Volume 287, number 1,2, 53–56 FEBS 09974 August 1991

ADONIS 001457939100679E

Molecular cloning of the Ecotin gene in *Escherichia coli*

Hae Ryun Lee[1], Jae Hong Seol[1], Ok Mae Kim[1], Cheol Soon Lee[1], Se Won Suh[2], Yeong-Man Hong[3], Keiji Tanaka[4], Akira Ichihara[4], Doo Bong Ha[1] and Cin Ha Chung[1]

Departments of [1]Molecular Biology and [2]Chemistry, College of Natural Sciences, Seoul National University, Seoul 151-742, Korea, [3]APRO Life Science Institute, Tokushima 772, Japan and *[4]Institute for Enzyme Research, Tokushima University, Tokushima 770, Japan*

Received 24 May 1991

The nucleotide sequence of a 876 bp region in *E. coli* chromosome that encodes Ecotin was determined. The proposed coding sequence for Ecotin is 486 nucleotides long, which would encode a protein consisting of 162 amino acids with a calculated molecular weight of 18 192 Da. The deduced primary sequence of Ecotin includes a 20-residue signal sequence, cleavage of which would give rise to a mature protein with a molecular weight of 16 099 Da. Ecotin does not contain any consensus reactive site sequences of known serine protease inhibitor families, suggesting that Ecotin is a novel inhibitor.

Ecotin; Protease inhibitor; Signal sequence; *eti* gene

1. INTRODUCTION

Ecotin is a periplasmic protein in *Escherichia coli*, that is capable of inhibiting trypsin and other pancreatic serine proteases including chymotrypsin and elastase [1]. It also inhibits rat mast cell chymase and human plasma urokinase [1]. However, it does not inhibit any of the nine soluble endoproteases (named protease Do, Re, Mi, Fa, So, La, Ti, Ci and Pi) nor the chymotrypsin-like (protease I) and trypsin-like (protease II) esterases in *E. coli* [1–5]. Therefore, it has been suggested that Ecotin is unlikely involved in the control of intracellular protein breakdown but may play a role in protecting against proteases in the natural environment of *E. coli*, such as the mammalian gastrointestinal tract [1].

The 38 kDa Ecotin consists of two identical subunits of 18 kDa, and contains one intramolecular disulfide bond [1]. One dimeric inhibitor binds two trypsin molecules to form a mixed tetramer, in which trypsin molecules are inhibited completely [1]. It is stable at low pH and high temperature, and therefore shares general physical properties of polypeptide inhibitors [1,6,7]. In an attempt to elucidate the structural features of Ecotin and its mode of interaction with target proteases, we determined the nucleotide sequence of the Ecotin gene, which is henceforth referred to as the *eti* gene.

2. MATERIALS AND METHODS

2.1. *Cloning of the* eti *gene*

The DNA clones producing Ecotin were obtained by immunoscreening of an *E. coli* genomic library in λgt11 (Clontech). Screening of the library was carried out using anti-Ecotin antiserum as described [8]. Preparation of phage DNA and other DNA manipulations were done by following the standard procedures [9]. Ecotin was purified as described previously [1]. Anti-Ecotin antiserum was obtained by injecting the purified inhibitor to albino rabbits.

The 6.6 kb insert DNA carrying the *eti* gene in one of the positive phage clones was ligated into the *Eco*RI site of Bluescript KS$^+$ plasmid (pBS). The insert DNA was trimmed to 2.2 kb with *Eco*RI and *Cla*I, and ligated into the respective sites of the plasmid. The 2.2 kb DNA was further trimmed to 0.8 kb using *Eco*RI and *Stu*I, and ligated into the *Eco*RI and *Eco*RV sites of pBS. The resulting plasmids were referred to as pBS2.2 and pBS0.8, respectively.

2.2. *Protein and DNA sequencing*

The purified Ecotin was degraded with an automated gas-phase protein sequencer (Applied Biosystems, model 470A) according to operation program 02RPTH provided for the sequencer. PTH-amino acids liberated were identified by HPLC using an Ultrasphere ODS column (2 × 250 mm, Beckman) at 49°C. Solvents used for the elution of PTH-amino acids were composed of acetonitrile/0.1% trifluoroacetic acid adjusted to pH 4.9 at a ratio of 1:10 or 1:1 (v/v). The eluates were monitored simultaneously at 269 and 322 nm. DNA sequencing was carried out by the dideoxy chain-termination method [10] with a T7 Sequencing kit (Pharmacia).

3. RESULTS AND DISCUSSION

To isolate DNA encoding Ecotin, we screened an *E. coli* genomic DNA library in λgt11 and obtained 3 positive clones. The 6.6 kb DNA insert from one of the clones was trimmed to 2.2 kb fragment, and ligated into pBS as summarized in Fig. 1. *E. coli* JM109 cells transformed with pBS2.2 were capable of overproducing Ecotin as examined by immunoblot analysis using anti-Ecotin antibody (Fig. 2). The plasmid pBS0.8 containing *Eco*RI-*Stu*I digest of 2.2 kb fragment also similarly overproduced the inhibitor when transformed into the same cell (data not shown). Thus, it appears

Correspondence address: C.H. Chung, Department of Molecular Biology, College of Natural Sciences, Seoul National University, Seoul 151-742, Korea. Fax: (82) (2) 872 1993.

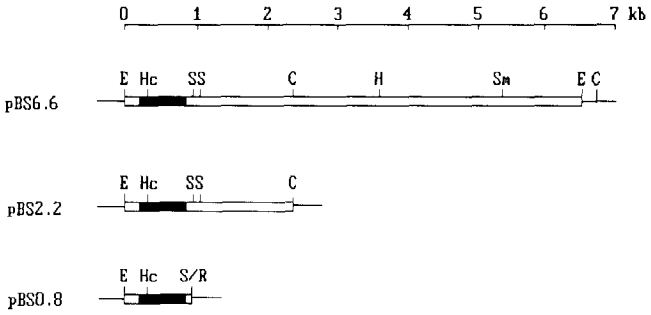

Fig. 1. Restriction maps of plasmids containing the *eti* gene. The 6.6 kb insert DNA of a positive phage clone was ligated into *Eco*RI site of Bluescript KS$^-$ (pBS) to generate pBS6.6. The resulting plasmid was digested with *Eco*RI and *Cla*I, and the 2.2 kb DNA was ligated into the respective sites of pBS. After digesting pBS2.2 with *Eco*RI and *Stu*I, 0.8 kb fragment was ligated into *Eco*RI and *Eco*RV sites. This construction generates pBS0.8, in which the *Stu*I and *Eco*RV sites are no longer available. The lines show pBS and the bars indicate insert DNA, in which the shaded portions represent the coding region. Restriction sites are indicated by the following abbreviations: E, *Eco*RI; Hc, *Hin*cII; S, *Stu*I; C, *Cla*I; H, *Hin*dIII; Sm, *Sma*I; R, *Eco*RV.

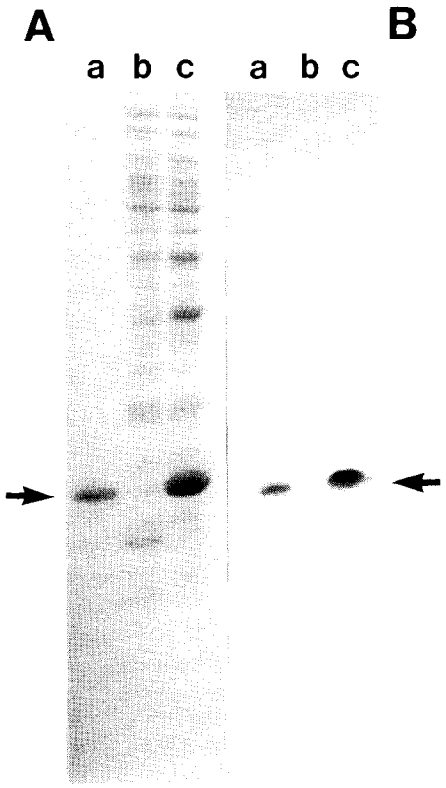

Fig. 2. Immunoblot analysis of Ecotin in *E. coli* cells. Extracts were prepared from *E. coli* that had been grown in Luria broth until an A_{600} of 0.5. The extracts were electrophoresed in 7–18% polyacrylamide gradient slab gels containing SDS [17], and the proteins in the gels were stained with Coomassie R250 (A) or transferred onto nitrocellulose membranes (B). The membranes were incubated with anti-Ecotin antiserum and then with anti-rabbit IgG conjugated with alkaline phosphatase [18]. Lanes: (a) 2 μg of the purified Ecotin; (b) JM109; (c) JM109/pBS2.2. The arrows indicate where Ecotin migrated.

that the 0.8 kb DNA contains a full-length DNA sequence for the expression of the *eti* gene.

The nucleotide sequence of a 876 bp DNA including the entire *eti* gene is shown in Fig. 3. The coding region is 486 nucleotides long and corresponds to 162 amino acids. The most likely initiation codon is followed by the sequence AGGA, which is a consensus ribosome binding site in other *E. coli* promotors [11]. Putative sequences homologous to the −35 and −10 regions of *E. coli* promotor elements are located upstream of the ATG start codon [12]. No other open reading frame was found in either strand.

The deduced amino acid sequence of Ecotin reveals the presence of two Cys residues that likely form a disulfide bridge, in accord with our previous findings [1]. The molecular weight of Ecotin was calculated to be 18 192 Da, which is also in agreement with the observed size of the purified protein on polyacrylamide gel containing sodium dodecyl sulfate [1]. However, the amino acid composition of Ecotin derived from the DNA se-

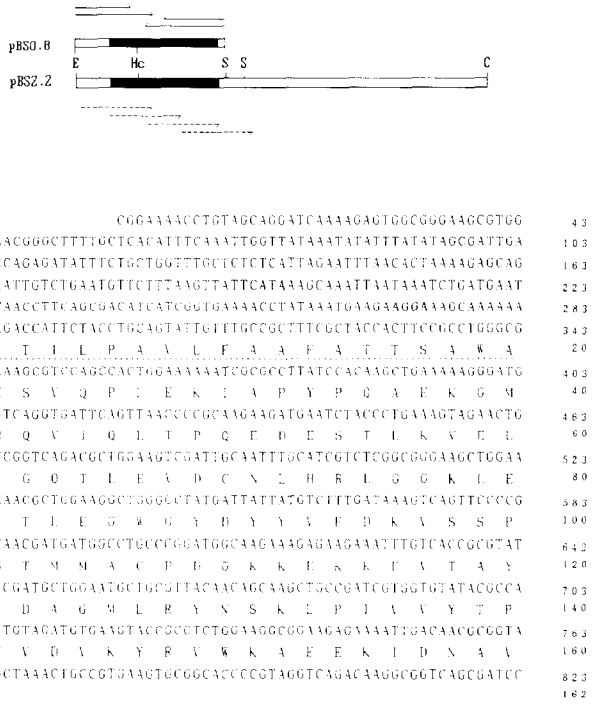

Fig. 3. Nucleotide sequence of the *eti* gene and its flanking regions. Upper panel shows the sequencing strategies. The sequences of the 5′- and 3′-ends of pBS0.8 were determined by direct sequencing of pBS0.8 using primers of T_3 and T7 promotor of pBS (solid arrows). Clones used for sequencing were obtained by making deletions from the *Eco*RI site of pBS2.2 and then sequences as above (dashed lines). Lower panel shows the nucleotide sequence of the *eti* gene and the deduced amino acid sequence. The boldface sequences with underlines from 5′-end represent putative −35, −10 and Shine-Dalgarno regions, respectively. The dotted line shows the signal sequence and the solid line indicates the amino acid sequence obtained by Edman degradation of the purified Ecotin protein.

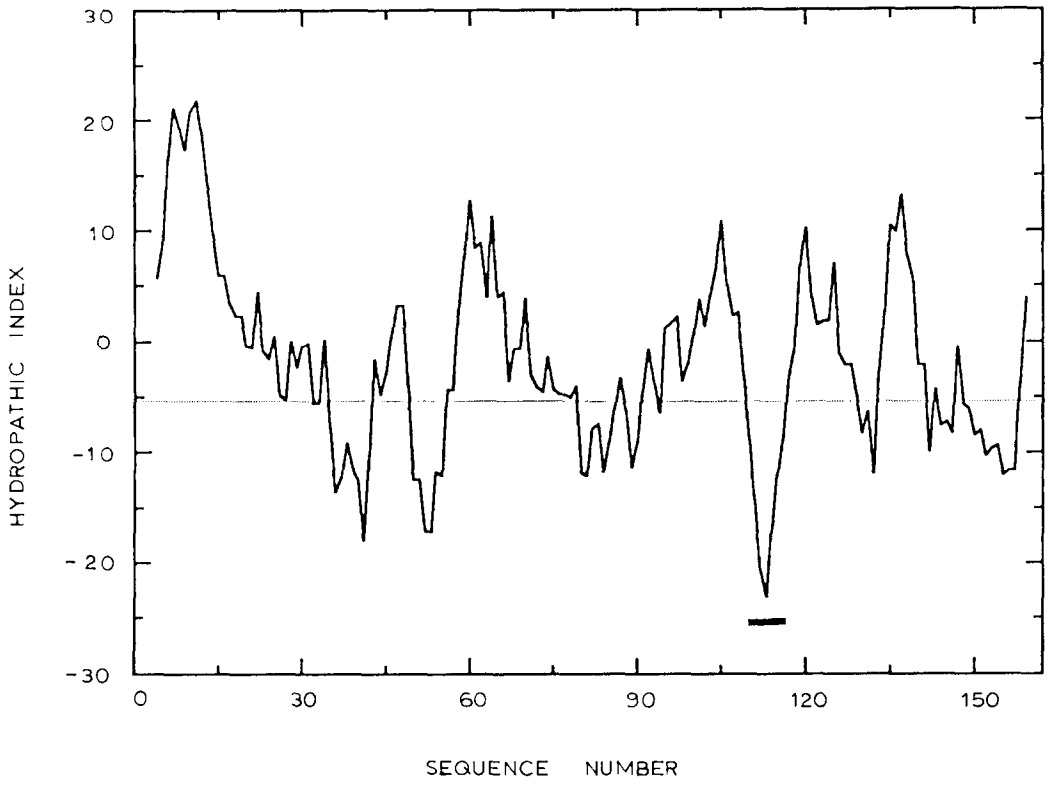

Fig. 4. Hydropathy pattern of Ecotin. Hydrophobicity and hydrophilicity were calculated along the deduced amino acid sequence of Ecotin with a window of 7 residues. The bar indicates a hydrophilic domain that could serve as a potential antigenic site in Ecotin (KKEKK of 111–115 residues in Fig. 3).

quence differs appreciably, particularly the content of Val and Arg, from that previously reported for the isolated protein [1]. Because of this disagreement, the N-terminal sequence of the purified Ecotin was determined to verify that we had indeed cloned the gene for Ecotin. The sequence of 30 amino acids (underlined in Fig. 3) was identical to that deduced from the DNA sequence. Thus, the amino acid composition reported previously must be erroneous for reasons that are unclear.

Since Ecotin is a periplasmic protein, the N-terminal signal sequence can be assumed to exist for its secretion across the cytoplasmic membrane [13,14]. The N-terminus of the deduced amino acid sequence (Fig. 3) indeed displays the common features of prokaryotic signal sequences [14]. A basic amino acid (Lys) resides within the N-terminal region, followed by a central hydrophobic core. In addition, the C-terminal end of the signal sequence contains a potential processing site [14], which was confirmed by sequencing the purified Ecotin protein from its N-terminus (Fig. 3). Thus, the signal sequence appears likely to consist of 20 amino acid residues.

The hydropathy of Ecotin was calculated along the deduced amino acid sequence (162 residues) according to the method of Kyte and Doolittle [15]. As shown in Fig. 4, a distinct hydrophobic domain is evident in the N-terminal region of Ecotin that corresponds to the central core of the signal sequence. Of particular interest is the presence of a strong hydrophilic domain that could serve as a potential antigenic site of the protein [16]. This domain includes a stretch of highly charged amino acids, Lys-Lys-Glu-Lys-Lys (Fig. 3).

On computer analysis using the data bases of EMBL/GDB (European Molecular Biology Laboratory), NBRF/PDB (National Biomedical Research Foundation) and PRF/SEQDB (Protein Research Foundation), no obvious overall sequence homology of Ecotin with previously reported proteins could be found. Ecotin neither contains any consensus reactive site sequences of known serine protease inhibitor families [6,7]. Thus, Ecotin seems to be a novel protein inhibitor.

Acknowledgements: This work was supported by grants from The Ministry of Education and Korea Science and Engineering Foundation through SRC for Cell Differentiation.

REFERENCES

[1] Chung, C.H., Ives, H.E., Almeda, S. and Goldberg, A.L. (1983) J. Biol. Chem. 258, 11032–11038.

[2] Goldberg, A.L., Swamy, K.H.S., Chung, C.H. and Larimore, F.S. (1982) Methods Enzymol. 80, 680–702.
[3] Hwang, B.J., Woo, K.M., Goldberg, A.L. and Chung, C.H. (1988) J. Biol. Chem. 263, 8727–8734.
[4] Pacaud, M. and Ureil, J. (1971) Eur. J. Biochem. 23, 435–442.
[5] Pacaud, M. and Richard, C. (1975) J. Biol. Chem. 250, 7771–7779.
[6] Laskowski, M. Jr. and Kato, I. (1980) Annu. Rev. Biochem. 49, 593–626.
[7] Barrett, A.J. and Salvesen, G. (1986) Proteinase Inhibitors, Elsevier, Amsterdam.
[8] Young, R.A. and Davis, R.W. (1983) Proc. Natl. Acad. Sci. USA 80, 1194–1198.
[9] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd edn., Cold Spring Harbor Laboratory, Gold Spring Habor, NY.
[10] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[11] Kozak, M. (1983) Microbiol. Rev. 47, 1–45.
[12] McClure, W.R. (1985) Annu. Rev. Biochem. 54, 171–204.
[13] Silhavy, T.J., Benson, S.A. and Emr, S.D. (1983) Microbiol. Rev. 47, 313–344.
[14] Oliver, D. (1985) Annu. Rev. Microbiol. 39, 615–648.
[15] Kyte, J. and Doolittle, R.F. (1982) J. Mol. Biol. 157, 105–132.
[16] Hopp, T.P. and Woods, K.R. (1981) Proc. Natl. Acad. Sci. USA 78, 3824–3828.
[17] Laemmli, U.K. (1970) Nature 227, 680–685.
[18] Blake, M.S., Johnston, K.H., Russell-Jones, G.J. and Gotschlich, E.C. (1984) Anal. Biochem. 136, 175–179.

NOTE ADDED IN PROOF

During the preparation of this manuscript, a report of McGrath et al., of cloning of the same Ecotin gene appeared in J. Biol. Chem. (1991) 266, 6620–6625, but the cloning strategy of the gene reported was different from that described in present work.