Volume 17 Number 21 1989 Nucleic Acids Research

Overproduction and crystallization of *FokI* restriction endonuclease

Keiko Kita, Hirokazu Kotani, Nobutsugu Hiraoka, Teruya Nakamura and Kazuo Yonaha[1]

Central Research Laboratories, Takara Shuzo Co., Ltd, Otsu 3-4-1, Shiga 520-21 and [1]Department of Agricultural Chemistry, University of the Ryukyus, Nishihara, Okinawa 903-01, Japan

Received August 22, 1989; Revised and Accepted September 20, 1989

ABSTRACT

To overproduce FokI endonuclease (R.FokI) in an Escherichia coli system, the coding region of R.FokI predicted from the nucleotide sequence was generated from the FokI operon and joined to the tac promoter of an expression vector, pKK223-3. By introduction of the plasmid into E. coli UT481 cells expressing the FokI methylase gene, the R.FokI activity was overproduced about 30-fold, from which R.FokI was purified in amounts sufficient for crystallization. The removal of a stem-loop structure immediately upstream of the R.FokI coding region was essential for overproduction.

INTRODUCTION

FokI endonuclease (R.FokI), classified as a type IIS restriction endonuclease, recognizes an asymmetric pentanucleotide, and cleaves outside its recognition sequence as indicated by the arrows here: $^{5'}$GGATG(N)$_{9}\downarrow^{3'}$ (1). Because of its $_{3'}$CCTAC(N)$_{13}\uparrow_{5'}$ unusual cleavage specificity some workers have suggested that it can be used differently from other type II enzymes. Unique 5'-protruding ends liberated by digestion with R.FokI permit the efficient simultaneous ligation of several restriction fragments. Some kinds of vector were constructed for this purpose: gene synthesis and DNA segment replacement with the use of such vectors have been reported (2,3). R.FokI cleaves at a set distance from the recognition sequence. With the use of R.FokI and the adaptor-primer that contains the R.FokI recognition sequence and the ssDNA moiety complementary to the targe. DNA, ssDNA and dsDNA can be cleaved at any desired position (4,5).

We have isolated the genes for the FokI restriction-modification system and determined the complete nucleotide sequence (6). The system form an operon in which

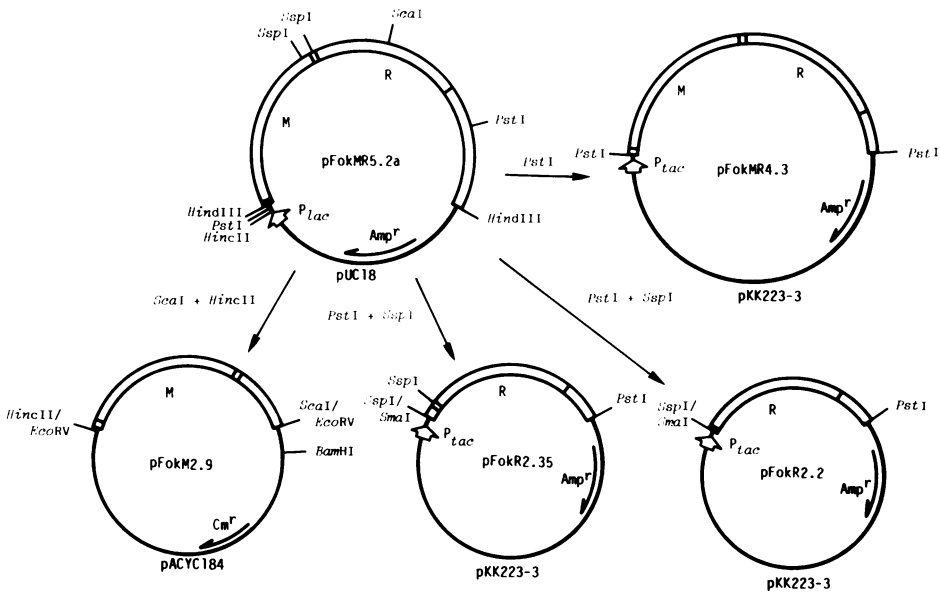

Figure 1 Structures of plasmids

pFokMR4.3 was constructed by insertion of the 4.3-kb PstI fragment that carried the genes for FokI methylase and endonuclease into the PstI site of pKK223-3. pFoKM2.9 carried the FokI methylase gene inserted into the EcoRV site of pACYC184. pFokR2.35 and pFokR2.2 were constructed by replacement of a small SmaI-PstI fragment of pKK223-3 by the SspI-PstI fragments (2.35 kb and 2.2 kb, respectively) carrying the FokI endonuclease gene. The open box and the line show the insert and the vector moiety, respectively. Symbols: Amp^r, ß-lactamase gene; Cm^r, chloramphenicol acetyltransferase gene; R, FokI endonuclease gene; M, FokI methylase gene; P_{lac}, lac promoter; P_{tac}, tac promoter.

R.FokI gene is downstream from the FokI methylase (M.FokI) gene. The expression of the system is regulated by a promoter upstream of the M.FokI gene. We have constructed a plasmid, pFokMR5.2a, that carries the FokI restriction-modification gene inserted downstream of the lac promoter, and purified R.FokI from E. coli cells carrying the plasmid. The active form of R.FokI is monomeric and the molecular weight is 66,000. The amino acids of the N-terminal have not yet been sequenced.

Here, we constructed an overproducing strain and purified enough R.FokI to crystallize it. The N-terminal amino acid sequence of R.FokI was analyzed.

MATERIALS AND METHODS

Plasmids and strain

The E. coli host strains, JM109 (7) and UT481 [lon, Δ(lac pro), thyA, met, supD, rm/F'traD36, proAB, lacI^q, ZΔM15] (8), were from our collection. Plasmids were obtained as follows: pKK223-3 (carries a tac promoter) was from Pharmacia P-L Biochemicals, and pACYC184 (9,10) and pFokMR5.2a (carries genes for FokI restriction and modification (6)) were from our collection. Plasmids pFokMR4.3, pFokM2.9, pFokR2.35, and pFokR2.2 were constructed from pFokMR5.2a for this experiment. Their structures are given in Fig. 1.

Enzymes and chemicals

Restriction endonucleases and T4 DNA ligase were products of Takara Shuzo Co., Ltd., and were used according to the manufacturer's instructions. Protein standards for SDS-polyacrylamide gel electrophoresis (SDS-PAGE), the protein assays, and isoelectric focusing were from Bio-Rad and SDS-polyacrylamide gradient gels were from Daiichi Pure Chemicals.

Enzyme activity

R.FokI activity in the crude extract was measured as follows. Cells (JM109[pFokMR5.2a], JM109[pFokM2.9, pFokR2.2], and UT481[pFokM2.9, pFokR2.2]) were treated with lysozyme and disrupted by sonication. Debris was removed by centrifugation (10^5 x g for 1 hr) and the supernatant was assayed.

R.FokI was diluted with 10 mM potassium phosphate buffer, pH 7.5, 10 mM 2-mercaptoethanol (ME), and 5% glycerol, and its activity was assayed by incubation for 1 hr at 37°C in 50 μl of a reaction mixture containing 10 mM Tris-HCl, pH 7.5, 7 mM $MgCl_2$, 7 mM ME, 60 mM NaCl, 0.01% bovine serum albumin (BSA), and 1 μg of λ-DNA. One unit is defined as the amount of enzyme that cleaves 1 μg of λ-DNA in 1 hr under these conditions.

Growth of cells

For protein and activity analysis of the E. coli cells carrying overproducing plasmids, cells were grown at 37°C in 50 ml of LB-medium (1% tryptone, 0.5% yeast extract, and 1% NaCl) containing 100 μg/ml ampicillin, 30 μg/ml chloramphenicol, or both. When the cell density reached A_{600} = 0.6,

isopropyl-ß-D-thiogalactopyranoside (IPTG) was added to the concentration of 2 mM and incubation was continued at 37°C for 3 hr.

For purification of R.FokI, cells of E. coli UT481[pFokM2.9, pFokR2.2] were grown aerobically at 37°C in 40 l of LB-medium containing 100 µg/ml ampicillin and 30 µg/ml chloramphenicol. When the cell density reached A_{600} = 0.6, IPTG was added to 0.2 mM. After 3 hr of induction at 37°C, cells (120 g) were harvested, washed with 0.15 M NaCl in 10 mM Tris-HCl, pH 7.5, and stored frozen.

Purification of R.FokI

From 120 g of cells, R.FokI was purified as described elsewhere (6). The cell-free extract was treated with polyethyleneimine and chromatographed on a P-cellulose and then a DEAE-cellulose column. The activity, recovered in the flow-through fraction, was adsorbed to a hydroxylapatite column (Clarkson) and eluted with a linear gradient of potassium phosphate buffer, pH 7.5 (0.1 to 0.5 M). The active fraction obtained was further fractionated on a heparin-Sepharose column (Pharmacia) and eluted with a linear gradient of KCl (0 to 1 M). The activity was eluted at 0.22-0.32 M KCl.

Analysis of N-terminal amino acid sequences

R.FokI from E. coli JM109[pFokMR5.2a] was blotted from SDS-polyacrylamide gel to Polyvinylidene difluoride membranes (Millipore) by the method of Matsudaira (11). R.FokI from both Flavobacterium okeanokoites and E. coli UT481[pFokM2.9, pFokR2.2] was desalted and analyzed. The protein was degraded sequentially with a protein analyzer (Applied Biosystems model 470A) equipped with an on-line HPLC apparatus (model 120A).

Isoelectric focusing

Isoelectric focusing was performed at 15°C on 1% agarose gels containing 4% ampholytes (Bio-Rad, Biolyte 8/10) in a Resolute HMP Chamber (FMC Corp.) by the procedure recommended by this manufacturer. The enzyme was desalted and samples of about 18 µg were put onto the gel with marker proteins. The samples were focused at a constant voltage of 500 V for 2 hr. The gel containing the marker proteins and samples was fixed with

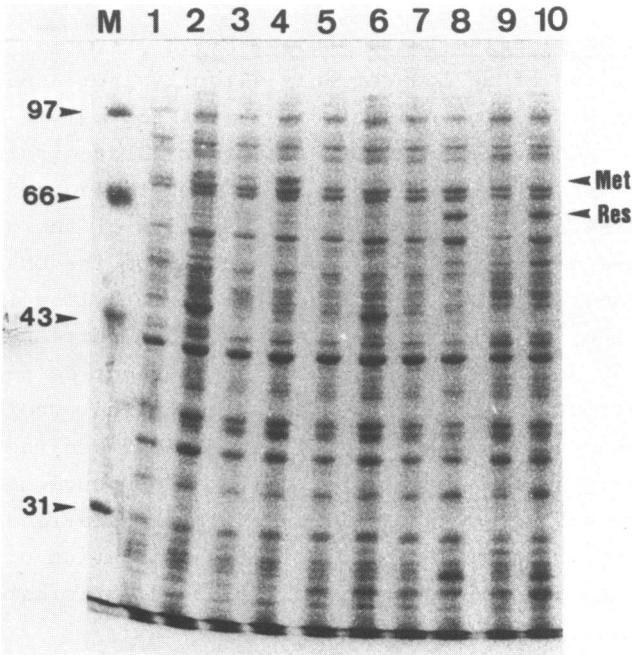

Figure 2 Overexpression of endonuclease in E. coli cells
Cells carrying overproducing plasmids were grown as described in Materials and Methods. Total proteins were extracted from cells harvested before induction (lanes 1, 3, 5, 7, and 9) and after (lanes 2, 4, 6, 8, and 10). Extract from an equivalent number of cells was put onto a lane, separated by electrophoresis on SDS-polyacrylamide gel (10%), and stained with Coomassie brilliant blue. By the side of the marker (M) lane, the molecular weights (x 10^{-3}) of the markers are given. The protein bands corresponding to the FokI methylase (Met) and the FokI endonuclease (Res) are indicated by arrows.

trichloroacetic acid and stained with Coomassie brilliant blue. The remainder of the gel was cut into 30 segments. Slices of gels were placed in 2 ml of distilled water overnight, and the pH of the water was measured at 15°C.

Other methods

Protein was assayed by the method of Bradford (12) with use of a kit from Bio-Rad. SDS-PAGE was done by the method of Laemmli (13). DNA ligation, transformation, and plasmid isolation were done as described by Maniatis et al. (14).

RESULTS

Construction of plasmids overproducing R.FokI

The addition of an inducer only slightly increased the production of R.FokI in cells of JM109[pFokMR5.2a], in which FokI methylase and endonuclease genes were under the control of the lac promoter (Fig. 2, lanes 1 and 2). To overproduce R.FokI, the plasmid pFokMR4.3 was constructed by insertion of the genes for FokI methylase and endonuclease downstream from the DNA fragment containing the tac promoter. Production of methylase increased after induction, but the yield of endonuclease increased slightly (Fig. 2, lanes 3 and 4). Analysis of the nucleotide sequence of the FokI restriction-modification gene showed that there was a stem-loop structure followed by a T cluster between the end of the methylase gene and the beginning of the endonuclease gene (Fig. 3). We suspected that this putative translational termination structure interfered with the initiation of translation of the endonuclease, so a plasmid deficient in that structure was constructed.

We devised a two-plasmid system in E. coli with use of the compatible plasmids pKK223-3 and pACYC184. Plasmid pFokM2.9 carries the FokI methylase gene inserted into the EcoRV site of plasmid pACYC184. The methylase was expressed constitutively under the control of its own promoter. pFokR2.2 carries the endonuclease gene without the stem-loop structure, and pFokR2.35 carries the endonuclease gene and the complete stem-loop structure under the control of the tac promoter. The plasmids constructed were introduced into E. coli cells, proteins were separated by SDS-PAGE, and the R.FokI activity was measured. JM109[pFokM2.9, pFokR2.35] production was on the same level as that of JM109[pFokMR5.2a] and JM109[pFokMR4.3]; the production of endonuclease increased greatly in cells carrying pFokM2.9 and pFokR2.2 (Fig. 2, lanes 5-10). Enzyme activity in the crude extract was compared. In JM109[pFokM2.9, pFokR2.2] cells, R.FokI activity was 9-fold that in JM109[pFokMR5.2a] cells, and in UT481 [pFokM2.9, pFokR2.2] cells, it was 30-fold.

Induction of R.FokI activity

The effect ofthe concentrations of IPTG on the induction of R.FokI activity was examined. UT481[pFokM2.9, pFokR2.2] cells

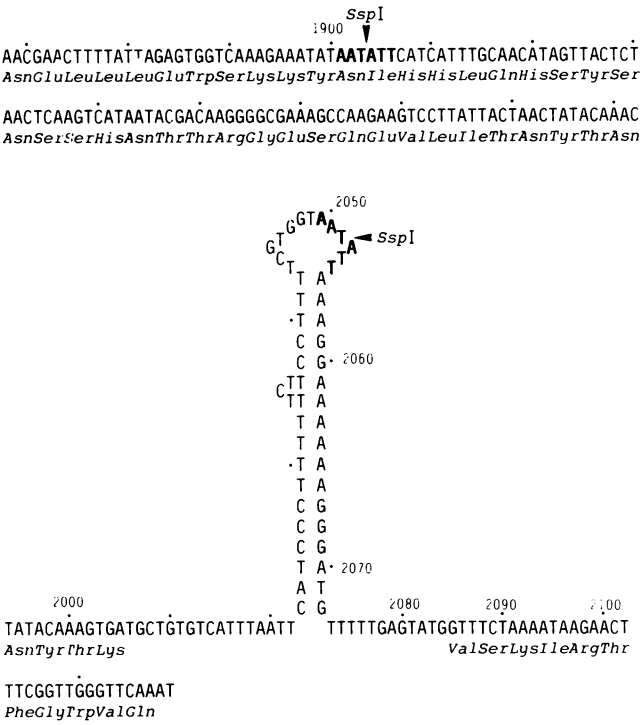

Figure 3 Nucleotide sequence and putative secondary structure upstream of the FokI endonuclease gene

Nucleotide number was the same as indicated in the previous paper (6). Two SspI sites, which was used for generation of the DNA fragment containing the R.FokI gene, indicated as bold characters.

were grown to A_{600} = 0.6 and treated with 0.2 or 2 mM IPTG. After induction, some of the cells were harvested every hour and their endonuclease production was analyzed by SDS-PAGE. With either concentration of IPTG, production was maximum after 3 hr of induction, and there was little difference in the level of production with different concentrations. The growth of cells carrying both pFokM2.9 and pFokR2.2 was compared with the growth of those carrying no plasmid. Inhibition was slight in the recombinant cells. JM109[pFokR2.2] cells without pFokM2.9 were not viable.

Purification and crystallization of R.FokI

Table I summarizes the purification scheme by which 300 mg

Table I Purification of FokI endonuclease

Steps	Total activity (units)	Total protein (mg)	Specific activity (units/mg)	Purification (fold)	Yield (%)
Crude extract	1,200,000	12,320	97	1	100
Phosphocellulose	990,000	498	1,988	20	82.5
DEAE-cellulose	500,000	234	2,137	22	41.7
Hydroxylapatite	1,485,000	313	4,744	49	123.8
Heparin-Sepharose	1,680,000	299	5,619	58	140

Proteins was assayed as described in the text with BSA as the standard.

of R.FokI was purified from 120 g of UT481[pFokM2.9, pFokR2.2] cells. Figure 4 shows the results of SDS-PAGE. The purified sample gave a single protein band. This fraction was dialyzed against distilled water, and put on a thin-layer agarose gel; then isoelectric focusing was done. A single band was observed and the pI value of the R.FokI was 8.9 (Fig. 5). The purified

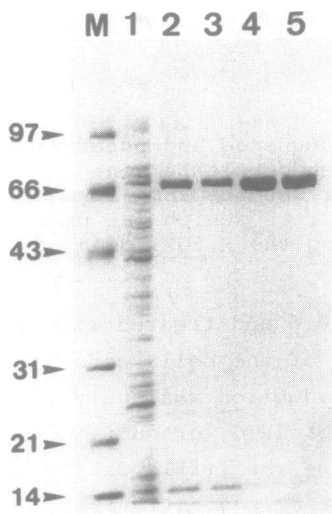

Figure 4 SDS-polyacrylamide gel electrophoresis of FokI endonuclease

Proteins were electrophoresed on a SDS-polyacrylamide gradient gel (10-20%) and stained with Coomassie brilliant blue. Lanes 1 to 5 correspond to fractions of crude extract, P-cellulose, DEAE-cellulose, hydroxylapatite, and heparin-Sepharose, respectively. By the side of the marker (M) lane, the moiecular weight (x 10^{-3}) of markers are given.

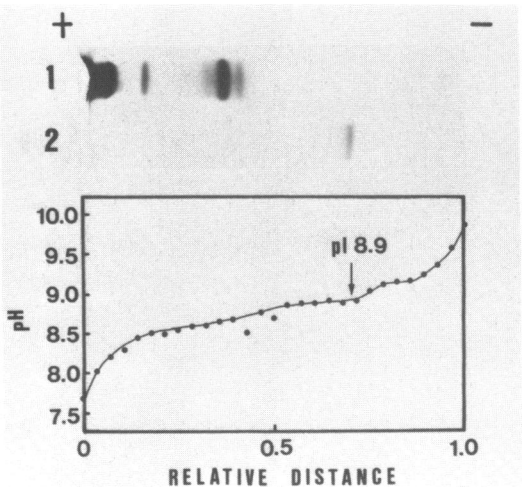

Figure 5 Isoelectric focusing of FokI endonuclease
Upper panel: Standard proteins (lane 1) and FokI endonuclease (lane 2) were focused on a thin-layer agarose plate and stained with Coomassie brilliant blue. Lower panel: Gels were sliced and pH values were measured. The arrow indicates the position taken to be the isoelectric pH.

enzyme was dissolved in 10 mM Tris-HCl, pH 7.5, containing 100 mM KCl, 10 mM ME, and 0.1 mM EDTA, and then ammonium sulfate powdered finely was gradually added to the enzyme solution (10.0 mg/ml) until a faint turbidity was obtained. The concentration of ammonium sulfate at the end was about 50% saturation. The pH of the solution was kept constant at about 7.2 with 1 N NaOH. On standing overnight at 4°C, crystal formation occurred. The crystal took a form of thin plates (Fig. 6)

N-terminal amino acid sequencing

The first eight amino acids of R.FokI at the final purification step were (Met,Val)-(Ser,Phe)-(Leu,Lys)-(Ile,Ser)-(Arg,Met)-(Thr,Val)-(Ser,Phe)-(Lys,Gly). This corresponded to the amino acid sequence, Met-Phe-Leu-Ser-Met-Val-Ser-Lys, that started at the nucleotide position 2070, and to the amino acid sequence, Val-Ser-Lys-Ile-Arg-Thr-Phe-Gly, that started at 2082; the ratio of the production of the two proteins was 1:2-3. The N-terminal amino acid sequence of R.FokI purified from E. coli JM109[pFokMR5.2a] and from F. okeanokoites was also analyzed. The first ten amino acids were X-X-X-Ile-Arg-Thr-Phe-Gly-X-Val (X

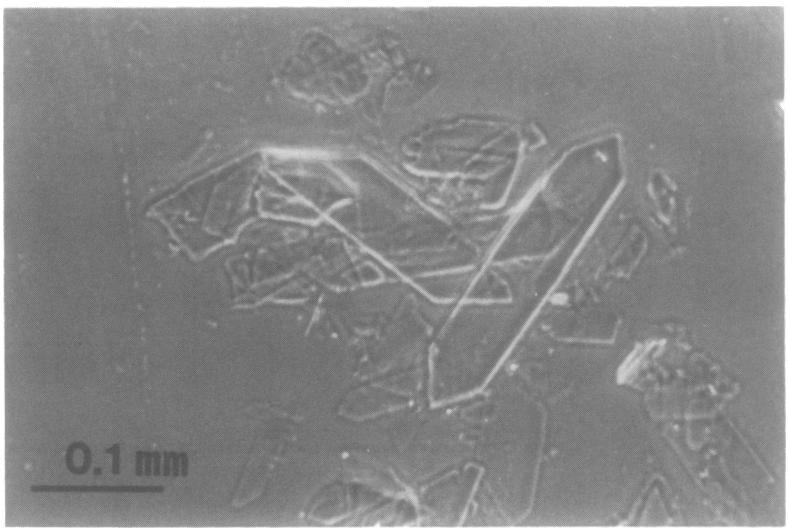

Figure 6 Crystals of FokI endonuclease

not identified) and Val-Ser-Lys-Ile-Arg-Thr-Phe-Gly-Trp-Val, respectively. These corresponded to the amino acid sequence that started at nucleotide position 2082.

Properties of the enzyme

Of the conditions tested, enzyme activity was maximum at 37°C in 10 mM Tris-HCl, pH 8.0, 5 mM $MgCl_2$, 60 mM NaCl, and 0.01% BSA. The addition of 7 mM ME or 0.01% Triton X-100 did not increase the activity. Without BSA, activity decreased to 40%. Activity was greatest at 37-42°C. The pH optimum was 7.5-8.5. R.FokI was incubated in 10 mM potassium phosphate buffer, pH 7.5, 10 mM ME, and 5% glycerol at different temperatures, and the activity remainings was assayed at different times. Incubation at 42°C for 1 hr caused no loss of activity, but after incubation at 50°C for 10 min, the activity decreased to 3%. There was no loss of activity after storage at 20°C for 1 week or at 4°C for 2 months.

DISCUSSION

We assumed that the weak translation of the endonuclease gene in the cloned FokI restriction-modification gene was caused by the stem-loop structure upstream of the endonuclease gene.

Weak translation caused by mRNA folding has been reported for other restriction-modification systems. A stem-loop structure downstream of the endonuclease gene seems to interfere with EcoRV translation initiation (15,16). A 14-bp hairpin structure within the endonuclease coding sequence probably attenuates TaqI transcription or translation (17). In the TaqI system, overproduction was greatest when the hairpin structure was removed. Here, removal of the stem-loop structure increased the translation efficiency of the R.FokI gene, and the R.FokI gene was expressed efficiently under the control of the tac promoter.

Possible explanations for the high yield (140%) by our purification for R.FokI include the use of the lon$^-$ E. coli strain as the host cells and the removal of substances that interfere with enzyme activity by hydroxylapatite chromatography. The pI of R.FokI (8.9) was the highest of several restriction endonucleases (EcoRI, 6.4 (18), 6.3 (19); RsrI (18), 7.0; and AatII, 5.5 (20)). The high pI explains the recovery of R.FokI activity in the pass-through fraction of DEAE-cellulose. An abundance of highly basic amino acids in R.FokI might account for this high pI, so we investigated the proportions of basic and acidic amino acids. EcoRI, for which the amino acid composition and the pI have been published, was studied in the same way. In R.FokI, 16.5% of the amino acid residues were acidic and 12.4% were basic; in EcoRI, these figures were 14.8% and 12.6%, respectively. The difference was not significant.

The purified enzyme gave a single band on SDS-PAGE and isoelectric focusing gel electrophoresis, but analysis of the N-terminal amino acid sequence showed that the sample was a mixture of two proteins. One started at valine and the other contained an additional five amino acids at the N-terminal. The N-terminal amino acid of the endonucleases purified from F. okeanokoites and from E. coli JM109[pFokMR5.2a] was valine. The protein that has an additional five amino acids in the N-terminal and that was produced in the overproducing strain was probably an artifact of the gene manipulation. The following could explain the production of the artifact. In the FokI restriction-modification operon R.FokI is translated from ATG at 2082 and is processed to become mature R.FokI by removal of the

methionine at the N-terminal. Enzymatic activity that removes the N-terminal residue of methionine when it precedes valine has been observed in both prokaryotes and eukaryotes (21). In the overproducing strain, a stable secondary structure around 2070 is destroyed and the ATG at 2070 can be used as an initiation site in addition to the ATG at 2082. Initiation codons at 2070 and 2082 are accompanied by putative ribosomal-binding sequences as follows: AAAGGAAAAAGGGATG, at 2070; and AAGGGATGTTTTTGAGTATG, at 2082 (the underlined nucleotides are complementary to 16s-rRNA). It is reasonable to think that in pFokR2.2, translation starts from the initiation codon at either 2070 or 2082. Although the protein translated from ATG at 2082 is processed as described above, newly synthesized protein is not processed and thus accumulates. The results obtained here suggest that the insertion of a strong promoter just upstream of ATG at 2082 would bring about overproduction of the homogeneous R.FokI protein that had valine at the N-terminal.

It was not known whether the protein that had an additional five amino acids at the N-terminal of R.FokI had the same catalytic properties as R.FokI or not. To check this, it would be necessary to separate the two proteins, but such separation would be difficult, because the differences in the molecular weights and isoelectric points are small.

Several restriction endonucleases were purified and crystallized from genetically engineered overproducing strains (15,17,22-27). Isolation of several hundred milligrams of R.FokI will make possible its physical characterization by X-ray diffraction.

REFERENCES

1. Sugisaki, H. and Kanazawa, S. (1981) Gene 16, 73-78.
2. Vermersch, P.S. and Bennett, G.N. (1987) Gene, 54, 229-238.
3. Mandecki, W. and Bolling, T.J. (1988) Gene, 68, 169-173.
4. Szybalski, W. (1985) Gene, 40, 169-173.
5. Kim, S.C., Podhajska, A.J., and Szybalski, W. (1988) Science, 240, 504-506.
6. Kita, K., Kotani, H., Sugisaki, H., and Takanami, M. (1989) J. Biol. Chem. 264, 5751-5756.
7. Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) Gene, 33, 103-119.
8. Ishino, Y., Shinagawa, H., Makino, K., Tsunasawa, S., Sakiyama, F. and Nakata, A. (1986) Mol. Gen. Genet. 204, 1-7.

9. Chang, A.C.Y. and Cohen, S.N. (1978) J. Bacteriol. 134, 1141-1156.
10. Rose, R.E. (1988) Nucleic Acids Res. 16, 355.
11. Matsudaira, P. (1987) J. Biol. Chem. 262, 10035-10038.
12. Bradford, M.M. (1976) Anal. Biochem. 72, 248-254.
13. Laemmli, U.K. (1970) Nature 227, 680-685.
14. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Lab.) Cold Spring Harbor, New York.
15. Bougueleret, L., Schwarzstein, M., Tsugita, A. and Zabeau, M. (1984) Nucleic Acids Res. 12, 3659-3676.
16. Bougueleret, L., Tenchini, M.L., Botterman, J. and Zabeau, M. (1985) Nucleic Acids Res., 13, 3823-3839.
17. Barany, F. (1988) Gene, 65, 167-177.
18. Aiken, C., and Gumport, R.I. (1988) Nucleic Acids Res. 16, 7901-7916.
19. Rubin, R.A. and Modrich, P. (1977) J. Biol. Chem. 252, 7265-7272.
20. Sato, H. and Yamada, Y. (1988) Abstracts of Papers, 110th Meeting of the Agricultural Chemical Society of Japan, Chubu Branch, Gifu, October, Abstract No. 10 (in Japanese).
21. Tsunasawa, S., Stewart, J.W. and Sherman, F. (1985) J. Biol. Chem. 260, 5382-5391.
22. Rosenberg, J.M., Boyer, H.W. and Greene, P.J. (1981) in Gene Amplification and Analysis (Chirikjian, J.G., ed.) vol. 1, pp. 131-164, Elsevier, N.Y.
23. Chen, S.-C., Kim, R., King, K., Kim, S.-H. and Modrich, P. (1984) J. Biol. Chem. 256, 11571-11575.
24. Botterman, J. and Zabeau, M. (1985) Gene 37, 229-239.
25. Kelly, S., Kaddurah-Daouk, R. and Smith, H.O. (1985) J. Biol. Chem. 260, 15339-15344.
26. Chandrasegaran, S., Smith, H.O., Amzel, M.L. and Ysern, X. (1986) Prot. Struct. Funct. Genet. 1, 263-266.
27. D'Arcy, A., Brown, R.S., Zabeau, M., van Resandt, R.W. and Winkler, F.K. (1985) J. Biol. Chem. 260, 1987-1990.