Cloning and nucleotide sequence of the gene encoding the Ecal DNA methyltransferase

Vladimir Brenner†, Pál Venetianer and Antal Kiss*  
Institute of Biochemistry, Biological Research Center, Hungarian Academy of Sciences, PO Box 521, 6701 Szeged, Hungary

Received November 10, 1989; Accepted December 4, 1989  
EMBL accession no. X17111

ABSTRACT

The gene coding for the GGTNACC specific Ecal DNA methyltransferase (M.Ecal) has been cloned in E. coli from Enterobacter cloacae and its nucleotide sequence has been determined. The ecalM gene codes for a protein of 482 amino acids (M.; 51,111). It was determined that M.Ecal is an adenine methyltransferase. M.Ecal shows limited amino acid sequence similarity to other adenine methyltransferases. A clone that expresses Ecal methyltransferase at high level was constructed.

INTRODUCTION

DNA methyltransferases (MTases) fall in three groups according to the methylated base they generate: C5mC-MTases, N4mA-MTases and N6mA-MTases. The genes coding for several MTases have been cloned and sequenced (reviewed in 1). Comparison of the deduced amino acid sequences revealed that C5mC-MTases share extensive regions of homology (2). Similarly, N6mA-MTases were found to contain conserved sequences, most of which are unrelated to the conserved sequences of C5mC-MTases (3–5). Sequence similarity can be detected between M.PvaII, an N4mA-MTase, and N6mA-MTases (6) and there is one motif which seems to be present in all DNA MTases sequenced to date (7,8).

Comparison of the derived amino acid sequences helped to identify the target recognizing domain of multispecific C5mC-MTases (9) and yielded data strongly suggesting the involvement of certain regions in the transmethylation reaction (10), in target recognition (11) and in S-adenosyl-methionine binding (8).

Here we report the cloning from Enterobacter cloacae and characterization of a gene encoding a sequence specific DNA methyltransferase. The cloned methyltransferase modifies Ecal sites in the DNA of the cloning host E. coli, therefore we think it is identical with the methyltransferase of the Ecal restriction-modification system (12). We present evidence that M.Ecal is an adenine methyltransferase. The clone carrying the ecalM gene does not express Ecal endonuclease.

MATERIALS AND METHODS

Strains and Media

Enterobacter cloacae DSM 30056 (12) was obtained from J. Collins. For most plasmid cloning experiments E. coli ER1398 (13) was used as host. Cloning with the vector pER23S (ATG) (T. Lukácsovicz, unpublished) was done in E. coli JM107 (14) or ER1398 (pVH1). Plasmid pVH1 (15) carries the kanamycin resistance marker and the lacF2 gene. M13 phage was grown in JM107. Bacteria were grown in LB medium (16) at 37°C.

Enzymes and chemicals

Restriction endonucleases were either prepared in this institute or were purchased from New England Biolabs. DNA polymerase I large fragment was from Vepex (Szeged) and modified T7 DNA polymerase (Sequenase) from United States Biochemical Corp. Deoxyadenosine 5'-α-(35S)thiotriphosphate was purchased from Amersham.

Cloning methods

Preparation of bacterial genomic DNA, isolation of plasmids, transformation of E. coli, restriction mapping, agarose gel electrophoresis and subcloning of DNA fragments were done by standard procedures (16,17).

Determination of the nucleotide sequence

DNA fragments were cloned in mp18 and mp19 phage vectors (18) and were sequenced by the chain termination method (19,20) using deoxyadenosine 5'-α-(35S)thiotriphosphate (21) and either DNA polymerase I large fragment or Sequenase.

SDS-polyacrylamide gel electrophoresis of proteins

Cells were sedimented by centrifugation, then dissolved in 1/5–1/3 volume of sample buffer (22). The extracts were heated to 100°C for 2 min, then run in a 10% SDS-polyacrylamide gel (22). Gels were stained with Coomassie Brilliant Blue R.

* To whom correspondence should be addressed

†Present address: Institute of Microbiology, Department 112, Videnušká 1083, Prague 4, 14200, Czechoslovakia
RESULTS AND DISCUSSION

Cloning of the ecaIM gene

The ecaIM gene was cloned by a method which selects for a 'self-modifying' clone (23,1). This method works best if the vector contains many recognition sites in essential parts of the plasmid. Since the available plasmid vectors did not contain Ecal sites, we first constructed a suitable vector. A 326 bp AalI fragment of lambda phage DNA (nucleotides 5101–9394, ref. 24) that contains four Ecal sites was cloned in the 829 bp AalI site of pUC18 plasmid (18) to yield plasmid pVB40 (Fig. 1). The 0.9 kb BamHI fragment between the BamHI site at 5505 in the lambda sequence and the BamHI site in the multiple cloning site of pUC18 was deleted. The resulting plasmid (pVB44) containing only one BamHI site was used to clone the Ecal methyltransferase gene.

Enterobacter cloacae DNA was partially digested with Sau3A1 and ligated to pVB44 which had been cleaved with BamHI and dephosphorylated with bacterial alkaline phosphatase. The ligated DNA was transformed into E. coli ER1398. Plasmid DNA isolated from a mixed liquid culture of approximately 30,000 AmpR transformants was digested with completion to Ecal endonuclease, then the digested DNA was used to transform ER1398.

Most AmpR clones contained plasmids which were smaller than the vector. The appearance of these clones is probably due to the fact that the Ecal sites of pVB44 are in a non-essential region of the plasmid. Presumably, the DNA fragments resulting from Ecal digestion are circularized in vivo at low frequency (25) and the largest fragment carrying the origin of replication and the AmpR marker can be rescued using this type of selection. We found two clones which contained plasmids larger than the vector and the lambda-derived DNA region seemed to be intact. DNA isolated from these clones was resistant to Ecal digestion suggesting it carries the ecaIM gene. The two plasmids were identical by restriction analysis. One of the plasmids named pEcal was chosen for further investigation. In the course of this work pEcal DNA was digested with several restriction enzymes known to be sensitive to either adenine or cytosine methylation and there was no indication that pEcal carried any modification other than Ecal, thus we think that the cloned enzyme is the methyltransferase of the Ecal restriction-modification system. The clone carrying pEcal did not show phage restriction in vivo and no Ecal endonuclease could be detected in partially purified cell extracts.

In the approximately 2.2 kb insert of pEcal there is a 1.6 kb HindIII fragment. This HindIII fragment was recloned in the plasmid vector pVB40. The resulting plasmids (pEca4 and pEca5) contain the HindIII fragment in opposite orientation (Fig. 1). pEca4 DNA is completely resistant to Ecal digestion (even if no IPTG was added to the culture) whereas pEca5 is only partially resistant (not shown). The methylation properties of pEca4 and pEca5 suggest that the coding region is located in the HindIII fragment but the ecaIM promoter lies outside of the fragment. This would explain why the expression is dependent on the orientation.

Figure 1. Map of the DNA region coding for the Ecal methyltransferase. The coding sequence is indicated by the thick arrow.
A: pEca4 (striped segment: lambda DNA AalI fragment, empty segment: pUC18).
B: Restriction map of the 1.6 kb HindIII fragment cloned in pEca4.
Restriction sites which were used to subclone fragments for sequencing are shown.

Figure 2. Nucleotide sequence of the ecaIM gene and the amino acid sequence of the methyltransferase protein. The DPPY and F-G-G blocks are underlined. Dots indicate a Shine – Dalgarno site.
Nucleotide sequence

The nucleotide sequence of the 1.6 kb HindIII fragment was determined (Fig. 2). Only one open reading frame was found, it begins with ATG at 33 and ends at 1389 with TAA. The ATG at 33 is followed by other potential (in-frame) start codons downstream in the sequence (the closest is at 155).

To test if the first ATG is the functioning start codon of the EcaI methyltransferase, a pair of plasmids (pEca6 and pEca7) was constructed using the expression vector pER23S(−ATG) (Fig. 3). DNA fragments cloned in the SalI site of pER23S(−ATG) are transcribed from the ribosomal RNA B gene P2 promoter, but translation can only occur if the cloned gene contains an initiator codon. Expression of the cloned gene can be regulated via the lac operator which is located downstream of the promoter (T. Lukacsovich, pers. comm., Fig. 3). The ecaIM gene contains a unique BspRI site (nucleotide 67) which separates the first and second ATG codons (Fig. 2). The large BspRI-HindIII fragment containing the second but not the first ATG was inserted between the filled-in SalI site and the HindIII site of pER23S(−ATG). The resulting plasmid was named pEca6 (Fig. 3). To construct the other plasmid (pEca7), the ends of the HpaII-HpaII fragment containing the whole open reading frame were filled in by DNA polymerase I large fragment then it was ligated into the filled-in SalI site of pER23S(−ATG). Plasmids pEca6 and pEca7 contain no EcaI site, therefore total DNA was isolated from the clones and tested for EcaI methylation. Figure 4 demonstrates that DNA isolated from JM107 (pEca6) was digestible with EcaI, whereas DNA from JM107 (pEca7) was resistant, suggesting that translation starts at the first ATG. Translation cannot start further upstream, because there is an in-frame stop codon (TAA) at position 6. The reading frame beginning at 33 predicts a protein of 452 amino acids (M.: 51,111). A protein of the expected size can be detected, upon IPTG induction, in cells containing pEca7 (Fig. 5). The level of M.EcaI production directed by pEca7 (Fig. 5) seems to be sufficiently high to facilitate purification of the enzyme.

The sequenced region upstream of the open reading frame does not contain structures resembling a typical E. coli promoter. The lack of such structure may explain why the level of expression of the ecaIM gene is different in pEca4 and pEca5.

Comparison with other DNA methyltransferases

M.EcaI is the first DNA methyltransferase with GGTNACC specificity for which the amino acid sequence has become known. The analysis of the deduced amino acid sequence revealed that it contains the motif F.G.G which, in a more or less conserved form, seems to be present in all DNA-MTases (7, 8) and has been suggested to be part of the S-adenosyl-methionine binding site (8). The identification of this motif in M.EcaI gives further support to the notion that this conserved element must be responsible for a general step in the methylation process, possibly for binding S-adenosyl-methionine. The EcaI sequence does not show the motifs specific for C5mC-MTases (2), it does contain, however, the sequence DPPY common to adenine methyltransferases (26-42). This motif, in a modified form, was also found in N4mC-methyltransferases, therefore it was
suggested to be characteristic for DNA-MTases modifying extracyclic nitrogen (6, 42). The lack of resemblance to C5mC-MTases and the presence of the typical DPPY motif suggested that M.EcaI would be an adenine methyltransferase.

**M.EcaI is an adenine methyltransferase**

We have indirect experimental evidence suggesting that M.EcaI is indeed an adenine methyltransferase. The sequenced region contains a unique EcaI site which partially overlaps with a HpaII site: GGTTACCGG (at position 14). Methylation, whether C5 or N4, of either cytosine in the sequence CCGG, is known to block cleavage by HpaII (43, 44). To test if methylation by M.EcaI protects the overlapping HpaII site from HpaII cleavage, a 640 bp EcoRI-EcoRI fragment was isolated from pEca4. This fragment contains 584 bp of the 1.6 kb HindIII fragment (with the 5'-end of the ecaM gene) and the multiple cloning site of pUC18 (approximately 50 bp).

The appearance, upon HpaII digestion, of an approximately 560 bp fragment indicates (Fig. 6) that EcaI methylation did not block HpaII cleavage at the overlapping site, thus M.EcaI must be an adenine methyltransferase.

**ACKNOWLEDGEMENTS**

We thank Drs. J. Collins, E. Raleigh, U. Günthert, T. Lukácsovich for strains and E. Csorba for her excellent technical assistance.

**REFERENCES**