analyses at lower protein concentrations gave the same results (data not shown).

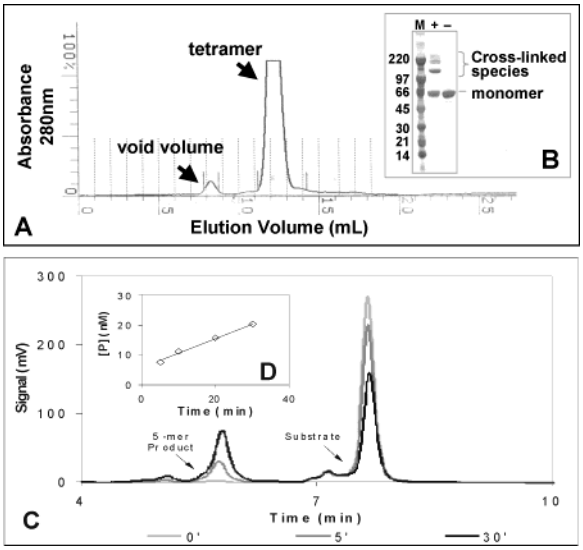

FIGURE 2: Size exclusion chromatography, cross-linking, and catalytic activity of the tetrameric N-terminal domain. (A) Size exclusion chromatography profile of the RNase E N-terminal domain. The elution of the protein at 12−13 mL on an analytical HR10/30 Superdex 200 column (Amersham Pharmacia Biotech) corresponds to a species with an apparent molecular mass of approximately 290 kDa (equation of best fit calibration curve is $\log(MW) = -0.211V + 5.11$, where MW is the molecular mass in kilodaltons and *V* is the elution volume in milliliters. (B) Chemical cross-linking reaction. Lane M corresponds to the marker track, and lanes + and − correspond to the N-terminal domain in the presence and absence of the DMS chemical cross-linker, respectively. The monomeric form of the N-terminal domain at ~66 kDa is labeled, as are the cross-linked species. Oligomers corresponding to a dimer (at >97 kDa), trimer (at ~220 kDa), and tetramer (at >220 kDa) are seen. (C) Cleavage of a decanucleotide substrate by the N-terminal domain. The pentanucleotide product was separated from the substrate using ion-pair reverse phase chromatography and the column calibrated as described previously (*44*). (D) Time course showing the rate of product release by the endonucleolytic cleavage of 5′P-RNA10-FAM by the N-terminal domain. The assay conditions are described in Materials and Methods. The straight line is a best fit to all the data.

Nanoflow-electrospray mass spectrometry of the N-terminal domain and its RNA complex was carried out under conditions in which noncovalent interactions were maintained, allowing preservation of oligomers and association with ligands (*51*). The spectrum for the N-terminal domain (Figure 4A) showed a series of charge states, which are centered around 7200 *m*/*z*, and that represent one stable species. Each monomer of the N-terminal domain has a theoretical mass of 61 884 Da. The measured mass of 248 281 ± 31 Da corresponds to a tetramer (theoretical mass of 247 537 Da) with additional mass that can be attributed to the binding of water and ions, which is normally observed in intact noncovalent complexes. The tetramer was the only species observed. However, the tetrameric complex could be dissociated in the collision cell, which revealed that essentially all monomers were intact and without modification.

Feng *et al*. (*52*) have reported that the N-terminal domain of RNase E will bind specifically to 2′-*O*-methylated substrates, derived from RNAI and related to the RNA 10-mer ligands we have used in this study. They also found

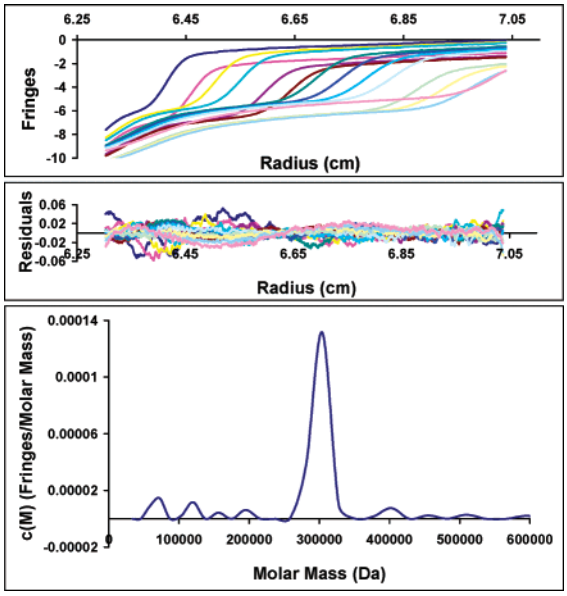

FIGURE 3: Analytical ultracentrifugation analysis of the N-terminal domain. The top panel shows the raw sedimentation velocity data, obtained at 52 000 rpm and at 5 min intervals (3 mg/mL). The data are presented as fringe displacements (which is proportional to the protein concentration) plotted against the distance from the axis of rotation. Each measurement is separately colored. The center panel shows the residuals after fitting to a continuous concentration vs molecular mass, $c(M)$, model. The rmsd for the fit was 0.012, and the sedimentation coefficient was 7.88 S. The bottom panel shows the distribution of molecular masses obtained from the sedimentation velocity data. Data analysis was performed using the program Sedfit (*46*).

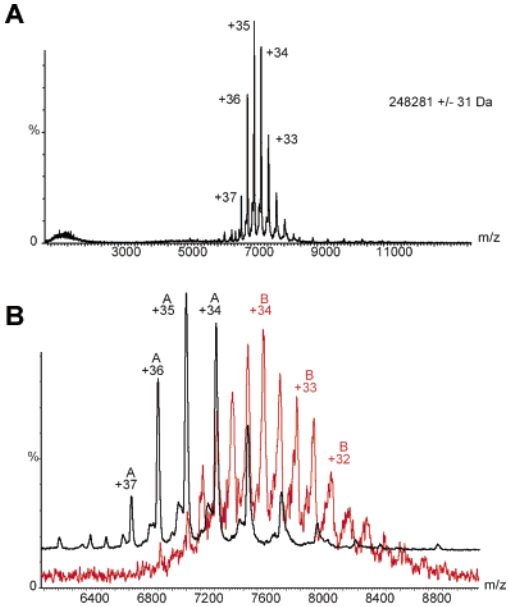

FIGURE 4: Nanoflow-ESI mass spectra of the N-terminal domain. (A) ESI mass spectrum of the N-terminal domain under nondissociating conditions. Individual charged species of the tetrameric form are resolved by the mass/charge ratio (m/z), and five abundant charged states are labeled. Ordinate axis is the percent abundance relative to the maximum peak height (+35 species). (B) The top spectrum (black line) is shown vertically offset from zero for clarity and corresponds to an expansion from m/z 6300−9000 of the spectrum in panel A. The charge states are labeled as series A (+37 to +34). The bottom spectrum (red line) was recorded for a solution of the N-terminal domain after addition of 5′P-2′*O*-RNA10-FAM. This spectrum has equivalent charge states at higher masses because of the RNA bound to the tetramer (series B, +34 to +32 labeled).

that the modification does not affect specific cleavage at nonmodified bonds. To test RNA binding, a synthetic oligoribonucleotide with a 5′-monophosphate (5′P-2′*O*-RNA10-FAM) was added in a slight stoichiometric excess over the protein, and at concentrations exceeding the estimated dissociation constant. Figure 4B shows the superposition of the spectra for the N-terminal domain in the absence (trace A, black) and presence (trace B, red) of RNA. The latter spectrum shows additional peaks corresponding to the binding of RNA to the protein. By comparison of the acquired spectra with simulations of the possible complexes, populations of one, two, three, and four bound RNA molecules were observed [these are observed in the fine structure of the peaks in series B (red), Figure 4B]. For example, one of the shifts of the +34 state is 347 m/z units, corresponding to a mass difference of 11 798 Da. This particular shift may be attributed to the binding of three RNA molecules. Other shifts can be attributed to one, two, or four bound RNA molecules. These data are consistent with a tetramer being able to bind up to four molecules of RNA. That a distribution of bound species is observed reflects the binding equilibrium and possible loss of RNA during the mass spectrometry analysis; it also suggests that binding might be independent (i.e., noncooperative). Finally, we note that the presence of RNA does not appear to affect the ability of the N-terminal domain to remain a tetramer.

Crystals of the native N-terminal domain were obtained, but heavy atom derivatives are not yet in hand. However, the potential molecular symmetry of the N-terminal domain can be derived from the diffraction pattern of the native crystals. The volume of the unit cell can accommodate four protomers of the N-terminal domain with 48% solvent content, which is a value typical for macromolecular crystals. The self-rotation function of the native crystal data indicates that the N-terminal domain molecules may be related by imperfect D_2 (222 point) symmetry (Figure 5A). Three dyad axes are observed that are mutually perpendicular and are thus likely to be intersecting. There are two such sets, and within both sets, the three mutually perpendicular dyad axes are at different strengths. These features may indicate that the subunits of the tetramer comprise two or more domains that are organized around slightly misaligned tetrameric axes.

For further structural analysis of the shape of the tetramer, the N-terminal domain was analyzed by X-ray solution scattering (Figure 5B). The solution scattering profiles were used to estimate the radius of gyration ($R_g = 51.0 \pm 0.3$ Å) and the maximum particle dimension ($D_{max} = 160 \pm 5$ Å); these are consistent with the proposed tetrameric state of the protein.

Although the scattering profiles can provide a shape for RNase E, they cannot provide information about the molecular symmetry; however, molecular shapes can be calculated with the constraint that they have D_2 symmetry, as inferred from the crystallographic observations. These calculations produce a 3D arrangement of spherical scattering centers that are in accordance with the experimental 1D scattering curve. The smooth line in Figure 5B represents the scattering profile obtained from the shape reconstruction of a tetramer with D_2 symmetry. The corresponding molecular shape is shown in Figure 5C; here the scattering spheres are displayed using a space filling representation to produce the overall molecular

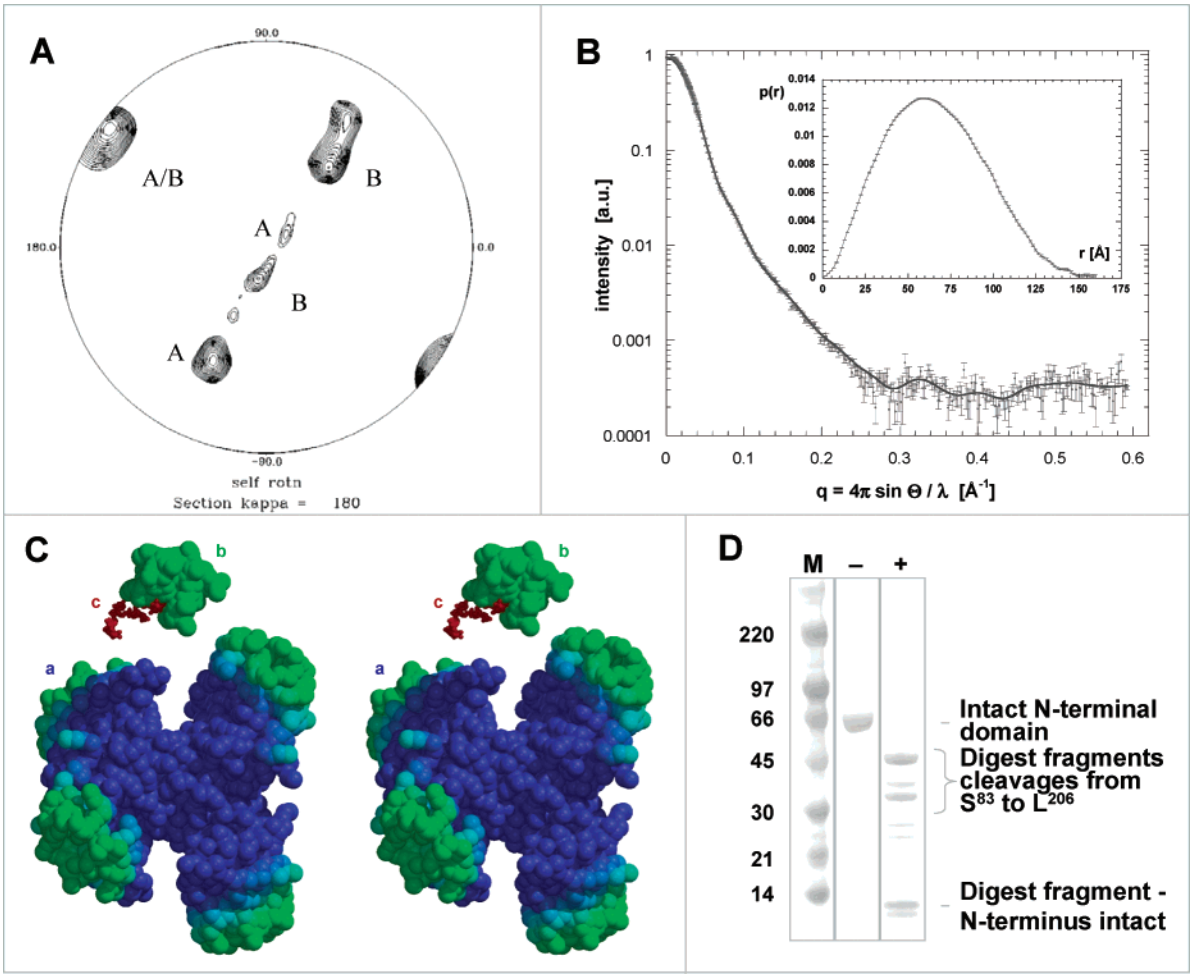

FIGURE 5: Molecular symmetry and shape of the N-terminal domain. (A) Self-rotation function of data from N-terminal domain native crystals. The self-rotation function of the *P*1 crystal data (Table 1) is represented as a polar stereographic plot for a κ of 180° (corresponding to dyad axes) as a function of the other pair of polar geometry rotation angles (but restricted to one hemisphere; POLARRFN). The contours show peaks representing the angular position of possible noncrystallographic dyad axes. These 2-fold symmetry axes form mutually orthogonal sets that are denoted with the letters A and B. The 2-fold axes within a set (A or B) are likely to be intersecting; however, the three 2-fold axes are at different heights, so the tetramer may have imperfect D_2 symmetry. Sets A and B may be related by imperfect symmetry of the pseudo-monoclinic lattice. Contours are set at 70% of the maximal peak height. (B) The X-ray scattering profiles are shown for the N-terminal domain. The smooth line represents the scattering profile of the tetramer with D_2 symmetry (see panel C). The probability distribution for the distances between scattering centers within the molecule is shown in the inset. The area under $p(r)$ was scaled to unity. (C) A stereoscopic view (crossed-eye version) of a possible solution for the N-terminal domain molecular envelope (a, shown blue to green) based on the solution scattering profile and D_2 symmetry constraint. The residues of the tetramer are modeled as identical spheres (shown with a 4 Å diameter) and colored with a blue-to-green gradient, corresponding to low-to-high radius from the center of the model. Also shown for comparison is a space filling model of the Rho helicase S1 domain (residues 46−118, green, b) and a stick model of a bound RNA 6-mer (C_6, red, c) (PDB entry 2A8V). Panel C was produced with Molscript (*58*); the gradient was colored with the pdb_depth program. (D) Proteolytic cleavage pattern for the N-terminal domain partially digested with chymotrypsin. Lane M corresponds to the marker track, and lanes + and − correspond to samples of the N-terminal domain in the presence and absence of chymotrypsin, respectively. Digestion was initiated by addition of the protease and the sample incubated for 1.5 h prior to SDS−PAGE analysis. Amino-terminal sequencing revealed that the larger digest fragments start at the sequence partway through or after the S1 domain, as labeled.

shape. This shape was the most frequently recurring and physically plausible among more than 30 independent restoration calculations, and the constraint of D_2 particle symmetry reduces considerably the number of possible shape models that are in accord with the experimental data. A striking feature of the restored shapes is the appearance of four small and clearly identifiable lobes in the periphery of the main protein core. However, from the shape reconstructions, it is not possible to tell whether these lobes are separated by deep clefts or smaller crevices from the core domain. In any case, our protease digestion experiments, described below, favor an easily accessible domain which prompted our choice of the model shown in Figure 5C. The radius of the model is approximately 70 Å, which is in excellent agreement with the radius calculated from the sedimentation velocity data (66 Å).

Protease digestion patterns (Figure 5D) indicate that the S1 RNA-binding subdomains are readily liberated, either in part or completely, while the remainder of the protein is comparatively more resistant to protease. This suggests that the S1 domains are exposed and linked flexibly to the core of the protein. For this reason, we propose that they must lie exposed on the periphery of the tetrameric assembly where they may contact the RNA. This is represented pictorially in Figure 5C, where a model S1 domain (in green, molecule b) is compared with the N-terminal domain tetramer (colored with a blue-green gradient, corresponding to short to long radius from the center of the model). It can be seen that the

peripheral portions of the tetramer, in green, could in principle accommodate the S1 domains. This hypothesis awaits further testing.

DISCUSSION

Using a number of techniques, we have shown that the 62 kDa N-terminal domain of *E. coli* RNase E forms a homotetramer both in solution and, apparently, in the crystalline state, and confirm that the tetramer is catalytically active.

Nanoflow-electrospray mass spectrometry has been used to characterize the N-terminal domain under conditions that maintained noncovalent intermolecular interactions. This permitted the direct observation of the tetrameric species of the N-terminal domain and its stable complex with an RNA substrate analogue stabilized with 2′-*O*-methyl groups (Figure 4). These data are consistent with the binding of one oligonucleotide per protomer of the tetrameric assembly.

A combination of X-ray solution scattering data and D_2 symmetry constraints has allowed us to develop a physically plausible shape reconstruction for the N-terminal domain tetramer in solution (Figure 5C). There is a central volume (Figure 5C, blue residues) in which regions of the different protomers cannot be distinguished. The subdomains apparent at the periphery in the reconstruction (Figure 5C, green residues) are more likely to each derive from the protomers individually and not from any interface between them. It is these peripheral subdomains that would be expected to be the most susceptible to proteolytic attack. Proteolysis of the N-terminal domain was shown to produce N-terminally truncated products with partial or complete removal of the S1 domains (Figure 5D), and it is therefore likely that the S1 domains are located here in these peripheral subdomains. Figure 5C (parts a and b) illustrates that a typical S1 domain, derived from a crystal structure, could be physically accommodated in each peripheral subdomain of the N-terminal domain reconstruction.

Previous studies have suggested that RNase E N-terminal domain contains a binding pocket that can discriminate between triphosphate and monophosphate 5′ termini. The phosphate-binding pocket has been proposed to interact with the 5′ terminus of the substrate some distance from the active site (*53*). Recent studies on the cleavage of oligonucleotide substrates (in fact, the same one as studied here) have revealed that the distance from the active site to the 5′-phosphate binding site is no greater than the length spanned by five nucleotides (*42*, *43*, *52*). Figure 5C shows the S1-like domain with a bound 6-mer oligonucleotide (24 Å long), of which two nucleotides lie within the binding site and a further four protrude in a 3′ direction [Figure 5C (part c); PDB entry 2A8V]. As discussed earlier, an S1 domain of this type could be accommodated in the peripheral subdomain regions of the N-terminal domain reconstruction. The 3′ end of the protruding oligonucleotide illustrates an approximate position at which cleavage may occur in an oligonucleotide bound to RNase E. Unlike the oligonucleotide in the example shown, the 5′ terminus of an RNase E substrate might project from the S1 binding site, where it would be recognized by C-terminal elements of the S1 domain. The figure shows that it would be unlikely, physically, for the small RNA substrate to be cleaved by an adjacent protomer. However, larger substrates might bridge adjacent protomers.

Various homologues of the *E. coli* RNase E N-terminal domain have been identified, including those of the RNase G family. The finding that the *E. coli* RNase E N-terminal domain adopts a tetrameric arrangement as its active form suggests that the same quaternary structure might also occur in other members of the RNase E/RNase G family. Rott *et al*. (*54*) have reported that the RNase E from *Synechocystis* elutes from an S200 column with an apparent molecular mass of 250−300 kDa, which is consistent with our proposal that RNase E homologues are also likely to form tetramers.

RNase E is the central component of the degradosome, and its C-terminal domain is the scaffold for the assembly of the remaining components of the complex. We suggest that these other components will be organized around a tetrameric RNase E core. Reports vary with regard to the compositional stoichiometry of the degradosome (*1*, *24*), and it is possible that the assembly around the RNase E tetramer is heterogeneous so that each C-terminal domain may not have the same number of components bound.

The tetrameric organization of the RNase E N-terminal domain suggests that longer RNA substrates could be bound simultaneously to neighboring protomers, while smaller RNA species, like the 10-mer studied here, might be bound independently. Recently, Feng *et al*. (*52*) provided the first evidence of RNase E acting via a scanning mechanism. Previously, Coburn and Mackie (*1*) had proposed that a hypothetical dimeric organization of RNase E could permit the interplay of binding and catalytic sites to affect a processive mode of substrate cutting. In their model, the binding of the 5′ end of RNA (namely, the 5′-monophosphate) could occur at one protomer, with cleavage of the tethered substrate proceeding at the neighboring protomer. This tethered, processive model was supported by evidence that segments of RNase E N- and C-terminal domains are sites of self-interaction (*28*). It is now clear, however, that the tetrameric assembly of the N-terminal domain would also facilitate this processive activity.

ACKNOWLEDGMENT

We thank Charles Hill (University of Cambridge) and Chris Adams (University of Leeds) for careful synthesis of the RNA and Len Packman and Mike Weldon (PNAC Facility, University of Cambridge) for protein analyses and advice. Vladimir Kaberdin kindly provided expression plasmids and helpful suggestions. We thank the staff of the Advanced Light Source, particularly Thomas Earnest and Gerry McDermott, for use of facilities and Keith Henderson for advice during crystallographic data collection and processing. We thank Kenji Mizuguchi and Jukka Aurikko for invaluable discussions and Carmen San Martin for the pdb_depth program.

REFERENCES

1. Coburn, G. A., and Mackie, G. A. (1999) *Prog. Nucleic Acid Res. Mol. Biol. 62*, 55−108.
2. Regnier, P., and Arraiano, C. M. (2000) *BioEssays 22*, 235−244.
3. Apirion, D., and Lassar, A. B. (1978) *J. Biol. Chem. 253*, 1738−1742.
4. Bessarab, D. A., Kaberdin, V. R., Wei, C.-L., Liou, G.-G., and Lin-Chao, S. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 3157−3161.
5. Li, Z., Pandit, S., and Deutscher, M. P. (1999) *EMBO J. 18*, 2878−2885.

6. Li, Z., and Deutscher, M. P. (2002) *RNA 8*, 97−109.
7. Lundberg, U., and Altman, S. (1995) *RNA 1*, 327−334.
8. Lin-Chao, S., Wei, C. L., and Lin, Y. T. (1999) *Proc. Natl. Acad. Sci. U.S.A. 96*, 12406−12411.
9. Ehretsmann, C. P., Carpousis, A. J., and Krisch, H. M. (1992) *Genes Dev. 6*, 149−159.
10. Mackie, G. A. (1992) *J. Biol. Chem. 267*, 1054−1061.
11. Lin-Chao, S., Wong, T.-T., McDowall, K. J., and Cohen, S. N. (1994) *J. Biol. Chem. 269*, 10797−10803.
12. McDowall, K. J., Lin-Chao, S., and Cohen, S. N. (1994) *J. Biol. Chem. 269*, 10790−10796.
13. Cohen, S. N. (1995) *Cell 80*, 829−832.
14. Tomcsányi, T., and Apirion, D. (1985) *J. Mol. Biol. 185*, 713−720.
15. Lin-Chao, S., and Cohen, S. N. (1991) *Cell 65*, 1233−1242.
16. Kaberdin, V. R., Walsh, A. P., Jakobsen, T., McDowall, K. J., and von Gabain, A. (2000) *J. Mol. Biol. 301*, 257−264.
17. Mackie, G. A. (1998) *Nature 395*, 720−723.
18. Mackie, G. A. (2000) *J. Biol. Chem. 275*, 25069−25072.
19. McDowall, K. J., and Cohen, S. N. (1996) *J. Mol. Biol. 255*, 349−355.
20. Lopez, P. J., Marchand, I., Joyce, S. A., and Dreyfus, M. (1999) *Mol. Microbiol. 33*, 188−199.
21. Ow, M. C., Liu, Q., and Kushner, S. R. (2000) *Mol. Microbiol. 38*, 854−866.
22. Leroy, A., Vanzo, N. F., Sousa, S., Dreyfus, M., and Carpousis, A. J. (2002) *Mol. Microbiol. 45*, 1231−1243.
23. Carpousis, A. J., Van Houwe, G., Ehretsmann, G., and Krisch, H. M. (1994) *Cell 76*, 889−900.
24. Miczak, A., Kaberdin, V. R., Wei, C.-L., and Lin-Chao, S. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 3865−3869.
25. Py, B., Higgins, C. F., Krisch, H. M., and Carpousis, A. J. (1996) *Nature 381*, 169−172.
26. Carpousis, A. J. (2002) *Biochem. Soc. Trans. 30*, 150−155.
27. Xu, F., and Cohen, S. N. (1995) *Nature 374* 180−183.
28. Vanzo, N. F., Li, T. S., By, P., Blum, E., Higgins, C. F., Raynal, L. C., Krisch, H. M., and Carpousis, A. J. (1998) *Genes Dev. 12*, 2770−2781.
29. Bateman, A., Birney, E., Cerruti, L., Durbin, R., Etwiller, L., Eddy, S. R., Griffiths-Jones, S., Howe, K. L., Marshall, M., and Sonnhammer, E. L. (2002) *Nucleic Acids Res. 30*, 276−280.
30. Lee, K., and Cohen, S. N. (2003) *Mol. Microbiol. 48*, 349−360.
31. Kaberdin, V. R., Miczak, A., Jakobsen, J. S., Lin-Chao, S., McDowall, K. J., and von Gabain, A. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 11637−11642.
32. Aravind, L., and Koonin, E. V. (2001) *Methods Enzymol. 341*, 3−28.
33. Baginsky, S., Sheiman-Kotler, A., Liveanu, V., Yehudai-Resheff, S., Bellaou, M., Settlage, R. E., Shabanowitz, J., Hunt, D. F., Schuster, G., and Gruissem, W. (2001) *RNA 7*, 1464−1475.
34. Kanai, A., Oida, H., Matsuura, N., and Doi, H. (2003) *Biochem. J. 372*, 253−261.
35. Wachi, M., Umitsuki, G., and Nagai, K. (1997) *Mol. Gen. Genet. 253*, 515−519.
36. Lee, K., Bernstein, J. A., and Cohen, S. N. (2002) *Mol. Microbiol. 43*, 1445−1456.
37. Ow, M. C., Perwez, T., and Kushner, S. R. (2003) *Mol. Microbiol. 49*, 607−622.
38. Wachi, M., and Nagai, K. (2001) *Methods Enzymol. 342*, 55−63.
39. Condon, C., and Putzer, H. (2002) *Nucleic Acids Res. 30*, 5339−5346.
40. Bycroft, M., Hubbard, T. J., Proctor, M., Freund, S. M., and Murzin, A. G. (1997) *Cell 88*, 235−242.
41. Diwa, A. A., Jiang, X., Schapira, M., and Belasco, J. G. (2002) *Mol. Microbiol. 46*, 959−969.
42. Tock, M. R., Walsh, A. P., Carroll, G., and McDowall, K. J. (2000) *J. Biol. Chem. 275*, 8726−8732.
43. McDowall, K. J., Kaberdin, V. R., Wu, S. W., Cohen, S. N., and Lin-Chao, S. (1995) *Nature 374*, 287−290.
44. Redko, Y., Tock, M. R., Adams, C. J., Kaberdin, V. R., Grasby, J. A., and McDowall, K. J. (2003) *J. Biol. Chem*. (in press).
45. Loo, J. A. (2000) *Int. J. Mass Spectrom. 200*, 175−186.
46. Schuck, P. (2000) *Biophys. J. 78*, 1606−1619.
47. Laue, T., Shaw, B. D., Ridgeway, T. M., and Pelletier, S. L. (1992) in *Analytical Ultracentrifugation in Biochemistry and Polymer Science* (Harding, S. E., Rowe, A. J., and Horton, J. C., Eds.) pp 90−125, The Royal Society of Chemistry, Cambridge, U.K.
48. Semenyuk, A. V., and Svergun, D. I. (1991) *J. Appl. Crystallogr. 24*, 537−540.
49. Svergun, D. I., Petoukhov, M. V., and Koch, M. H. J. (2001) *Biophys. J. 80*, 2946−2953.
50. Collaborative Computational Project No. 4 (1994) *Acta Crystallogr. D50*, 760−763.
51. Sobott, F., Hernandez, H., McCammon, M. G., Tito, M. A., and Robinson, C. V. (2002) *Anal. Chem. 74*, 1402−1407.
52. Feng, Y., Vickers, T. A., and Cohen, S. N. (2002) *Proc. Natl. Acad. Sci. U.S.A. 99*, 14746−14751.
53. Spickler, C., Stronge, V., and Mackie, G. A. (2001) *J. Bacteriol. 183*, 1106−1109.
54. Rott, R., Zipor, G., Portnoy, V., Liveanu, V., and Schuster, G. (2003) *J. Biol. Chem. 278*, 15771−15777.
55. Otwinowski, Z., and Minor, W. (1997 *Methods Enzymol. 276*, 307−326.
56. Higgins, D., Thompson, J., and Gibson, T. J. (1994) *Nucleic Acids Res. 22*, 4673−4680.
57. Gouet, P., Courcelle, E., Stuart, D. I., and Metoz, F. (1999) *Bioinformatics 15*, 305−308.
58. Kraulis, P. J. (1999) *J. Appl. Crystallogr. 24*, 946−950.

BI0351099