

The DNA sequence recognized by the *EcoDXX1* restriction endonuclease

Andrzej PIEKAROWICZ and Jon D. GOGUEN

Institute of Microbiology, Warsaw University; and
Department of Microbiology, University of Alabama in Birmingham

(Received July 30/October 7, 1985) — EJB 85 0845

EcoDXX1 is a type-I restriction enzyme coded for by the plasmid pDXX1. Like other type-I restriction endonucleases, the enzyme catalyses the modification of susceptible DNA. We have determined the DNA sequence recognised by *EcoDXX1* to be:



where N can be any nucleotide. This sequence has an overall structure very similar to previously determined type-I sequences.

The type-I DNA restriction and modification (R-M) systems of Enterobacteriaceae are the best characterised systems from the genetic and physiological point of view (for recent reviews, see [1–3]).

The DNA recognition sequence of three type-I enzymes, *EcoK*, *EcoB* and *EcoA*, had been determined [4–9]. These sequences, although different, have a common structure in that they consist of a specific tetranucleotide and a specific trinucleotide sequence separated by a spacer of non-specific sequence, which is 6, 7 or 8 base pairs long. Recently it has been shown that *Salmonella* type-I restriction enzymes recognize specific sequences, which have a structure similar to those of *Escherichia coli* type-I restriction enzymes (T. A. Bickle, unpublished work).

In this paper we describe the DNA sequence recognised by type-I restriction enzyme *EcoDXX1* coded for by the plasmid pDXX1 [10] and show that this sequence has the characteristic structure of the known type-I recognition sequences.

MATERIALS AND METHODS

Purification of *EcoDXX1*

The purified *EcoDXX1* enzyme was obtained as described previously [10].

In vitro DNA methylation

A typical reaction contained in a volume of 0.25 ml 100 mM Tris/HCl, pH 7.5, 0.25 mM EDTA, 6.5 mM MgCl₂, 14 mM 2-mercaptoethanol, 3 μM [methyl-³H]AdoMet (Amersham, typically 70 Ci/mmol), and 10 μg DNA. The reaction was initiated by the addition of 4 μg *EcoDXX1*

Correspondence to A. Piekarowicz, Instytut Mikrobiologii, Uniwersytet Warszawski, Nowy Swiat 67, PL-00-046 Warszawa, Poland.

Abbreviation. AdoMet, S-adenosyl-L-methionine.

Enzymes. (IUB Recommendations 1984). Restriction endonuclease *EcoDXX1* (EC 3.1.21.3); restriction endonucleases *Hae*III, *Hha*I, *Hind*III, *Hin*II and *Stu*I (EC 3.1.24.4).

enzyme and incubated at 37°C for 5 h. Reactions were terminated by the addition of 70 μl glycerol and 25 μl phenol and the DNA was separated from low-molecular-mass radioactive material by chromatography on a 6-ml bed of Bio-Gel A-0.5 m prepared in a disposable 10-ml pipette. The labelled DNA was precipitated with ethanol and redissolved in a small volume of 10 mM Tris/HCl, pH 8.0, 0.1 mM EDTA.

Other methods

DNA fragments were separated on either agarose or polyacrylamide gels depending on the expected size of the fragments. DNA was detected following staining with ethidium bromide and ³H-labelled DNA after drying the gel and fluorography according to Chamberlain [11] or by excision of DNA bands from the gel, solubilizing with NCS tissue solubilizer and determination of the ³H content by liquid scintillation counting [12].

RESULTS

The strategy that we used was developed for the determination of type-III restriction enzyme recognition sites [13, 14] and was also used for the determination of the *EcoA* recognition sequence [9]. This method is based on mapping sites for the enzyme on DNA molecules of known sequence. A computer is then used to search the sequence for small sequences present in these regions of the DNA and absent from all others. The presumptive recognition sequences that are found in this way are then verified by checking that the enzyme acts at the predicted positions in other sequenced DNA molecules.

We began these experiments with φX174 DNA [15]. This DNA was methylated by *EcoDXX1* as described in Materials and Methods. Samples of the methylated φX174 DNA was digested with the restriction enzymes *Hha*I, *Hin*II and *Hae*III. The DNA fragments were separated by electrophoresis through 2% (w/v) or 5% (w/v) acrylamide gels, which were

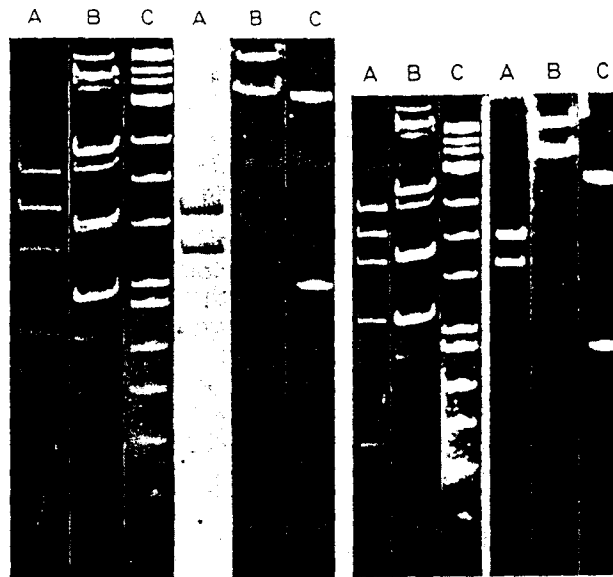


Fig. 1. Localization of *EcoDXX1* sites on ϕ X174 DNA. A sample of ϕ X174 DNA was methylated with *EcoDXX1* as described in Materials and Methods, aliquots of the DNA were then digested with the enzymes *HaeIII*, *HhaI* and *HinfI*, the digests were run in 2% acrylamide/0.7% agarose or 5% acrylamide/0.7% agarose gels, which were stained with ethidium bromide and photographed and then dried and fluorographed. The left panels show the stained gels, the right panels the fluorographs (contact print for B and C). (A) *HaeIII* digestion, (B) *HhaI* digestion, (C) *HinfI* digestion

Table 1. The regions of ϕ X174 and M13mp7 DNAs methylated by *EcoDXX1*

Substrate DNA	Digested with	Minimum number of methylations sites	Location of source for methylation sequence of	
			sites	substrate
ϕ X174	<i>HhaI</i> , <i>HinfI</i> <i>HaeIII</i>	2	156–204	[15]
			3731–4148	
M13mp7	<i>HhaI</i> , <i>HaeIII</i> <i>FnuDII</i>	2	2711–3400	[16]
			6740–7011	

stained and photographed and then subjected to fluorography. The results of these experiments are shown in Fig. 1. In a similar type of experiment M13mp7 DNA was methylated by *EcoDXX1*, digested with *HaeIII*, *HhaI* and *FnuDII* and analysed for the presence of the methylated fragments (data not shown). The results of these analyses are summarised in Table 1.

The results of the mapping have shown that the two methylated regions in ϕ X174 DNA are located between base pairs 156 and 204 and base pairs 3731 and 4148, and in M13mp7 DNA between base pairs 2711 and 3400 and base pairs 6740 and 7011. We then asked the computer to find sequences present in these four regions of ϕ X174 and M13mp7 DNAs that were absent from the rest of these DNAs. The computer has found that no unhyphenated sequence between 4 and 12 base pairs long was consistent with the data. Testing all possible hyphenated sequences with a single hyphen 1–12 base pairs long, only three possible candidates were found, all of them of the general type-I form. These sequences were: (a) ATCAN₇ATTC, (b) TCAN₇ATCA, and (c) ATCAN₈TTC.

The computer analysis of the several DNAs sequences has shown that all three sequences are present in the simian virus 40 sequence [17]; only sequence (c) is present in the pBR322 sequence [18], and in λ DNA [19] sequence (a) is not present

while sequence (b) is present six times at positions 13210, 21061, 32437, 33406, 34611, and 46748. The facts that only fragment *HindIII* B of SV40 DNA (Fig. 2), carrying sequence (a) and (b), is methylated, that pBR322 DNA is not methylated by *EcoDXX1* and that λ DNA is methylated at the predicted restriction fragments (Fig. 3) suggest that the correct recognition sequence for *EcoDXX1* is sequence (b).

The DNA sequence in the region of the ten *EcoDXX1* sites that we have found is shown in Fig. 4. An examination of these sequences showed that the only conserved elements are the indicated trinucleotide and tetranucleotide sequences. All of the positions in the 7-base-pair spacer, as well as the flanking nucleotides, can vary.

The sequence data have shown that one site on the ϕ X174 DNA contains a *HinfI* site. Assuming that modification puts one methyl group into the specific trinucleotide sequence and a second into the tetranucleotide, and because (Fig. 1) after cleavage of the methylated ϕ X174 DNA with *HinfI* the fragment between 53 and 204 is methylated while the fragment from 205 to 286 is not, we can argue that the methyl group must be on one of the two adenosyl residues in the tetranucleotide part of the sequence, in the lower strand as drawn in Fig. 4. (*EcoDXX1* modification methylates adenosyl residues, unpublished results.)



Fig. 2. Mapping of the *EcoDXX1* site on simian virus 40 (SV40) DNA. A sample of SV40 DNA was methylated with *EcoDXX1* as described in Materials and Methods, an aliquot of the DNA was then digested with *HindIII*, the digest run on a 2% acrylamide gel and analysed as described in the legend to Fig. 1. The left panel shows the stained gel, the right panel the fluorograph (contact print). Only the second fragment extending from position 3920 to 5089 in the SV40 DNA sequence is labelled (the nucleotide numbering is from Fiers et al. [17])

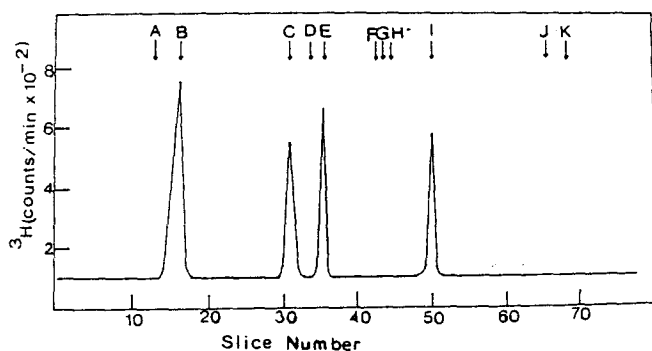


Fig. 3. Confirmation of the *EcoDXX1* recognition sequence. Phage λ DNA was methylated with *EcoDXX1* and the DNA was purified as described in Materials and Methods. The DNA was then digested with *HindIII* and *StuI* and the digest was run on a 1% agarose gel, which was stained with ethidium bromide and photographed. Each band of DNA was excised from the gel and solubilized with NCS tissue solubilizer and the ^3H content of each was determined by liquid scintillation counting. Letters in the figure indicate the λ DNA fragments obtained after double digestion with *HindIII* and *StuI* restriction enzymes

If one assume that, as with other type-I recognition sequences, the *EcoDXX1* methylates both strands, then the second methyl group should be on the adenosyl residue in the trinucleotide of the sequence in the top strand as drawn in Fig. 4.

DISCUSSION

The results presented in this paper show that the structure of the *EcoDXX1* recognition sequence is very similar to the other type-I sites that have been sequenced, a specific

ϕ X174	(1)	ATAATCAGCGTCACATTCAGAA
		TATTAGTGGCACTGTAAGTCTT
	(2)	GGCATCAACTAACGATTCGTGTC
		CCGTAGTTGATTGCTAAGACAG
M13mp7	(1)	ACAATCAATAGAAAATTCATAT
		TGTTAGTTATCTTTAAGTATA
	(2)	ACCATCAATATGATATTCACC
		TGGTAGTTATACTATAAGTTGG
λ	(1)	GTTTTCAGGTGGCATTTCAGCC
		CAAAAGTCCACCGTAAGTCGG
	(2)	GCGGTCACTGCGCATTCTTCG
		CGGCAGTACAGCGTAAGAAGC
	(3)	TTTTTCATGGTGTATTCTTCG
		AAAAAGTACCACAATAAGAAGC
	(4)	TCTCTCATTAACATATTCAGGC
		AGAGACTAATTGTATAAGTCCG
	(5)	TCGTTCAAGTAAGATTCGGAA
		AGCAAGTTCATTTCTAAGCCTT
	(6)	GCATTCATCCATCCATTCCGTA
		CCTAAGTAGCTAGGTAAGCCAT

Fig. 4. The *EcoDXX1* recognition sequence. The regions of DNA containing the ten *EcoDXX1* sites discussed in the text are shown with the specific parts of the sequences boxed. The *HinI* site within the ϕ X174 sequence is underlined

trinucleotide, a specific tetranucleotide and a spacer 5–8 base pairs long [4–9] (T. A. Bickle, unpublished work). These results, together with the results of the characterization of the *EcoDXX1* enzyme [10], allow the final conclusion that the restriction and modification system coded for by the *E. coli* plasmid pDXX1 belongs to the classical type-I R-M systems.

These results also allow the conclusion that the generalization about type-I ATP-hydrolysing restriction enzymes; that they are a very homogeneous group, coded by allelic genes located on the chromosome and with very similar subunit structures and reaction mechanisms, is not entirely correct. It seems that this group of restriction enzymes is more heterogeneous than it was thought to be until recently. The first example of such heterogeneity is the *EcoA* enzyme, which differs from the classical enzymes in the sizes and structures of all of the subunits and in important features of the reaction mechanism [9]. Another example is the *EcoDXX1* enzyme, which differs from the classical enzymes, being coded by plasmid genes which are not allelic to the chromosomal genes coding for other type-I enzymes (Skrzypek and Piekarowicz, unpublished work), and in the sizes of its subunits [10].

We wish to thank Roy Curtiss III for arranging the visit of A. P. and for his kind provision of facilities used in part of this project. A. P. was supported by National Institute of Dental Research grant 2 P50 DE02670. J. D. G. was supported by an individual postdoctoral fellowship from the National Institute of Health by genetic training grant T32 GM 07090 also from N. I. H.

REFERENCES

1. Yuan, R. (1981) *Annu. Rev. Biochem.* 50, 285–315.
2. Endlich, B. & Linn, S. (1981) in *The Enzymes* vol. 14A, 3rd edn, pp. 137–156, Academic Press, New York.

- Bickle, T. A. (1982) in *The Nucleases* (Linn, S. & Roberts, R. J., eds) pp. 85–108, Cold Spring Harbor Laboratory, Cold Spring Harbor.
4. Kan, N. C., Lautenberger, J. C., Edgell, M. H. & Hutchinson, C. A. III (1979) *J. Mol. Biol.* 130, 191–209.
 5. Ravetch, J. V., Horiuchi, K. & Zinder, N. D. (1978) *Proc. Natl Acad. Sci. USA* 75, 2266–2270.
 6. Lautenberger, J. A., Kan, N. C., Lackey, D., Linn, S., Edgell, M. H. & Hutchinson, C. A., III (1978) *Proc. Natl Acad. Sci. USA* 75, 2271–2275.
 7. Lautenberger, J. A., Edgell, M. H., Hutchinson, C. A., III & Godson, G. N. (1979) *J. Mol. Biol.* 131, 871–875.
 8. Sommer, R. & Schaller, H. (1979) *Mol. Gen. Genet.* 168, 331–335.
 9. Suri, B., Shepherd, J. C. W. & Bickle, T. A. (1984) *EMBO J.* 3, 575–579.
 10. Piekarowicz, A., Goguen, J. & Skrzypek, E. (1985) *Eur. J. Biochem.* 152, 387–393.
 11. Chamberlain, J. P. (1979) *Anal. Biochem.* 98, 132–135.
 12. Brzeziński, R. & Piekarowicz, A. (1982) *J. Mol. Biol.* 154, 615–627.
 13. Hadi, S. M., Bächli, B., Shepherd, J. C. W., Ineichen, K., Yuan, R. & Bickle, T. A. (1979) *J. Mol. Biol.* 134, 655–666.
 14. Piekarowicz, A., Bickle, T. A., Shepherd, J. C. W. & Ineichen, K. (1981) *J. Mol. Biol.* 146, 167–172.
 15. Sanger, F., Air, G. M., Burrell, B. G., Brown, N. L., Coulson, A. R., Fiddes, J. C., Hutchinson, C. A., III, Slocombe, P. M. & Smith, M. (1977) *Nature (Lond.)* 265, 687–695.
 16. Messing, J., Crea, R. & Seeburg, P. H. (1981) *Nucleic Acids Res.* 9, 3015–3022.
 17. Fiers, W., Contreras, R., Haegeman, G., Rogiers, R., Van de Voorde, A., Van Heuverswyn, H., Van Herrweghe, J., Volckaert, G. & Ysebaert, M. (1978) *Nature (Lond.)* 273, 113–120.
 18. Sutcliffe, J. G. (1979) *Cold Spring Harbor Symp. Quant. Biol.* 43, 77–90.
 19. Sanger, F., Coulson, A. R., Hong, G. F., Hill, D. F. & Paterson, G. B. (1982) *J. Mol. Biol.* 162, 729–773.