JOURNAL OF BACTERIOLOGY, Sept. 1991, p. 5589–5591
0021-9193/91/175589-03$02.00/0

Vol. 173, No. 17

Escherichia coli orfE (Upstream of *pyrE*) Encodes RNase PH

KAREN A. OST AND MURRAY P. DEUTSCHER*

Department of Biochemistry, University of Connecticut Health Center, Farmington, Connecticut 06030

Received 28 May 1991/Accepted 25 June 1991

RNase PH from extracts of *Escherichia coli* was purified to homogeneity and subjected to NH_2-terminal sequencing. Comparison of this sequence with all open reading frames in the GenBank data base revealed at least 95% identity to an unidentified open reading frame (*orfE*) upstream of *pyrE* at 81.7 min on the *E. coli* chromosome. Clones of *orfE* overexpress RNase PH activity, verifying that *orfE* encodes this ribonuclease. We suggest that *orfE* be renamed *rph*.

RNase PH is a phosphorolytic exoribonuclease that removes nucleotide residues following the -CCA terminus of tRNA (3, 5). Action on tRNA precursors is favored over that on mature tRNA by about 100-fold (5). In addition to its degradative activity, RNase PH also carries out a synthetic reaction in which nucleoside diphosphates are substrates for the addition of nucleotides onto RNA molecules. Chain lengths of up to 40 additional nucleotides have been observed (13). In *Escherichia coli* cells lacking RNases I, D, T, and BN and polynucleotide phosphorylase (5), a mutation in RNase PH leads to temperature-sensitive growth. Temperature-resistant revertants and transductants regain RNase PH activity, suggesting that RNase PH is essential for growth in cells lacking other exoribonucleases (14).

In order to clone the gene encoding RNase PH, the enzyme was purified to homogeneity and its NH_2-terminal sequence was determined. We show here that this sequence has at least 95% identity with the product of the *orfE* gene at 81.7 min on the *E. coli* chromosome (16, 17). We also show that clones containing *orfE* overexpress RNase PH activity, supporting the conclusion that *orfE* encodes RNase PH.

Purification of RNase PH. RNase PH was purified from an S100 extract of the *E. coli* strain 18-11 II^+ PNP^- (5) through a number of chromatographic steps, including Affi-Gel Blue, Q-Sepharose, S-Sepharose, Ultrogel HA, and Poly(U)-Sepharose. In fractions from the Poly(U)-Sepharose column, one protein, a 33-kDa band on sodium dodecyl sulfate (SDS)–12.5% polyacrylamide gel electrophoresis (PAGE), coeluted with RNase PH activity. This band, which was well separated from nearby bands, was blotted onto Immobilon-P for N-terminal sequencing by using the Genie electrophoretic blotter apparatus with 25 mM Tris-glycine buffer (pH 8.3)–20% methanol. Transfer was at 30 V for 90 min. The blot was washed and stained by the procedure of LeGendre and Matsudaira (10).

N-terminal sequence analysis. Gas-phase sequencing was used to determine 21 amino-terminal residues of RNase PH. These were compared with the translation products of all the nucleotide sequences contained in the GenBank data base after they were translated in all six reading frames by the program TFastA (6). A match of at least 20 of 21 amino acids was found (Fig. 1) to an unidentified open reading frame, *orfE*, located at 81.7 min on the *E. coli* chromosome (15–17). One residue of the amino acid sequence was difficult to identify because of its low recovery in the sequencing procedure, although arginine was its likely identity. Nevertheless, the high degree of similarity between the RNase PH protein sequence and that of the *orfE* translation product strongly suggested that RNase PH is encoded by the *orfE* gene.

RNase PH activity of *orfE* clones. From the published restriction map of the *orfE-pyrE* region (9, 16, 17), it was possible to identify a lambda phage from the *E. coli* genomic library (8) which would contain this portion of the chromosome. This clone, no. 572, was digested with *Bam*HI, and the 6.1-kb fragment containing the *orfE* gene was ligated into the *Bam*HI site of the low-copy-number plasmid pHC79 (7). One clone (pORFE-1), in which *orfE* and *tet* transcriptions were oriented in the same direction, was used for subsequent analysis. A 3.1-kb *Bam*HI-*Eco*RI fragment containing *orfE* was also subcloned into the *Bam*HI-*Eco*RI site of the high-copy-number vector pUC18 (19); this plasmid is referred to as pORFE-2. Restriction analysis of the insert in pORFE-2 showed that it corresponded exactly to the published restriction map of *orfE* (9, 16, 17) and to the computer-generated restriction map based on the nucleotide sequence of this region.

The RNase PH activity of the RNase-deficient strain 18-11 (5), carrying pORFE-1 and pORFE-2, was measured by using tRNA-CCA-[^{3}H]C_n as substrate (4) (Table 1). Despite the fact that strain 18-11 contains polynucleotide phosphorylase, which also catalyzes a P_i-dependent degradative reaction (11) under standard assay conditions, most of the activity observed with this substrate is due to RNase PH (5). Thus, cells containing the parental vectors exhibited relatively low P_i-dependent activity such that overexpression of RNase PH could be measured. As shown in the experiment whose data are presented in Table 1, the *orfE* clones overexpress P_i-dependent activity against tRNA-CCA-[^{3}H]C_n about sixfold, whereas no increase was observed with [^{3}H]poly(A), a substrate for polynucleotide phosphorylase. These findings indicate that the *orfE* clones overexpress RNase PH but not polynucleotide phosphorylase. Interestingly, even though pUC18 has a higher copy number than pHC79 (7, 18), in various experiments the clones exhibited similar levels of overexpression (3- to 10-fold). This may be due to selection against high levels of RNase PH in these cells.

To further substantiate that *orfE* encodes RNase PH, pORFE-2 was cut with *Cla*I and *Mlu*I to remove about 450 nucleotides from the upstream and N-terminal regions of *orfE* (15, 16). The ends were blunted with mung bean nuclease, and a kanamycin resistance cassette from plasmid pUC4K (18) was inserted. The resulting plasmid, pORFE-3, did not increase RNase PH activity in strain 18-11. These

* Corresponding author.

RNase PH: Met X Pro Ala Gly Arg Ser Asn Asn Gln Val Arg Pro Val Thr Leu Thr Arg Asn Tyr Thr
* * * * * * * * * * * * * * * * * * * *
OrfE: Met Arg Pro Ala Gly Arg Ser Asn Asn Gln Val Arg Pro Val Thr Leu Thr Arg Asn Tyr Thr

FIG. 1. N-terminal sequence of RNase PH and the *orfE* gene product. RNase PH purified from strain 18-11 II^+ PNP^- (2) as described in the text was subjected to gas-phase sequencing on an Applied Biosystems model 470A sequencer. On the basis of the intensity of the RNase PH band when stained with Coomassie brilliant blue, the gel was estimated to contain approximately 200 pmol of RNAse PH protein. The sequence shown for the N-terminal sequence of *orfE* comes from the GenBank data base (accession number X00781). X means the residue could not be identified with certainty. Asterisks indicate identity.

data show that the 3.1-kb *Bam*HI-*Eco*RI fragment carrying *orfE* is directly responsible for RNase PH overexpression.

The data presented here identify the open reading frame, *orfE*, located upstream of *pyrE* at 81.7 min, as the gene encoding RNase PH. We suggest that this gene be called *rph*. This region of the chromosome was initially investigated because it was thought that it might be involved in the regulation of *pyrE* expression (16, 17). S1 analysis of RNA synthesis in this region showed that transcription of *pyrE* initiated from a promoter upstream of *orfE* and that the two genes were present in an operon (15, 16). It was also shown that actual translation of the region upstream of *pyrE* was essential for proper pyrimidine regulation, although *lacZ* could substitute for the OrfE (RNase PH) protein (1, 2). At present, it is not clear why a ribonuclease gene and a pyrimidine biosynthesis gene would be transcribed together in one operon, or, furthermore, why the translation of RNase PH would regulate the level of transcription of *pyrE*.

Some discrepancies have been observed in the molecular masses of OrfE (RNase PH) when determined by different methods. First, with minicell analyses, *orfE* was reported to encode two protein species of 30 and 32 kDa on SDS-PAGE, which both disappeared when *orfE* was interrupted (15, 16). In contrast, we find that active RNase PH migrates as a single band of about 33 kDa on SDS-PAGE (data not shown). The reason for these differences is not known, but they could be due to limited proteolysis of the protein. Second, sequencing of the *orfE* gene gave a molecular mass of 25,497 Da for the OrfE protein (16). This difference between the molecular masses obtained on SDS-PAGE and predicted from the DNA sequence is not understood. Many proteins do run anomalously on SDS-PAGE, but some type of protein modification also could be involved. Finally, on the basis of gel filtration, RNase PH elutes as a protein of approximately 45 kDa (5). This raises the possibility that the native protein is a dimer or is associated with another component. Further work will be necessary to resolve these issues. Nevertheless, the identification of *orfE* as the gene encoding RNase PH will allow for additional studies at the protein as well as the DNA level.

TABLE 1. Overexpression of RNase PH from *orfE*-containing plasmids

Plasmid	Relative P_i-dependent sp act[a] with substrate:	
	tRNA-CCA-[^{3}H]C_n	[^{3}H]poly (A)
pHC79	1.0	1.0
pORFE-1	6.0	0.95
pUC19	1.0	1.0
pORFE-2	5.9	1.3
pORFE-3	0.95	ND

[a] Extracts were prepared from strain 18–11 cells containing the indicated plasmid grown to an A_{550} of 1 in YT medium (12) with 50 μg of ampicillin per ml. Acid-soluble radioactivity from each substrate was measured in the presence and absence of 15 mM $NaPO_4$, pH 7.5, as described previously (5). Data are presented as the relative P_i-dependent specific activity compared with that of the parental vector, which was set to 1.0. ND, not determined.

We thank Juris Ozols and George Korza for the N-terminal sequencing and Asis Das for providing the Kohara lambda phage 572.

This work was supported by the NIH grant GM-16317.

REFERENCES

1. **Bonekamp, F., H. D. Andersen, T. Christensen, and K. F. Jensen.** 1985. Codon defined ribosomal pausing in *Escherichia coli* detected by using the *pyrE* attenuator to probe the uncoupling between transcription and translation. Nucleic Acids Res. **13**:4113–4123.
2. **Bonekamp, F., K. Clemmesen, O. Karlstrom, and K. F. Jensen.** 1984. Mechanism of UTP-modulated attenuation at the *pyrE* gene of *Escherichia coli*: an example of operon polarity control through the coupling of translation to transcription. EMBO J. **3**:2857–2861.
3. **Cudny, H., and M. P. Deutscher.** 1988. 3′ Processing of tRNA precursors in ribonuclease-deficient *Escherichia coli*. J. Biol. Chem. **263**:1518–1523.
4. **Deutscher, M. P., and R. K. Ghosh.** 1978. Preparation of synthetic tRNA precursors with tRNA nucleotidyltransferase. Nucleic Acids Res. **5**:3821–3830.
5. **Deutscher, M. P., G. T. Marshall, and H. Cudny.** 1988. RNase PH: an *Escherichia coli* phosphate-dependent nuclease distinct from polynucleotide phosphorylase. Proc. Natl. Acad. Sci. USA **85**:4710–4714.
6. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12**:387–395.
7. **Hahn, B., and J. Collins.** 1980. A small cosmid for efficient cloning of large DNA fragments. Gene **11**:291–298.
8. **Kohara, Y., K. Akiyama, and K. Isono.** 1987. The physical map of the whole *E. coli* chromosome: application of a new strategy for rapid analysis and sorting of a large genomic library. Cell **50**:495–508.
9. **Lee, J. S., G. An, J. D. Friesen, and K. Isono.** 1981. Cloning and the nucleotide sequence of the genes for *Escherichia coli* ribosomal proteins L28 (rpmB) and L33 (rpmG). Mol. Gen. Genet. **184**:218–223.
10. **LeGendre, N., and P. T. Matsudaira.** 1989. Purification of proteins and peptides by SDS-PAGE, p. 49–69. *In* P. T. Matsudaira (ed.), A practical guide to protein and peptide purification for microsequencing. Academic Press, Inc., San Diego, Calif.
11. **Littauer, U. Z., and H. Soreq.** 1982. Polynucleotide phosphorylase, p. 517–553. *In* P. D. Boyer (ed.), The enzymes: nucleic acids, vol. 15, part B. Academic Press, Inc., New York.
12. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
13. **Ost, K. A., and M. P. Deutscher.** 1990. RNase PH catalyzes a synthetic reaction, the addition of nucleotides to the 3′ end of RNA. Biochimie **72**:813–818.
14. **Ost, K. A., and M. P. Deutscher.** Unpublished data.
15. **Poulsen, P., J. T. Andersen, and K. F. Jensen.** 1989. Molecular and mutational analysis of three genes preceding *pyrE* on the *Escherichia coli* chromosome. Mol. Microbiol. **3**:393–404.

16. **Poulsen, P., F. Bonekamp, and K. F. Jensen.** 1984. Structure of the *Escherichia coli pyrE* operon and control of *pyrE* expression by a UTP modulated intercistronic attenuation. EMBO J. **3**:1783–1790.
17. **Poulsen, P., K. F. Jensen, P. Valentin-Hansen, P. Carlsson, and L. G. Lundberg.** 1983. Nucleotide sequence of the *Escherichia coli pyrE* gene and of the DNA in front of the protein-coding region. Eur. J. Biochem. **135**:223–229.
18. **Vieira, J., and J. Messing.** 1982. The pUC plasmids, an M13mp17-derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene **19**:259–268.
19. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of M13mp18 and pUC19 vectors. Gene **33**:103–119.