TWO SEQUENCE-SPECIFIC DEOXYRIBONUCLEASES FROM *RHODOSPIRILLUM RUBRUM*

M. G. C. DUYVESTEYN, A. DE WAARD and H. VAN ORMONDT

Laboratory for Physiological Chemistry, Sylvius Laboratories, University of Leiden, Wassenaarsweg 72, 2333 AL Leiden, The Netherlands

Received 19 June 1980

1. Introduction

The usefulness of sequence-specific endonucleases in solving fundamental problems in current molecular biology is well established. The catalog of enzymes has increased steadily (reviewed [1,2]). We report here the isolation and partial characterization of two enzymes from *Rhodospirillum rubrum*, endonucleases R.*Rru*I and R.*Rru*II. The first enzyme is unique in that it recognizes and cleaves the hexanucleotide sequence AGT↓ACT in DNA molecules of various origin. The latter enzyme (endo R.*Rru*II) recognizes the nucleotide sequence CC($\frac{A}{T}$)GG as does endo R.*Eco*RII; however, it cleaves the DNA within the site like the isoschizomer endo R.*Bst*NI (CC↓($\frac{A}{T}$)GG).

2. Materials and methods

2.1. *DNA and enzymes*

DNA from phage λ and pBR322 was prepared as in [11,14]. SV40 and Ad5 DNA were gifts from P. J. Abrahams and P. J. van den Elsen, respectively. Endo R.*Bst*NI was purchased from New England Biolabs, and *Hin*dIII and *Pst*I were gifts of H. S. C. Lupker-Wille. The preparation of T4 polynucleotide kinase and the 5′-labeling of DNA were as in [12].

2.2. *Growth of* R. rubrum

The strain of *R. rubrum* S1 was donated by A. H. M. de Wit of the Biophysical Laboratory of this University (Director Professor L. N. M. Duysens) and originated from the Laboratory of Dr M. D. Kamen at San Diego. It was grown anaerobically (nitrogen/5% CO_2) with illumination from 2 fluorescent light tubes at 40 cm distance (~1000 Lux) and from a visible light bulb. The aqueous medium contained (g/l): bactopeptone (5); yeast extract (Difco) (2); sodium glutamate (1); D,L-malic acid (1.5); sodium acetate (1.7); sodium succinate (1.5) (autoclaved together at pH 7); K_2HPO_4 (3.5); KH_2PO_4 (2.7) (autoclaved as a mixture); NaCl (4); NH_4Cl (1) (autoclaved as a mixture); plus the following additions (mg/l final conc.): nitrilo triacetic acid (200); $MgSO_4 \cdot 7\ H_2O$ (600); $CaCl_2 \cdot 2\ H_2O$ (70); ammonium molybdate (0.16); nicotinamide (1); thiamine hydrochloride (0.5); biotin (0.04) and 1 ml of a stock solution of trace metals containing (g/l stock soln.): $ZnSO_4$ (10.8); $MnSO_4$ (1.5); $FeSO_4 \cdot 7\ H_2O$ (7); $CuSO_4 \cdot 5\ H_2O$ (0.4); $Co(NO_3)_2$ (0.2); $Na_2B_4O_7 \cdot 10\ H_2O$ (0.2); EDTA (2.5); acidified with a few drops of HNO_3 to prevent precipitation. The final medium was adjusted to pH 7.0 with a sterile solution of 10% KOH. The cells were grown for 64 h from an inoculum (1:20) and subsequently harvested and stored at −20°C.

2.3. *Enzyme purification*

Frozen cells (17 g) were thawed and sonicated in a stainless steel beaker (surrounded by ice-water) in 185 ml 10 mM Tris–HCl (pH 7.4) containing 2 mM mercaptoethanol. The temperature was not allowed to rise above 8°C. The turbidity (A_{590}) of the suspension dropped to 25% of its original value in 6 min. The broken cells were spun down in a Sorvall centrifuge (SS 34 rotor) at 13 000 rev./min for 20 min at 4°C. The wine-red supernatant was made 0.1 M with respect to NaCl. A solution of polyethylene imine at pH 7.8 (Polymin P, BASF, Ludwigshafen) was stirred in to 1% final conc. The turbid solution was centrifuged as before. The supernatant

was treated with ammonium sulphate to 70% saturation and the resulting precipitate was dissolved in 20 ml 20 mM potassium phosphate (pH 7.4)–2 mM mercaptoethanol and dialysed overnight against 4 l this buffer (1 change). The precipitate resulting from the polyethylene imine step was extracted 3 times in succession with buffered 0.6 M NaCl as in [3] and also dialysed.

2.3.1. Purification of endonuclease R. *Rru*I

The combined dialysed solutions were freed from a red precipitate by centrifugation and applied to a 20 × 2.5 cm phosphocellulose column (Whatman PII) equilibrated with dialysis buffer, which was then washed with 100 ml same buffer. The pass-through contained some sequence-specific nuclease (endo R. *Rru*I) contaminated with exonucleases. An appreciable portion of the same endonuclease was adsorbed to the column material and was eluted very early (at ~0.04 M KCl) when a 460 ml linear gradient of 0–0.55 M KCl in dialysis buffer containing 10% glycerol was applied. These early eluting fractions were essentially uncontaminated by exonucleases but could be purified further (and concentrated) on a heparin–Sepharose column (5 × 1 cm) with a 60 ml gradient of 0.05–0.35 M KCl in 10 mM Tris–HCl (pH 7.4)–2 mM mercaptoethanol. The assays for specific endonuclease activity (endo R. *Rru*I and II) were routinely done by gel electrophoretic examination [4] with bacteriophage λ DNA as a substrate.

2.3.2. Purification of endonuclease R. *Rru*II

At the far end of the 0–0.55 M KCl gradient used for the elution of the nucleases from the phosphocellulose column a second sequence-specific deoxyribonuclease appeared (endo R. *Rru*II). Care was taken to keep the final concentration of KCl in the standard assay below 0.12 M. The active fractions were dialysed overnight against 2 l 10 mM Tris–HCl (pH 7.4) containing 2 mM mercaptoethanol and 10% glycerol. This enzyme was then chromatographed on DEAE-cellulose (10 × 1 cm column of Whatman DE 52) with a 50 ml gradient of 0–0.40 M KCl in dialysis buffer. It eluted at ~0.25 M KCl and was subsequently concentrated by adsorption onto a 30 × 4 mm column of DEAE-cellulose and elution with 0.45 M KCl in dialysis buffer.

2.4. *Determination of cleavage specificity*

2.4.1. Endonuclease R. *Rru*I

This was done primarily by incubating the plasmid pBR322 with the enzyme. Preliminary experiments were done which located the only cleavage site of endo R. *Rru*I on this DNA in the area between the *Pst*I site at position 3612 and the *Hin*dIII site at position 29 (see [5]). The *Rru*I digest of pBR322 DNA was labelled with T4 polynucleotide kinase and [γ–^{32}P]ATP (Radiochemical Centre, spec. act. >2000 Ci/mmol) and subsequently cleaved with endo R. *Hin*dIII and endo R. *Pst*I. The resulting 2 fragments (~500 and 230 nucleotides in size) were sequenched with the technique in [6]. A small portion of the labelled fragments was set aside for treatment with pancreatic DNase and subsequent two-dimensional separation on DEAE-cellulose thin-layer plates [7]. A similar 'wandering spot' analysis was done of labelled fragments originating from the 4 cleavage sites of endo R. *Rru*I in bacteriophage λ DNA. A small aliquot of the oligonucleotide collection was treated with snake venom exonuclease to yield the terminal [^{32}P]mononucleotide.

2.4.2. Endonuclease R. *Rru*II

The specificity of recognition by endo R. *Rru*II was established mainly by incubating Simian virus 40 (SV40) DNA with this enzyme for which there are ⩾ 16 cleavage sites. The resulting fragments were labelled with T4 polynucleotide kinase and [γ–^{32}P]ATP (2000 Ci/mmol). A number of the labelled fragments were cleaved with a secondary restriction endonuclease to yield singly-labelled fragments amenable to sequence analysis [6]. Other labelled fragments were processed with pancreatic DNase to give ^{32}P-labelled oligonucleotides to be characterized by the wandering spot analysis [7]. At a later stage a direct comparison was made between cleavage patterns in the *Hpa*I C fragment (map positions 4.3–25.5) of adenovirus 5 (Ad 5) DNA produced by endo R. *Rru*II and endo R. *Bst*NI [1].

3. Results

3.1. *Properties of endonuclease R.* Rru*I*

The purified *Rru*I enzyme was shown to cleave the following DNA preparations: λ DNA at 4 sites,

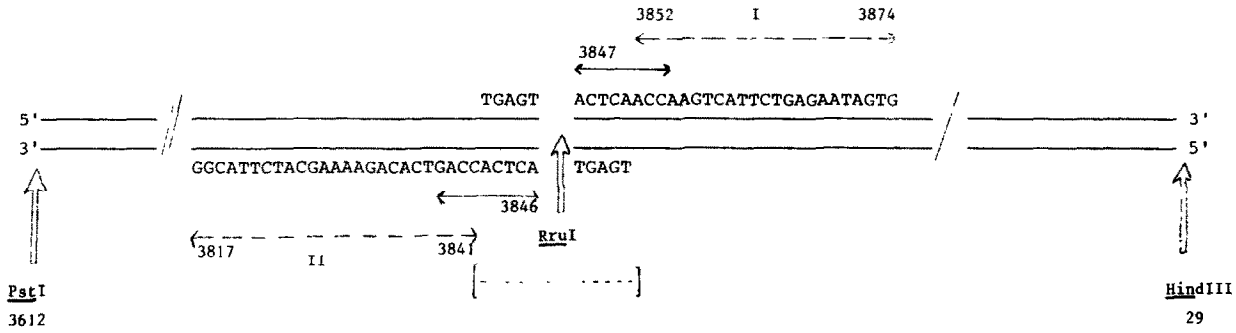

Fig.1. Results of nucleotide sequence determination around the endo R. *Rru*I cleavage site in pBR322. The stretches of DNA I and II that were sequenced according to [6] are indicated by the dashed lines (––). Original autoradiographs are documented in fig. 2). The shorter stretches that were sequenced by the 'wandering spot' procedure (documented in fig. 3a,b) are indicated by the drawn lines (——); The endo R. *Rru*I cleavage site resides in the centre of a symmetrical decanucleotide which is indicated by the symbol (- - -) coordinates 3842–3851).

adenovirus 2 DNA at 5 sites and the plasmid pBR322 at one site. Neither bacteriophage θX174 RF DNA nor Simian virus 40 (SV40) DNA were cleaved. The nucleotide sequence recognized and cleaved by endo R. *Rru*I was determined using the approaches in section 2. The results are schematically drawn in fig. 1, while fig.2 and 3 present the data of the sequence determination according to [6] and to the wandering spot method, respectively. We found the two sequences on either side of the endo R. *Rru*I cleavage site (with the help of the computer program SEARCH [8]) to be located at positions 3852–3874 and 3841–3817. In between these two sequences the pBR322 DNA structure was found to contain a symmetrical decanucleotide TGAGTACTCA. The results of the wandering spot analyses with pBR322 confirmed these results and permitted conclusions as to the actual site of cleavage. One of the 2 labelled fragments terminating at the *Rru*I site was shown in fig. 3a to have the structure 5'-ACTCAACCA (positions 3847–3855) which partially overlaps with sequence I (cf. fig. 1). The other labelled fragment has the structure 5'-ACTCACCAG (fig. 3b) which is the sequence on the opposite strand from coordinate 3846–3838. Thus the actual cleavage site resides between coordinates 3846–3847 within the symmetrical decanucleotide TGAGT↓ACTCA. From these experiments it was impossible to decide whether the recognition site of endo R. *Rru*I is a symmetrical hexa-, octa-, or decanucleotide. This question was resolved in the following way.

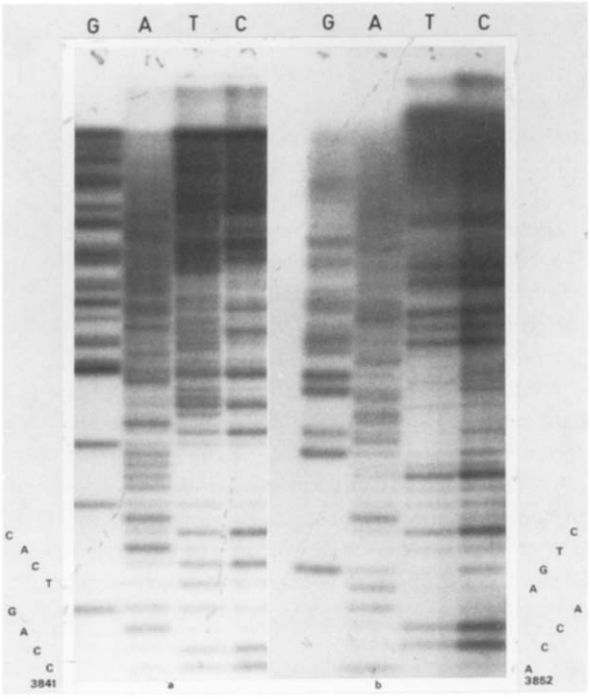

Fig. 2. Autoradiograph of gel showing the nucleotide sequence (a) from coordinate 3841–3817 (II) (and further) and (b) that from coordinate 3852–3874 (I) (and further) of pBR322 DNA. The bands nearer to the endo R. *Rru*I cleavage site were unclear. Consult also fig. 1.

Bacteriophage λ DNA was cleaved with endo R. *Rru*I and labelled with T4 polynucleotide kinase and [γ–^{32}P]ATP. A second sample of λ DNA was labelled solely at its termini by omitting cleavage with endo R. *Rru*I. After secondary cleavage of both labelled DNA preparations with endo R. *Mbo*II, 8 labelled fragments from the first degradation were selected which did not originate from the labelled

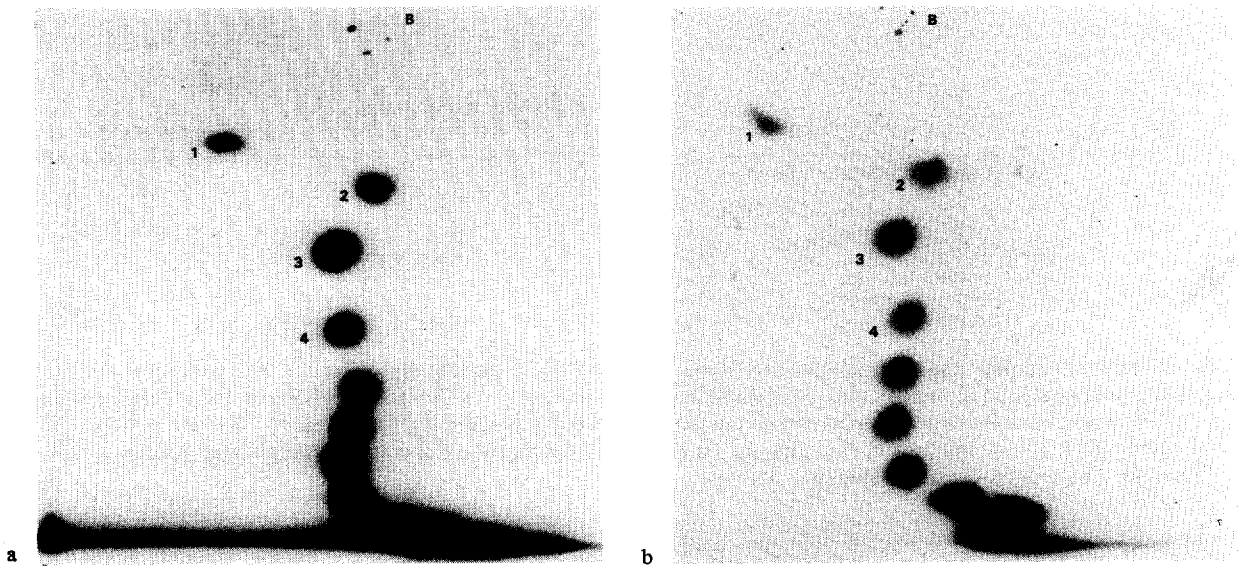

Fig.3. Autoradiographs of homochromatograms showing the nucleotide sequences at both sides of the endo R. *Rru*I cleavage site ('wandering spot' method') in pBR 322 DNA.
(a) Sequence of stretch between coordinates 3847–3855 (ACTCAACCA..). Spot 1, pApC as determined by snake venom exonuclease digestion and place occupied on chromatogram; pot 2, pApCpT; spot 3, pApCpTpC; spot 4, pApCpTpCpA; etc.
(b) Sequence of stretch between coordinates 3846–3838 (ACTCACCAG..). Spot 1, pApC as determined by snake venom exonuclease digestion and place occupied on chromatogram; spot 2, pApCpT; spot 3, pApCpTpC; spot 4, pApCpTpCpA; etc.

termini of the λ DNA molecule. The pooled 8 fragments were partially degraded with pancreatic DNase to give the wandering spot pattern presented in fig.4. It becomes clear that the specificity of the recognition site ends with pApCpT, so that its structure must be AGT↓ACT, a hexanucleotide.

3.2. *Properties of endonuclease R.* Rru*II*

This enzyme cleaves the DNA of bacteriophage λ (many sites) and θX174 RF (twice) while it also cleaves the viral DNAs of adenovirus 5 and SV40 (at > 30 and 16 sites, respectively). The cleavage patterns obtained by treating the *Hpa*I C fragment of adenovirus 5 DNA with endo R. *Rru*II and *Bst*NI recognition site CC($\frac{A}{T}$)GG) are identical (fig. 5). SV40 DNA was cleaved with endo R. *Rru*II, and the 5′-termini of the resulting fragments were labelled with T4 kinase and [$\gamma-^{32}$P]ATP. Some of these fragments were pooled and subjected to hydrolysis with pancreatic DNase and venom exonuclease. The 5′-termini of the fragments were either pA or pT, while the 5′-sequences deduced from the wandering spot pattern were AGG or TGG, after which the homology ended (not shown). A single fragment was selected and treated in the same way (fig. 6). Its 5′-ends are, according to the two-dimensional pattern, 5′-AGGATT and 5′-TGGGGG. and are located at positions 4812–4817 and 5012–5007; they coincide with *Eco*RII sites at 4810 and 5010 [13]. In another experiment, 5′-labelled *Rru*II fragments of SV40 DNA were treated with endo R. *Alu*I and fractionated, after which 4 of the resulting fragments with only one labelled 5′-end were subjected to sequence analysis according to [6]. The sequences deduced from the autoradiographs were located in the SV40 DNA with the help of a computer [8] and were found to be next to the *Eco*RII sites at positions 882, 3926 and 4810. The 36-long sequence read from gel 4 could be either next to the *Eco*RII site at position 95, or to the *Eco*RII site at 150, since SV40 DNA has a remarkable repetition in this region. However, since this fragment was obtained by secondary cleavage with *Alu*I we can exclude the *Eco*RII site at position 150 as a possible end.

All the above evidence demonstrates that endo R. *Rru*II attacks the same site as *Eco*RII, that is CC($\frac{A}{T}$)GG. However, *Eco*RII cleaves the DNA next to this site: ↓CC($\frac{A}{T}$)GG [1,2] whereas *Rru*II cuts

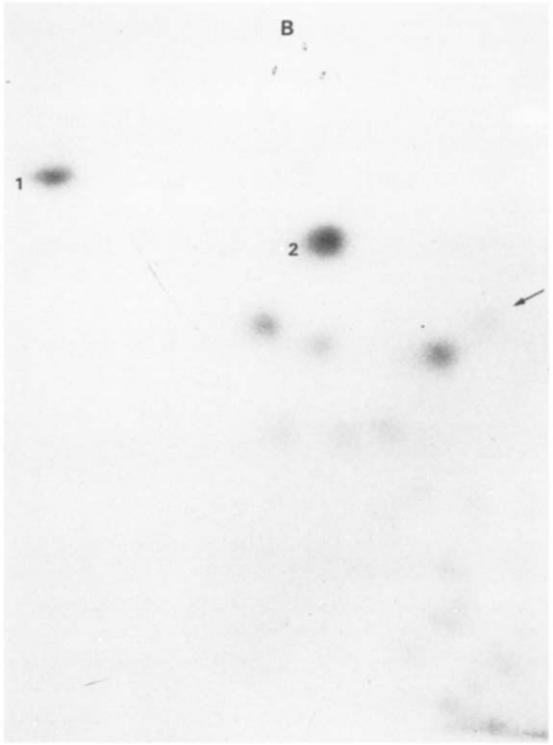

Fig. 4. Autoradiograph of homochromatogram of λ DNA digest by endo R. *Rru*I. Spot 1, pApC as shown by snake venom exonuclease digestion and place occupied on chromatogram; spot 2, pApCpT. The higher homologues (tetranucleotides) terminate at their 3′-ends in C, A, G and T, respectively, indicating that these last symbols are not part of the endo R. *Rru*I recognition sequence.

within the site CC↓($\frac{A}{T}$)GG. Another enzyme with this specificity is endo R. *Bst*NI [1,2].

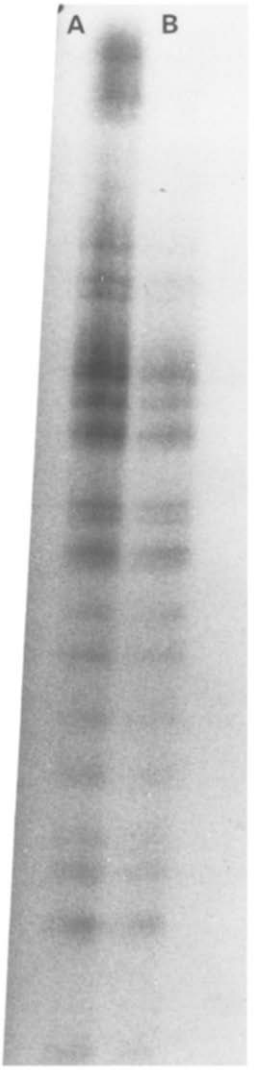

Fig.5. Comparison of cleavage patterns obtained by treating Ad5 *Hpa*I C (coordinates 4.3–25.5) with endo R. *Bst*NI and endo R. *Rru*II (see section 3.2). The bands at the top were visible in lane B on the original photograph.

4. Discussion

Extracts obtained by sonication of cells of *Rhodospirillum rubrum* contain two sequence-specific deoxyribonucleases both of which have been purified and characterized. Endonuclease R. *Rru*I was shown to cleave a symmetrical decanucleotide present in the plasmid pBR322 (coordinates 3842–3851), but subsequent experiments on the structure of 4 recognition sites present in bacteriophage λ DNA (fig.4), showed that the high level of symmetry in the cleavage site of pBR322 was coincidental. The total evidence pointed to a unique hexanucleotide, AGT↓ACT as the recognition sequence for endo R. *Rru*I. The length of this sequence is compatible with the observed frequency of cleavage in large DNA molecules such as from bacteriophage λ or from adenovirus type 2. The structure of the recognition sequence is also in line with its single occurrence in pBR322 [8] and its absence from the DNAs of bacteriophage θX174 and Simian virus 40 [10].

We can exclude the possibility that the endo R. *Rru*I recognition sequence is an isochizomer of endo R. *Rsa*I from *Rhodopseudomonas sphaeroides* [1,2] which cleaves the tetrameric sequence GTAC. This sequence occurs 3 times in pBR322 DNA (which

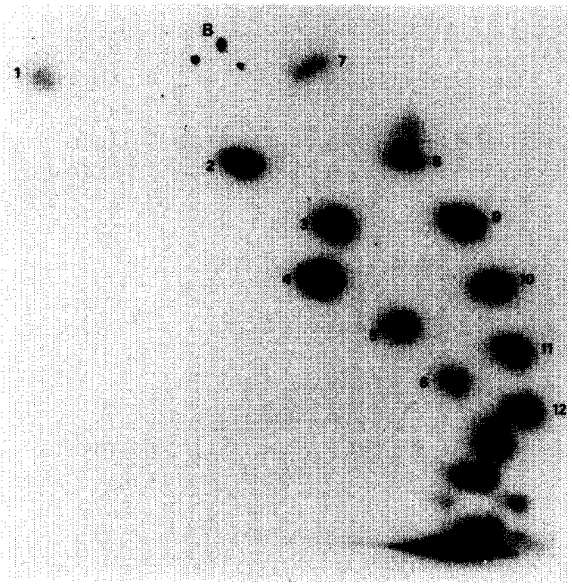

Fig.6. Wandering spot analysis of oligonucleotides labelled at two endo R. *Rru*I cleavage sites (positions 4812–4817 and 5012–5007) in SV40 DNA. Spot 1, pA; spot 2, pApG; spot 3, pApGpG; spot 4, pApGpGpA; spot 5, pApGpGpApT; spot 6, pApGpGpApTpT; spot 7, pT; spot 8–12, homologues having 1–5 additional G residues. B, blue marker.

is cleft only once by endo R. *Rru*I). In addition, the sequence GTAC occurs 11 times both in bacteriophage θX174 RF DNA and in SV40 [10] which are resistant to endo R. *Rru*I. It is interesting to note that the single endo R. *Rru*I cleavage site in pBR322 occurs in the ampicillin resistance gene [11] and constitutes an insertion site for blunt end ligation.

The yield of the second sequence-specific endonuclease (endo R. *Rru*II) recognizing CC↓$(^{A}_{T})$GG (as does endo R. *Bst*NI) was rather low so that for preparative purposes *R. rubrum* does not seem to be the source of choice for an enzyme of said specificity. In contrast to endo R. *Eco*RII, endo R. *Rru*II (and endo R. *Bst*NI) are not inhibited by meC methylation.

Acknowledgements

The authors thank A. H. M. de Wit of the laboratory of Biophysics at Leiden University for a gift of *R.rubrum*. Mrs H. S. C. Lupker-Wille kindly prepared endo R. *Hin*dIII and endo R. *Pst*I, while B. M. M. Dekker helped with the nucleotide sequence determination. This work was supported in part through the Foundation of Chemical Research in the Netherlands (SON) and by the Netherlands Organization for the Advancement of Pure Research (ZWO).

References

[1] Roberts, R. J. (1980) Gene 8, 329–343.
[2] Roberts, R. J. (1980) Nucleic Acids Res. 8, 63–80.
[3] Duyvesteyn, M. and De Waard, A. (1980) FEBS Lett. 111, 423–426.
[4] Sugden, B., De Troy, B., Roberts, R. J. and Sambrook, J. (1975) Anal. Biochem. 68, 36–46.
[5] Sutcliffe, J. G. (1978) Nucleic Acids Res. 5, 2721–2728.
[6] Maxam, A. M. and Gilbert, W. (1977) Proc. Natl. Acad. Sci. USA 74, 560–564.
[7] Tu, C.-P. D., Jay, E., Bahl, C. P. and Wu, R. (1976) Anal. Biochem. 74, 73–93.
[8] Staden, R. (1977) Nucleic Acids Res. 4, 4037–4051.
[9] Sutcliffe, J. G. (1978) Cold Spring Harbor Symp. Quant. Biol. 43, 77–90.
[10] Fuchs, C., Rosenvold, E. C., Honigman, A. and Szybalski, W. (1978) Gene 4, 1–23.
[11] Bolivar, R., Rodriguez, R. L., Greene, P. J., Betlach, M. C., Heijneker, H. L., Boyer, H. W., Crosa, J. H. and Falkow, S. (1977) Gene 2, 95–113.
[12] Van Ormondt, H., Maat, J., De Waard, A. and Van der Eb, A. J. (1978) Gene 4, 309–328.
[13] Reddy, V. B., Thimmappaya, B., Dhar, R., Subramanian, K. N., Zain, B. S., Pan, J., Ghosh, P. K., Celma, M. L. and Weissman, S. M. (1980) Science 200, 494–502.
[14] Maat, J. and Lupker-Wille, H. S. C. (1978) Biochim. Biophys. Acta 520, 512–520.