Arch. Pharm. Res. Vol.21, No.3, pp. 291-297, 1998

Purification and Characterization of Cop, a Protein Involved in the Copy Number Control of Plasmid pE194

Jin-Hwan Kwak[1], Jungho Kim[2], Mu-Yong Kim[3] and Eung-Chil Choi[4]

[1]School of Bioscience and Food Technology, Handong University, Pohang, [2]Department of Biochemistry, McGill University, Montreal, Canada, [3]LG Chem Research Park, LG Chemical Ltd., Taejon and [4]College of Pharmacy, Seoul National University, Seoul 151-742, Korea

(Received February 12, 1998)

Cop protein has been overexpressed in *Escherichia coli* using a T7 RNA polymerase system. Purification to apparent homogeneity was achieved by the sequential chromatography on ion exchange, affinity chromatography, and reverse phase high performance liquid chromatography system. The molecular weight of the purified Cop was estimated as 6.1 kDa by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). But the molecular mass of the native state Cop was shown to be 19 kDa by an analytical high performance size exclusion chromatography, suggesting a trimer-like structure in 50 mM Tris-HCl buffer (pH 7.5) containing 100 mM NaCl. Cop protein was calculated to contain 39.1% α-helix, 16.8% β-sheet, 17.4% turn, and 26.8% random structure. The DNA binding property of Cop protein expressed in *E. coli* was preserved during the expression and purification process. The isoelectric point of Cop was determined to be 9.0. The results of amino acid composition analysis and N-terminal amino acid sequencing of Cop showed that it has the same amino acid composition and N-terminal amino acid sequence as those deduced from its DNA sequence analysis, except for the partial removal of N-terminal methionine residue by methionyl-aminopeptidase in *E. coli*.

Key words : Cop protein, pE194, Isoelectric point, N-terminal sequencing, Circular dichroism spectrum

INTRODUCTION

The regulation of gene expression can be performed through a variety of mechanisms. One of them, well characterized in bacteriophages, is exerted through the interaction of a repressor protein with its operator target (Harrison and Aggarwal, 1990; Steitz, 1990; Pabo and Sauer, 1984). However, informations on plasmid-encoded repressor proteins are not so well documented (Gruss and Ehrich, 1989; Kues and Stahl, 1989; Nivick, 1989; Viret *et al.*, 1991). The replication of plasmid pE194 and several other small plasmids, mostly from gram-positive bacteria, is believed to proceed by a rolling circle model (Byeon and Weisblum, 1990; Sozhamannan *et al.*, 1990), in which a site specific nuclease, the Rep protein, defines the plus replication origin and produce a specific single-stranded nick that allows initiation of synthesis of the leading (i.e, plus) strand. The replication control region of plasmid pE194 contains two tandem open reading frames that specify the proteins Cop and RepF. Regulation of pE194 replication involves the two genes: *repF* that codes for the site specific initiator protein RepF and its expression is repressed by 65 nt countertranscript RNA ; and *cop*, placed upstream of *repF*, transcribed in the same mRNA and that specifies for a small protein Cop (Kwak and Weisblum, 1994). Cop was known to bind to proposed operator of 28-bp inverted complimentary repeat sequence in the cop-repF promoter region and acts as a repressor involved in the regulation of the synthesis of RepF (Kwak and Weisblum, 1994).

In this study, Cop protein was hyperproduced in *E. coli* expression system. Cop was purified, and its terminal amino acid sequence was partially determined. Further biochemical and biophysical characteristics of Cop, such as amino acid composition analysis and circular dichroism spectra, were determined also. Cop protein showed the α helix-turn-α helix motif, a typical structure of many DNA-binding proteins, and shared homology with a variety of repressors in the domains that shows the α helix-turn-α helix motif. Cop

Correspondence to: Jin-Hwan Kwak, School of Bioscience and Food Technology, Handong University, Namsong-ri 3, Buk-ku, Pohang 791-949, Korea

Table I. Bacterial strain and plasmids used in this study

Strain/Plasmid	Description and/or derivation	Reference or source
Bacteria		
E. coli BL21 (DE3)	Contained IPTG-inducible T7 RNA polymerase, and used as host for plasmid pJHK102	(Studier, 1990)
Plasmid		
pET3a	Used for overexpression of cloned gene under T7 promoter control in *E. coli* BL21 (DE3), from Novagen	(Studier, 1990)
pJHK102	Used for overexpression of Cop protein. Derived from pET3a by insertion of *cop* gene	(Kwak, 1994)

appeared to be a trimer of identical subunit. The general architecture of Cop and its interactions with DNA resembled that of Cro repressor of bacteriophage λ. Cop seemed to be one of the smallest repressor protein so far described.

MATERIALS AND METHODS

Bacterial strain and plasmids

The strain and plasmids used in this study are listed in Table I.

Growth medium and overexpression of Cop protein

The *cop* gene was cloned in plasmid pET3a, and then Cop protein was hyperproduced in *E. coli* BL21 (DE3). For large-scale preparation, four liter culture of *E. coli* BL21 (DE3) containing pJHK102 was grown in freshly prepared M9 minimal medium supplemented with 1% NZ amine and 4% glucose at 37°C. After overnight culture, the expression of Cop protein was induced by adding IPTG to a final concentration of 0.4 mM for 2 hr, and then rifampicin (200 μg/ml) was added to the culture to inhibit *E. coli* RNA polymerase. The cell was further incubated for 3 hr, and harvested by centrifugation. The cell pellets were washed twice with ice-cold solution (50 mM Tris-HCl (pH 8.0), 1 mM EDTA and 100 mM NaCl) and then stored frozen at -20°C until used.

Purification of Cop protein

The frozen cell paste was resuspended in ice-cold solution (50 mM Tris-HCl (pH 8.0), 1 mM EDTA and 100 mM NaCl) and digested with lysozyme (800 μg per gram of cell). The cells were disrupted with a sonicator and centrifugated at 8,000 g for 40 min. The supernatant was collected and dialyzed with buffer A (20 mM Tris-HCl (pH 8.0), 1 mM EDTA, 5 mM DTT, and 5% ethylene glycol) containing 50 mM KCl. After dialysis, this sample was applied to a DEAE-Sepharose fast-flow column equilibrated with buffer A containing 50 mM KCl. The column was washed with buffer A containing 50 mM KCl until absorbance at 280 nm returned to its preloading value. Proteins were eluted with gradient buffer A solution (50 to 500 mM KCl). Cop protein was eluted in flow-through fraction (Fig. 1 (A)). Cop protein in column fractions was detected by SDS-PAGE. The fractions containing Cop protein were pooled and directly applied to a Heparin-Sepharose affinity column equilibrated with buffer A containing 50 mM KCl. The column was washed with buffer A containing 50 mM KCl until the absorbance at 280 nm returned to its preloading value. The step gradient buffer A solution (50 to 500 mM KCl) was used for the elution of Cop protein. Cop protein was consistently eluted at 500 mM KCl (Fig. 1 (B)). Cop protein was traced by SDS-PAGE. Finally, HPLC on a reverse-phase column was used for the purification of Cop protein (Fig. 1(C)). The fractions containing Cop protein were loaded directly on a μBondaPak™ C_{18} column and eluted for 32 min at the flow rate of 1.2 ml/min with a step gradient: Initial to 2 min, 10% RP buffer B; 2 to 20 min, 10~50% RP buffer B; 20 to 25 min, 50~80% RP buffer B; 25 to 30 min, 80% RP buffer B; 30 to 32

Table II. Amino acid composition of Cop protein

Amino acid	Expected[a]	observed[b]
Ala	4.0	4.4
Asp	5.0[c]	4.7
Glu	8.0[d]	7.2
Phe	1.0	1.0
Gly	2.0	3.2
Ile	5.0	4.3
Lys	9.0	8.1
Leu	4.0	3.8
Met	1.0	1.5
Arg	3.0	2.9
Ser	2.0	2.0
Thr	4.0	4.1
Val	4.0	3.3
Tyr	3.0	2.5
Total	55	

[a]Calculated values from the deduced sequence of cop gene.
[b]Calculated values when Phe residue is assumed as 1.0.
[c]This value is the summation of the number of Asp and Asn.
[d]This value is the summation of the number of Glu and Gln.

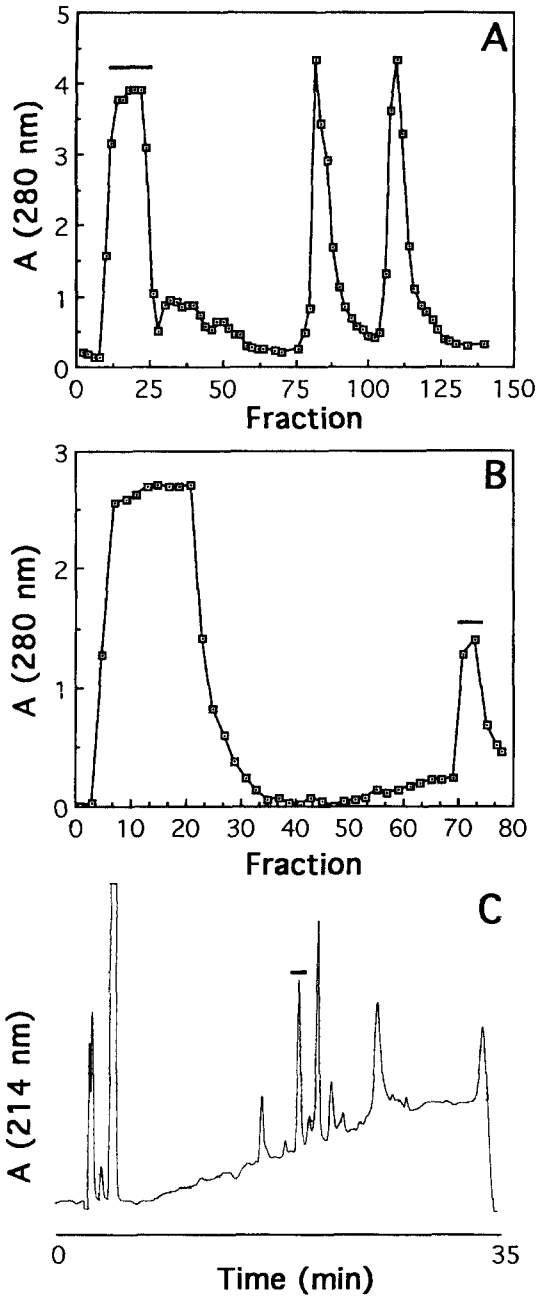

Fig. 1. Purification of Cop protein overexpressed in *E. coli* BL21 (DE3) cells. Panel (A), purification profile of crude extract on DEAE-Sepharose fast-flow column; Panel (B), chromatography of peak fraction from DEAE-Sepharose column on Heparin-Sepharose column; Panel (C), purification profile of peak fraction from Heparine-Sepharose column on a RP-HPLC (reverse-phase column of HPLC). The fractions in which Cop protein was present are indicated by horizontal bars (-).

min, 80~90% RP buffer B. The RP buffer A was 0.1% TFA in water, and RP buffer B was 0.1% TFA in acetonitrile. Cop protein peak in the RP-HPLC chromatogram was detected by SDS-PAGE and identified by the N-terminal amino acid sequencing. The purified protein was freeze-dried with Speed-Vac concentrator, and stored at -70°C. The purification yield was 1.91 mg/liter. The concentration of protein was measured by using Micro BCA (bicinchoninic acid) protein assay reagent (Pierce, USA) with bovine serum albumin (Pierce, USA) as a standard (Smith *et al.*, 1985).

Gel electrophoresis

The molecular weight of denatured Cop protein was determined by SDS-PAGE using the Bio-Rad SDS-PAGE Standards™ containing lysozyme (molecular weight, 14,400), soybean trypsin inhibitor (21,000), carbonic anhydrase (31,000), ovalbumin (45,000), bovine serum albumin (66,200), and phosphorylase b (97,400).

High-performance size exclusion chromatography

To determine the molecular weight of native state Cop protein, the purified Cop was loaded onto a Bio-Sil™ SEC-125 HPLC size exclusion chromatography column (600×7.5 mm, Bio-Rad, USA) which had been pre-equilibrated with 50 mM Tris-HCl, 100 mM NaCl (pH 7.5). The column was developed with same buffer at a flow rate of 0.8 ml/min. The elution rate was compared with those of the following gel filtration molecular weight markers (alcohol dehydrogenase, 150,000; carbonic anhydrase, 29,000; aprotinin, 6,500; Sigma Chemical Co., St. Louise, USA). Waters corporation Model 434 HPLC system (Tunable Absorbance Detector contained) was employed. The solvents were delivered by a Waters Model 510 HPLC pump. The delivery condition was programmed by a Waters Automated Gradient Controller™.

Characterizations of Cop

Predictions of the secondary structure of Cop was performed with the aid of the program PEPLOT (Gribskov *et al.*, 1986), included in the GCG computer programs of University of Wisconsin.

For circular dichroism (CD) spectra experiment, Cop protein (20 μg/ml) was dialyzed against 10 mM Tris-HCl, (pH 7.5). Circular dichroism spectrum was recorded with a JASCO J-600 spectropolarimeter (Japan Spectroscopic Co. LTD., Japan). The circular dichroism data was analyzed for secondary structure prediction by an algorism provided by the manufacturer.

Gel retardation assay

The DNA fragment (312 bp) for this assay was obtained by PCR using primer 1 (5' CCC CTA GGT GTC CAT TGT CCA TTG 3') , primer 2 (5' CAG CAT GCC AGT GCT GTC TAT 3'), and pE194 plasmid DNA as a template (Kwak and Weisblum, 1994). The resultant PCR product was purified with Magic Prep Column™ (Promega, USA), and then 5'-end labeled with [γ-^{32}P] ATP (specific activity, 6000 Ci/mM, Amersham). Gel

retardation assay was performed as described by Garner and Revzin (Garner and Revzin, 1981). The purified Cop protein (0.8 to 80 ng) was added to 20 µl of the reaction mixture (10 µg labeled DNA probe, 10 mM Hepes (pH 7.9), 60 mM KCl, 1 mM EDTA, 1 mM DTT, 1 mM Tris-HCl (pH 7.9), and 10% glycerol). The reaction mixture was incubated for 15 min at room temperature, and then 1 µg of bovine serum albumin and 2 µg of poly (dI-dC) were added to disrupt non-specific complexes. The resultant mixture was analyzed on a 5% polyacrylamide gel (0.5× Tris-borate-EDTA buffer).

Isoelectric focusing

Isoelectric focusing of Cop protein was performed by HPLC using a Mono-P HR 5/20 column (Pharmacia fine chemicals, USA). The column was equilibrated according to the instructions given by the manufacturer. Ethanolamine-CH_3COOH (25 mM) at pH 9.4 was used as starting buffer. A gradient (pH 9.4 to 4.0) was then developed in the column by applying polybuffer 96-CH_3COOH (pH 6.0). The sample (250 µg) was dissolved with starting buffer (250 µl) and applied to the column. After injection, the column was washed with starting buffer and finally eluted with polybuffer 96-CH_3COOH. The flow rate was 1.2 ml/min.

Amino acid composition analysis

Amino acid composition analysis was performed according to the modified method reported by manufacturer with an Applied Biosystems amino acid analyzer (Applied Biosystems Inc., CA, USA). In the analysis, Cop protein was hydrolyzed and derivatized with Model 420A Derivatizer, separated with Model 130A Separation System, and analyzed with Model 920A Data Analysis Module.

N-terminal amino acid sequence analysis

Amino acid sequencing of protein was carried out according to the modified reported by Applied Biosystems Inc. Edman degradation was performed in an Applied Biosystems model 471A Protein/Peptide Sequencer (Applied Biosystems Inc., CA, U.S.A.).

RESULTS AND DISCUSSION

Purification of Cop protein

To purify large quantities of Cop protein, *E. coli* BL 21 (DE3) containing pJHK102 was grown in four-liter culture as described in Materials and Methods. Ion-exchange chromatography (DEAE-Sepharose in a radial-flow column) (Fig. 1(A)), affinity chromatography (Heparin-Sepharose) (Fig. 1(B)), and reverse-phase HPLC (Fig. 1(C)) were used sequentially for the purification of Cop. Cop protein was identified on the basis of Coomassie blue staining of an SDS-PAGE and amino acid sequencing data. After the third step, Cop protein was free of any contaminants detectable by SDS-PAGE, RP-HPLC, and N-terminal amino acid sequencing. The collected HPLC fraction was used for *in vitro* characterization.

Determination of molecular weight of Cop protein

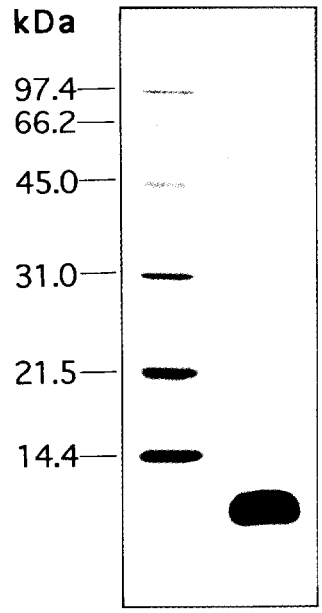

Fig. 2. SDS-PAGE of Cop protein purified from a RP-HPLC. Purified protein was loaded on 15% polyacrylamide gel in the presence of 0.1% SDS.

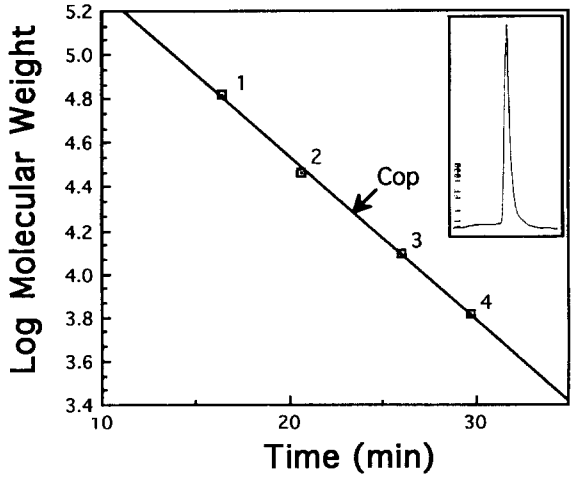

Fig. 3. Determination of molecular weight of Cop protein by HPLC using Bio-Sil™ SEC-125 column. Molecular weight standards; 1. bovine serum albumin (66 kDa); 2. carbonic anhydrase (29 kDa); 3. cytrochrome c (12.4 kDa); 4. aprotinin (6.5 kDa). The arrow indicated Cop protein. An inset shows an elution profile of purified Cop protein on the high performance size exclusion chromatography.

To determine the molecular weight of denatured and native Cop protein, SDS-PAGE and high performance size exclusion chromatography were performed, respectively. The purified Cop protein migrated as a single band upon polyacrylamide gel electrophoresis in the presence of SDS (Fig. 2). From the migration pattern of standard proteins, its molecular mass could be estimated as 6.1 kDa. On the other hand, the molecular mass of native Cop protein in Tris-HCl buffer (pH 7.5) containing 100 mM NaCl was determined by analytical high performance size exclusion chromatography on Bio-Sil™ SEC-125 column (Fig. 3). The protein was eluted as an apparently symmetrical peak. When compared with the filtration pattern of standard proteins, the molecular mass of native Cop protein was estimated to be 19 kDa. This result showed that Cop might behave as a trimer-like structure in 50 mM Tris-HCl buffer (pH 7.5) containing 100 mM NaCl.

Characteristics of Cop protein

Predictions of the secondary structure of Cop by the PEPLOT program showed the existence of a clear helix-turn-a helix motif, the typical one of DNA-binding proteins (Pabo and Sauer, 1984), as shown in Fig. 4.

An additional way to characterize the native state of the purified Cop protein would be to analyze its structural integrity. For this study, the native structure of the purified Cop by circular dichroism spectroscopy was evaluated. The circular dichroism spectrum of Cop protein in the far UV region is shown in Fig. 5. Based on this spectrum, Cop is calculated to contain 39.1%

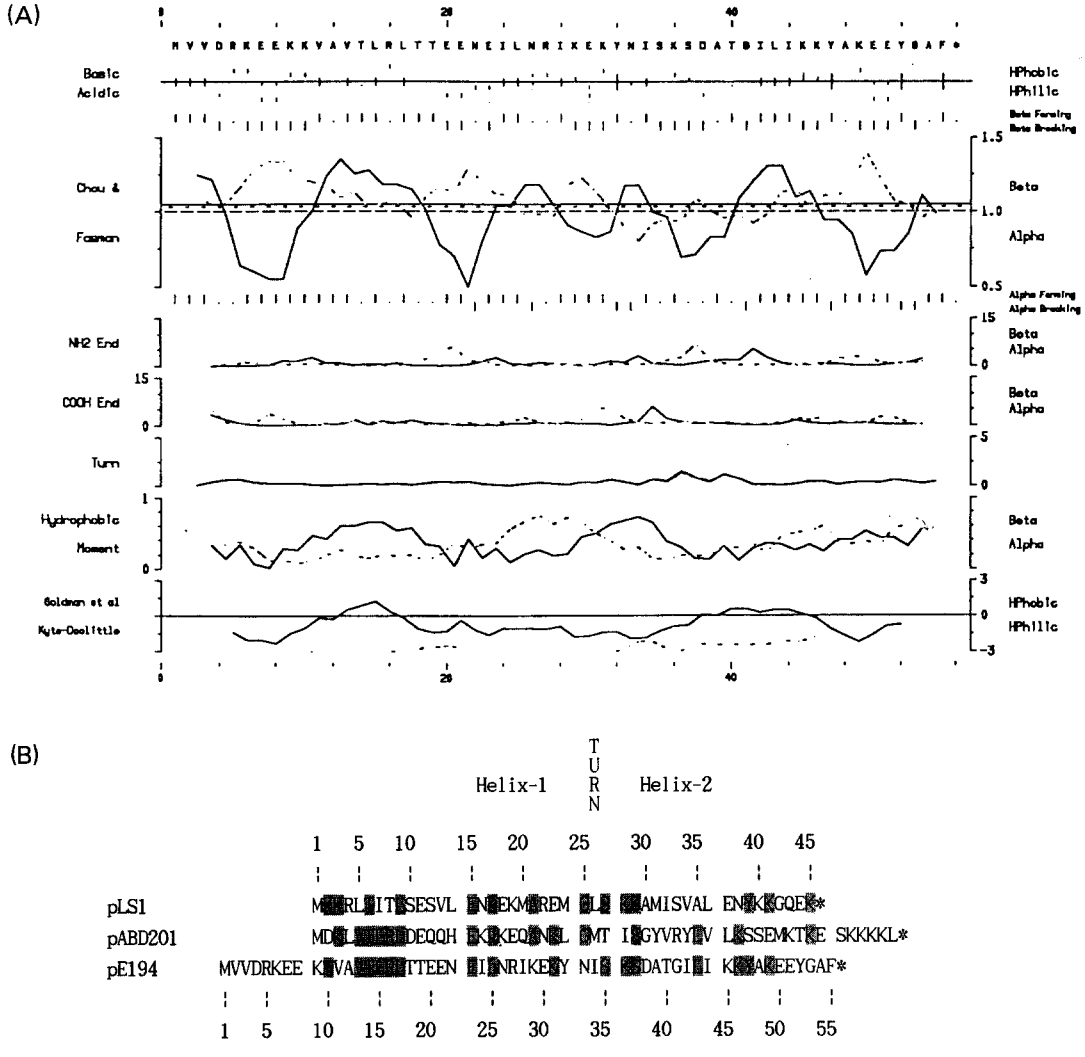

Fig. 4. Computer-derived predictions on secondary structure of Cop. Panel (A), Drawing of the output of PEPLOT program. The upper part shows the sequence and the sequence schematic of Cop with the characteristics of the residues indicated on the left and right parts; Panel (B), Comparison of the amino acids sequences of Cop and Cop-like putative proteins of plasmids pLS1 (RepA) and pADB201. Residues identities are shadow-boxed. In these comparison no gaps were allowed.

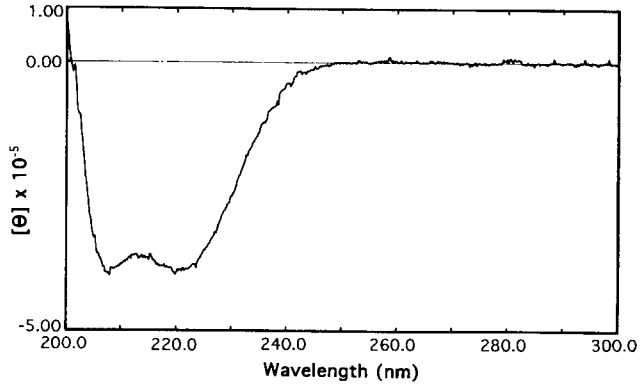

Fig. 5. Circular dichroic spectrum of Cop protein. The circular dichroic spectrum of Cop was recorded in 10 mM Tris-HCl buffer (pH 7.5) in a cell with a pathlength of 1.0 cm. The concentration of Cop was 20 μg/ml.

α-helix, 16.8% β-sheet, 17.4% turn, and 26.8% random structure.

Gel retardation assay

Cop protein was shown to bind to the promoter region of *cop* gene (Kwak and Weisblum, 1994). To confirm the DNA-binding property of the recombinant Cop protein expressed in *E. coli*, a 312-bp radiolabeled DNA fragment encompassing a binding site for Cop was used in DNA binding assay with the purified Cop protein. The resolution of protein-DNA complexes by the neutral PAGE revealed a significant band shift of the 312-bp labeled fragment when Cop protein was present in the assay (Fig. 6). This result indicated that the DNA binding activity of Cop expressed in *E. coli* was preserved during the expression and purification process.

Isoelectric focusing and amino acid analysis of Cop

The purified Cop protein was further characterized by isoelectrofocusing, amino acid composition analysis, and N-terminal amino acid sequencing. The isoelectric point of Cop was determined to be 9.0 by isoelectric focusing on a Mono-P HR 5/20 column.

The analysis of amino acid composition of the acid hydrolysate of Cop showed that it had the same amino acid composition as that deduced from its DNA sequence (Kwak and Weisblum, 1994).

The direct sequencing of Cop preparation by Edman degradation yielded the mixture of the sequence MVVDRKE... and the sequence VVDRKEE..., which was identical to that expected from the nucleotide sequence (Fig. 7).

These results confirmed that the correct Cop protein was produced in *E. coli* and Cop protein was not contaminated by any other *E. coli* proteins. However, the result of N-terminal sequencing indicated that about 15% of Cop expressed in *E. coli* was processed

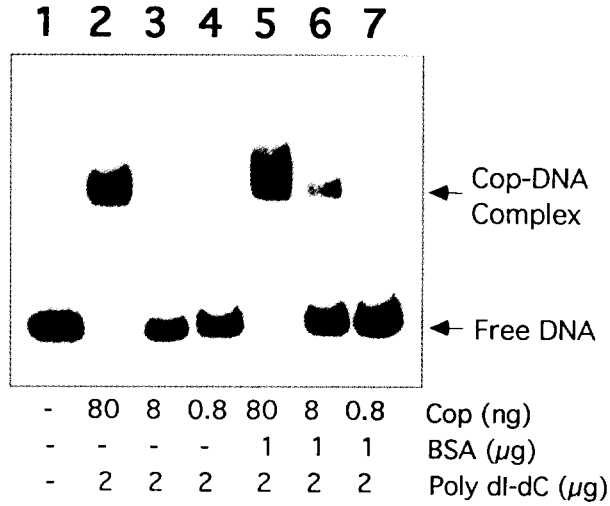

Fig. 6. Gel retardation analysis of complexes between *cop-repF* region DNA and Cop protein. To identify the DNA binding activity of Cop protein in the *cop-repF* region, selected DNA sequence was amplified by PCR, 5'-[^{32}P] end labeled, and incubated with Cop protein. Reaction mixtures containing Cop protein were further incubated with poly (dI-dC) to dissociate nonspecific Cop-DNA complexes, and fractionated by 5% PAGE to determine whether complexes were formed. All lanes have equal amounts (10 μg) of labeled DNA. Cop protein, BSA, and poly (dI-dC) were added to the reaction mixtures as indicated. Dashed lines (-) indicates that nothing was added.

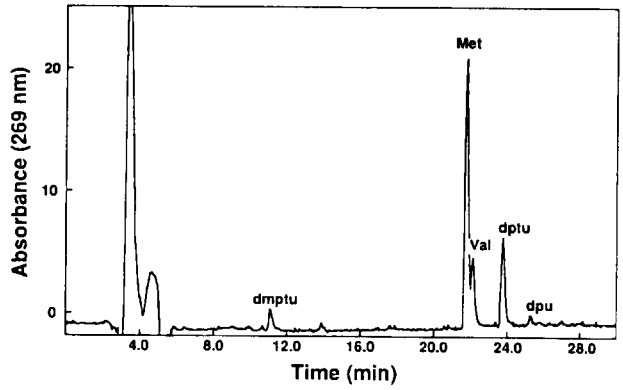

Fig. 7. The N-terminal sequence analysis of purified Cop protein. The PTH-amino acids were shown by the corresponding three-letter amino acid code. dmptu, dptu, and dpu are the byproducts of sequencing reagents and used for internal standards in this experiment.

to remove the N-terminal Methionine residue by methionyl-aminopeptidase in *E. coli*. This observation could be supported by the suggestion of the other report (Hirel *et al.*, 1989) that extension of N-terminal methionine excision from *E. coli* proteins was governed by the side-chain length of the penultimate amino acid.

Further studies, such as X-ray crystallography and NMR spectroscopy using a good amount of Cop protein, would be necessary to establish the mechanism of interaction between DNA and regulatory protein,

and the mechanism of regulation of plasmid DNA replication.

REFERENCES CITED

Byeon, W. H. and Weisblum, B., Replication genes of plasmid pE194-*cop* and *repF*: Transcripts and encoded proteins. *J. Bacteriology,* 172, 5892-5900 (1990).

Garner, M. M. and Revzin. A., A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: application to components of the *Escherichia coli* lactose operon regulatory system. *Nucleic Acid Res.,* 9, 3047-3060 (1981).

Gribskov, M., Burgess, R. R. and Devereux, J., PEPPLOT, a protein secondary structure analysis program for the UWGCG sequence analysis software package. *Nucleic Acids Res.*, 14, 327-334 (1986).

Gruss, A. and Ehrich, S. D., The family of highly interrelated single-stranded deoxyribonucleic acid plasmids. *Microbiol. Rev.*, 53, 231-241 (1989).

Harrison, S. C. and Aggarwal, A. K., DNA recognition by proteins with the helix-turn-helix motif. *Ann. Rev. Biochem.,* 59, 933-969 (1990).

Hirel, P.-H., Schmitter, J.-M., Dessen, P., Fayat, G. and Blanquet , S., Extent of N-terminal methionine excision from *E. coli* proteins is governed by the side chain length of the penultimate amino acid. *Proc. Natl. Acad. Sci.,* USA, 86, 8247 (1989)

Kuse, U. and Stahl, U., Replication of plasmids in gram-negative bacteria. *Microbiol. Rev.,* 53, 491-516 (1989).

Kwak, J.-H. and Weisblum, B., Regulation of plasmid pE194 replication: Control of *cop-repF* operon transcription by Cop and of *repF* translation by contertranscript RNA. *J. Bacteriology,* 176, 5044-5051 (1994).

Novick, R. P., Staphylococcal plasmids and their replication. *Annu. Rev. Microbiol.*, 43, 537-565 (1989).

Pabo, C. O. and Sauer, R. T., Protein-DNA recognition. *Ann. Rev. Biochem.,* 53, 293-321 (1984).

Smith, P. K., Krohn, R. I., Hermanson, G. T., Mallia, A. K., Gartner, F. H., Provenzano, M. D., Fujimoto, E. K., Goeke, N. M., Olson, B. J. and Klenk, D. C, Measurement of protein using bicinchoninic acid. *Anal. Biochem.,* 150, 76-85 (1985)

Sozhamannan, S., Dabert, P., Moretto, V., Ehrilich, S. D. and Gruss, A., Plus-origin mapping of single-stranded DNA plasmid pE194 and nick site homologies with other plasmids. *J. Bacteriology,* 172, 4543-4548 (1990).

Steitz, T. A., Structural studies of protein-nucleic acid interaction: the sources of sequence-specific binding. *Quarterly Rev. of Biophysics,* 23, 205-280 (1990).

Studier, F. W., Rosenberg, A. H., Dunn, J. J., and Dubendorff, J. W., Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.,* 185, 60-89 (1990).

Viret, J.-F., Bravo, A. and Alonso, J. C., Recombination-dependent concatemeric plasmid replication. *Microbiol. Rev.,* 55, 675-683 (1991).