RELAXED SPECIFICITY OF ENDONUCLEASE *Bam*H1 AS DETERMINED BY IDENTIFICATION OF RECOGNITION SITES IN SV 40 AND pBR 322 DNAs

V. A. KOLESNIKOV, V. V. ZINOVIEV, L. N. YASHINA, V. A. KARGINOV, M. M. BACLANOV and E. G. MALYGIN

Institute of Research in Molecular Biology, Glavmicrobioprom, Novosibirsk 90, USSR

Received 6 July 1981

1. Introduction

The specificity of restriction nucleases, which is very high under conventional conditions, may become 'relaxed' in certain media (e.g., at decreased ionic strength, in the presence of organic solvents) [1]. The 'condition-relaxed' recognition sites usually correspond to shortened or degenerate sequences derived from the 'canonic' ones [2–7]. The reasons for relaxation are unclear, although they are of great interest for understanding the mechanism of restriction nuclease action in general.

Up to now, the structure of 'relaxed' sites has been determined for only 3 endonucleases: *Eco*RI, *Bsu*1, *Bst*1 [2,3,7].

This communication is concerned with the determination of the structure of 'relaxed' restriction sites for the endonuclease *Bam*H1. These sites in SV 40 and pBR 322 DNAs appeared to be GGATC and GPuATCC (cf. canonic site GGATCC).

2. Materials and methods

Restriction endonucleases *Bam*H1, *Msp*1, *Bsp*1 and *Sau*3A were isolated as in [8–11]. SV 40 DNA was generously supplied by R. A. Gybadullin. DNA of pBR 322 and λC1857 were isolated as in [12,16]. Modified conditions of reaction for *Bam*H1: 0.02 M Tris–HCl, (pH 8.5), 5 mM $MgCl_2$, 6 mM 2-mercaptoethanol, 20% glycerol, 8% dimethylsulfoxide, 0.5 pmol/μl of SV 40 or pBR 322 DNA, 0.2 units/μl of enzyme. The incubations were performed for 2–4 h at 37°C. Digestion with other enzymes was performed as in [8–11]. The fragments resulting from digestion of pBR 322 DNA with *Msp*1 were resolved by electrophoresis on 4% polyacrylamide gel and isolated as in [13].

3. Results and discussion

To determine the sequence of nucleotides recognized by *Bam*H1 under the modified reaction conditions (*Bam*H1*-specificity, by analogy with *Eco*RI* [2]), we used two DNAs with known sequences: SV 40 and pBR 322 [14,15]. The fragments resulting from digestion of λC1857 DNA with *Bsp*1 were used as length standards [16]. Other fragment lengths were estimated from their mobilities in gels by comparison with these standards and with *Sau*3A-fragments of SV 40 and pBR 322 DNA as illustrated in fig.1.

Patterns of the *Bam*H1* digestion were analysed by computer program based on the approach proposed in [17], but adapted for the analysis of the products of incomplete hydrolysis. The program looks for the complete or partial palindromic sequences potentially recognizable by restriction endonuclease in DNA with known primary structure. The recognition sites are deduced by scanning the original DNA sequence for the occurence of patterns which are repeated at intervals corresponding to experimental fragment lengths. This program takes into account that the biochemical procedure gives approximate values for these fragment lengths. According to our data (see also [17]), mean relative errors for the fragment lengths are chosen to be 10% in the interval 100–1000 basepairs, and 15% outside of this interval.

As shown for endonucleases *Eco*RI, *Bsu*1 and *Bst*1, part of the normal recognition sequences are preserved in relaxed recognition sequences [2,3,7]. This allows one to limit the search, taking into account only degenerated *Bam*H1-recognition sites produced in the program by changing one or several nucleotides in the normal sequence GGATCC.

Computer analyses of the digestion patterns of SV

Published by Elsevier/North-Holland Biomedical Press
00145793/81/0000–0000/$02.50

40 and pBR 322 DNAs revealed two recognition sites:

$$GGATC\begin{matrix}C\\T\\A\\G\end{matrix} \quad \text{and} \quad G\begin{matrix}G\\A\end{matrix}ATCC \qquad \text{(Table 1)}$$

Information concerning the new cleavage sites had to be directly confirmed. For this purpose we isolated two fragments of pBR 322 DNA cleaved by endonucleases *Msp*1 and *Bsp*1. Fragment Msp-622 is located between positions 3902 and 162 of the pBR 322 genome, and the fragment Msp/Bsp-365 is located between positions 3045 and 3410 [15]. Fragment Msp-622 was ^{32}P-labeled at both the 5′-ends, fragment Msp/Bsp-365 was labeled only at the 5′-end of the *Msp*1-site.

^{32}P-Label was introduced using [γ-^{32}P]ATP and polynucleotidekinase [13]. Fragment Msp/Bsp-365 was subjected to: (i) Maxam–Gilbert sequencing procedure; (ii) *Bam*H1*-hydrolysis.

Sizing of the products obtained from the Msp/Bsp-365 fragment by action of *Bam*H1 under modified conditions (fig.2) showed that the fragment was cut at the indicated positions of the sequences G$^{\downarrow}$GATCT and A$^{\downarrow}$GATCC. The fragment Msp/Bsp-365 contains the sequences TGATCT and CGATCT having 4 internal nucleotides identical with that of the canonic *Bam*H1 site; these sequences are completely resistant to hydrolysis.

Msp-622 fragment was digested with *Bam*H1* and *Sau*3A. A comparison of the patterns (not shown) confirmed that *Bam*H1* cleaved the sequences

Table 1
The predicted and experimental fragment lengths (in base pairs) after incomplete *Bam*H1*-hydrolysis of pBR 322 and SV 40 DNAs

	DNA pBR 322 length		DNA SV 40 length	
N	Expt.	Pred.	Expt.	Pred.
1	3550	3618	3600	3660
2	3250	3314	3230	3226
3	3000	3070	2950	2990
4	2670	2691	2500	2632
5	2340	2326	2300	2236
6	2100	2128	2000	2000
7	1950	1944	1670	1660
8	1730	1739	1550	1566
9	1600	1558	1330	1330
10	1430	1449	1260	1264
11	1350	1374	1110	1160
12	1300	1292	1080	1076
13	1000	1049	700	670
14	930	927	500	490
15	870	866	400	396
16	730	728		

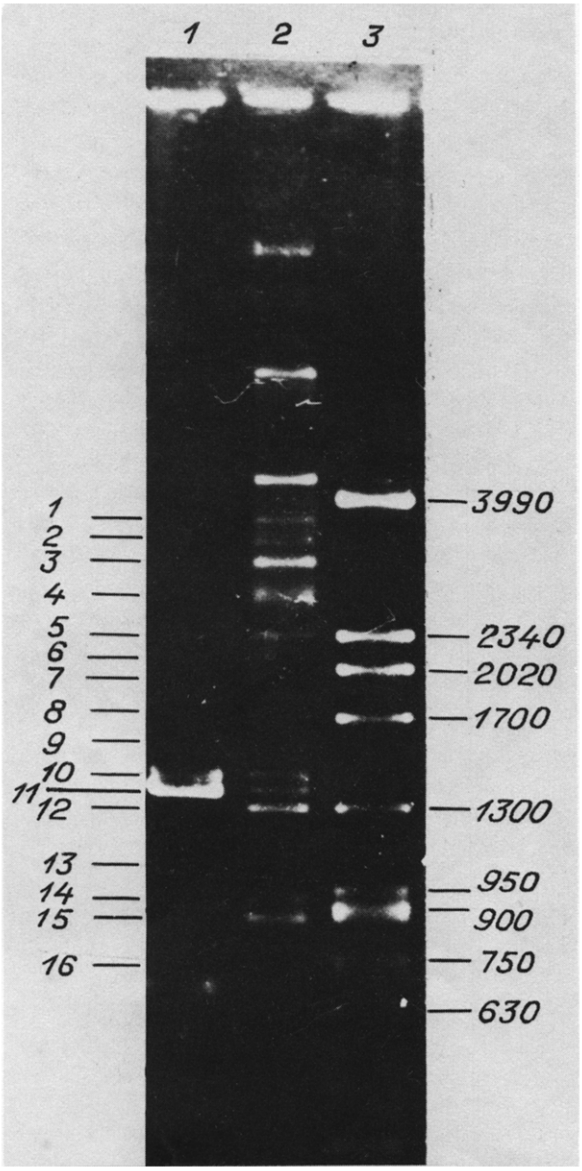

Fig.1. Sizing the fragments of pBR 322 DNA generated by *Sau*3A (1) or by *Bam*H1* (2) and of λC1857 DNA generated by *Bsp*1 (3). Numbers 1–16 refer to bands of track 2. The lengths in nucleotide pairs of the λC1857 DNA fragments product by *Bsp*1 are shown. Conditions of electrophoresis are described in [6].

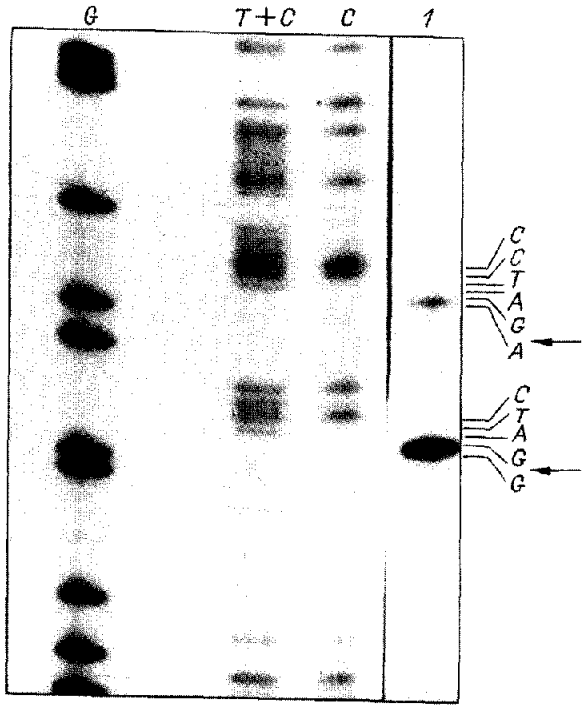

Fig.2. Autoradiograph of a sequencing gel of Msp/Bsp-365 fragment of pBR 322 DNA after cleavage by Maxam–Gilbert method [13] and by *Bam*H1* (track 1). Right, *Bam*H1*-recognition site and position of the cleavage are indicated.

GGATCT and AGATCC, but not the sequence TGATCT.

Fig.1 shows that the mobility of the *Bam*H1* fragment 11 is the same as that of the *Sau*3A-fragment of pBR 322 DNA cleaved at positions GGATCT (1666) and TGATCC (3115) (numbers refer to the first nucleotide in the sequence). The site TGATCC is complementary to the site GGATCA. The lengths of some *Bam*H1*-fragments are the same as those of the *Mbo*1-fragments of SV 40 DNA [6] due to cleavage of GGATCA. Thus, 2 of 3 computer-predicted changes at the terminal positions of the site GGATCC are experimentally confirmed. Taking into account the results of calculations and experiments with SV 40 and pBR 322 DNA it is possible to formulate *Bam*H1*-specificity as $G^{\downarrow}GATC$ and $G^{\downarrow}PuATCC$. Relaxation of the specificity of endonucleases *Eco*RI, *Bsu*1 and *Bst*1 involves loss of the selectivity to both terminal nucleotides of the canonic site [2,3,7]. In the case of *Bam*H1, only one of the six nucleotides of the canonic site may be substituted under 'relaxed' conditions. The first possibility is the substitution of the 3′-terminal N for C (GGATCN), the second possibility, substitution of Pu for G at the second position of the site (GPuATCC). However, these 2 substitutions may not occur simultaneously.

According to the symmetric recognition model [18], each of the 2 identical subunits of the restriction enzyme interacts with $n/2$ nucleotides of the 2-fold rotationally-symmetrical nucleotide sequences. If the TCC-half of the site interacts with *Bam*H1 under the standard conditions of the reaction, only one of the 2 enzyme subunits interacts under relaxed conditions with the specific trinucleotide TCC, whereas the other one interacts with the degenerated trinucleotide TCN or TPyC. The sequences ($N \neq G$) GATC ($N \neq C$) are completely resistant to *Bam*H1* activity in our experiments. Thus in the case of *Bam*H1* activity there is minimal specificity reduction, i.e., loss of 1 of the 6 specific contacts with DNA.

A comparison of the band intensities (fig.2) indicates that GG-bonds are more sensitive to the cleavage, so the hydrolytic center also has a certain specificity.

The generality of the proposed formula of *Bam*H1*-specificity can be checked by examining other substrates, containing other recognition sites which are absent in SV 40 and pBR 322 DNAs. Most striking in this connection is the fact that in the presence of high glycerol concentration endonuclease *Bam*H1 cuts φX 174 DNA within the sequence $GAA(N)_5TCC$ unrelated to the canonic recognition site [19].

References

[1] Roberts, R. J. (1980) Gene 8, 329–343.
[2] Polisky, B., Greene, P., Garfin, D. E., Mc Carty, B. J., Goodman, H. M. and Boyer, H. W. (1975) Proc. Natl. Acad. Sci USA 72, 3310–3314.
[3] Heininger, K., Hörz, W. and Zachan, H. G. (1977) Gene 1, 291–303.
[4] Hsu, M. T. and Berg, P. (1978) Biochemistry 17, 131–138.
[5] Tikchonenko, T. I., Karamov, E. V., Zavizion, B. A. and Naroditsky, B. S. (1978) Gene 4, 195–212.
[6] Malyguine, E., Vannier, P. and Yot, P. (1980) Gene 8, 163–177.
[7] Clarke, C. M. and Hartley, B. S. (1979) Biochem. J. 177, 49–62.
[8] Wilson, G. A. and Young, F. E. (1975) J. Mol. Biol. 97, 123–125.
[9] Grachev, S. A. and Netesov, S. V. (1981) Bioorgan. Khim. 7, 628–630.
[10] Kiss, A., Sain, B., Csordas-Toth, E. and Venetianer, P. (1977) Gene 1, 323–329.

[11] Susseubach, J. S., Monfoort, C. H., Schiphof, R. and Subberingh, E. (1976) Nucleic Acids Res. 3, 3193–3202.
[12] Kravchenko, V. V., Vassilenko, S. K. and Gileva, I. P. (1980) Dokl. Akad. Nauk SSSR 252, 1255–1257.
[13] Maxam, A. M. and Gilbert, W. (1977) Proc. Natl. Acad. Sci. USA 74, 560–564.
[14] Reddy, V. B., Thimappaya, B., Dhar, R., Subramanian, K. N., Zain, B. S., Pan, I., Ghosh, P. K., Celma, M. L. and Weisman, S. M. (1978) Science 200, 494–502.
[15] Sutcliff, I. G. (1979) Cold Spring Harbor Symp. Quant. Biol. 43, 77 90.
[16] Kravchenko, V. V., Vassilenko, S. K. and Grachev, M. A. (1979) Gene 7, 181–195.
[17] Gingeras, T. R., Milazzo, J. P. and Roberts, R. J. (1978) Nucleic Acids Res. 5, 4105–4127.
[18] Smith, H. O. (1979) Science 205, 455–462.
[19] George, J., Blakesley, R. W. and Chirikjian, J. G. (1980) J. Biol. Chem. 255, 6521–6524.