

Overexpression, purification and crystallization of *Bam*HI endonuclease

William E. Jack, Lucia Greenough, Lydia F. Dorner, Shuang-yong Xu, Teresa Strzelecka¹, Aneel K. Aggarwal¹ and Ira Schildkraut*

New England Biolabs, Inc., 32 Tozer Road, Beverly, MA 01915 and ¹Department of Biochemistry and Molecular Biophysics, Columbia College of Physicians and Surgeons, 630 W 168th St., New York, NY 10032, USA

Received December 28, 1990; Revised and Accepted March 16, 1991

ABSTRACT

The type II restriction endonuclease *Bam*HI has been expressed in *E. coli*, producing 100-fold more enzyme than the wild type *Bacillus amyloliquefaciens* H strain. This high yield has facilitated purification to homogeneity of large amounts of the enzyme, along with its crystallization in a form which diffracts to at least 1.9 Å in X-ray analysis.

INTRODUCTION

Type II restriction endonucleases are accessible examples of proteins which interact site-specifically with nucleic acids. These enzymes require only a divalent cation and a short DNA recognition sequence for action, and produce an easily assayed product. With several restriction endonucleases site-specific DNA binding is observed in the absence of the cation cofactor (absence of cleavage), allowing these endonucleases to serve as prototypes for sequence specific binding interactions (1,2). Although many restriction endonucleases have been purified for use as reagents in molecular biology, few have been biochemically characterized. Published crystal structures are available for only the *Eco*RI endonuclease (3), although several other structures are in progress (4; J. Anderson, personal communication).

The *Bam*HI endonuclease was one of the first type II restriction endonucleases discovered (5) and has been the subject of considerable biochemical and kinetic characterization (6). The endonuclease from *Bacillus amyloliquefaciens* H cleaves the duplex DNA sequence 5'-GGATCC-3' between adjacent guanine residues in both strands to generate a 4 nucleotide 5' extension (5). The *Bam*HI endonuclease coding sequence predicts a 213 amino acid protein with a molecular weight of 24,570 daltons (7). Gel filtration and ultracentrifugation studies have shown the existence in solution as homodimeric and tetrameric species (6). In combination with kinetic studies, these values indicate that the dimer is a catalytically active form of the endonuclease.

To facilitate our studies of the interactions of *Bam*HI endonuclease with its substrate we have constructed an over-producing strain of *E. coli*, developed a purification scheme

allowing us to purify large amounts of the native enzyme and its variants to levels of purity which are suitable for X-ray crystal analysis as well as detailed biochemical studies on synthetic oligonucleotide substrates. The purified enzyme has been crystallized, and we report here preliminary diffraction parameters.

MATERIALS AND METHODS

Enzymes, oligonucleotides, reagents

All restriction endonucleases and nucleic acid modifying enzymes were from New England Biolabs. Oligonucleotides used for site-directed mutagenesis and dideoxy DNA sequencing were from the Organic Synthesis Division of New England Biolabs, and were synthesized on an Applied Biosystems DNA synthesizer model 380B. NaDodSO₄ polyacrylamide gradient gels (10–20%) were from Integrated Separation Systems and were run in Tris glycine buffer. Protein standards for gel filtration analysis were from Pharmacia.

Strains and media

Luria-Bertani (LB) medium and LB agar have been described previously (8). Where cells contained plasmids coding for ampicillin resistance media was supplemented with 0.1 mg/ml ampicillin. *E. coli* strain ADK21 is a derivative of K802 (*lac3 galK2 galT22 metB1 hsdR2 supE44 mcrB1 mcrA1*; ref. 9) which expresses the *Bam*HI methylase constitutively. The methylase gene (*bamHIM*) is carried in single copy on a lambda prophage (lambda *imm*⁴³⁴ *ind*⁻ *bamHIM*) within the cell (L.G. and W.E.J., unpublished).

Plasmid preparation and manipulation

Plasmid DNAs were prepared on a small scale by minor variations of a boiling protocol (10) and on a large scale by alkaline lysis followed by CsCl banding (11). ADK21 cells were made competent by CaCl₂ treatment of cells (11).

* To whom correspondence should be addressed

Expression plasmid construction

Plasmid pPR594, a derivative of pTAC12 (12), was constructed and generously provided by P. Riggs (New England Biolabs, Beverly, MA). pPR594 carries the β -lactamase gene, *lacI* gene, *lac* ribosome binding site, *rrnB* transcription terminator (13), and *NcoI* and *HindIII* sites downstream of the P_{tac} promoter and *lac* ribosome binding site (Figure 1). Overexpression of the *BamHI* endonuclease was accomplished by replacing the initiation codon of the endonuclease gene with that within the *NcoI* site of pPR594, thus utilizing the *lac* ribosome binding site for translation. This step was facilitated by oligonucleotide-directed mutagenesis (14) to create an *EcoRI* site at codons 6 and 7, followed by ligation of a double-stranded synthetic oligonucleotide which introduced an *NcoI* site overlapping the initiation codon (Figure 1). These alterations, and others introduced to alter codon usage, were translationally silent. An *NcoI* to *BglII* DNA fragment containing the entire *bamHIR* gene was then inserted into the *NcoI* and *BamHI* sites of a pPR594 derivative which contained a 700 bp *HindIII/BamHI* fragment, creating the overexpression plasmid pAEK14 (Figure 1). The identity of sequences which were modified in the construction were confirmed by dideoxy chain termination sequencing using double stranded templates (15).

Assays of *BamHI* unit activity

One unit of *BamHI* is defined as the amount of enzyme which produces a complete digest of 1 μ g of lambda DNA at 37°C in a reaction volume of 50 μ l in 60 minutes in *BamHI* reaction buffer (0.1 M NaCl, 6 mM Tris-HCl (pH 7.8), 6 mM MgCl₂, 1 mM dithiothreitol). Protein concentrations were determined using the Bradford assay with bovine serum albumin as a standard (16).

Assays of contaminating exonuclease and endonuclease activity

Exonuclease activity was quantified via release of acid soluble radiolabel from a uniformly labeled substrate. Samples were incubated in *BamHI* reaction buffer with 1 μ g of [³H] labeled *E. coli* DNA (200,000 cpm/ μ g) in 50 μ l at 37°C for 16 hrs. DNA was then precipitated by adding herring sperm DNA to a final concentration of 1.3 mg/ml and TCA to 5% final concentration, mixing and incubating on ice for 10 minutes. After centrifugation for 5 minutes at 4°C and 16,000 \times g, liberated [³H] nucleotide in the supernatant was quantified by liquid scintillation counting.

Nonspecific endonuclease activity was detected by incubating samples overnight at 37°C in 50 μ l *BamHI* reaction buffer containing 1 μ g covalently closed circular (RFI) Φ X174, which lacks *BamHI* sites. Nonspecific nicking and cleavage resulted in the appearance of RFII and RFIII forms, respectively, which were resolved by agarose gel electrophoresis.

BamHI endonuclease purification

ADK21 cells carrying pAEK14 were grown at 37°C to 65 Klett units. *BamHI* endonuclease expression was induced by addition of isopropylthiogalactoside (IPTG) to 0.4 mM and incubation was continued for 3 hours. Cells were collected by centrifugation and stored at -70°C.

80 grams of cells were thawed and suspended at 4°C in 400 ml of 10 mM potassium phosphate buffer (pH 6.9), 1 mM DTT, 1 mM EDTA, 1 mM EGTA (Buffer A) containing 0.15 M NaCl and 1 mM phenylmethylsulfonyl fluoride. All subsequent procedures were done at 4°C. The cells were sonicated and the debris removed by centrifugation at 20,000 \times g for 1 hour. The

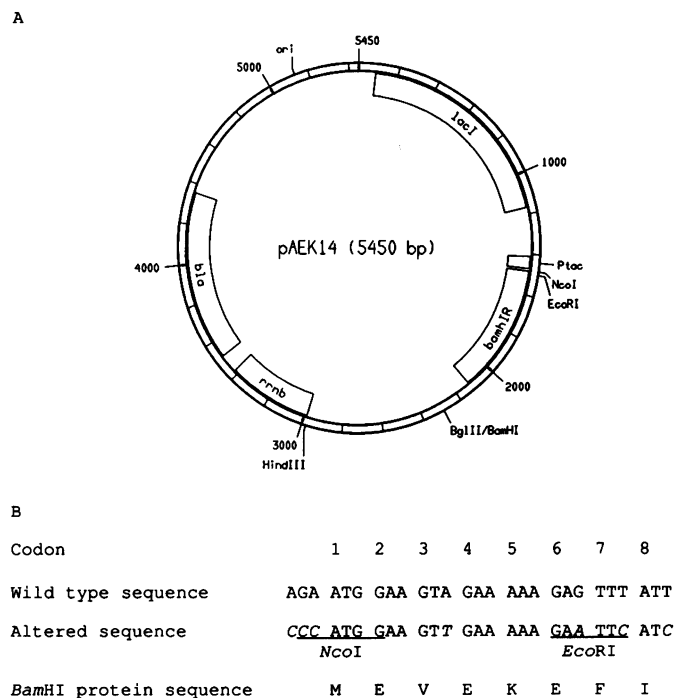


Figure 1. (A) Map of pAEK14, a plasmid constructed for the overexpression of *BamHI* endonuclease. The *bamHIR* gene is oriented clockwise under P_{tac} promoter control. *Bla*, *lacI* are the genes coding for β -lactamase and *Lac* repressor, respectively. Transcriptional terminators (*rrnB*) were introduced distal to the *bamHIR* gene. Numbers are in base-pairs. (B) The first 8 codons of *bamHIR* gene. Translationally silent alterations (shown as italic letters) were introduced as described in Materials in the 3rd, 6th, 7th, and 8th codons during pAEK14 construction. The *NcoI* and *EcoRI* sites are underlined.

supernatant was applied to a 5 \times 8 cm phosphocellulose column (Whatman P11) equilibrated in Buffer A containing 0.15 M NaCl and 10% (v/v) glycerol.

After washing to remove unbound material, the activity was eluted with a 1500 ml gradient from 0.15 to 1.0 M NaCl. *BamHI* endonuclease activity eluted between 0.40 and 0.55 M NaCl. Peak tubes were pooled and loaded directly onto a 2.5 \times 9 cm hydroxylapatite column (Calbiochem) equilibrated in Buffer A containing 0.5 M NaCl and 10% glycerol. The column was washed and eluted with a 500 ml linear gradient of potassium phosphate (pH 6.9) from 0.01 to 0.5 M. Peak endonuclease activity eluted between 0.15 and 0.25 M potassium phosphate. After dialysis against Buffer A containing 0.15 M NaCl and 10% glycerol, the enzyme was applied to a 2.5 \times 11 cm Heparin-Sepharose column (Pharmacia) equilibrated in the same buffer, and eluted with a 500 ml gradient from 0.15 to 1.0 M NaCl. The active fractions, which eluted between 0.7 and 0.8 M NaCl, were pooled and dialysed against 50 mM KCl, 10 mM Tris-HCl (pH 7.8), 1 mM DTT, 1 mM EDTA, 1 mM EGTA, 10% glycerol (Buffer B) and loaded onto a 2.5 \times 5 cm Q-Sepharose column (Pharmacia) equilibrated in Buffer B. The column was washed and eluted with a 500 ml gradient from 0.05 to 1.0 M NaCl. The peak of restriction endonuclease activity, which eluted between 0.4 and 0.5 M KCl, was pooled and stored at -70°C.

Beginning with the hydroxylapatite column, representative fractions were also assayed for contaminating exonuclease and endonuclease activities, as described above.

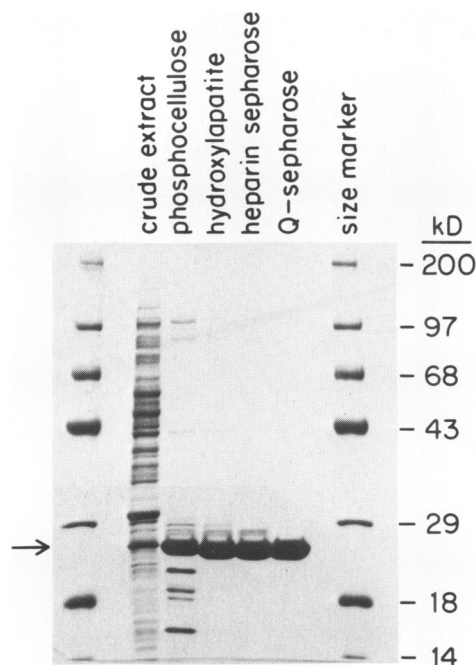


Figure 2. Purification of wild type *Bam*HI endonuclease. Protein samples collected from each purification step were dissolved in Laemmli protein sample buffer, heated for 2 min at 90°C and loaded into a 10–20% NaDodSO₄-polyacrylamide gradient gel. Electrophoresis was carried out in a buffer containing 25 mM Tris, 192 mM glycine, 0.1% NaDodSO₄ at 100 volt (constant voltage) for 4 hours. Arrow indicates the *Bam*HI protein (24,500 daltons). kd, kilodaltons.

Crystallization

The *Bam*HI protein was concentrated about ten-fold to 25 mg/ml by centrifugation in a Centricon chamber (Amicon) and stored at –70°C. Crystallization experiments were performed by the hanging drop vapor diffusion method (17,18). Large crystals were reproducibly obtained by mixing 1 μ l of *Bam*HI protein (~25 mg/ml in 0.5 M KCl, 20 mM KPO₄ (pH 6.9), 1 mM DTT, 10% glycerol) with 1 μ l of ‘seeding’ solution (12% PEG 8000, 0.25 M KCl, 5% glycerol, 0.2% NaN₃, a crushed crystal) and equilibrating the resulting drop over a reservoir containing 12% PEG 8000, 20 mM KPO₄ (pH 6.9), 10% glycerol, 0.5 M KCl at 20°C.

X-ray analysis

The crystals were initially characterized by a series of screened and screenless precession pictures (18) taken on a Elliot GX-6 rotating anode X-ray generator. The full diffraction limit of the crystal was assessed on a Xuong-Hamlin area detector (19) operating on a Rigaku RU200 rotating anode X-ray generator.

Aggregation state analysis

Gel filtration was carried out on a Superose 12 column (10 \times 300 mm, Pharmacia) using 150 mM KCl, 20 mM Tris-HCl (pH 7.8), 5 mM b-mercaptoethanol, 5% glycerol. 100 μ l of *Bam*HI protein (50 μ g protein in 0.45 M NaCl, 10 mM Tris-HCl (pH 7.8), 1 mM DTT, 1 mM EDTA, 1 mM EGTA) was loaded, and elution detected by monitoring UV absorbance at 254 nm. Molecular weight standards used were chymotrypsinogen A (25,000 daltons), ovalbumin (43,000 daltons), bovine serum albumin (67,000 daltons).

Table 1. *Bam*HI purification

Purification step	Total protein (mg)	Total activity (units)	Specific activity (units/mg)	Purification (-fold)	Yield (%)
Crude extract	4200	7×10^8	1.7×10^5	1	100
Phosphocellulose	340	4×10^8	1.2×10^6	7	57
Hydroxylapatite	180	3×10^8	1.7×10^6	10	43
Heparin-sepharose	60	1×10^8	2.0×10^6	12	17
Q-sepharose	48	9×10^7	2.1×10^6	13	13

RESULTS

Overproduction of *Bam*HI endonuclease

Crude extracts of *Bacillus amyloliquefaciens* H contain about 3.0×10^4 units of *Bam*HI endonuclease per gram wet cell paste. Following induction with IPTG, ADK21 [pAEK14] cells yielded about 5×10^6 units per gram wet cell paste, an increase of 100-fold over the native source. The enzyme represented about 5% of the total cellular protein and was easily visualized by Coomassie brilliant blue staining of denaturing polyacrylamide gels (Figure 2). *Bam*HI endonuclease production in uninduced cells was still significant, yielding about 10% of the activity seen in the induced state. This basal level of expression caused a 10^3 decrease in transformation into cells lacking the *Bam*HI methylase.

Purification of wild type *Bam*HI endonuclease

After four chromatographic steps (Table 1) the *Bam*HI endonuclease appeared as a single polypeptide on Coomassie blue staining of NaDodSO₄-PAGE (Figure 2), indicating >99% homogeneity of the preparation. The preparation was free of detectable non-specific endo- and exonuclease activities as determined by incubation overnight of 250 units of purified *Bam*HI with 1 μ g of substrate DNA in 50 μ l as described in methods.

Comparison of *Bam*HI endonuclease isolated from native and recombinant sources

A major consideration before using the recombinant *Bam*HI endonuclease for physical characterization was verifying that it behaved as the wild type protein. *Bam*HI endonuclease isolated from the recombinant had a specific activity of 2.1×10^6 u/mg (Table 1), comparing favorably with the specific activity of 2.5×10^6 u/mg derived for the enzyme isolated from *Bacillus amyloliquefaciens* (H.J. Benner, personal communication). Correct translation initiation and protein processing was verified by amino terminal sequencing of the recombinant protein (7) and (L.D. and J. Benner, personal communication). Previous experiments with *Bam*HI endonuclease isolated from the native organism showed the protein to exist as a dimer at 0.5 M NaCl (6). Consistent with this result, we also find a single protein peak corresponding to 45,000 dalton dimeric species on FPLC gel filtration on Superose 12 (data not shown). As observed with the native protein, endonuclease purified from the recombinant displays star activity in the presence of Mn²⁺ when the enzyme is in large excess (100 units *Bam*HI to cleave 1 μ g lambda DNA, Figure 3; reference 20). In addition, the recombinant showed the same star activity as the native protein in the presence of 30%

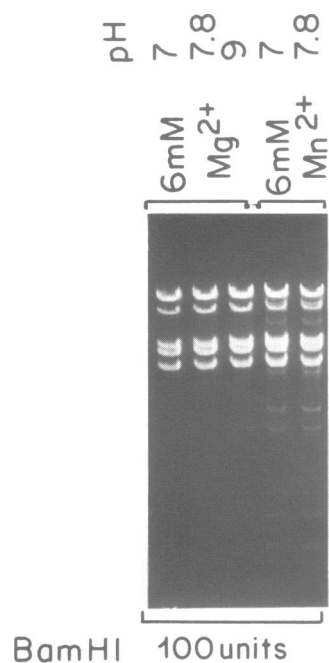


Figure 3. *Bam*HI cleavage reactions in the presence of Mg^{2+} or Mn^{2+} . Reactions were performed on lambda DNA under standard assay conditions using 100 units of *Bam*HI endonuclease and cleavage was assayed by agarose gel electrophoresis. As indicated above the lanes, samples contained either Mg^{2+} or Mn^{2+} as the divalent cation. pH values in the individual reactions were also varied, as indicated above the lanes. Glycerol concentration was less than 10% in all cleavage reactions.

glycerol (data not shown). Star activity, however, was not detected in Mg^{2+} buffer at pH 7, 7.8, or 9 (Figure 3). In all assays we have performed, *Bam*HI endonuclease purified from the recombinant was indistinguishable from that from the native organism.

Crystallization

Crystals were formed with several precipitants (PEG 8000, PEG 20,000 and ammonium sulfate) and appeared in a few days. The reproducibility of large crystals was greatly improved by 'seeding' techniques. The morphology of the crystals was commonly thin plates (up to 1.0 mm \times 0.6 mm \times 0.12 mm, see Figure 4A), although cubes (up to 0.2 mm on edge, see Figure 4B) were also observed. Cocrystals between *Bam*HI and DNA have also been prepared and will be described elsewhere (T.S., L.D., I.S. and A. A., manuscript in preparation).

Preliminary X-ray analysis

Crystals diffracted to at least 1.9 Å in resolution and occurred in monoclinic (Figure 5) and orthorhombic forms. The monoclinic form belongs to space group C2, with unit cell constants $a = 76.4$ Å, $b = 46.0$ Å, $c = 69.4$ Å, and $\beta = 110.5^\circ$. The crystal volume per unit molecular weight, V_m , assuming a monomer of *Bam*HI in the crystallographic asymmetric unit is 2.5 Å³/dalton, close to the values observed for globular proteins (21). The orthorhombic form belongs to space group C222₁, with unit cell constants $a = 46.7$ Å, $b = 76.6$ Å, and $c = 143.6$ Å, and the calculation of V_m values (2.6 Å³/dalton) again argues for the existence of a monomer of *Bam*HI per crystallographic asymmetric unit. *Bam*HI binds to DNA as a dimer and exists in solution as dimer and tetrameric forms

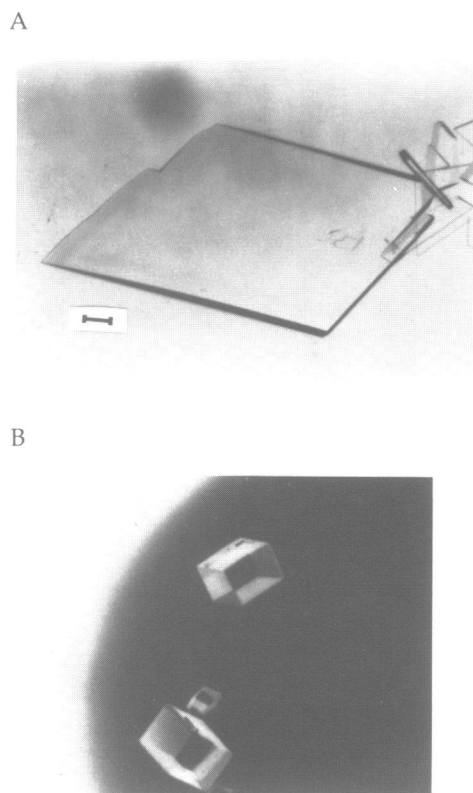


Figure 4. (A) A plate like crystal of *Bam*HI endonuclease protein. The scale bar corresponds to 100 microns. (B) Cube shaped crystals of *Bam*HI observed occasionally.

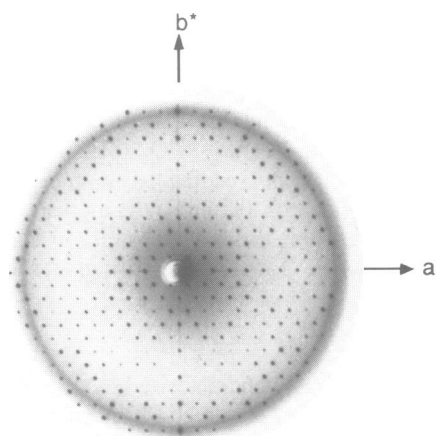


Figure 5. A 12° screened precession photograph of the $hk0$ zone of the monoclinic form of *Bam*HI crystals taken on an Enraf-Nonius precession camera. X-rays were produced by a Rigaku RU200 rotating anode generator and selected for Cu-K α by a nickel filter. The crystal to film distance was 10.0 cm.

depending on salt concentration (6). Presumably, the crystallographic two-fold axes of both space groups coincide with the two-fold axis of the dimer of the protein. X-ray data sets on the monoclinic crystals (1.95 Å) and several heavy atom derivatives (3.0–3.5 Å) have been collected on the Xuong-Hamlin area detector, and determination of the three dimensional structure of the enzyme is well underway.

DISCUSSION

In order to perform structural studies we have created a recombinant which produces >100 fold more *Bam*HI endonuclease than the original *Bacillus amyloliquefaciens* H strain. This has allowed us to purify large amounts of homogeneous *Bam*HI endonuclease with properties identical to that isolated from the original host strain. Crystals have been grown from the purified protein which produce high quality X-ray diffraction patterns. These patterns are being used to determine the three-dimensional structure of the native protein and of the DNA-protein complex.

The *Bam*HI recognition sequence and the sites of cleavage possess two-fold rotational symmetry. One way to exploit such sequence symmetry is for the *Bam*HI endonuclease to interact symmetrically as a dimer with the two half sites. Consistent with this idea, the purified *Bam*HI endonuclease exists as a dimer in solution. This observation, coupled with the simple kinetic behavior (6,22), indicates that the dimer is a catalytically active form of the enzyme. In this respect it resembles the *Eco*RI endonuclease which also recognizes a two-fold symmetric site and has been shown to interact as two equivalent monomers (3,23).

The relatively high basal level of expression of pAEK14 has also been used to select *Bam*HI endonuclease variants with catalytic defects. The transformation efficiency of the *Bam*HI overexpression plasmid pAEK14 in *E. coli* strain lacking the gene coding for the *Bam*HI methylase is extremely low, presumably reflecting lethal expression of the endonuclease. *Bam*HI variants with reduced cleavage activities were isolated by mutagenizing the *Bam*HIR gene on the pAEK14 plasmid, introducing the mutagenized plasmid into an *E. coli* strain lacking the methylase and selecting for surviving transformants. Such studies have identified Glu77, Asp94, and Glu113 residues as being critical for catalytic activity (24).

Type II restriction endonucleases require Mg^{2+} as a cofactor for cleavage activity. Mn^{2+} can replace Mg^{2+} in the *Eco*RI cleavage reaction, but the endonuclease displays decreased specificity in the presence of Mn^{2+} , also known as *Eco*RI star activity (25). *Bam*HI endonuclease shows a similar relaxation of specificity in buffers which contain Mn^{2+} (20). *Bam*HI star activity (also termed *Bam*HI.1) has also been observed in the presence of organic perturbants such as glycerol, ethylene glycol, ethanol, or dioxane (6,26,27). The mechanism underlying star activity is unknown, but presumably involves altered DNA-protein interactions. Variants have also been isolated which differ in their sensitivity to Mn^{2+} perturbation (S.Y.X and I.S., manuscript in preparation), allowing us to explore the nature of the altered interaction both in vitro and through crystallographic studies.

ACKNOWLEDGEMENTS

We thank Joan Brooks for providing the *Bam*HIR sequence, Paul Riggs for the pPR594 plasmid, Jack Benner for advise on the FPLC gel filtration experiment and amino acid sequence determinations, Laurie Mazzola and Barton Slatko for DNA sequence determinations, Rebecca Kucera for stimulating discussions, Joan Brooks, Elisabeth Raleigh and Richard Roberts for critical reading of the manuscript. Part of the work was supported by a grant from National Institute of Health (GM 44006) to A.K.A.

REFERENCES

1. Terry, B.J., Jack, W.E., Rubin, R.A. and Modrich, P. (1983) *J. Biol. Chem.* 258, 9820–9825.
2. Frankel, A.D., Ackers, G.K. and Smith, H.O. (1985) *Biochemistry* 24, 3049–3054.
3. Kim, Y.C., Grable, J.C., Love, R., Greene, P.J. and Rosenberg, J.M. (1990) *Science* 249, 1307–1309.
4. D'Arcy, A., Brown, R.S., Zabeau, M., van Resandt, R.W. and Winkler, F.K. (1985) *J. Biol. Chem.* 260, 1987–1990.
5. Roberts, R.J., Wilson, G.A. and Young, F.E. (1977) *Nature* 256, 82–84.
6. Nardone, G. and Chirikjian, J.G. (1987) in *Gene Amplification and Analysis* (Chirikjian, J. ed), Elsevier Science Publisher, New York, 5, 147–184.
7. Brooks, J.E., Nathan, P.D., Landry, D., Szynter, L.A., Waite-Rees, P., Ives, C.C., Mazzola, L.M., Slatko, B.E. and Benner, J.S. (in press, *Nucleic Acids Res.*).
8. Miller J.H. (1972) *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
9. Woods, W.B. (1966) *J. Mol. Biol.* 16, 118–133.
10. Holmes, D.S. and Quigley, M. (1981) *Anal. Biochem.* 114, 193–197.
11. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) *Molecular Cloning*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
12. Amann, E., Brosius, J. and Ptashne, M. (1983) *Gene* 25, 167–178.
13. Brosius, J. (1984) *Gene* 27, 151–160.
14. Kunkel, T.A., Roberts, J.D. and Zakour, R.A. (1987) *Methods Enzymol.* 154, 367–382.
15. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA* 74, 5463–5467.
16. Bradford, M. (1976) *Anal. Biochem.* 72, 248–254.
17. McPherson, A. (1982) *Preparation and Analysis of Protein Crystals*, John Wiley and Sons, New York.
18. Blundell, T. and Johnson, L.N. (1976) *Protein Crystallography*, Academic Press, New York.
19. Xuong, N.-H., Freer, S.T., Hamlin, R., Nielsen, C. and Vernon, W. (1978) *Acta Cryst.* A34, 289–296.
20. Malygin, E.G. and Zinoviev, V.V. (1989) *Sov. Sci. Rev. Physiochem. Biol.* 9, 87–142.
21. Matthews, B.W. (1968) *J. Mol. Biol.* 33, 491–497.
22. Hensley, P., Nardone, G., Chirikjian, J.G. and Wastney, M.E. (1990) *J. Biol. Chem.* 265, 15300–15307.
23. Lu, A.L., Jack, W.E. and Modrich, P. (1981) *J. Biol. Chem.* 256, 13200–13206.
24. Xu, S.-Y. and Schildkraut, I. (1991) *J. Biol. Chem.* 266, 4425–4429.
25. Hsu, M. and Berg, P. (1978) *Biochemistry* 17, 131–138.
26. George, J. and Chirikjian, J. G. (1982) *Proc. Natl. Acad. Sci. USA*, 79, 2432–2436.
27. George, J., Blakesley, R.W., Chirikjian, J.G. (1980) *J. Biol. Chem.* 255, 6521–6524.