Volume 235, number 1,2, 241-246 FEB 06147 August 1988

*Asn*I: a novel class II restriction endonuclease from *Arthrobacter* sp., strain N-CM, recognizing 5′-AT/TAAT-3′

Bernhard U. Rexer, Michael Jarsch, Christine Sagmeister, Barbara Glück, Gerhard Berger and Christoph Kessler*

Boehringer Mannheim GmbH, Biochemical Research Center, Department of Molecular Biology, Nonnenwald 2, D-8133-Penzberg, FRG

Received 7 April 1988; revised version received 7 June 1988

A new class II restriction endonuclease, *Asn*I, with a novel sequence specificity was isolated from the Gram-positive eubacterium *Arthrobacter* species, strain N-CM. *Asn*I recognizes the unambiguously defined palindromic hexanucleotide

5′-ATTAAT-3′
3′-TAATTA-5′,

consisting of A- and T-residues. The novel enzyme in the presence of Mg^{2+} cleaves specifically both strands as indicated by the arrows. The staggered cuts generate 5′-protruding ends with single-stranded 5′-TA-3′ dinucleotide extensions. The novel enzyme may be a useful tool for cloning experiments by complementation of the few enzymes such as *Pst*I and *Pvu*I cutting only once in the Amp^r-gene of plasmids pBR322 and pBR328.

Restriction endonuclease; Hexanucleotide recognition sequence; T/A-nucleotide; 5′-Protruding terminus; Klenow enzyme; T_4 polynucleotide kinase; T_4 DNA polymerase; T_4DNA ligase

1. INTRODUCTION

More than 650 class II restriction endonucleases have been isolated from eu- and archaebacteria [1,2]. These enzymes represent more than 100 different sequence specificities. Most important are enzymes recognizing unequivocally defined tetra- and hexanucleotide palindromes. Out of the 16 possible tetranucleotides, 11 sequences are currently covered by corresponding enzymes.

From the 64 possible hexanucleotides, 49 enzyme specificities are known so far [3]. Most of these enzymes recognize sequences rich in G- and C-residues. With *Ssp*I and *Dra*I as well as its isoschizomer *Aha*III only tree enzymes are known recognizing hexanucleotide palindromes consisting only of A- and T-residues [4–6].

We have discovered a further enzyme, *Asn*I, recognizing only A- and T-residues within the hexanucleotide palindrome 5′-AT/TAAT-3′. We report here the isolation and characterization of the novel enzyme enlarging the set of class II restriction endonucleases cutting at AT-rich hexanucleotides.

We designated the new class II enzyme in accordance with the proposal of Smith and Nathans [7] as *Asn*I to distinguish it from the series of enzymes isolated from different strains of *Achromobacter* species [1].

2. MATERIALS AND METHODS

2.1. *Bacterial strain and culture conditions*

Arthrobacter sp., strain N-CM, belongs to the irregular, non-

Correspondence address: C. Kessler, Boehringer Mannheim GmbH, Abtlg BB-BRM, Nonnenwald 2, 8122 Penzberg, FRG

Published by Elsevier Science Publishers B.V. (Biomedical Division)

sporing Gram-positive rods [8]. Cells were cultured aerobically at 25°C in a mixture of 10 g/ml malt extract and 10 g/l yeast extract at a pH value between 6.8 and 7.0.

2.2. *Enzymes*

Restriction endonucleases *Asp*700I, *Ava*I, *Bam*HI, *Bgl*I, *Eco*RI, *Eco*RV, *Hin*dII, *Nde*I, *Pst*I, *Pvu*I, *Pvu*II, *Sal*I and *Sca*I, calf intestinal alkaline phosphatase (CIAP), T_4 DNA polymerase (T_4 DNAP), T_4 DNA ligase (T_4 DNA*Lig*), T_4 polynucleotide kinase (T_4 PNK), terminal transferase (TdT) and Klenow enzyme (*Pol*IK) were from Boehringer Mannheim. The enzymes were employed according to the manufacturer's specifications.

2.3. *DNAs, nucleotides and reagents*

Phage λ*c*I857*S*am7 and M13mpl8RF DNA, plasmids pUC18, pBR322 and pBR328, DNA molecular mass markers III and IV, dATP, dGTP, dCTP and dTTP were from Boehringer Mannheim. Ad-2 DNA was from BRL. [γ-^{32}P]ATP (~3000 Ci/mmol) and [α-^{32}P]ddATP (~3000 Ci/mmol) were from Amersham, heparin-Sepharose CL-6B, Sephadex G25 from P-L Pharmacia, and cellulose Pll was from Whatman.

2.4. *Isolation and purification of* Asn*I*

*Asn*I was purified according to standard procedures [3] including fractination on Sephadex G25, cellulose Pll and heparin-Sepharose CL-6B.

2.5. *Enzyme assay*

For determination of *Asn*I activity, varying amounts of the final enzyme preparation were incubated with 1 μg λ*c*I857*S*am7 DNA in 25 μl incubation mixture (10 mM Tris-HCl, pH 8.0/37°C; 100 mM NaCl; 10 mM β-mercaptoethanol; 10 mM $MgCl_2$) for 1 h at 37°C. The reactions were terminated by addition of 5 μl cold stop solution. 25 μl of each reaction mixture was resolved by electrophoresis as above.

One unit of *Asn*I is defined as the amount of enzyme which digests 1 μg of λ*c*I857*S*am7 DNA within 1 h at 37°C under the stated assay conditions.

2.6. *DNA labelling*

Restriction fragments were 5′-endlabelled by treatment with 0.05 units of CIAP per pmol 5′-termini [9] and rephosphorylation with 1 unit T_4 PNK per pmol 5′-termini in the presence of 50 μCi [γ-^{32}P]ATP (~3000 Ci/mmol) [10]. 3′-endlabelling was achieved with 25 units TdT per pmol 3′-termini and 50 μCi [α-^{32}P]ddATP (~3000 Ci/mmol) as described by [11].

2.7. *Determination of cleavage positions*

Determination of cleavage positions within both strands of *Asn*I cleavage sites using the chemical method was according to [12] and [13]. Nucleotide sequence analysis was performed following the chemical method described in [10]. The A/G-reaction was modified according to [14]. Treatment with T_4 DNAP was performed as described in [13].

2.8. *Computer analysis*

Computer data were generated on a Microvax II (Digital Equipment Corporation) under VAX/VMS using the UWGCG software package [23]. In addition, computer programs were used which are analogous to those described in [24]. Sequence data for Ad-2 DNA [15] and for λ*c*I857*S*am7 DNA [16] were obtained from the EMBL databank [17], Entries AD2 and LAMBDA. The additional nucleotide sequences used for computer analysis were determined by [18,22] for pUC18 DNA and M13mp18RF DNA, by [19,20,22] for pBR322 DNA, and by [21,22] for pBR328 DNA.

3. RESULTS

3.1. *Optimal conditions for enzyme activity*

The activity of *Asn*I is optimal to pH 8.0 (37°C) and salt concentrations of 100 mM NaCl. The enzyme is strictly dependent on Mg^{2+} with an optimum at 10 mM $MgCl_2$, but does not require *S*-adenosyl-methionine or ATP for activity.

Highly purified *Asn*I is free of contaminating site-specific or unspecific endonucleases as well as exonucleases, since either λ*c*I857*S*am7 DNA or pBR322 DNA digested with at least 20 units *Asn*I for 16 h (320-fold excess of enzyme) gave sharp bands without any smearing after electrophoresis in agarose gels (not shown). Functional purity is shown by ligation with 0.1 units T_4 DNA*Lig* per μg *Asn*I-fragments which yields in >95% ligation products >95% of the ligation products can be recut with *Asn*I.

3.2. *Determination of recognition sequence*

The recognition sequence of *Asn*I was determined by mapping its recognition sites on pBR322 DNA at position 3539, on pBR328 DNA at position 2450, and on pUC18 DNA at positions 516, 635 and 1870 by double digestions with *Asn*I and various restriction enzymes cutting around these sites.

With pBR322 after linearization of the plasmid with *Asn*I, double digestions were performed with a series of restriction endonucleases including *Bam*HI, *Bgl*I, *Eco*RV, *Hin*dIII, *Nde*I, *Pst*I, *Pvu*I, *Pvu*II and *Sal*I (fig.1). For all tested enzymes, the resulting fragment sizes are in agreement with *Asn*I recognition at the hexonucleotide 5′-ATTAAT-3′ sequence at position 3539 (table 1). The position of the single *Asn*I recognition site on pBR328 DNA at position 2450 was determined in an analogous way with *Pst*I (pos. 2520), *Pvu*I (pos. 2647) and *Asp*700I (pos. 2874) double digestions.

In addition, the three *Asn*I recognition sites on pUC18 DNA at positions 576, 635 and 1870 were mapped with the help of *Bgl*I (pos. 245, 1813),

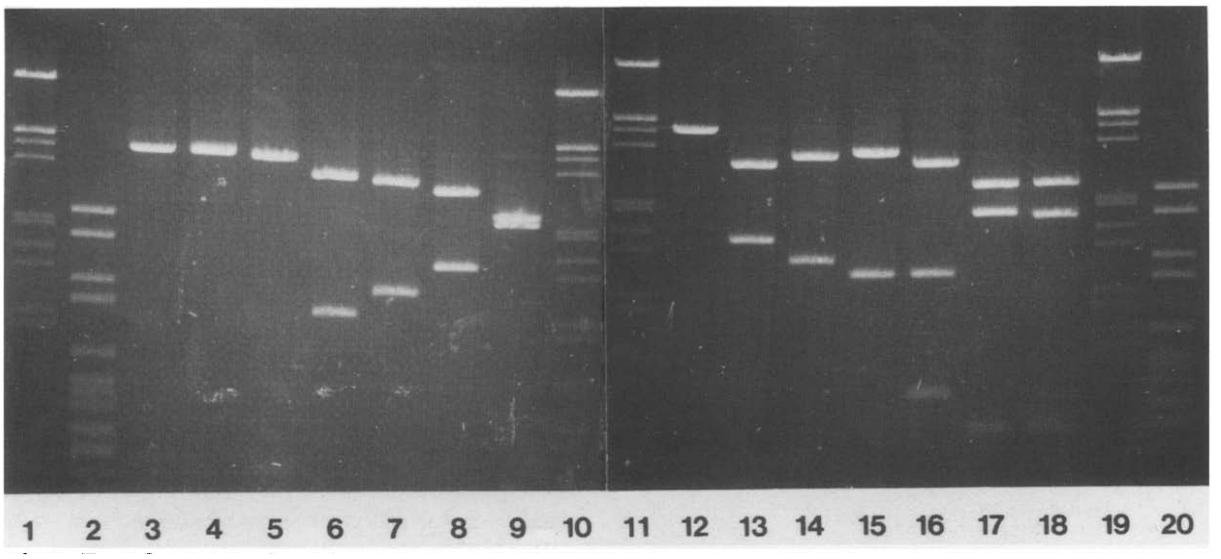

Fig. 1. Double digestions of pBR322 DNA with *Asn*I and additional restriction enzymes. pBR322 was cut with *Asn*I (lanes 3, 12), and the resulting linearized DNA was additionally cleaved with *Pst*I (lane 4), *Pvu*I (lane 5), *Eco*RV (lane 6), *Bam*HI (lane 7), *Sal*I (lane 8), *Ava*I (lane 9), *Pvu*II (lane 13), *Nde*I (lane 14), *Hin*dII (lane 16), and *Bgl*I (lane 18), respectively. pBR322-*Hin*dII fragments (3256, 1107 bp) are in lane 15; pBR322-*Bgl*I fragments (2319, 1810, 234 bp) are in lane 17. λcI857Sam7-*Eco*RI/*Hin*dIII fragments (lanes 1, 10, 11, 19), and pBR328-*Bgl*I/*Hin*fI fragments (lanes 2, 20) were used as DNA molecular mass markers.

*Pst*I (pos. 411), *Eco*RI (pos. 450) and *Pvu*I (pos. 276, 2066) (not shown). The cleavage sites on pUC18 DNA were localized independently by determination of the exact cut positions within the respective recognition sequences at positions 577, 636 and 1871.

Table 1

Mapping of the single *Asn*I recognition sequence on pBR322 DNA at position 3539 by double digestions with additional enzymes

Second enzyme	Positions of recognition sites of second enzyme	Experimentally observed fragment sizes of double digestions (bp)	Computer-derived fragment sizes of double digestions (bp)
*Pst*I	3609	4300, (<100)[c]	4290, 73
*Pvu*I	3736	4200, 200	4165, 198
*Eco*RV	187	3400, 1000	3351, 1012
*Bam*HI	375	3200, 1200	3165, 1198
*Sal*I	651	2900, 1500	2889, 1474
*Ava*I	1425	2200, 2100	2248, 2115
*Pvu*II	2066	2900, 1500	2891, 1472
*Nde*I	2296	3100, 1200	3120, 1243
*Hin*dII	651, 3907[a]	2900, 1100, 350	2887, 1107, 369
*Bgl*I	929, 1163, 3482[b]	2300, 1700, 250, (<100)[d]	2319, 1758, 234 52

[a]Resulting fragment sizes: 3256, 1107 bp
[b]Resulting fragment sizes: 2319, 1810, 234 bp
[c]Decrease of 4363 bp fragment size
[d]Decrease of 1810 bp fragment size.

The *Asn*I recognition sequence 5′-ATTAAT-3′ was confirmed by comparison of the experimentally observed *Asn*I fragment pattern obtained by digestion of λcI857Sam7 DNA at recognition sites 3418, 19481, 23083, 23692, 24986, 26762, 27202, 30619, 33221, 33735, 33800, 34686, 36274, 43409, 46407, 46852 and 48377, Ad-2 DNA at positions 31900, 33079, 35425 and 35936 and M13mp18RF DNA at 4131, 4135, 4238, 4628, 6046, 6105 and 6966 with the corresponding computer data derived by analysis of the known nucleotide sequences with the program MAP of the UWGCG program package [23].

3.3. *Determination of the cleavage positions*

The exact cleavage position of *Asn*I was determined in parallel by two approaches, the chemical [10,12] and enzymatic sequencing method [13]. First, for chemical sequencing ^{32}P-endlabelled DNA fragments containing *Asn*I sites were prepared from plasmid pUC18. Four aliquots of each of these fragments were subjected to the various chemical sequencing reactions. Another

aliquot was treated with *Asn*I. After electrophoretic separation on 5% (w/v) polyacrylamide gels, containing 8 M urea, and autoradiography, the size of the *Asn*I fragments was analysed by comparison with the sequencing ladders as shown in fig.2.

In detail, the cut positions within the *Asn*I recognition sequence were determined as follows: plasmid pUC18 was linearized by *Pvu*I. This fragment was labelled at the 3'-ends with TdT and [α-^{32}P]ddATP. After a second cleavage with *Eco*RI a 1.6 kb fragment was isolated after electrophoretic separation on an agarose gel. In a second experiment pUC18 DNA was linearized by *Pst*I. The resulting fragment was 3'-endlabelled with TdT and [α-^{32}P]ddATP, or 5'-endlabelled with T_4 PNK and [γ-^{32}P]ATP. After recutting with *Pvu*I, 1.6 kb fragments were isolated after electrophoretic separation on an agrose gel. In the next experiment pUC18 DNA was linearized by *Eco*RI. The resulting fragment was either 3'-endlabelled with TdT and [α-^{32}P]ddATP, or 5'-endlabelled with T_4 PNK and [γ-^{32}P]ATP. After recutting with *Pvu*I 1.6 kb fragments were isolated after electrophoretic separation on an agarose gel.

Each of these isolated endlabelled DNA fragments were divided into five aliquots. Four of these aliquots are subjected to base-specific cleavage reactions according to [12]. One aliquot was treated with *Asn*I. Reactions products were analyzed by electrophoresis on 5% (w/v) polyacrylamide gels, containing 8 M urea, and subsequent autoradiography. Interpretation of the results followed the rules described in [13].

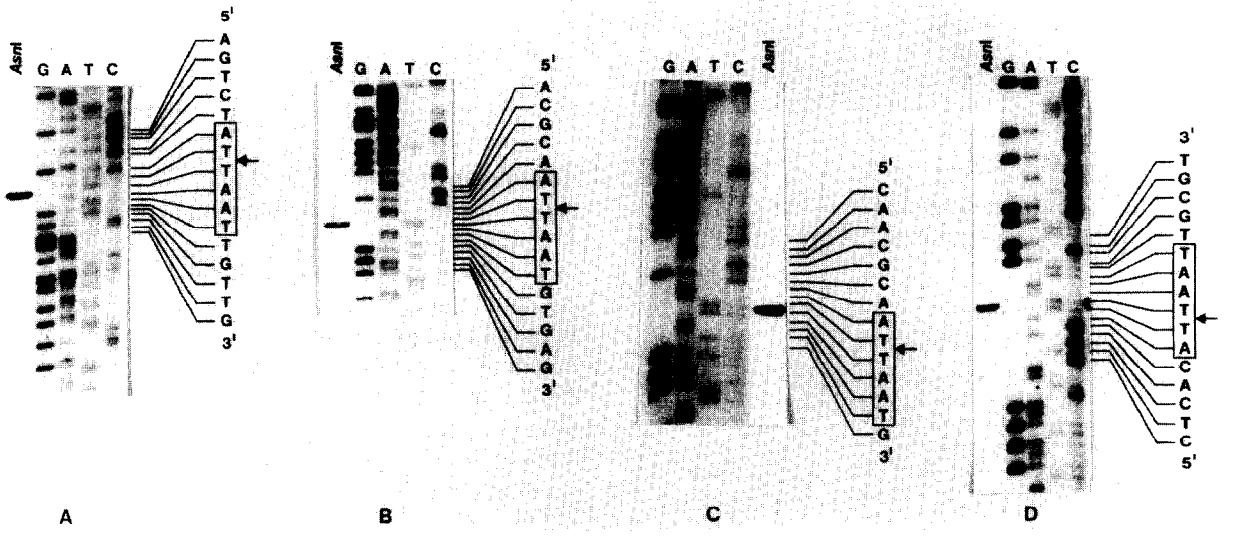

Fig. 2. Autoradiograph of the sequencing ladders around *Asn*I recognition sites on plasmid pUC18 at positions 1870 (A) and 576 (B,C,D). To analyze the *Asn*I site at position 1870 of pUC18 DNA, the plasmid was cut with *Pvu*I, subsequently 3'-endlabelled with TdT and [α-^{32}P]ddATP, and recut with *Eco*RI. An endlabelled *Pvu*I-*Eco*RI fragment (A) was isolated by low melting temperature agarose electrophoresis. For the analysis of the *Asn*I site at position 576 pUC18 DNA was either cut with *Pst*I, 3'-endlabelled with TdT and [α-^{32}P]ddATP, or cut with *Eco*RI, either 5'-endlabelled with T_4 PNK and [γ-^{32}P]ATP, or 3'-endlabelled with TdT and [α-^{32}P]ddATP. All samples were recut with *Pvu*I. For the sequencing ladder a suitable 3'-endlabelled *Pst*I-*Pvu*I fragment (B), a 3'-endlabelled *Eco*RI-*Pvu*I fragment (C), and a 5'-endlabelled *Eco*RI-*Pvu*I fragment (D) was prepared as above. The isolated fragments were divided into aliquots which were subjected to chemical sequencing reactions according to [10] with the modification of the A + G reaction described in [14]. An additional aliquot of each fragment was treated with *Asn*I. The resulting samples were analyzed by electrophoresis on 5% (w/v) polyacrylamide sequencing gels, containing 8 M urea, followed by autoradiography. Lanes G, A, T, C represent the four chemical sequencing reactions of a particular DNA fragment, lane *Asn*I represents the respective *Asn*I-treated sample. The interpretation followed the rules described in [13]. The cut sites of *Asn*I at positions 1871 (A) and 577 (B,C,D) were deduced from the size of the corresponding bands of the sequencing ladder. During the base-specific chemical sequencing reactions a particular base is eliminated leaving a phosphate group at both flanking 5'- and 3'-DNA termini. In the case of the 5'-endlabelled sample (D) this results in a slightly faster electrophoretic mobility as compared with the corresponding enzymatically created fragment which has lost only its 3'-terminal phosphate group leaving 5'-phosphorylated fragment termini [26].

Cleavage positions of *Asn*I on pUC18 DNA determined by chemical sequencing are as follows:

```
                576
                . ↓
*5'EcoRI---CTCACAT TAATTGCGT-3'

                  ↓
*3'-PstI---GAGTGTAAT TAACGCA-5'

                  ↓
*3'-EcoRI--GAGTGTAAT TAACGCA-5'

                1870
                . ↓
      5'-AGTCTAT TAATTGTTG----PvuI-3'*
```

The specificity of *Asn*I is concluded as:

```
     ↓
5'-AT TA-AT-3'
3'-TA-AT TA-5'.
       ↑
```

These results were confirmed by the enzymatic sequencing approach described in [13]. A M13mp18-derivative containing an insert with an *Asn*I site was used as template for enzymatic sequencing reactions starting with the universal sequencing primer. In parallel, an universal sequencing primer, ^{32}P-endlabelled with T_4 PNK and [α-^{32}P]ATP, was annealed to the same template, and a partial double-stranded DNA was created by treatment with *Pol*IK and all four dNTPs. This double-stranded DNA was used as substrate for *Asn*I to produce a 5'-endlabelled DNA fragment comparable to the sequencing ladder. Samples were analyzed by electrophoresis and subsequent autoradiography as described above (fig.3). The cleavage position of *Asn*I determined by enzymatic sequencing on M13 DNA is as follows:

```
                ↓
*5'---CCGGTAT TAATTCATG-3'
```

Again, the specificity of *Asn*I is concluded as:

```
     ↓
5'-AT TA-AT-3'
3'-TA-AT TA-5'.
       ↑
```

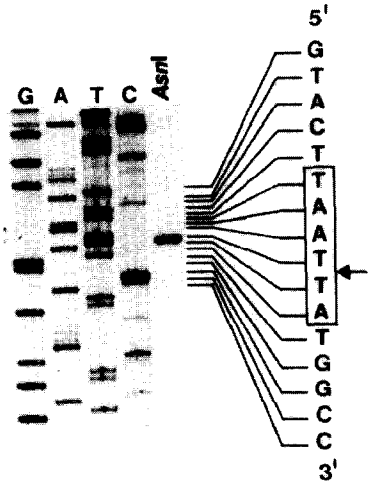

Fig. 3. Autoradiograph of the sequencing ladder around the *Asn*I recognition site of M13mp18 DNA. Preparation of single-stranded M13mp18 DNA and sequencing reactions were done according to [27]. A partial double-stranded, 5'-endlabelled DNA as substrate for *Asn*I was prepared and *Pol*IK extension of a 5'-endlabelled universal primer annealed to single-stranded M13mp18 DNA. Lanes G, A, T, C represent the four base-specific sequencing reactions, lane *Asn*I represents the respective *Asn*I-treated sample. The interpretation followed the rules described in [13]. The fragment resulting from *Asn*I cleavage has the identical size compared with the band of the first T in the sequencing ladder of the *Asn*I recognition sequence ATTAAT. This band of the T sequence ladder represents a fragment ending with 5'...AT-3', and therefore implicates the cleavage mode: AT↓TAAT.

4. DISCUSSION

With *Asn*I from *Arthrobacter* species, strain N-CM, a new class II restriction endonuclease has been isolated in high purity recognizing the novel palindromic hexanucleotide recognition sequence 5'-AT/TAAT-3'.

The new enzyme complements class II restriction endonucleases recognizing AT-rich hexanucleotide sequences. Out of the eight possible hexanucleotide palindromes existing completely of A- and T-residues, corresponding enzymes for only two sequences 5'-AATATT-3' (*Ssp*I [4]) and 5'-TTTAAA-3' (*Dra*I [5], *Aha*III [6]) were known to date. In contrast, for most of recognition sequences rich in G- and C-residues enzymes have been discovered. For the eight possible hexa-

nucleotide palindromes composed completely of G- and C-residues many more class II enzymes have been discovered so far. Only the 5′-CGCGCG-3′ sequence has not yet been realized by a class II restriction endonuclease.

*Asn*I may be a useful tool for the construction of recombinant DNA. The single *Asn*I cleavage site in the commonly used cloning vectors pBR322 and pBR328 is located in the structural gene for ampicillin resistance. Cloning into the unique *Asn*I site should be facilitated because in contrast to single-cutting enzymes *Pvu*I, *Sca*I or *Asp*700I, creating blunt ends, *Asn*I generates fragments with 5′-protruding AT-dinucleotides. Compared with the 3′-protruding fragment ends obtained by linearization of pBR322 of pBR328 DNA with *Pst*I, *Asn*I fragment termini can both be efficiently 5′-labelled with T_4 polynucleotide kinase and 3′-endlabelled with either Klenow enzyme or T_4 DNA polymerase.

Finally, as with *Ssp*I, *Dra*I and its isoschizomer *Aha*III, *Asn*I may be applied for analysis of AT-rich DNA sequence stretches, or as a rare cutting class II restriction endonuclease for genomes of high G:C-content.

Acknowledgements: We wish to thank L. Rüssmann for establishment of the various computer programs and S. Metzger-Bergel for typing the manuscript.

REFERENCES

[1] Kessler, C. and Höltke, H.-J. (1986) Gene 47, 1–153.
[2] Roberts, R.J. (1987) Nucleic Acid Res. 15, r189–r217.
[3] Kessler, C. (1987) Class II Restriction Endonucleases. in: Cytogenetics (Obe, G. and Basler, A. eds.) pp. 225–279, Springer, Berlin.
[4] Grandoni, R.P. and Schildkraut, I., unpublished.
[5] Purvis, I.J. and Moseley, B.E.B. (1983) Nucleic Acids Res. 11, 5467–5474.
[6] Whitehead, P.R. and Brown, N.L. (1985) Arch. Microbiol. 141, 70–74.
[7] Smith, H.O. and Nathans, D. (1973) J. Mol. Biol. 81, 419–423.
[8] Jones, D. and Collins, M.D. (1986) Irregular, Nonsporing Gram-Positive Rods. in: Bergey's Manual of Sympomatic Bacteriology, vol. 2 (Sneath, P.H.A. et al. eds) pp.1261–1434, Williams and Wilkins, Baltimore.
[9] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1980) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, pp.133, 134 and 453.
[10] Maxam, A.M. and Gilbert, W. (1980) Methods Enzymol. 65, 499–560.
[11] Roychoudhury R. and Wu, R. (1980) Methods Enzymol. 65, 43–60.
[12] McConnel, D.J., Searcy, D.G. and Sutcliffe, J.G. (1978) Nucleic Acids Res. 5, 1729–1739.
[13] Brown, N.L. and Smith, M. (1980) Methods Enzymol. 65, 391–404.
[14] Gray, G.E., Sommer, R., Polke, C., Beck, E. and Schaller, H. (1978) Proc. Natl. Acad. Sci. USA 75, 50–53.
[15] Gingeras, T.R., Sciaky, D., Gelinas, R.E., Jiang, B.-D., Yen, C.E., Kelly, M.M., Bullock, P.A., Parsons, B.L., O'Neill, K.E. and Roberts, R.J. (1982) J. Biol. Chem. 275, 13457–13491.
[16] Sanger, F., Coulson, A.R., Hong, G.F., Hill, D.F. and Petersen, G.B. (1982) J. Mol. Biol. 162, 729–773.
[17] Hamm, G.H. and Cameron, G.N. (1986) Nucleic Acids Res. 14, 5–15.
[18] Buchmann, A.R., Burnett, L. and Berg, P. (1980) The SV40 nucleotide sequence. in: DNA Tumor Viruses. (Tooze, J. ed.) pp.779–829, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
[19] Sutcliffe, J.G. (1979) Cold Spring Harbor Symp. Quant. Biol. 43, 77–90.
[20] Peden, K.W.C. (1983) Gene 22, 277–280.
[21] Soberon, X., Covarrubias, L. and Bolivar, F. (1980) Gene 9, 287–305.
[22] Pfeiffer, F. and Gilbert, W. (1988) Protein Sequences and Data Analysis, in press.
[23] Devereux, J., Haeberli, P. and Smithies, O. (1984) Nucleic Acids Res. 12, 387–395.
[24] Gingeras, T.R., Milazzo, J.P. and Roberts, R.J. (1978) Nucleic Acids Res. 5, 4105–4127.
[25] O'Farrell, P.H., Kutter, E. and Nakanishi, M. (1980) Mol. Gen. Genet. 179, 421–435.
[26] Tomizawa, J.-J., Ohmori, H. and Bird, R.E. (1977) Proc. Natl. Acad. Sci. USA 74, 1865–1869.
[27] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 560–564.