Structure of *Escherichia coli* RNase E catalytic domain and implications for RNA turnover

Anastasia J. Callaghan[1], Maria Jose Marcaida[1], Jonathan A. Stead[2], Kenneth J. McDowall[2], William G. Scott[3] & Ben F. Luisi[1]

The coordinated regulation of gene expression is required for homeostasis, growth and development in all organisms. Such coordination may be partly achieved at the level of messenger RNA stability[1], in which the targeted destruction of subsets of transcripts generates the potential for cross-regulating metabolic pathways. In *Escherichia coli*, the balance and composition of the transcript population is affected by RNase E, an essential endo-ribonuclease that not only turns over RNA but also processes certain key RNA precursors[2–10]. RNase E cleaves RNA internally, but its catalytic power is determined by the 5′ terminus of the substrate, even if this lies at a distance from the cutting site[11–14]. Here we report crystal structures of the catalytic domain of RNase E as trapped allosteric intermediates with RNA substrates. Four subunits of RNase E catalytic domain associate into an interwoven quaternary structure, explaining why the subunit organization is required for catalytic activity. The subdomain encompassing the active site is structurally congruent to a deoxy-ribonuclease, making an unexpected link in the evolutionary history of RNA and DNA nucleases. The structure explains how the recognition of the 5′ terminus of the substrate may trigger catalysis and also sheds light on the question of how RNase E might selectively process, rather than destroy, specific RNA precursors.

E. coli RNase E is one of the largest members of a highly conserved RNase family[15]. A catalytic core is contained within the amino-terminal half[16–18] of RNase E, which in length and sequence closely resembles its paralogue, RNase G[9,19] (Fig. 1a and Supplementary Fig. S1). The carboxy-terminal half is largely unstructured and poorly conserved[20], but in *E. coli* and related proteobacteria this portion organizes and coordinates the activities of the multienzyme RNA degradosome complex[21].

The crystal structure that we have determined at 2.9 Å is the first for a member of the RNase E/G family, and it reveals that the catalytic domain of RNase E forms a homotetramer with a molecular mass of roughly 260 kDa (Fig. 1). The isolated RNase E protomer is composed of two globular portions, which we refer to as the 'small' and 'large' domains (Fig. 2). The tetramer is organized as a dimer-of-dimers, and the arrangement of the domains within each dimer resembles the blades and handles of an open pair of scissors (Fig. 1b). At the scissor junction point, intra-domain linkers cross and coordinate a zinc ion through a pair of cysteine residues in the conserved CPxCxGxG motif[22] (Fig. 1f). The shared-metal site helps to organize the dimer–dimer interface, thus explaining why mutation of the coordinating cysteine residues disrupts the quaternary structure[22].

The 'large' domain (residues 1–400) can be divided into sub-domains corresponding to established folds (Fig. 2b and Supplementary Fig. S2). One is structurally related to the RNase H

endoribonuclease family, but in RNase E this subdomain seems to have a structural rather than a catalytic function because none of the active-site residues is present. Embedded in the RNase H fold is an S1 domain, which is a widely occurring RNA-binding structural motif (residues 36–118; Fig. 2), and a second fold that participates in the recognition of the 5′ terminus of the bound RNA (residues 119–215; Fig. 2). The remaining component of the large domain is similar to the repetitive structural element within the endo-deoxy-ribonuclease, DNase I (residues 280–400; red in Fig. 2). Their structural congruence has mechanistic implications, as discussed below.

The three RNAs that crystallized together with RNase E contain the targeted sequence 5′-ACAGUAUUUG-3′ found at the 5′ end of the antisense regulator of ColE1-type plasmid replication, RNA I[23] (Supplementary Fig. S3). The RNAs have 2′-*O*-methyl groups, which permit binding but prevent degradation in solution studies. The RNA molecule is accommodated in a channel and is contacted on one side by the S1 domain, 5′ sensor, DNase I-like and RNase H-like subdomains of one protomer (protomer B in Fig. 3a), and on the other side by residues of the DNase I-like subdomain from the partner protomer (protomer A in Fig. 3a).

Whereas the 10-mer and 13-mer RNAs are bound entirely by one tetramer, the 15-mer RNA is long enough for it to be shared between two different tetramers in the crystal lattice (Fig. 4a), and this RNA extends from the 5′ sensing pocket of one tetramer into the binding channel and catalytic site of a symmetry-related tetramer. Two molecules of 15-mer RNA pass each other in antiparallel orientations and form a highly distorted duplex. We suggest that the 15-mer complex might mimic the association of RNase E with RNA substrates in which the sites of primary cleavage are separated from a single-stranded 5′ terminus by a structured segment. A model of such a hairpin is shown in Fig. 4b.

The group on the 5′ terminus of the RNA substrate can affect the catalytic rate constant of RNase E by orders of magnitude[12] but does not seem to have a marked effect on substrate binding[24,25]. A 5′-hydroxy group or triphosphate cap impedes activity, whereas the catalytic power is greatest for substrates with a 5′-terminal mono-phosphate. The RNAs that crystallize together with the catalytic domain of RNase E contain a 5′-terminal monophosphate, and this is accommodated in a pocket formed between the 5′ sensor and RNase H-like subdomains (Fig. 3). The terminal phosphate is engaged by a semicircular ring of hydrogen-bonding donors from the side chain and peptide amide of Thr 170 and the guanidino group of Arg 169. The interaction of Arg 169 is consolidated by a hydrogen bond to the peptide backbone of Gly 124 in the neighbouring strand. The terminal base forms a hydrophobic contact with the side chain of Val 128 (Fig. 3c). Replacement of Arg 169 or Val 128 decreases the

[1]Department of Biochemistry, University of Cambridge, 80 Tennis Court Road, Cambridge CB2 1GA, UK. [2]Astbury Centre for Structural Molecular Biology, University of Leeds, Leeds LS2 9JT, UK. [3]Department of Chemistry and Biochemistry and the Center for the Molecular Biology of RNA, Sinsheimer Laboratories, University of California at Santa Cruz, Santa Cruz, California 95064, USA.

enhancement of cleavage seen for substrate with 5′ monophosphate, and eliminating the hydrogen bond between the terminal phosphate and Thr 170 by replacement with Val inactivates the enzyme (J.A.S., B.F.L. and K.J.M., unpublished observation). Arg 373 from the RNase H-like subdomain bridges the phosphate at positions 2 and 5, to help dock the 5′ terminus and redirect the RNA towards the catalytic site. Replacement of Arg 373 by either Ala or Asp decreases catalytic activity tenfold, with only a modest effect on RNA-binding affinity (Supplementary Information).

We identified a sequestered magnesium ion, presumably hydrated, that is in proximity to the RNA phosphate backbone (Fig. 3b). This metal is coordinated by the carboxylates of Asp 303 and Asp 346, which are conserved throughout the RNase E/G family (Supplementary Information). The catalytic mechanism of RNase E probably involves the in-line nucleophilic attack of the scissile phosphate by an activated hydroxyl group generated from a water molecule that is coordinated to the magnesium. It can be envisaged that the carboxylates of Asp 303 and Asp 346 might act as general bases to activate the attacking water. The conservative replacement of either of these residues with the isosteric and polar Asn decreased catalytic

activity at least 25-fold while having little effect on RNA binding. The neighbouring Asn 305 supports the orientation of Asp 303 through hydrogen bonding, and the elimination of this interaction by the replacement of Asn 305 with either Asp or Leu caused a decrease in activity (Supplementary Information, Table S1). Asp 303, Asp 346 and Asn 305 are on the surface of the subdomain of RNase E that resembles DNase I, and the catalytic sites of these two different enzymes roughly coincide (Fig. 2c, d).

The first base after the cleavage site is partitioned into a hydrophobic pocket on the surface of the S1 domain (Fig. 3b). The pocket is formed on one side by Phe 67, with which the base's aromatic face makes an alternative stacking contact, and on the other side by the aliphatic portion of the Lys 112 side chain. The latter interaction orients the lysine so that its terminal amino group hydrogen bonds with the scissile phosphate and presumably supports the charge built up in the catalytic transition state. Phe 57, also from the S1 domain, helps to consolidate this interaction with the RNA by supporting Phe 67. We find that the individual replacement of any of these residues by Ala decreased activity at least 50-fold; moreover, this does not seem to be due to a marked decrease in affinity for the substrate
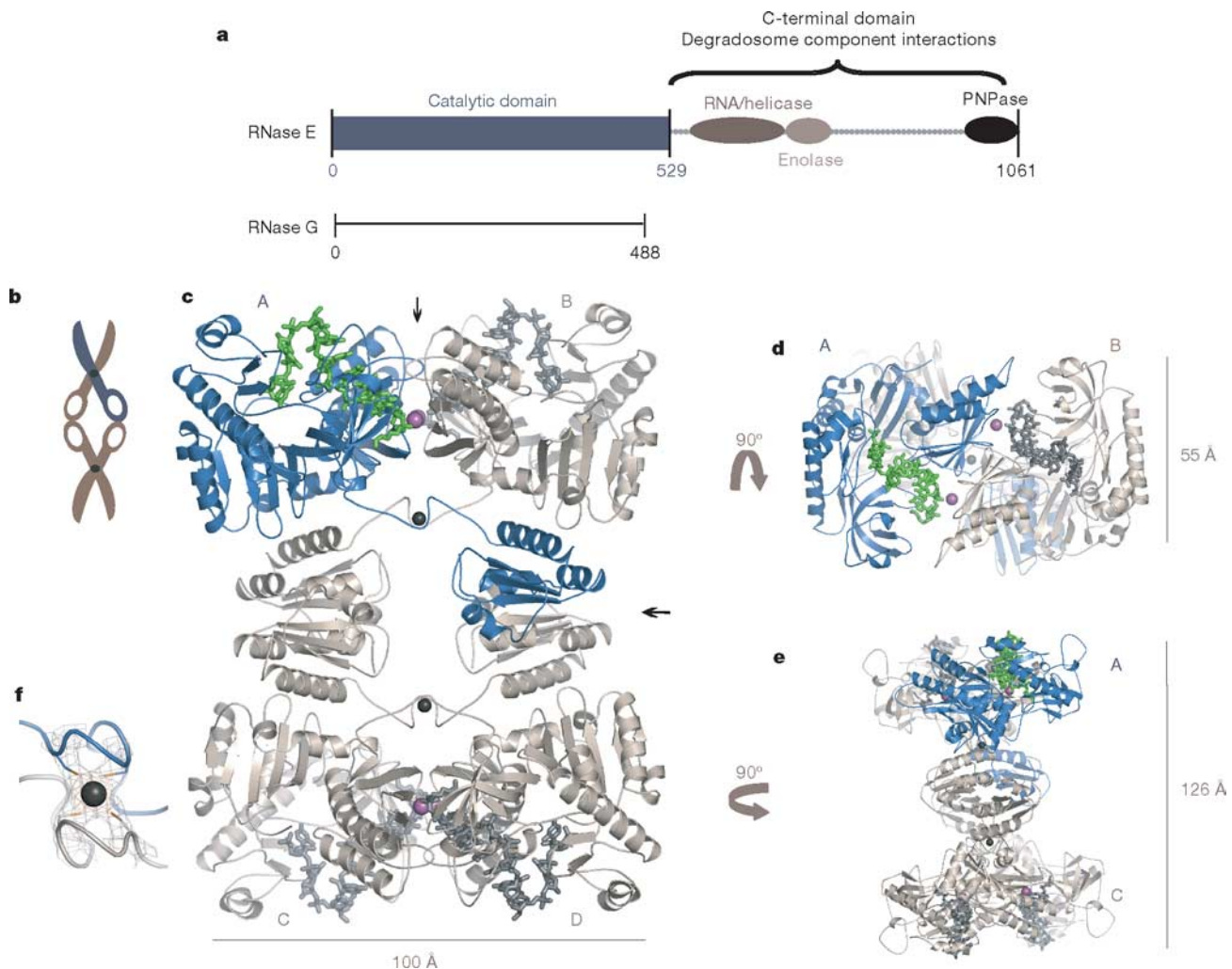

Figure 1 | The structure of RNase E catalytic domain. a, Schematic representation of the primary structure of RNase E and paralogue RNase G. The crystallized N-terminal catalytic domain is represented by the blue box. The degradosome-scaffolding C-terminal domain (grey line) has binding regions for RNA/helicase (brown), enolase (beige) and PNPase (black). **b,** Scissor representation of the RNase E tetramer. **c,** The homotetramer, showing protomer A (blue) with bound RNA (green), and the remaining three protomers B, C and D (beige) with bound RNA (grey). The principal dimers are the pairs A/B and C/D. The zinc and magnesium ions are shown as dark grey and magenta spheres, respectively. **d, e,** Homotetramer as viewed from above (**d**) and rotated through 90° (**e**). **f,** The zinc-link coordination site at the scissor junction, formed by two cysteine residues from protomer A (blue) and two from protomer B (beige). Electron density ($2mF_o - DF_c$ coefficients) is shown at 8σ (orange) and 3σ (grey).
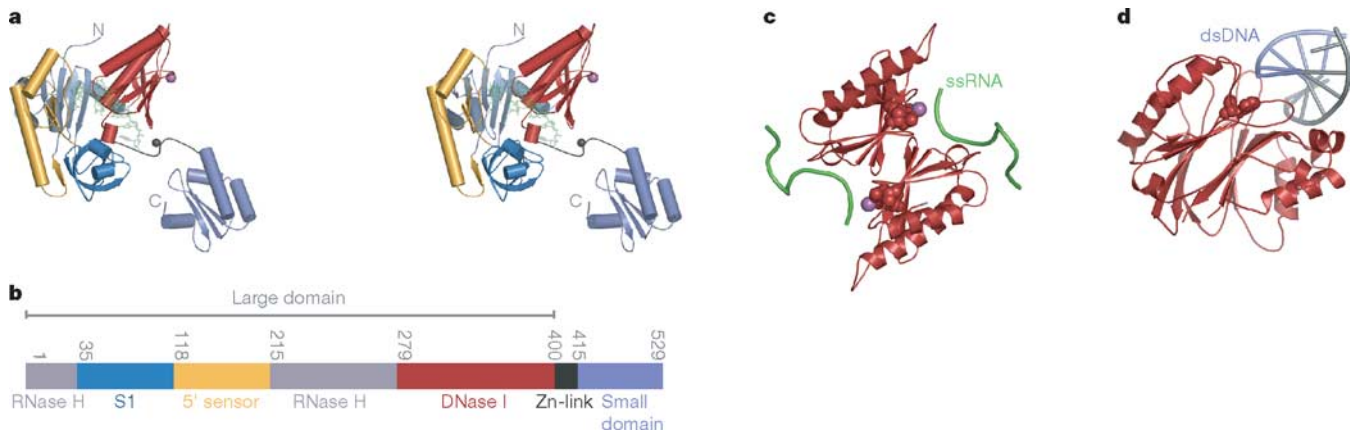
Figure 2 | RNase E isolated protomer–RNA complex. a, Stereoscopic view of the isolated protomer–RNA complex, showing the subdomains S1 (blue), RNase H (grey), 5′-sensor (gold), the small domain (violet) and DNase I (red). The single-stranded RNA (green) is semi-transparent for clarity. Zinc and magnesium ions are grey and magenta, respectively. **b,** Linear representation of the subdomain boundaries. **c, d,** Dimerization of the DNase I subdomains in the principal dimer (**c**) and the similar fold of DNase I structure (**d**). Shown are the active site residues (space-filling representation), single-stranded RNA (ssRNA, green) and the double-stranded DNA (dsRNA, grey).
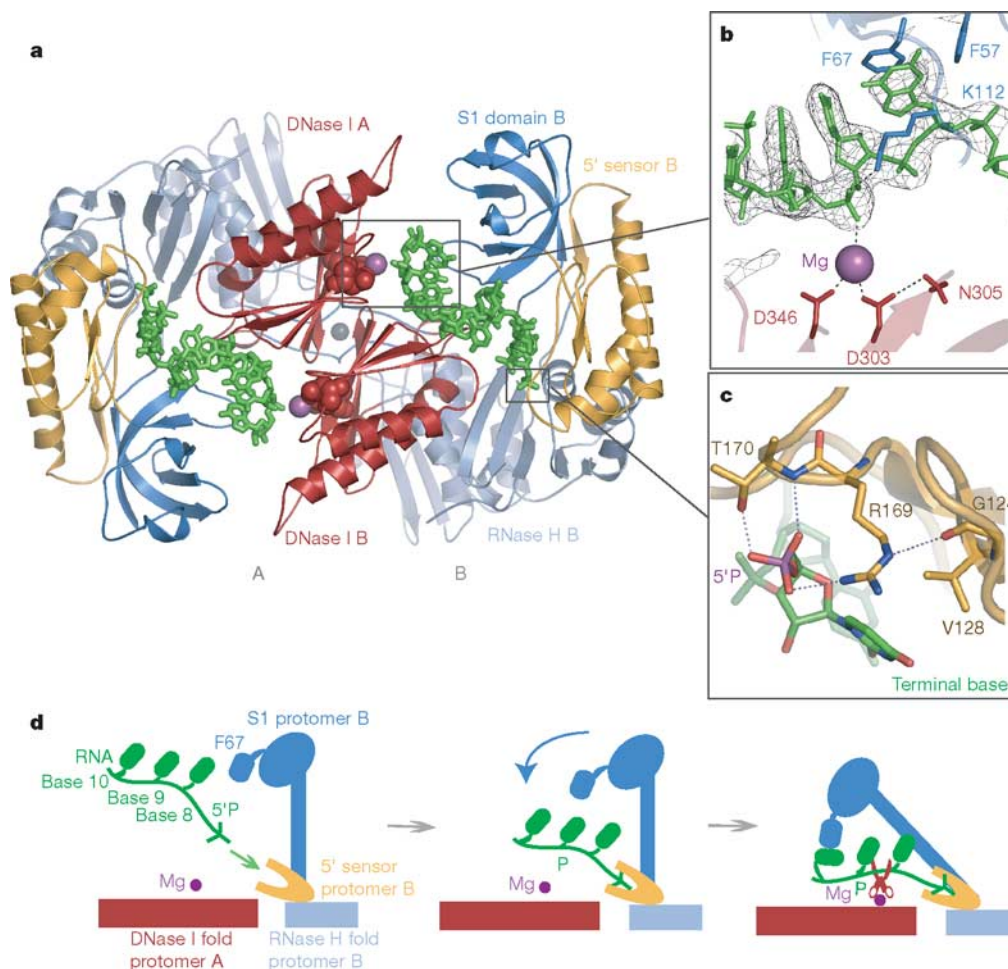

Figure 3 | The RNA-binding channel. a, The principal dimer viewed along the twofold symmetry axis. The RNA is contacted by the S1, 5′ sensor, RNase H-like and different parts of the DNase I-like subdomains of protomers A and B. **b,** The proposed catalytic site of RNase E. The magnesium ion (magenta) interacts with a phosphate of the RNA (green), and Asp 346 and Asp 303 of protomer A (red). The $2F_o - F_c$-style electron density map (black) is at 2.5σ. **c,** The 5′ sensing pocket. RNA (green) and 5′ sensor residues (gold) are shown. The terminal phosphate of the RNA (purple) is engaged by a semicircular ring of hydrogen-bond donors, Thr 170 and Arg 169. **d,** A 'mouse-trap' model for communication between the 5′ sensing pocket and the site of catalysis. On binding the 5′ phosphate, the 5′ sensor (gold) induces the S1 domain (blue) to clamp down on the RNA (green) near the scissile phosphate (green P). This orients the RNA for cleavage at the Mg ion (magenta) in the DNase I domain (red). The proposed movement of the S1 domain is exaggerated for illustrative purposes.
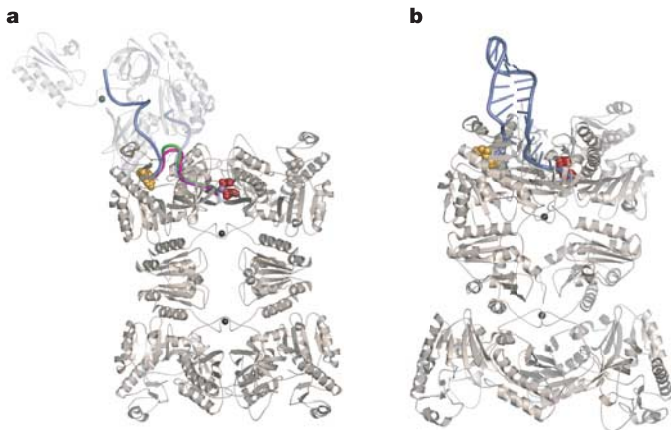
a   b

Figure 4 | **RNA recognition by RNase E catalytic domain. a**, An overlay of RNase E tetramer in complex with bound 10-mer (magenta), 13-mer (green) and 15-mer (violet) RNA. Catalytic residues (red) and 5′ sensor pocket (gold) in space-filling representation. The 15-mer RNA extends away from the principal dimer into a symmetry-related tetramer (semi-transparent beige) and, reciprocally, from the symmetry-related tetramer (semi-transparent violet) back into the principal dimer (beige). Only one active site with RNA bound is shown for simplicity. **b**, RNase E bound to an RNA hairpin model (violet) extrapolated from the RNase E/15-mer RNA complex crystal structure.

(Supplementary Table S1). Perhaps the sandwiching of the base between the Phe 67 and Lys 112 distorts the RNA to favour the transition state.

In all the complexes, the RNA follows the same stacking arrangement over the surface of the RNA-binding channel, and only one of the bases forms a hydrogen bond with the protein; Lys 106 contacts the exocyclic oxygen of the base at position −1 with regard to the scissile phosphate. This direct contact might explain the preference for G or U at this position[26]. Aside from that contact, there is no sequence recognition as such, so it seems that the preference of RNase E for A + U-rich substrates[27–29] arises mainly through the recognition of the RNA conformation.

Our crystal structure shows that the catalytic site and the 5′ sensing site are physically separated, yet they must somehow communicate so that the status of the 5′ end of the substrate is relayed to the active site. It seems most likely that the communication is mediated through allosteric change in the protein. In support of this proposal, we observe that the binding of RNA induces compaction of the protein using neutron solution scattering with contrast variation (A.J.C., J.G. Grossmann, M.J.M. and B.F.L., unpublished observation). On the basis of the crystal structure and the solution experiments, we propose a 'mouse-trap'-like mechanism that explains how 5′ sensing is linked to catalysis through an allosteric change in the protein (Fig. 3d). The engagement of the 5′ monophosphate of the RNA organizes the 5′ sensor pocket, which in turn interacts with the S1 domain and causes it to clamp down on the RNA downstream, thus orienting the phosphate backbone to favour the in-line attack on the scissile phosphate by the magnesium-coordinated, activated hydroxyl group. Thus, catalysis proceeds through classical induced-fit.

The question naturally arises as to why RNase E does not indiscriminately cleave all RNA with a 5′ monophosphate, including the structured substrates that it processes, such as ribosomal RNA[9] and RNase P precursors[7]. The answer may lie in the mouse-trap mechanism, which requires that the 5′ end is engaged in the sensing site and that the portion to be cleaved is simultaneously engaged at the active site: structured RNA segments may not be able to satisfy all of these requirements simultaneously.

RNase E/G proteins are important in global gene regulation and cellular homeostasis through their preferential degradation of transcripts and their processing of key RNA species. Their homologues in

plants are also implicated in the general control of gene expression. The crystal structure of *E. coli* RNase E opens the possibility of manipulating this control, either for therapeutic intervention or for the purpose of examining the complex interaction of different components in dynamic cellular systems.

METHODS

Structure determination. The catalytic domain of RNase E and a selenomethionine derivative (residues 1–529 with an N-terminal hexahistidyl affinity tag) were prepared by the method described previously[24,30]. For the crystallization trials, the purified protein was mixed in the ratio 1:1.2 with HPLC-purified synthetic RNA containing a 5′-monophosphate and 2′-O-methyl groups and concentrated to 10–15 mg ml^{-1}. Crystals of the RNase E catalytic-domain–RNA complex appeared after 2–4 weeks in 5–20% w/v poly(ethylene glycol) 8000, 0.1 M Tris-HCl pH 7.5–8.0, and 10–50 mM magnesium formate at 20 °C, and in all cases the complex had a hexagonal prismatic habit. Crystals were cryo-protected by serial transfer into crystallization buffer containing 5–30% v/v ethylene glycol, and then frozen by plunging into liquid nitrogen. Diffraction data were collected at 100 K at stations ID 14-4, 23-1 and 29-1 (European Synchrotron Radiation Facility, Grenoble); station 14.2 (Synchrotron Radiation Source, Daresbury); and Howard Hughes Medical Institute station 8.8.2 (Advanced Light Source, Berkeley). The structure was solved by using multiple-wavelength anomalous scattering and multiple isomorphous replacement, as detailed in Supplementary Information. Structural figures were made using PYMOL (DeLano Scientific) and NUCCYL (http://www.biosci.ki.se/groups/ljo/software/nuccyl.html).

Site-directed mutagenesis and RNase E assays. The method of site-directed mutagenesis of RNase E catalytic domain and expression of the mutant polypeptides in *E. coli* were described previously[22]. The sequences of the pairs of primers used to introduce the specific mutations are provided in Supplementary Table S3. Each recombinant polypeptide was purified with Ni^{2+}-nitrilotriacetate spin columns. The eluate was dialysed against 20 mM Tris-HCl pH 7.9, 0.5 M NaCl, 10 mM dithiothreitol, 10 mM MgCl$_2$, 0.5 mM EDTA, 5% (v/v) glycerol, divided into aliquots and stored frozen at −70 °C. Protein concentrations were determined with a Bradford protein assay (Bio-Rad), and confirmed by SDS–PAGE.

The cleavage of RNA was assayed with a continuous system similar to that described previously[25]. The 5′ monophosphorylated substrate used here had Dabcyl-dT (D-dT) at position 5 and fluorescein (Fl) at the 3′ end. The sequence was 5′-GGGA(D-dT)CAGUAUUU-Fl, which is based on the RNase E cleavage site at the 5′ end of RNA I[23]. The reaction conditions were those optimized previously[24]. The concentrations of substrate and enzyme were 500 nM and 8 nM, respectively. Mutants with greatly decreased activity were also assayed at a concentration of 40 nM. The increase in fluorescence was measured with a Perkin Elmer LS50B fluorimeter at 37 °C. The excitation and emission wavelengths were 488 and 523 nm, respectively. Each mutant was assayed in triplicate. The system was calibrated with substrates cleaved to completion by RNase E and RNase A.

For the electrophoretic mobility-shift assays, a 5′ monophosphorylated substrate with sequence 5′-ACAGUAUUUG-3′, with protective 2′-O-methyl groups at all positions and fluorescein at the 3′ end, was used[22]. The optimized cleavage conditions[24] were used to assay RNA binding. Binding was assayed by determining the amount of unbound substrate remaining in each reaction relative to a 5′-fluorescein-labelled pentadeoxyribonucleotide (5′-GACTA-3′) that served as an internal control.

1. Grunberg-Manago, M. Messenger RNA stability and its role in control of gene expression in bacteria and phages. *Annu. Rev. Genet.* **33**, 193–227 (1999).
2. Kuwano, M. *et al.* Gene affecting longevity of messenger RNA: a mutant of *Escherichia coli* with altered mRNA stability. *Mol. Gen. Genet.* **154**, 279–285 (1977).
3. Ono, M. & Kuwano, M. A conditional lethal mutation in an *Escherichia coli* strain with a longer chemical half life of mRNA. *J. Mol. Biol.* **129**, 343–357 (1979).
4. Bernstein, J. A., Lin, P.-H., Cohen, S. N. & Lin-Chao, S. Global analysis of *Escherichia coli* RNA degradosome function using DNA microarrays. *Proc. Natl Acad. Sci. USA* **101**, 2758–2763 (2004).
5. Mudd, E. A. & Higgins, C. F. *Escherichia coli* endoribonuclease RNase E: autoregulation of expression and site-specific cleavage of mRNA. *Mol. Microbiol.* **9**, 557–568 (1993).
6. Apirion, D. & Lassar, A. B. A conditional lethal mutant of *Escherichia coli* which affects the processing of ribosomal RNA. *J. Biol. Chem.* **253**, 1738–1742 (1978).

7. Lundberg, U. & Altman, S. Processing of the precursor to the catalytic RNA subunit of RNase P from *Escherichia coli*. *RNA* **1**, 327–334 (1995).
8. Ow, M. C. & Kushner, S. R. Initiation of tRNA maturation by RNase E is essential for cell viability in *Escherichia coli*. *Genes Dev.* **16**, 1102–1115 (2002).
9. Li, Z., Pandit, S. & Deutscher, M. P. RNase G (CafA protein) and RNase E are both required for the 5′ maturation of 16S ribosomal RNA. *EMBO J.* **18**, 2878–2885 (1999).
10. Kim, K.-S. & Lee, Y. Regulation of 6S RNA biogenesis by switching utilization of both sigma factors and endoribonucleases. *Nucleic Acids Res.* **32**, 6057–6068 (2004).
11. Bouvet, P. & Belasco, J. G. Control of RNase E-mediated RNA degradation by 5′-terminal base pairing in *E. coli*. *Nature* **360**, 488–491 (1992).
12. Mackie, G. A. Ribonuclease E is a 5′-end-dependent endonuclease. *Nature* **395**, 720–723 (1998).
13. Mackie, G. A. Stabilization of circular rpsT mRNA demonstrates the 5′ end dependence of RNase E action *in vivo*. *J. Biol. Chem.* **275**, 25069–25072 (2000).
14. Coburn, G. A. & Mackie, G. A. Degradation of mRNA in *Escherichia coli*: an old problem with some new twists. *Prog. Nucleic Acid Res.* **62**, 55–108 (1999).
15. Avarind, L. & Koonin, E. V. A natural classification of ribonucleases. *Methods Enzymol.* **341**, 3–28 (2001).
16. McDowall, K. J. & Cohen, S. N. The N-terminal domain of the *rne* gene product has RNase E activity and is non-overlapping with the arginine-rich RNA binding site. *J. Mol. Biol.* **255**, 349–355 (1996).
17. Kido, M. *et al.* RNase E polypeptides lacking a carboxyl-terminal half suppress a *mukB* mutation in *Escherichia coli*. *J. Bacteriol.* **178**, 3917–3925 (1996).
18. McDowall, K. J., Hernandez, R. G., Lin-Chao, S. & Cohen, S. N. The *ams-1* and *rne-3071* temperature-sensitive mutations in the *ams* gene are in close proximity to each other and cause substitutions within a domain that resembles a product of the *Escherichia coli mre* locus. *J. Bacteriol.* **175**, 4245–4249 (1993).
19. Wachi, M., Umitsuki, G., Shimizu, M., Takada, A. & Nagai, K. *Escherichia coli cafA* gene encodes a novel RNase, designated as RNase G, involved in processing of the 5′ end of 16S rRNA. *Biochem. Biophys. Res. Commun.* **259**, 483–488 (1999).
20. Callaghan, A. J. *et al.* Studies of the RNA degradosome-organising domain of the *Escherichia coli* RNase E. *J. Mol. Biol.* **340**, 965–979 (2004).
21. Carpousis, A. J. The *Escherichia coli* RNA degradosome: structure, function and relationship in other ribonucleolytic multienzyme complexes. *Biochem. Soc. Trans.* **30**, 150–155 (2002).
22. Callaghan, A. J. *et al.* The 'Zn-link': a metal-sharing interface that organizes the quaternary structure and catalytic site of the endoribonuclease, RNase E. *Biochemistry* **44**, 4667–4675 (2005).
23. Lin-Chao, S. & Cohen, S. N. The rate of processing and degradation of antisense RNA I regulates the replication of ColE1-type plasmids *in vivo*. *Cell* **65**, 1233–1242 (1991).
24. Redko, Y. *et al.* Determination of the catalytic parameters of the N-terminal half of *E. coli* ribonuclease E and the identification of critical functional groups in RNA substrates. *J. Biol. Chem.* **278**, 44001–44008 (2003).
25. Jiang, X. & Belasco, J. G. Catalytic activation of multimeric RNase E and RNase G by 5′-monophosphorylated RNA. *Proc. Natl Acad. Sci. USA* **101**, 9211–9216 (2004).
26. Kaberdin, V. R. Probing the substrate specificity of *Escherichia coli* RNase E using a novel oligonucleotide-based assay. *Nucleic Acids Res.* **31**, 4710–4716 (2003).
27. Cohen, S. N. & McDowall, K. J. RNase E: still a wonderfully mysterious enzyme. *Mol. Microbiol.* **23**, 1099–1106 (1997).
28. McDowall, K. J., Lin-Chao, S. & Cohen, S. N. A + U content rather than a particular nucleotide order determines the specificity of RNase E cleavage. *J. Biol. Chem.* **269**, 10790–10796 (1994).
29. Lin-Chao, S., Wong, T. T., McDowall, K. J. & Cohen, S. N. Effects of nucleotide sequence on the specificity of *rne*-dependent and RNase E-mediated cleavages of RNA I encoded by the pBR322 plasmid. *J. Biol. Chem.* **269**, 10797–10803 (1994).
30. Callaghan, A. J. *et al.* Quaternary structure and catalytic activity of the *Escherichia coli* ribonuclease E amino-terminal catalytic domain. *Biochemistry* **42**, 13848–13855 (2003).