The existence of a third gene whose product, C protein (controller), stimulates expression of the endonuclease gene was shown for the PvuII and BamHI modification–restriction systems [1, 2]. Until the cell accumulates a sufficient amount of C protein, there occurs virtually no expression of the endonuclease gene. It was demonstrated for the PvuII system that the product of the C gene activates the endonuclease gene transcription [1]. Products of C genes of the PvuII and BamHI systems were also shown to be interchangeable [2].

On analyzing the nucleotide sequence preceding the EcoRV endonuclease gene, we found two ATG codons that give rise to two open reading frames (ORF1 and ORF2) ending at the same point inside the endonuclease gene. Two DNA fragments corresponding to ORF1 and ORF2 were cloned, and the homogenous products of proteins encoded by them were found to be DNA-binding proteins. A specific DNA sequence (C box) recognized by the proteins was determined with DNAse I footprinting. The C box CCCATTTGGGTTATCCCATTTTGGG is located inside ORF1 and, similar to the PvuII C box consisting of tandem repeats of 11 nucleotides, is divided by four nucleotides. In its turn each of the repeats contains inverted repeats of four terminal nucleotides. The EcoRV C box sequence differs both from the PvuII C box sequence and from the proposed consensus sequence of C boxes in other modification–restriction systems.

**Key words:** modification–restriction systems, C.EcoRV, DNA-binding proteins, gel-shift assay, footprinting

**MATERIALS AND METHODS**

*Escherichia coli* BL21(DE3) strain (Novagen, USA), pET21d (Novagen) and pGL74 [3] plasmids, and restriction endonuclease NeoI (New England BioLabs, USA) were used in the experiments. The other enzymes (restriction endonucleases BspLU11I [4], Bli7361 [5], and XhoI, DNA polymerases Taq and Pfu, polynucleotide kinase, and phage T4 DNA-ligase) were prepared by us. The plasmids were isolated by alkaline lysis [6].

**Oligonucleotides** were synthesized by Syntol (Russia). The following oligonucleotides were used:

(B10): 5'-ATGACCATGATTACG-3';

(C1): 5'-ACCGTGCACTTGCCA-3';

(3C2): 5'-GCGGGTCTGCACTGCAAGAAGAAGCC-3';

(3C9): 5'-AATACATGTCTGTAAGAGAAA-3';

(3C10): 5'-GCATTAATTCTCGAGGAACGAAGAC-3';

T7-primer: 5'-TAATACGACTCTATAGGG-3'.

---


**Original Russian Text Copyright © 2003 by Zheleznaya, Kainov, Yunusova, Matvienko.**

**Regulatory C Protein of the EcoRV Modification–Restriction System**

L. A. Zheleznaya¹, D. E. Kainov¹, A. K. Yunusova¹, and N. I. Matvienko²*

¹Institute of Theoretical and Experimental Biophysics, Russian Academy of Sciences, Pushchino, Moscow Region, 142290 Russia; fax: (0967) 790–553

²Institute of Protein Research, Russian Academy of Sciences, Pushchino, Moscow Region, 142290 Russia; fax: (095) 924–0493; E-mail: nikmatv@vega.protres.ru

Received December 5, 2001

Revision received February 5, 2002

**Abstract**—The C gene product of the modification–restriction system PvuII binds to its own promoter (C box) and stimulates transcription of both the C gene and the endonuclease gene. According to our data the same regulatory mechanism is realized in the EcoRV system. It was found that upstream of the EcoRV endonuclease gene two ATG codons give rise to two open reading frames (ORF1 and ORF2) ending at the same point inside the endonuclease gene. Two DNA fragments corresponding to ORF1 and ORF2 were cloned, and the homogenous products of proteins encoded by them were found to be DNA-binding proteins. A specific DNA sequence (C box) recognized by the proteins was determined with DNAse I footprinting. The C box CCCATTTGGGTTATCCCATTTTGGG is located inside ORF1 and, similar to the PvuII C box consisting of tandem repeats of 11 nucleotides, is divided by four nucleotides. In its turn each of the repeats contains inverted repeats of four terminal nucleotides. The EcoRV C box sequence differs both from the PvuII C box sequence and from the proposed consensus sequence of C boxes in other modification–restriction systems.

*To whom correspondence should be addressed.*
The oligonucleotides were labeled using $[\gamma-^{32}\text{P}]\text{ATP}$ (Cluster Scientific Production Association, Russia) and phage T4 polynucleotide kinase as described by Maniatis et al. [6].

**Polymerase chain reaction (PCR)** was performed in two ways: using purified DNA as a template and directly from cell colonies. For this, part of the colony picked by the bacteriological loop was preliminary destroyed by boiling for 5 min in 50 µl water, then centrifuged for 1 min at 12,000g and 10 µl of the supernatant was used in the reaction. The reaction was performed in 100 µl of solution containing 1× ThermolPol-buffer (20 mM Tris-HCl, pH 8.8, 10 mM KCl, 10 mM (NH$_4$)$_2$SO$_4$, 2 mM MgSO$_4$, 0.1% Triton X-100), 200 µM deoxynucleoside triphosphate (each), 15 pmol oligonucleotides (each), 50 ng DNA template, and 2.5 units of the Taq- and Pfu-polymerases mixture (40 : 1 U/U). The reaction was performed in a Mastercycler (Eppendorf, Germany) with a temperature gradient on annealing. The temperature of annealing was chosen in a preliminary experiment.

**Analysis of PCR products** and other DNAs was made by electrophoresis using 1-1.5% agarose gels in 1× TBE-buffer (0.089 M Tris, 0.089 M boric acid, 1 mM EDTA, pH 8.3).

**Purification of the PCR products** from the reaction components was done using preparative electrophoresis. A cut-out gel band containing the required DNA fragment was dissolved in five volumes of 5 M guanidine thiocyanate, 0.1 M Tris–HCl, pH 7.0, the solution was transferred onto a QIAquick Spin column (Qiagen, USA) and further purification was carried out following the manufacturer’s instructions.

**DNA cleavage** with restriction endonucleases was done under the conditions advised for the endonucleases.

**Ligation** was performed in a volume of 20 µl containing 7 mM Tris-HCl, pH 7.6, 10 mM MgCl$_2$, 5 mM dithiothreitol, and 1 mM ATP. The molar ratio of the vector 5′-ends to the insertion was 1 : 1. Phage T4 DNA-ligase was added in such an amount that 5 units of the activity fell on 1 pmol of DNA 5′-ends. The sample was incubated for 12 h at 16°C. When the reaction was terminated, DNA-ligase was inactivated by heating at 70°C for 15 min.

**Cells competent for transformation** of E. coli BL21 (DE3) were prepared as described in [7].

**Transformation of competent cells** with the ligase mixture was done using a BioRad electroporator (USA) at field intensity of 15 kV/cm and pulse length of 1 msec.

**Gene expression was induced** with an addition to the cell culture of isopropyl-β-D-thiogalactopyranoside (IPTG) to the final concentration of 1 mM. Prior to this, cells were grown in LB mixture (1% bactotryptone, 0.5% yeast extract, 1% NaCl, pH 7.0) in the presence of ampicillin (100 µg/ml) with constant aeration at 37°C to the absorption ($A_{600}$) of 0.4. Following the induction the cells were incubated at 37°C for 4 h with vigorous stirring.

**Cell disruption and Ni-NTA-agarose chromatography.** After the induction the cells were collected by centrifugation, the pellet was suspended in a buffer for lysis (50 mM Tris-HCl, pH 7.5, 50 mM NaCl, 2 mM EDTA, 2 mM 2-mercaptoethanol), lysozyme was added to the final concentration of 1 mg/ml, and the mixture was incubated for 30 min on ice. Then the cells were disrupted using a UZDN-A ultrasonic disintegrator (Russia) for 2 min at 10°C with 10 sec sonications and 10 sec pauses. The lysate was centrifuged (30 min, 10,000g, 4°C) and the supernatant was transferred on the agarose column chelated with nickel-nitrilotriacetic acid—Ni-NTA-agarose (Qiagen). Such columns are designed to purify recombinant proteins with six histidine residues at their end that selectively bind to Ni ions. The protein was purified as recommended by the manufacturer. The protein concentration was determined using a spectrophotometer.

**Electrophoresis** of the proteins and lysates was done as described by Laemmli [8]. To analyze the total cell protein, the cell precipitate was also lysed according to Laemmli [8]. A mixture of Diaprot proteins (Dia M, Russia) were used as molecular mass markers.

**Gel shift.** To study the binding of C proteins with labeled DNA fragments, the gel-shift assay was employed [9]. Conditions of binding were chosen experimentally. Electrophoresis was done in 25 mM Tris-borate buffer, pH 7.6, at 10°C and 5 V/cm intensity.

**Footprinting.** The sequence recognized by C proteins was determined with footprinting using DNAse I [10]. DNAse I from calf pancreas (Koch Light, Germany) was used in the experiments. The solution of DNAse I (1 mg/ml) containing 150 mM NaCl and 50% glycerol was prepared from a dry preparation, stored at −20°C, and diluted before use to the required concentration in 1× MRB-buffer (10 mM Tris-HCl, pH 7.8, 10 mM MgCl$_2$, 50 mM NaCl). Labeled DNA fragments were hydrolyzed at room temperature in a solution (5 µl) containing 1× MRB-buffer, 0.5 pmol fragment, and 0.5 µl DNAse I. The DNAse I dilution and incubation time were chosen in preliminary experiments. The reaction was stopped by the addition of 3 µl of the stop-solution (deionized formamide containing 0.01% bromophenol blue, 0.01% xylene cyanole, 4 mM EDTA). Binding of the proteins with labeled DNA fragments was done under conditions similar to those in the gel-shift assay, and then the preparation was treated with DNAse I in the same way as the free fragment. The hydrolysis products were assayed using electrophoresis in 6% polyacrylamide gel containing 7 M urea and 1× TBE buffer at 50°C on a Macrophor instrument (LKB, Sweden). In parallel with the hydrolysis products the products of sequencing of the same fragment were also layered. The method of Sanger modified by Promega [11] with the use of the labeled oligonucleotide and Taq-polymerase was employed for sequencing.