Volume 3 no.7 July 1976 Nucleic Acids Research

# Two restriction endonucleases from Bacillus globiggi

Vincenzo Pirrotta

Dept. of Microbiology, Biozentrum of the University of Basel, Klingelbergstrasse 70, 4056 Basel, Switzerland.

Received 7 May 1976

## ABSTRACT

The sites of action of the restriction enzyme Bgl II on λ DNA are mapped. This enzyme recognises the sequence 5'...AGATCT...3' and makes staggered cuts producing sticky ends. In λ DNA, the second A in this sequence is methylated about 50 % of the time by a bacterial methylase absent in E. coli dam⁻. In contrast to Bgl II, Bgl I makes many cuts in λ DNA and produces 5' terminals which are not substrates for polynucleotide kinase.

## INTRODUCTION

Like several other Bacillus strains, Bacillus globiggi contains site specific endonucleases of the type generally referred to as class II restriction endonucleases. Wilson and Young first isolated from this strain two activities, Bgl I and Bgl II, and found that neither enzyme requires co-factors other than $Mg^{++}$ and that they cleaved λ DNA respectively into 22 and 5 fragments (Wilson and Young, personal communication). In this article I report further studies of these two enzymes, in particular, the mapping of the cleavage sites of Bgl II in λ DNA and the sequence recognised and cleaved by this enzyme.

Three interesting features were revealed by this work.

1) Although the recognition sequence of Bgl II is different from that of other Bacillus restriction enzymes such as BamH I[1] or Bst I[2], the enzyme also makes staggered cuts, producing sticky ends which are identical to those produced by BamH I.

2) The recognition sequence of Bgl II apparently corresponds to at least one of the sequences methylated by E. coli DNA methylases. The Bgl II cleavage sites in λ DNA are roughly 50 %

methylated.

3) In contrast to other known class II restriction enzymes, but similar to class I enzymes, Bgl I produces 5' terminals which are not substrates for polynucleotide kinase even after treatment with phosphatase.

## MATERIALS

DNA of phages λ, λimm434 and λdbio M30-7 nin 5 was isolated by heat induction of cI ts lysogens as previously described[3]. E. coli GM 48 dam$^-$ dcm$^-$ was obtained from Dr. N.R. Morris and lysogenised with λ cI 857 S7.

Bgl I and Bgl II enzymes were isolated from Bacillus globiggi and are a generous gift of Dr. B. Bächi.

T4 polynucleotide kinase was purified from T4 infected E. coli according to Richardson[4]. Bacterial alkaline phosphatase, snake venom phosphodiesterase and pancreatic DNase were obtained from Worthington, Sigma and Boehringer respectively.

$\gamma$-$^{32}$P-ATP (200-800 mci/μmole) was prepared by a modification of the method of Glynn and Chappell[5] and used without further purification.

## METHODS

Bgl I or II digestions were performed in buffer containing 0.05 M Tris pH 7.9, 0.05 M KCl, 0.01 M $MgCl_2$ and 0.01 M dithiothreitol. After digestion, the reaction mixture was analysed on 1 % agarose slab gels made up in 0.04 M Tris, 0.04 M potassium phosphate pH 7.8 and 0.1 mM EDTA. Electrophoresis proceeded at 60 v and 150 m amp.

To phosphorylate the fragments, the cleavage reaction mixture containing 20-100 μg λ DNA, was treated with 0.1 units bacterial alkaline phosphatase for 1 hr at 37°, then phenol extracted four times. The phenol was removed by repeated ether extraction and the ether by blowing a gentle air stream over the solution. 0.5 - 1 n moles of $\gamma$-$^{32}$P-ATP at 200-800 mci/μmole and 1-2 units of polynucleotide kinase were added and the mixture was incubated for 1 hr at 37°. The labeled fragments were purified by

one phenol extraction and filtration through Sephadex G-100. The purified labeled fragments were precipitated with ethanol, resuspended in 0.01 M Tris pH 7.5, 0.01 M $MgCl_2$ and digested with 1 µg pancreatic DNase for 1 hr at 37° in a capillary. The products were fractionated by electrophoresis on cellulose acetate pH 3.5 and in the second dimension by homochromatography on DEAE cellulose thin layers using homomix VI of Jay et al.[6] After autoradiography, the spots were cut out of the thin layer, eluted with 30 % triethylamine carbonate and subjected to partial digestion with venom phosphodiesterase (1 µl 0.1 mg/ml) for 5-60 min at room temperature. The venom digests were analysed by electrophoresis on aminoethyl cellulose paper (AE) at pH 3.5 or on DEAE cellulose paper pH 2 [7].

## RESULTS

### Bgl II sites in λ DNA

Bgl II cuts λ DNA into six fragments (fig. 1). The smallest of these, fragment F, is about 400 base pairs long and contains one of the λ sticky ends. It is usually present in lower than stoichiometric amounts but the band is restored to normal strength by briefly heating the sample to 70° before applying it to the agarose gel.

The other sticky end was located by terminally labeling intact λ DNA with polynucleotide kinase before cleaving it with Bgl II. Only fragments C and F become labeled (fig. 2).

The differences in the band patterns produced by Bgl II cleavage of λ, λimm 434 or λdbio M30-7 nin 5 DNAs (fig. 1) permit the ordering of the fragments almost by inspection. Fragments B, D and E are absent in the λimm434 digest, indicating that they are contained in whole or in part in the immunity region. Fragment B, constituting about 27 % of λ DNA is too large to be contained inside the immunity region or even to contain the right end of the immunity region, since the boundaries of the latter are 73 and 79 percent units from the left end of λ DNA. Fragment B must therefore extend to the left of the immunity region. λdbio M30-7 contains a substitution of the

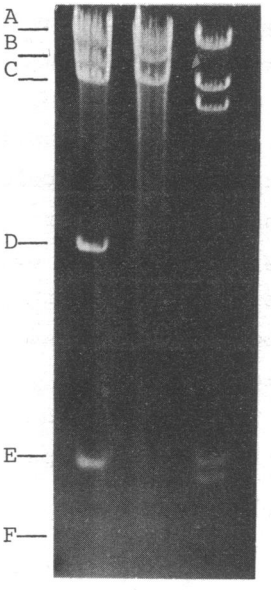


Fig. 1

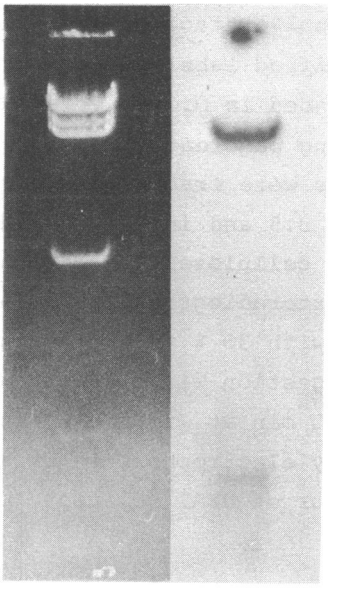

Fig. 2

Figure 1. Agarose gel electrophoresis of λ DNA digested with Bgl II. The gel contained 1% agarose and the bromphenol blue marker migrated just above fragment E. a)λ b) λimm434 c) λ dbio M30-7 nin 5

Figure 2. Gel eletrophoresis of λ DNA terminally labeled with polynucleotide kinase before Bgl II digestion. The gel was stained with ethidium bromide and photographed (left), then dried and autoradiographed (right).

region from att to the middle of the immunity region. In addition this DNA contains nin5, a 5.4 % deletion located between the P and Q genes. These deletions alter fragments B, C and D, but not fragment E, indicating that the latter is located at the right end of the immunity region. Fragment C is reduced in size by the nin5 deletion, although this is obscured in fig. 1 by the presence of a new band, due to the bio substitution, which migrates in the same position as the intact fragment C. Since fragment A, because of its size can only come from the

left arm of λ, fragment C is the only possible candidate to fill the region between the immunity region and the right end. I conclude that the order of the fragments in λ DNA must be, from left to right, F-A-B-D-E-C (fig. 3).

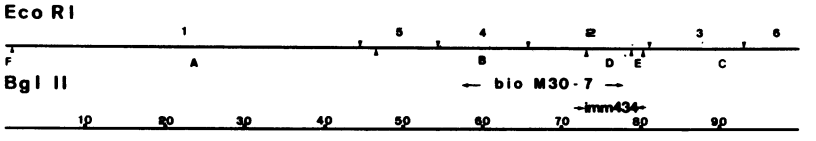


Figure 3. Cleavage sites in λ DNA for Eco R I and for Bgl II. Eco R I fragments are numbered from 1 to 6 in order of migration in agarose gels. Bgl II fragments are labeled A to F as in fig. 1. Map locations are estimated from the Eco R I sites according to Thomas and Davis[8].

λ repressor binds to DNA fragments containing the λ operators and retains them on a nitrocellulose filter, while fragments not bound to repressor pass through the filter. Fig. 4 shows that the repressor selects fragments B and D from a digest of λ DNA with Bgl II, while only a short fragment containing the $O_R$ operator is selected from a digest of λdbio M30-7 DNA. These results confirm the assignment of the fragments and the conclusion that fragment B contains the $O_L$ operator and fragment D the $O_R$ operator.

The cleavage sites can be located with greater precision using other restriction enzyme cleavage sites as reference points. Fig. 4 shows repressor binding fragments produced by Bgl II (fragment B and D) cut again by Eco Rl and compared with the Eco Rl fragment selected with repressor (fragment 2) and cut again with Bgl II. As expected from the known cleavage sites of Eco Rl[8], this enzyme does not cut the Bgl II fragment D, but it cleaves fragment B into three pieces, one of which corresponds to the Eco Rl fragment 3 and the other two are slightly

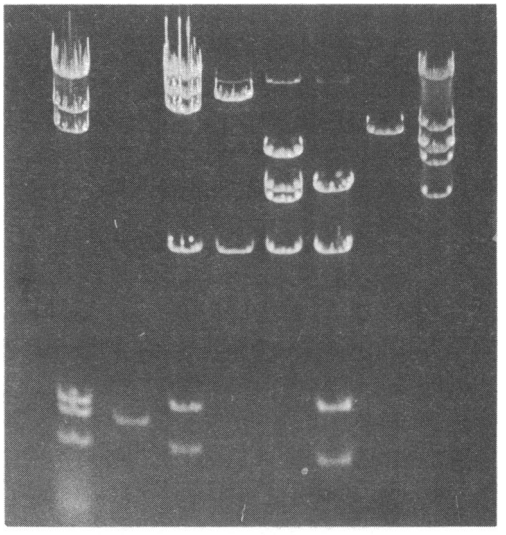


Figure 4. λ repressor binding fragments.
a) λ dbio M30-7 nin 5 DNA digested with Bgl II
b) repressor binding fragments from a)
c) λ DNA digested with Bgl II
d) repressor binding fragments from c)
e) repressor binding fragments from c) redigested with Eco RI
f) repressor binding fragments from h) redigested with Bgl II
g) repressor binding fragments from h)
h) λ DNA digested with Eco RI

larger than the Eco Rl fragment 6.

Bgl II on the other hand cuts the Eco Rl fragment 2 into four, generating again the Bgl II fragments D and E, a fragment containing the $O_L$ operator and a very small fragment of about 400 base pairs. The Bgl II fragment B contains therefore the Eco Rl cleavage sites 2 and 3 and comes within 400 base pairs of site 4.

Similar experiments using Hae III and Hind II place the right endpoint of the Bgl II fragment B about 90 base pairs to the right of the Hind II cleavage site in the $O_L$ operator.

Terminal labeling of Bgl I and Bgl II cleavage sites

λ DNA digested with Bgl I or with Bgl II was treated with

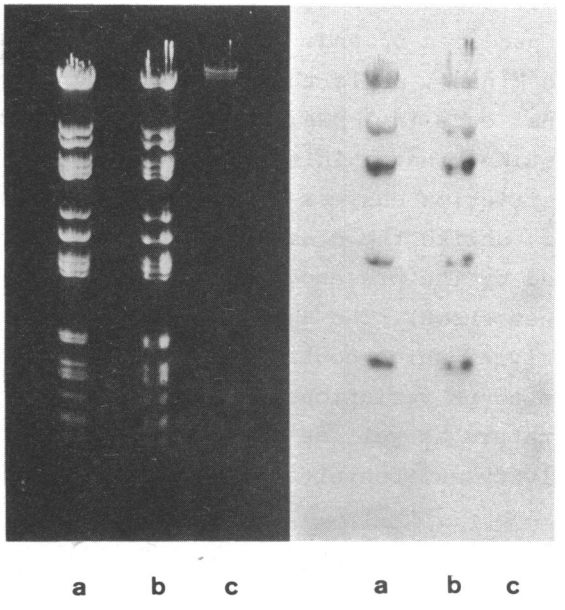


Figure 5. Agarose gel electrophoresis of λ DNA digested with Bgl I or Bgl II and terminally labeled. λ DNA was digested with Bgl I a) or Bgl II c) and then terminally labeled with polynucleotide kinase. b) shows the Bgl II digest retreated with Bgl I after terminal labeling. The panel to the left shows the ethidium bromide stained gel. To the right is the autoradiographed gel.

alkaline phosphatase and then labeled with polynucleotide kinase and $\gamma$-$^{32}$P-ATP. Fig. 5 shows agarose gel electrophoresis of the two labeled digests, stained with ethidium bromide or autoradiographed. Clearly, only some of the fragments produced by Bgl I are end labeled by the polynucleotide kinase, while all the Bgl II fragments are labeled. Furthermore, comparison of the two labeled digests shows that some of the labeled bands in the Bgl I digest correspond to some of the Bgl II bands and suggests that some of the Bgl I cutting sites correspond to the Bgl II sites. This hypothesis is verified by subjecting the Bgl II digest to Bgl I treatment. The pattern of fragments produced is identical to that of Bgl I alone. In addition the labeling

pattern is identical, indicating that the Bgl I preparation contains in fact two activities: one which is identical to that of Bgl II and produces 5' ends which can be phosphorylated by polynucleotide kinase, while the other, characteristic of Bgl I, makes additional cuts in λ DNA, whose 5' ends are not substrates for polynucleotide kinase. This surprising feature is reminiscent of class I restriction enzymes which also produce non-kinaseable ends[9]. However, unlike the class I restriction enzymes, Bgl I is not affected by the presence or absence of ATP or Ado Met (experiments not shown). The Bgl I activity is less durable than the Bgl II activity which contaminates it, and is in addition more sensitive to pH variations. Prolonged storage or exposure to room temperature as well as pH values below 6.5 inactivate the Bgl I activity and convert the enzyme preparation to pure Bgl II.

Sequence of the Bgl II cleavage sites

λ DNA was digested with Bgl II and terminally labeled with polynucleotide kinase. The labeled fragments were extensively digested with pancreatic DNase and then fractionated by electrophoresis on cellulose acetate pH 3.5, followed by homochromatography in the second dimension on a thin layer of DEAE cellulose.

The fingerprint pattern (fig. 6) shows only a few spots, consistent with a single 5' terminal sequence which begins to diverge only after the pentanucleotide. The spots were cut out from the chromatogram, eluted and analysed by partial digestion with venom phosphodiesterase, followed by electrophoresis on AE cellulose at pH 3.5 or DEAE cellulose pH 2. The results show that all the spots from the fingerprint share the same terminal sequences. An additional surprising result is that each spot from the fingerprint is resolved into two by the AE cellulose system (fig. 7). The two components which are not resolved in the fingerprint migrate too closely to differ by a nucleotide. They cannot be due to sequence isomerism since the 5' terminal is uniquely G and since the dinucleotide is already a double spot. Moreover, a single 5' end, obtained by cleaving a Bgl II fragment with Eco RI, produces the identical pattern of double spots, with each

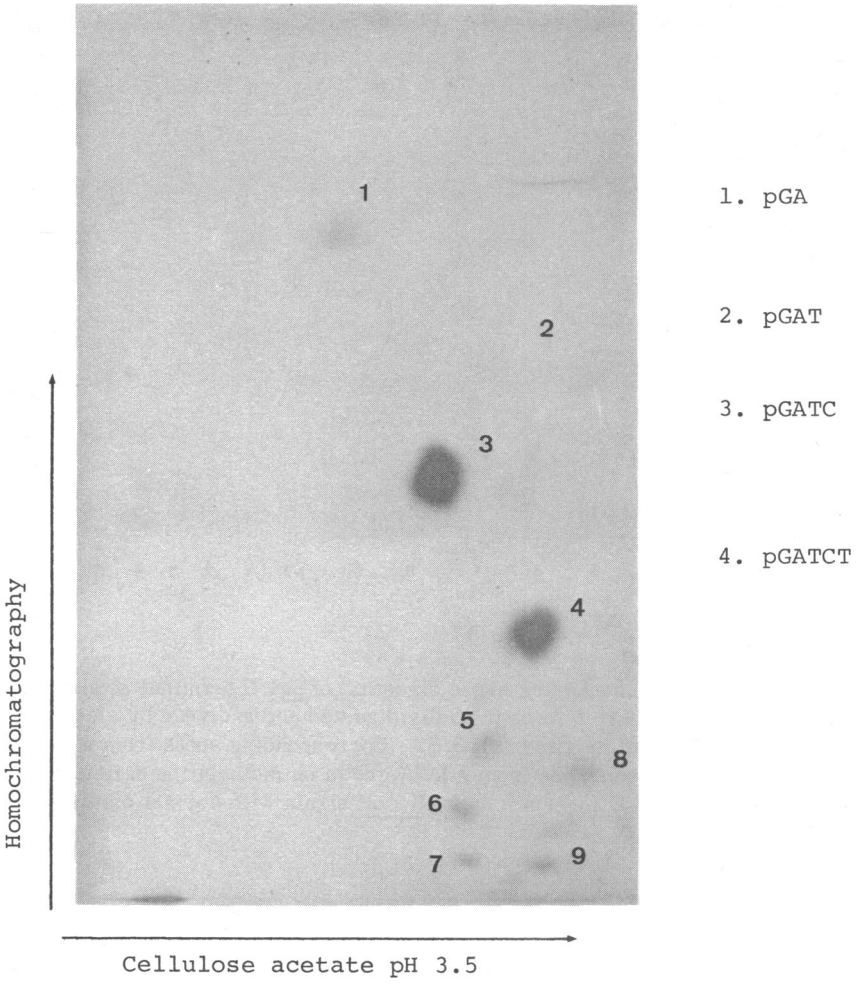


Figure 6. Fingerprint of Bgl II 5' terminal sequences. λ imm 434 DNA was used in this experiment. Sequences were deduced by partial digestion with venom phosphodiesterase and electrophoresis as shown in figure 7, or by two dimensional fingerprinting (not shown).

member of the pair having roughly equal strength. These results suggest that the doubling is due to some modification of the second nucleotide from the 5' end, for example methylation.

Aside from the doubling, the fingerprint pattern, together with the analysis of the partial venom digests, indicates that

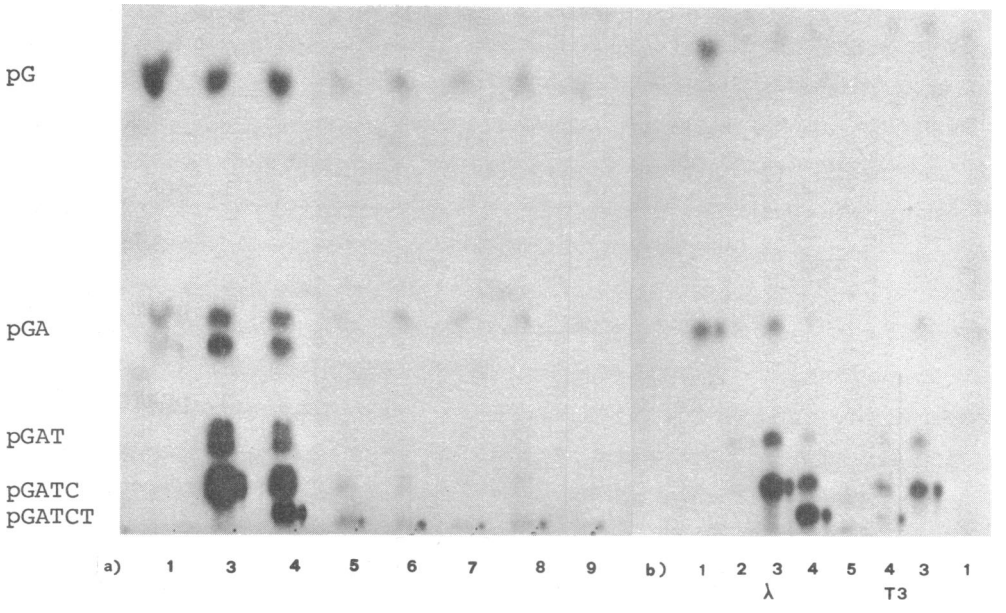


Figure 7. Partial venom phosphodiesterase digestion of Bgl II terminal sequences. a) Numbered spots from figure 6 were partially digested and analysed by electrophoresis on aminoethyl cellulose paper pH 3.5. Corresponding spots from a fingerprint of Bgl II 5' terminal sequences from λ DNA grown on methylation deficient E.coli dam dcm or from T3 DNA grown on an E.coli strain with normal methylation.

the sequence of the 5' termini is pGATCT.... Since only one 5' sequence is observed, each cleavage event must produce two identical 5' termini. The sequence recognised and cut by the enzyme could be 5'...AGATCGATCT...3' with a double strand cut in the middle. Or, since the first four nucleotides GATC have an axis of complementary symmetry in the middle, a simpler recognition sequence would be 5'...AGATCT...3' with staggered cuts.

The latter is the more likely possibility and is supported by the finding that the enzyme produces self complementary ends which, like the ends of Eco RI, BamH 1, Eco RII or Hind III can be used to join two fragments together (experiments not shown).

## Bgl II cleavage sites in λ are partially methylated

The doubling of the spots discussed in the previous section

could be explained if the Bgl II cleavage sites are partially methylated. To test this hypothesis, phage λ was grown on a strain of E. coli deficient in methylation of adenine[10]. An additional test was to use as substrate phage T3 DNA. This phage is known to induce an Ado Met hydrolysing activity[11]. Since Ado Met is the principal methyl donor for nucleic acids, T3 DNA is consequently free of methylated bases.

DNA extracted from λ grown on E. coli dam⁻ dcm⁻ and DNA extracted from phage T3 were treated with Bgl II, the kinase labeled and a fingerprint of the 5' terminal sequences obtained as before. Spots were further analysed by partial digestion with venom phosphodiesterase and electrophoresis on AE paper Ph 3.5 (fig. 7). As expected, the doubling of the spots is not present in either T3 or λ terminal sequences, confirming the interpretation that the Bgl II recognition sequence is normally methylated at the A residue on the 5' side of the cleavage, roughly 50 % of the time at any one site. The Bgl II enzyme is apparently able to cleave both the methylated and the unmethylated sequences.

## DISCUSSION

Bgl II is a useful addition to the repertoire of enzymes which dissect DNA into relatively large fragments. This enzyme produces sticky ends which make it possible to join these fragments to other fragments possessing the same sticky ends. The sequence recognised by the Bgl II enzyme is 5'...AGATCT...3' of which the
3'...TCTAGA...5'
central four nucleotides constitute the sticky ends. This sequence is in fact also the sequence of the sticky ends produced by the BamH I restriction enzyme[1] which recognises the sequence 5'...GGATCC...3'. It should be possible therefore to join frag-
3'...CCTAGG...5'
ments containing Bgl II ends to fragments containing BamH I ends. Due to the fact that the first and the sixth members of the hexanucleotide sequences recognised by these two enzymes are different, the product of such hybrid joining should not reconstitute a cleavage site for either enzyme and will therefore be

insensitive to either. Ligation of such hybrid molecules in the presence of both restriction enzymes should yield only the products of heteroaddition since Bam-Bam or Bgl-Bgl ends will be cleaved again by the enzymes present. This selection has been successfully applied to insert Bgl II fragments into plasmid DNA linearised by cleavage with BamH I (Pirrotta and Schedl, manuscript in preparation).

Different E. coli strains contain DNA methylase activities which show a certain degree of specificity. Methylases extracted from E. coli W, for example, can methylate DNA extracted from E. coli K, but are unable to further methylate E. coli W DNA, presumably because the target sites are already saturated with methyl groups[12]. However, the DNA of phage λ is apparently never completely saturated, possibly because it replicates faster than the methylases are able to act upon it. The methylating activity consists of at least two independent enzymes, one producing 6-Me adenine and the other 5-Me cytosine. The sequence results presented in this article indicate that the sequence recognised by the Bgl II enzyme is at least one of the sequences methylated by the adenine methylase of E. coli K. These results are consistent with those obtained by Hattman et al. (personal communication) who find that the sequences methylated in E. coli are 5'..NG$\overset{*}{A}$py$^{pypy}_{A\ A}$...3'of which the Bgl II sequence is a subset. The Bgl II enzyme cleaves methylated or unmethylated sites without apparent distinction, raising the question of which base the corresponding modification enzyme, if one exists, methylates to protect DNA from restriction. It is possible, as has been suggested by several workers, that many class II restriction enzymes are not members of a full fledged restriction-modification system such as those associated with Eco RK, Eco $R_I$ or Eco RII. In contrast, a cytosine methylase from E. coli K, which methylates the sequence NC$\overset{*}{C}$AGG and NC$\overset{*}{C}$TGG, renders these sites resistent to Eco RII cleavage[13]. The sequence recognised by Bgl II is also identical to that reportedly methylated by the phage P1 modification enzyme, yielding 5'...AG$\overset{*}{A}$TCT...3' (according to Brockes et al.[14]). The methylation observed in my experiments was not due

to phage P1. Two different E. coli strains, neither of them a P1 lysogen, were used as hosts for the growth of λ whose DNA turned out to be partially methylated in the Bgl II sequence. It is difficult at present to reconcile my results with those of Brockes et al.

I should point out in addition the curious case of the Dpn I and Dpn II restriction enzymes reported by Lacks and Greenberg[15]. Both of these enzymes recognise and cut the sequence GATC, the first enzyme when the sequence is not methylated, the second when it is methylated. It is interesting to note that this sequence is the same as the central part of the sequence recognised by Bgl II. Possibly this sequence is recognised by a methylase of unknown function which is widely spread in bacteria.

Finally, the Bgl I enzyme is one of an increasing number of anomalies among class II restriction enzymes. This enzyme cleaves DNA at specific sites by an unknown mechanism which produces 5' terminals which are not substrates for polynucleotide kinase even after treatment with alkaline phosphatase. The molecular structure of these terminals is being currently investigated and may provide a clue to the mechanism of cleavage of type I restriction enzymes.

REFERENCES

1) Roberts, R., personal communication
2) Catterall, J. and Welker, N., personal communication
3) Pirrotta, V. (1973) Nature New Biol. 244, 13-16
4) Richardson, C.C. (1971) in Procedures in Nucleic Acids Research, vol. II, pp. 815-828, Harper and Row
5) Glynn, I.M. and Chappell, J.B. (1964) Biochem. J. 90, 147-149
6) Jay, E., Bambara, R., Padmanabhan, R. and Wu, R. (1974) Nucleic Acids Res. 1, 331-353
7) Murray, K. (1973) Biochem. J. 131, 569-583
8) Thomas, M. and Davis, R.W. (1975) J. Mol. Biol. 91, 315-328
9) Murray, N.E., Batten, P.L. and Murray, K. (1973) J. Mol. Biol. 81, 395.407
10) Marinus, M.G. and Morris, N.R. (1973) J. Bacteriol. 114, 1143-1150
11) Gefter, M.L., Hausmann, R., Gold, M. and Hurwitz, J. (1966) J. Biol. Chem. 241, 1995-2005
12) Gold, M. and Hurwitz, J. (1963) Cold Spring Harbor Symp. Quant. Biol. 28, 149-156

13) Hughes, S.G. and Hattman, S. (1975) J. Mol. Biol. 98, 645-647
14) Brockes, J.P., Brown, P.R. and Murray, K. (1974) J. Mol. Biol. 88, 437-443
15) Lacks, S. and Greenberg, B. (1975) J. Biol. Chem. 250, 4060-4066