Purification and Characterization of *Escherichia coli* Endonuclease III from the Cloned *n*th Gene[†,‡]

Hitomi Asahara, Peter M. Wistort, Janet F. Bank, Ronald H. Bakerian, and Richard P. Cunningham*
Department of Biological Sciences, Center for Biochemistry and Biophysics, State University of New York at Albany, Albany, New York 12222

Received November 28, 1988; Revised Manuscript Received January 31, 1989

ABSTRACT: The gene which codes for endonuclease III of *Escherichia coli* has been sequenced. The *n*th gene was previously subcloned and defined as the gene which led to overproduction of endonuclease III when present on a multicopy plasmid and which created a deficiency in endonuclease III activity when mutated. The *n*th gene was sequenced and translated into a predicted polypeptide. The molecular weight (23 546), the amino-terminal amino acid sequence, and the amino acid composition of the polypeptide predicted from the nucleotide sequence are in excellent agreement with those same properties determined for the purified protein. Thus, the *n*th gene is the structural gene for endonuclease III. Inspection of the nucleotide sequence reveals that there is an open reading frame immediately upstream of the *n*th gene, suggesting that it might be part of an operon. There is a region of dyad symmetry which could form a hairpin stem and loop structure if transcribed into RNA characteristic of a ρ-dependent terminator downstream from the *n*th gene. The *n*th gene of *Escherichia coli* has been cloned onto a λP_L expression vector which yields approximately 300-fold overproduction of endonuclease III. We have purified the enzyme to apparent homogeneity using two chromatographic steps. Our purification scheme allowed the preparation of 117 mg of protein from 190 g of *E. coli* with a 70% yield. The purified protein has both AP endonuclease activity and DNA *N*-glycosylase activity. The protein has a Stokes radius of 2.25 nm, a sedimentation coefficient of 2.65 S, a molecular weight of 26 300 in the native state and 27 300 in the denatured state, and a frictional ratio of 1.13. The optical spectrum of the protein exhibits absorption peaks at 280 and 410 nm.

Endonuclease III from *Escherichia coli* has both an apurinic and/or apyrimidinic (AP)[1] endonuclease activity and a DNA *N*-glycosylase activity. The endonuclease activity cleaves 3′ to AP sites via a β-elimination reaction (Bailly & Verly, 1987; Kow & Wallace, 1987), leaving a baseless 2,3-didehydro-2,3-dideoxyribose residue. This residue is a poor substrate for DNA polymerase I and the Klenow fragment of DNA polymerase I (Mosbaugh & Linn, 1982; Katcher & Wallace, 1983), and further processing by exonuclease III or endonuclease IV may be required for efficient repair (Warner et al., 1980; Mosbaugh & Linn, 1982). The DNA *N*-glycosylase activity releases thymine residues damaged by ring saturation, fragmentation, or ring contraction including thymine glycol (Demple & Linn, 1980; Katcher & Wallace, 1983; Breimer & Lindahl, 1984; Higgins et al., 1987), 5,6-dihydrothymine (Demple & Linn, 1980), urea (Breimer & Lindahl, 1980, 1984; Katcher & Wallace, 1983), 5-hydroxy-5-methylhydantoin (Breimer & Lindahl, 1984, 1985), and methyltartronylurea (Breimer & Lindahl, 1984). The enzyme also releases an unidentified cytosine photoproduct (Doetsch et al., 1986). Endonuclease III incises damaged DNA at cytosines and thymines (Doetsch et al., 1986, 1987; Helland et al., 1986; Weiss & Duker, 1986, 1987) and also at guanines (Doetsch et al., 1987).

Enzymes with the same sequence specificity of cleavage of damaged DNA have been found in bovine and human cells (Helland et al., 1986; Doetsch et al., 1987) and in yeast (Gossett et al., 1988). These enzymes, like endonuclease III, are small (25 000–42 000 molecular weight) and have no requirement for divalent cations. In addition, endonucleases from a variety of eucaryotic organisms that recognize and cleave DNA damaged by oxidizing agents, ionizing radiation, and ultraviolet irradiation have been described (Van Lacker & Tomura, 1974; Bachetti & Benne, 1975; Kuhnlein et al., 1978; Breimer, 1983, 1986; Brent, 1983). An enzyme very similar to endonuclease III has been isolated from the Gram-positive eubacterium *Micrococcus luteus* (Jorgenson et al., 1987).

Thymine glycol and urea residues are found in DNA in vivo after cells are exposed to ionizing radiation or oxidative stress (Leadon & Hanawalt, 1983; Breimer & Lindahl, 1985). These lesions constitute a block to DNA replication (Ide et al., 1985; Rouet & Essigman, 1985; Clark & Beardsley, 1986) and thus must be considered as lesions capable of causing mutagenesis or cell death.

The fact that enzymes very similar to endonuclease III are found in a number of evolutionarily diverse species and that endonuclease III can initiate repair events at lesions which if left unrepaired could lead to mutagenesis or cell death suggests that this class of enzymes may be critical for maintaining the genetic and biological integrity of cells. Endonuclease III is the prototype for this class of enzymes, and the study of this enzyme may be of broad importance for understanding DNA repair in both procaryotes and eucaryotes.

A recombinant plasmid bearing the *n*th gene has been previously identified (Cunningham & Weiss, 1985). We have

[†] This research was supported by National Institutes of Health Grant GM33346.

[‡] The nucleic acid sequence in this paper has been submitted to GenBank under Accession Number J02857.

* Author to whom correspondence should be addressed.

[1] Abbreviations: AP, apurinic and/or apyrimidinic; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; HPLC, high-performance liquid chromatography; kbp, kilobase pair; PMSF, phenylmethanesulfonyl fluoride; PTH, phenylthiohydantoin; ORF, open reading frame; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; Tris·HCl, tris(hydroxymethyl)aminomethane hydrochloride; UV, ultraviolet.

0006-2960/89/0428-4444$01.50/0

subcloned the gene and determined its nucleotide sequence. Using the cloned copy of the gene and information from the DNA sequence, we have constructed a plasmid which overproduces the enzyme 300-fold. We have devised a simple, rapid purification scheme which allows for purification of hundreds of milligrams of the enzyme in high yield. The ability to easily purify the enzyme has allowed the detailed biochemical and biophysical characterization of the enzyme as an iron–sulfur protein (Cunningham et al., 1989).

Materials and Methods

Bacteria and Plasmids. *Escherichia coli* $\lambda N99_{CI857}$ was obtained from Dr. Laurence Grossman (John Hopkins). The plasmids pHUB2 (Bernard & Helinski, 1979), pLC9-9 (Clarke & Carbon, 1976), and pBR322 were from our laboratory collection.

Preparation and Manipulation of DNA. Plasmid DNA was prepared by a modified version of the alkaline lysis method of Birnboim and Doly (1979). Restriction enzymes were used in accordance with the suppliers' recommendations. The Klenow fill-in reaction and the ligation of restriction fragments by T4 DNA ligase have been described (Maniatis et al., 1982). The preparation and transformation of competent cells have been described (Rodriguez & Tait, 1983).

DNA Sequence Analysis. Restriction fragments of the *n*th gene were subcloned into the M13 cloning vectors mp18, mp19, um20, and um21. Nucleotide sequence was determined by the dideoxy chain termination method (Sanger et al., 1977) using either Klenow fragment of *E. coli* DNA polymerase I or a modified form of T7 DNA polymerase (Tabor & Richardson, 1987) and M13 universal sequencing primers. Nucleotide sequences were compiled and analyzed by using the IBI gel reader and IBI/Pustell DNA and protein sequence analysis system.

Protein Sequence and Amino Acid Composition Determinations. The N-terminal sequence analysis was performed by using an Applied Biosystems INC470A gas-phase protein sequenator which was connected on-line to an ABI120 (HPLC) PTH analyzer. Data were collected and yields analyzed on a Nelson Analytical 3000 Series chromatography system. Amino acid composition was determined by analysis of a 6 N HCl hydrolysate on a Dionex 500 ion-exchange HPLC system.

Substrates. Bacteriophage T4·AP [^{3}H]DNA was prepared as described (Cunningham & Weiss, 1985). Apurinic pBR322 DNA was prepared by the method of Lindahl and Andersson (1972). UV-irradiated pBR322 was prepared by irradiating pBR322 in 10 mM Tris·HCl (pH 8.0)/1 mM EDTA with a germicidal lamp at a fluence of 2 J m^{-2} s^{-1} to a final dose of 2000 J m^{-2}.

Endonuclease III Assays. The activity of the enzyme was based on the degradation of T4·AP [^{3}H]DNA to acid-soluble oligonucleotides as previously described (Cunningham & Weiss, 1985). The activity of the enzyme was also monitored by conversion of form I to form II apurinic or UV-irradiated pBR322 DNA by agarose gel electrophoresis or by monitoring the incision of damaged DNA fragments of defined sequence on sequencing gels (Helland et al., 1986).

Determination of Protein Concentration. Protein concentration was determined by the method of Bradford (1976).

Growth of Cells for Enzyme Purification. The medium used for growth of cells contained 33 g of Bacto-Tryptone, 20 g of yeast extract, 7.5 g of NaCl, and 3.5 mL of 10 N NaOH per liter. *E. coli* $\lambda N99_{CI857}$ carrying pHIT1 was grown in a 14-L New Brunswick Microfirm fermentor to an A_{600} of 2.0 at 32 °C. The temperature was then shifted to 42 °C for 2 h, and the cells were harvested by centrifugation at room temperature. The cell paste (190 g) was stored at −80 °C.

Enzyme Purification. All procedures were performed at 4 °C unless otherwise indicated. Frozen cells were suspended in 700 mL of 200 mM NaCl, 50 mM Tris·HCl, pH 7.6, 2.5 mM EDTA, 0.1 mM DTT, and 0.1 mM PMSF (extraction buffer) with the aid of a Waring blender. The cell suspension was sonicated for 2 × 5 min at maximum intensity with a Branson sonifier in a rosette cooling cell immersed in an ice bath. The sonicate was stirred for 2 h on ice, cell debris was removed by centrifugation at 48000*g* for 30 min, and the supernatant (fraction I, 770 mL) was saved.

Fraction I was treated with 77 mL of 5% poly(ethylenimine). Poly(ethylenimine) was prepared as a 10% solution in extraction buffer minus PMSF adjusted to a pH of 7.0 with concentrated HCl, dialyzed overnight against the same buffer, and then diluted to a final concentration of 5% with the same buffer. Poly(ethylenimine) was added to fraction I slowly with stirring, and after 1 h, the mixture was centrifuged at 48000*g* for 30 min. The supernatant was recovered and dialyzed overnight against two changes of 4 L of 100 mM potassium phosphate, pH 6.6, and 0.1 mM DTT (buffer A). The dialyzed protein was centrifuged at 48000*g* for 30 min to remove residual poly(ethylenimine) precipitated by the phosphate buffer. The supernatant (fraction II, 830 mL) was saved.

Fraction II was applied to a 2.6 × 39 cm column of S-Sepharose (Pharmacia) equilibrated with buffer A. The column was washed with 100 mL of buffer A, and the bound protein was eluted with a linear gradient (1000 mL) of zero to 0.43 M sodium chloride in buffer A. The most active fractions were pooled (fraction III, 83 mL).

Fraction III was adjusted to a concentration of 2 M ammonium sulfate and applied to a 1.6 × 35 cm column of phenyl-Sepharose CL-4B (Pharmacia) equilibrated with buffer A + 2 M ammonium sulfate. The column was washed with 60 mL of buffer A + 2 M ammonium sulfate, and the bound protein was eluted with a linear gradient (600 mL) of 2.0 to 0 M ammonium sulfate in buffer A. The most active fractions were pooled (fraction IV, 78 mL) and dialyzed against buffer A. Fraction IV is stable for several months at 4 °C. For long-term storage, the enzyme was dialyzed against 50% glycerol, 100 mM potassium phosphate, pH 6.6, 0.1 mM DTT, and 0.005% Triton X-100 and stored at −20 °C. The enzyme stored at −20 °C is stable for several years.

Polyacrylamide Gel Electrophoresis of Denatured Proteins. The method of Laemmli (1970) was used for SDS–PAGE. For the determination of denatured molecular weight, the following standard proteins were used for calibration: phosphorylase *b* (M_r 94000), albumin (M_r 67000), ovalbumin (M_r 43000), carbonic anhydrase (M_r 30000), trypsin inhibitor (M_r 20000), and α-lactalbumin (M_r 14400). The R_f for each protein was determined, and the denatured molecular weight for endonuclease III was derived from a standard calibration curve.

Gel Filtration of Native Proteins. Sephadex G-75 Superfine (Pharmacia) was equilibrated with a solution of 0.1 M potassium phosphate (pH 6.6). The column (1.5 × 45 cm) was operated at a flow rate of 1.25 mL cm^{-2} h^{-1}. Protein mixtures were applied in a volume of 1.5 mL. Endonuclease III was detected by enzymatic assay. Albumin (M_r 67000), ovalbumin (M_r 43000), chymotrypsinogen A (M_r 25000), and ribonuclease A (M_r 13700) were detected by the absorbance at 280 nm. The void volume of the column was determined by measuring the elution volume of Blue Dextran 2000. K_{av} for each protein was determined, and the molecular weight of

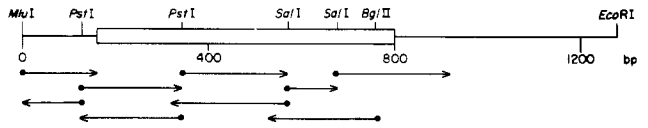

FIGURE 1: Strategy and restriction sites used for sequencing the *n*th gene. The restriction fragment shown is a 1280 bp *Mlu*I–*Eco*RI fragment. Arrows indicate the direction and extent of each sequence determination from internal restriction sites indicated by closed circles. The open bar represents the *n*th structural gene.

endonuclease III was derived from a standard calibration curve.

Sedimentation Analysis of Proteins. Proteins in a total volume of 0.1 mL were layered on linear sucrose gradients (4.8 mL) containing 200 mM NaCl, 50 mM potassium phosphate, pH 6.6, 10 mM EDTA, and 5–20% sucrose. Two micrograms of endonuclease III and 250 μg of bovine carbonic anhydrase (3.06 S) were layered on one gradient, and 250 μg of ovalbumin (3.66 S) was layered on a second gradient and centrifuged in a Beckman SW50.1 rotor at 45000 rpm for 24 h at 4 °C. Fractions were collected from the bottoms of the tubes. Ovalbumin and carbonic anhydrase were detected by protein assay, and endonuclease III was detected by enzymatic assay. The recovery of endonuclease III was approximately 60%.

Optical Spectrum. The optical spectrum was recorded in the 250–600-nm range with a Perkin-Elmer 552 spectrophotometer.

Results and Discussion

Sequencing the nth Gene. We have previously identified clones bearing two ColEl hybrid plasmids, PLC9-9 and pLC28-48, from the Clarke–Carbon collection which overproduce endonuclease III (Cunningham & Weiss, 1985). The *n*th gene, which causes overproduction of endonuclease III, was localized to a 1.3 kbp *Eco*RI–*Mlu*I restriction fragment from pLC9-9. The sequence of a portion of this restriction fragment, including the entire *n*th gene, was determined from overlapping DNA fragments. A detailed restriction map and the specific DNA fragments sequenced are shown in Figure 1. The DNA sequence of 780 nucleotides including the *n*th gene is shown in Figure 2. An open reading frame starting at nucleotide 85 with an ATG codon and ending at nucleotide 720 with a TGA stop codon was identified. The identified ORF has a codon usage consistent with the nonrandom codon usage found for a number of *E. coli* genes and would code for a protein of molecular weight 23546.

There is a potential ribosome binding site (Shine & Dalgarno, 1974) preceding the structural gene at nucleotides 73 through 77; however, no sequences similar to a consensus *E. coli* promoter (Hawley & McClure, 1983) were found preceding the gene. There is an open reading frame which overlaps with the N-terminus of the structural gene for endonuclease III and which has a codon usage bias consistent with known *E. coli* genes. This ORF is at least 61 amino acids long and shares 8 nucleotides with the *n*th gene. We found a region of dyad symmetry downstream of the endonuclease III gene which is capable of folding into a hairpin stem and loop structure with a free energy of −14.2 kcal mol^{-1} (Tinoco et al., 1973) if transcribed into RNA. This may be a ρ-dependent terminator of transcription. Verification of the proposal that endonuclease III is the terminal gene in an operon requires further experimental proof.

The gene product of the endonuclease III gene is a basic protein (+3 net positive charge), as might be expected of a DNA binding protein, with a predicted p*I* of 9.83 in agreement

```
          10         20         30         40         50         60
          *          *          *          *          *          *
ACC TGA TTG ATG AAA GAA TGA AAA AGC GCC GTG CTG AAG CAG CTG CAG AAC GTG CAT TGC

          70         80         90         100        110        120
          *          *          *          *          *          *
CAA CGG TGA AAC AGG GAA TGT CTG ATG AAT AAA GCA AAA CGC CTG GAG ATC CTC ACT CGC
              RBS          Met Asn Lys Ala Lys Arg Leu Glu Ile Leu Thr Arg

          130        140        150        160        170        180
          *          *          *          *          *          *
CTG CGT GAG AAC AAT CCT CAT CCC ACC ACC GAG CTT AAT TTC AGT TCG CCT TTT GAA TTG
Leu Arg Glu Asn Asn Pro His Pro Thr Thr Glu Leu Asn Phe Ser Ser Pro Phe Glu Leu

          190        200        210        220        230        240
          *          *          *          *          *          *
CTG ATT GCC GTA CTG CTT TCC GCT CAG GCG ACC GAT GTC AGT GTT AAT AAG GCG ACG GCG
Leu Ile Ala Val Leu Leu Ser Ala Gln Ala Thr Asp Val Ser Val Asn Lys Ala Thr Ala

          250        260        270        280        290        300
          *          *          *          *          *          *
AAA CTC TAC CCG GTG GCG AAT ACG CCT GCA GCG ATG CTT GAA CTG GGC GTT GAA GGG GTG
Lys Leu Tyr Pro Val Ala Asn Thr Pro Ala Ala Met Leu Glu Leu Gly Val Glu Gly Val

          310        320        330        340        350        360
          *          *          *          *          *          *
AAA ACC TAT ATC AAA ACG ATT GGG CTT TAT AAC AGC AAA GCA GAA AAT ATC ATC AAA ACC
Lys Thr Tyr Ile Lys Thr Ile Gly Leu Tyr Asn Ser Lys Ala Glu Asn Ile Ile Lys Thr

          370        380        390        400        410        420
          *          *          *          *          *          *
TGC CGT ATC TTG CTG GAG CAG CAT AAT GGC GAG GTT CCG GAA GAT CGT GCT GCG CTT GAA
Cys Arg Ile Leu Leu Glu Gln His Asn Gly Glu Val Pro Glu Asp Arg Ala Ala Leu Glu

          430        440        450        460        470        480
          *          *          *          *          *          *
GCC CTG CCC GGC GTA GGT CGT AAA ACA GCC AAC GTC GTA TTA AAC ACT GCA TTC GGC TGG
Ala Leu Pro Gly Val Gly Arg Lys Thr Ala Asn Val Val Leu Asn Thr Ala Phe Gly Trp

          490        500        510        520        530        540
          *          *          *          *          *          *
CCG ACT ATT GCT GTC GAC ACG CAC ATT TTC CGC GTT TGT AAT CGT ACT CAA TTT GCG CCG
Pro Thr Ile Ala Val Asp Thr His Ile Phe Arg Val Cys Asn Arg Thr Gln Phe Ala Pro

          550        560        570        580        590        600
          *          *          *          *          *          *
GGG AAA AAC GTC GAA CAG GTA GAA GAA AAG CTA CTG AAA GTG GTT CCA GCA GAG TTT AAA
Gly Lys Asn Val Glu Gln Val Glu Glu Lys Leu Leu Lys Val Val Pro Ala Glu Phe Lys

          610        620        630        640        650        660
          *          *          *          *          *          *
GTC GAC TGC CAC CAT TGG TTG ATC CTG CAC GGG CGT TAT ACC TGC ATT GCC CGC AAG CCC
Val Asp Cys His His Trp Leu Ile Leu His Gly Arg Tyr Thr Cys Ile Ala Arg Lys Pro

          670        680        690        700        710        720
          *          *          *          *          *          *
CGC TGT GGC TCT TGT ATT ATT GAA GAT CTT TGT GAA TAC AAA GAG AAA GTT GAC ATC TGA
Arg Cys Gly Ser Cys Ile Ile Glu Asp Leu Cys Glu Tyr Lys Glu Lys Val Asp Ile ---

          730        740        750        760        770        780
          *          *          *          *          *          *
AGA AAA GGG GTA ACA CCG ATT ACC CCA TTG ATA ACC TTT CTT TAT CCT CTT TTA AAA CAT
```

FIGURE 2: Nucleotide sequence of the *n*th gene and deduced amino acid sequence of endonuclease III. The DNA sequence of the antisense strand is shown; numbering is from the 5′ end. The proposed ribosome binding site is underlined and labeled. A region of dyad and inverted symmetry capable of forming a stem and loop structure is overlined with the center of symmetry indicated by a dot. The underlined amino acids are those confirmed by protein sequence analysis of endonuclease III.

with the experimentally determined value of approximately 10 (Radman, 1976). The gene product has an average hydropathy of −0.08 which suggests it is slightly more hydrophobic than the average soluble protein (Kyte & Doolittle, 1982).

A computer search using the algorithm of Lipman and Pearson (1985) did not reveal any proteins with extensive similarity to endonuclease III in the National Biochemical Research Foundation protein sequence data library release

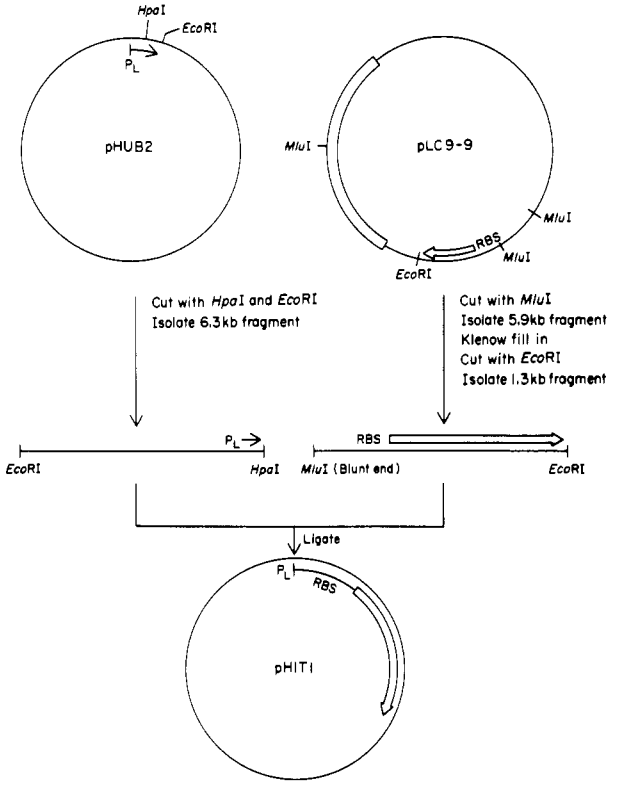

FIGURE 3: Cloning strategy for the endonuclease III gene. The endonuclease III gene (open arrow) with its ribosome binding site (RBS) was isolated from pLC9-9. The λ leftward promoter (P_L) was from plasmid pHUB2. When joined, these elements allow highly efficient transcription and translation of the endonuclease III gene.

10.0. We have also compared the amino acid sequence of endonuclease III to the sequences of exonuclease III (Saporito et al., 1988), endonuclease IV (Saporito & Cunningham, 1988), 3-methyladenine–DNA glycosylase I (Nakabeppu et al., 1984), 3-methyladenine–DNA glycosylase II (Kondo et al., 1986), and (formamido)pyrimidine–DNA glycosylase (Boiteux et al., 1987) of *E. coli* and endonuclease V–pyrimidine dimer–DNA glycosylase of bacteriophage T4 (Valerie et al., 1984; Radany et al., 1984) by the method of Smith (1981) and found no similarities of more than nine amino acids in which there were six matches and three mismatches with three related amino acids.

Cloning the nth Gene on a λP_L Expression Vector. The sequence of the *n*th gene showed a ribosome binding site upstream of the structural gene. The entire gene was on a 1.3 kbp *Mlu*I–*Eco*RI fragment which could be cloned into the plasmid pHUB2 which carries the λP_L promoter. The construction of a plasmid bearing the *n*th gene under control of λP_L is diagrammed in Figure 3. We started with a 5.9 kbp *Mlu*I fragment of pLC9-9 containing the *n*th gene. The ends of this fragment were filled in by using the Klenow fragment of *E. coli* DNA polymerase I, and the fragment was then cut with *Eco*RI. A 1.3 kbp fragment was isolated. This fragment had one blunt end and one *Eco*RI end and could be ligated into pHUB2 cut with *Hpa*I and *Eco*RI in only one orientation, which should fuse the *n*th gene and ribosome binding site to the λ promoter of pHUB2. The plasmid, pHIT1, resulting from the construction was transformed into an *E. coli* strain carrying a deleted λ prophage which expressed the λ_{CI857} gene product. After temperature induction, we measured a 300-fold overproduction of endonuclease III activity.

Purification of Endonuclease III. The purification of endonuclease III is summarized in Table I. We noticed that S-Sepharose fractions containing endonuclease activity had

Table I: Purification of *Escherichia coli* Endonuclease III

fraction	description	total protein (mg)	sp act. (units mg^{-1})	yield (%)
I	crude extract	13860	400	100
II	poly(ethylenimine) supernatant	7600	800	110
III	S-Sepharose peak fractions	274	18200	90
IV	phenyl-Sepharose peak fractions	117	33300	70

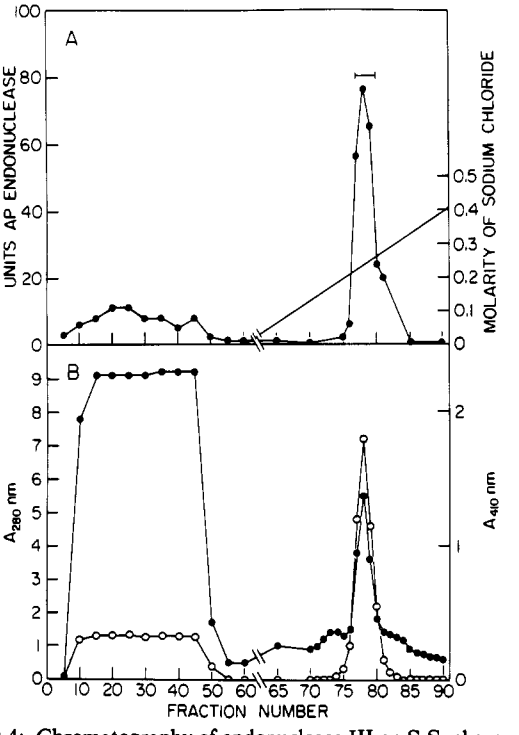

FIGURE 4: Chromatography of endonuclease III on S-Sepharose. (A) One-tenth of a nanoliter of the indicated fractions was assayed for AP endonuclease III under standard conditions. (●) AP endonuclease activity. The straight line indicates the concentration gradient of sodium chloride. The bar indicates the fractions pooled for fraction III. (B) (●) A_{280nm}; (○) A_{410nm}.

a yellow color, and we determined that these fractions had an optical absorption peak at 410 nm. Figure 4 shows an elution profile from S-Sepharose in which the endonuclease activity and the material absorbing at 410 nm coelute. The glycosylase activity of endonuclease III as measured by cleavage of UV-irradiated pBR322 also elutes at this position (data not shown). This purification step yields a 22-fold purification. We used hydrophobic interaction chromatography as a second chromatographic step. Figure 5 shows that the AP endonuclease activity and the material absorbing at 410 nm coelute from a phenyl-Sepharose column. The glycosylase activity of endonuclease III also elutes at this position (data not shown). The pooled peak fractions from the column (fraction IV) appear to contain a homogeneous preparation of endonuclease III since only a single protein band can be visualized by either silver staining (not shown) or Coomassie blue staining (Figure 6) of SDS–polyacrylamide gels.

Characterization of Endonuclease III. Since we purified endonuclease III with a different purification scheme than those used previously, we characterized several properties of the enzyme to ascertain that it was the enzyme previously described as endonuclease III. We used purified enzyme to determine the denatured molecular weight, the native molecular weight, and the sedimentation coefficient (Table II). The sedimentation coefficient was calculated according to

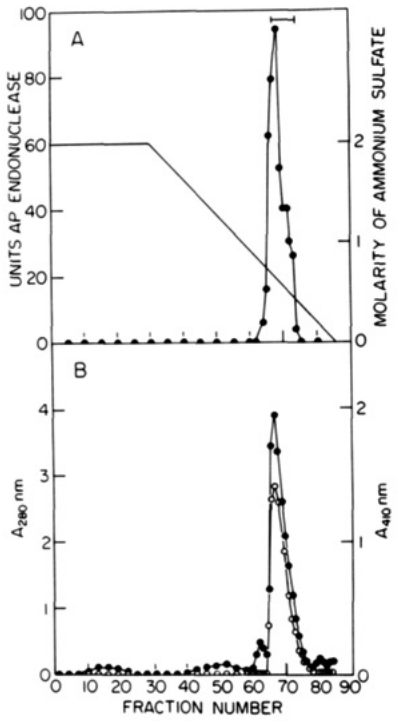

FIGURE 5: Chromatography of endonuclease III on phenyl-Sepharose. (A) One-tenth of a nanoliter of the indicated fractions was assayed for AP endonuclease activity under standard conditions. (●) AP endonuclease activity. The straight line indicates the concentration gradient of ammonium sulfate. The bar indicates the fractions pooled for fraction IV. (B) (●) A_{280nm}; (○) A_{410nm}.

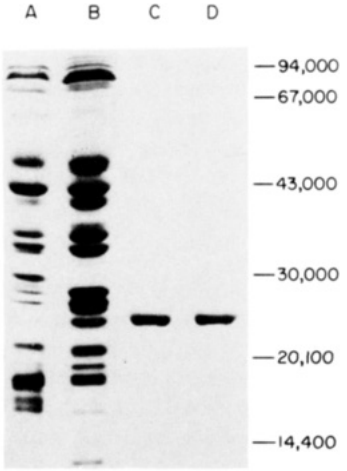

FIGURE 6: Sodium dodecyl sulfate–polyacrylamide gel of endonuclease III fractions. (Lane A) Crude extract (I); (lane B) poly(ethylenimine) supernatant (II); (lane C) S-Sepharose (III); (lane D) phenyl-Sepharose (IV). The position of molecular weight markers is indicated by arrows to the right of lane D.

Table II: Physical Parameters of Endonuclease III

parameter	value
Stokes radius, a (nm), gel filtration (Ackers, 1964)	2.15
sedimentation coefficient, $s_{20,w}$ (S), sucrose gradients (Martin & Ames, 1961)	2.65
frictional ratio, f/f_0, from a and M_r (Siegel & Monty, 1966)	1.13
molecular weight, M_r	
from K_{av}	26300
from a and s (Siegel & Monty, 1966)	23500
from SDS–PAGE (Laemmli, 1970; Weber & Osborn, 1969)	27300
partial specific volume, $\bar{v}$ ($cm^3\ g^{-1}$), from amino acid composition	0.736

Martin and Ames (1961) and is based on the assumption that the protein has a partial specific volume of 0.725 $cm^3\ g^{-1}$. We have calculated the partial specific volume of endonuclease III to be 0.736 $cm^3\ g^{-1}$ based on the amino acid composition. This difference results in less than 1% error in the $s_{20,w}$ of the enzyme. The values reported in Table II agree quite closely with those previously reported for endonuclease III (Radman, 1976; Gates & Linn, 1977; Breimer & Lindahl, 1984). In addition, we determined a Stokes radius of 2.15 nm and a frictional coefficient of 1.13 for the enzyme (Table II).

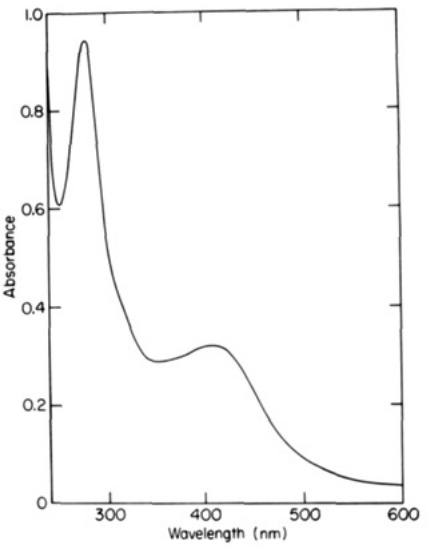

FIGURE 7: Absorption spectrum of endonuclease III. The sample was 0.5 mg mL^{-1} in 100 mM potassium phosphate, pH 6.6.

Table III: Amino Acid Composition of *E. coli* Endonuclease III

amino acid	composition from amino acid analysis[a]	composition from DNA sequence	amino acid	composition from amino acid analysis[a]	composition from DNA sequence
Ala	19	19	Ile	12	14
Arg	11	11	Leu	21	22
Asn		13	Lys	15	16
Asp		6	Met	2	2
Asx	17	(19)[b]	Phe	6	6
Cys	ND[c]	7	Pro	10	11
Gln		4	Ser	5	6
Glu		18	Thr	13	15
Glx	22	(22)[d]	Trp	ND[c]	2
Gly	10	10	Tyr	2	5
His	6	6	Val	16	18

[a] Rounded to nearest interger. [b] Sum of Asn and Asx. [c] ND, not determined. [d] Sum of Gln and Glu.

We also characterized the cleavage of UV-irradiated DNA by purified endonuclease III. Using restriction fragments of defined sequence, we determined that our purified enzyme had the same substrate specificity as reported previously for a less purified preparation of endonuclease III (Doetsch et al., 1987) (data not shown).

The optical spectrum of endonuclease III is given in Figure 7. The protein has two absorption peaks at 280 and 410 nm. The nature of the chromophore absorbing at 410 nm is described in the following paper (Cunningham et al., 1989).

The nth Gene Is the Structural Gene for Endonuclease III. The *n*th gene was defined by its ability to overproduce endonuclease III when present on a multicopy plasmid and by the deficiency in endonuclease III activity when mutated (Cunningham & Weiss, 1985). The *n*th gene could be the structural gene for endonuclease III or it could code for a regulatory factor necessary for the expression of endonuclease III. To determine if the identified open reading frame for the *n*th gene is the structural gene for endonuclease III, we determined the

N-terminal amino acid sequence of endonuclease III and its amino acid composition. The sequence of the first 10 amino acids (underlined in Figure 2) is in complete agreement with the predicted N-terminal sequence derived from the DNA sequence of the *nth* gene. Table III shows the excellent agreement between the amino acid composition of endonuclease III from purified protein and the composition predicted from the DNA sequence. These results establish the identified open reading frame as the structural gene for endonuclease III.

ACKNOWLEDGMENTS

We thank Brian J. Smith-White of Upjohn Laboratories for the determination of the N-terminal amino acid sequence and the amino acid composition of endonuclease III and Linda P. Welch for preparation of the manuscript.

REFERENCES

Ackers, G. W. (1964) *Biochemistry 3*, 723–730.
Bachetti, S., & Benne, R. (1975) *Biochim. Biophys. Acta 390*, 285–297.
Bailly, V., & Verly, W. G. (1987) *Biochem. J. 242*, 565–572.
Bernard, H.-U., & Helinski, D. R. (1979) *Methods Enzymol. 68*, 482–492.
Birnboim, H. C., & Doly, J. (1979) *Nucleic Acids Res. 7*, 1513–1523.
Boiteux, S., O'Connor, T. R., & Laval, J. (1987) *EMBO J. 10*, 3177–3183.
Bradford, M. M. (1976) *Anal. Biochem. 72*, 248–254.
Breimer, L. H. (1983) *Biochemistry 22*, 4192–4197.
Breimer, L. H. (1986) *Biochem. Biophys. Res. Commun. 134*, 201–204.
Breimer, L. H., & Lindahl, T. (1980) *Nucleic Acids Res. 8*, 6199–6210.
Breimer, L. H., & Lindahl, T. (1984) *J. Biol. Chem. 259*, 5543–5548.
Breimer, L. H., & Lindahl, T. (1985) *Biochemistry 24*, 4018–4022.
Brent, T. P. (1983) *Biochemistry 22*, 4507–4512.
Clark, J. M., & Beardsley, G. P. (1986) *Nucleic Acids Res. 14*, 737–749.
Clarke, L., & Carbon, J. (1976) *Cell 9*, 91–99.
Cunningham, R. P., & Weiss, B. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 474–478.
Cunningham, R. P., Asahara, H., Bank, J. F., Scholes, C. P., Salerno, J. C., Surerus, K., Münck, E., McCracken, J., Peisach, J., & Emptage, M. H. (1989) *Biochemistry* (following paper in this issue).
Demple, B., & Linn, S. (1980) *Nature (London) 287*, 203–208.
Doetsch, P. W., Helland, D. E., & Haseltine, W. A. (1986) *Biochemistry 25*, 2212–2220.
Doetsch, P. W., Henner, W. D., Cunningham, R. P., Toney, J. H., & Helland, D. E. (1987) *Mol. Cell. Biol. 7*, 26–32.
Gates, F. T., III, & Linn, S. (1977) *J. Biol. Chem. 252*, 2802–2807.
Gossett, J., Lee, K., Cunningham, R. P., & Doetsch, P. W. (1988) *Biochemistry 27*, 2629–2634.
Hawley, D., & McClure, W. (1983) *Nucleic Acids Res. 11*, 2237–2256.
Helland, D. E., Doetsch, P. W., & Haseltine, W. A. (1986) *Mol. Cell. Biol. 6*, 1983–1990.
Higgins, S. A., Frenkel, K., Cummings, A., & Teebor, G. W. (1987) *Biochemistry 26*, 1683–1688.
Ide, H., Kow, Y. W., & Wallace, S. S. (1985) *Nucleic Acids Res. 13*, 8035–8052.
Jorgenson, T. J., Kow, Y. W., Wallace, S. S., & Henner, W. D. (1987) *Biochemistry 26*, 6436–6443.
Katcher, H. L., & Wallace, S. S. (1983) *Biochemistry 22*, 4071–4081.
Kondo, H., Nakabeppu, Y., Kataska, H., Khara, S., Kawabata, S., & Sekiguchi, M. (1986) *J. Biol. Chem. 261*, 15772–15777.
Kow, Y. W., & Wallace, S. S. (1987) *Biochemistry 26*, 8200–8206.
Kuhnlein, U., Lee, B., Penhoet, E. E., & Linn, S. (1978) *Nucleic Acids Res. 5*, 951–960.
Kyte, J., & Doolittle, R. F. (1982) *J. Mol. Biol. 157*, 105–132.
Laemmli, U. K. (1970) *Nature (London) 270*, 680–685.
Leadon, S. A., & Hanawalt, P. C. (1983) *Mutat. Res. 112*, 191–200.
Lindahl, T., & Andersson, A. (1972) *Biochemistry 11*, 3618–3623.
Lipman, D. J., & Pearson, W. R. (1985) *Science 227*, 1435–1441.
Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) *Molecular cloning: a laboratory manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Martin, R. G., & Ames, B. N. (1961) *J. Biol. Chem. 236*, 1372–1379.
Mosbaugh, D. W., & Linn, S. (1982) *J. Biol. Chem. 257*, 575–583.
Nakkabeppu, Y., Miyata, T., Kondo, H., Iwanaga, S., & Sekiguchi, M. (1984) *J. Biol. Chem. 259*, 13730–13736.
Radany, E. H., Naumovski, L., Love, J. D., Gutekunst, K. A., Hall, D. W., & Friedberg, E. C. (1984) *J. Virol. 52*, 846–856.
Radman, M. (1976) *J. Biol. Chem. 251*, 1438–1445.
Rodriguez, R. L., & Tait, R. (1983) *Recombinant DNA techniques*, Addison-Wesley Co., New York.
Rouet, P., & Essigman, J. M. (1985) *Cancer Res. 45*, 6113–6118.
Sanger, F. S., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.
Saporito, S. M., & Cunningham, R. P. (1988) *J. Bacteriol. 170*, 5141–5145.
Saporito, S. M., Smith-White, B. J., & Cunningham, R. P. (1988) *J. Bacteriol. 170*, 4542–4547.
Shine, J., & Dalgarno, C. (1974) *Proc. Natl. Acad. Sci. U.S.A. 71*, 1342–1346.
Siegel, L. M., & Monty, K. J. (1966) *Biochim. Biophys. Acta 112*, 346–362.
Smith, T. F. (1981) *J. Mol. Biol. 146*, 195–197.
Tabor, S., & Richardson, C. C. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 4767–4771.
Tinoco, I., Jr., Borer, J. N., Dengler, B., Levine, M. D., Uhlenbeck, O. C., Crothers, D. M., & Gralla, J. (1973) *Nature (London), New Biol. 246*, 40–41.
Valerie, K., Henderson, E. H., & deRiel, J. K. (1984) *Nucleic Acids Res. 12*, 8085–8096.
VanLacker, J. L., & Tomura, T. (1974) *Biochim. Biophys. Acta 353*, 99–114.
Warner, H. R., Demple, B. F., Deutsch, W. A., Kane, C. M., & Linn, S. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 4602–4606.
Weber, K., & Osborn, M. (1969) *J. Biol. Chem. 244*, 4406–4412.
Weiss, R. B., & Duker, N. J. (1986) *Nucleic Acids Res. 14*, 6621–6631.
Weiss, R. B., & Duker, N. J. (1987) *Photochem. Photobiol. 45*, 763–768.