Volume 5 Number 6 June 1978 Nucleic Acids Research

A restriction enzyme Tha I from the thermophilic mycoplasma Thermoplasma acidophilum

D.J. McConnell, D.G. Searcy* and J.G. Sutcliffe

Biological Laboratories, Harvard University, Cambridge, MA 02138 and *Department of Zoology, University of Massachusetts, Amherst, MA 01003, USA

Received 27 April 1978

ABSTRACT

A type II restriction enzyme (Tha I) has been isolated from the thermophilic mycoplasma *Thermoplasma acidophilum*. A new method of general application was used to determine the DNA sequence cleaved by the enzyme. Tha I cuts DNA in the centre of the sequence CGCG. Single-stranded DNA is not a substrate. Tha I does not cut *T. acidophilum* DNA which is presumably modified. This is the first description of a restriction enzyme from a mycoplasma. Because Tha I is easily prepared in large amounts of approximately 10^5 units per gram of cells, it will be a valuable addition to the battery of restriction enzymes used in studies of DNA sequences. It is active at high temperatures and may therefore be useful for special purposes requiring more extreme conditions.

INTRODUCTION

Restriction enzymes of type II are valuable in the analysis of the structure and function of nucleic acid sequences and they are of interest from biological and biochemical points of view. The subject has been reviewed recently (1). Many restriction enzymes have been isolated from bacteria and a few from blue-green algae, but so far none has been reported from a mycoplasma. In many cases the specific nucleotide sequence recognised by a restriction enzyme is known and the enzyme hydrolyses both strands of the DNA at or near this sequence whenever it occurs. The sequences are usually 4, 5 or 6 base pairs long and frequently have a two-fold axis of symmetry.

The type II enzymes which recognise tetranucleotide sequences are expected a priori to cut DNA molecules on average once every 256 base pairs and these are especially useful for generating small DNA fragments for DNA sequencing by the methods of Sanger and Coulson (2) and Maxam and Gilbert (3). To date enzymes are known which are specific for 6 out of the 16 possible symmetric tetranucleotide sequences. In addition the Hind II and Hinf I enzymes cut at CTPyPuAC and GANTC respectively, sequences which also occur

© Information Retrieval Limited 1 Falconberg Court London W1V 5FG England

frequently. Other enzymes such as Hin1056 I, Sac III and Mnl I cut so frequently that the recognition sequences are likely to be tetranucleotides or similar to those of Hinf I and Hind II which have some variation in the bases at one or more positions. Together these enzymes already provide a powerful battery for dissecting DNA molecules but it would be useful to have enzymes which cut different sequences and to have better sources of some of the activities which are already known. We have tested extracts of the exotic mycoplasma Thermoplasma acidophilum and found that it contains a restriction enzyme Tha I which cuts at CGCG. The remarkable growth conditions of T. acidophilum (pH 0.7 to 2 and 59°C) and the high concentration of the restriction enzyme raise interesting questions concerning the biology of this organism and the functions of restriction enzymes.

MATERIALS AND METHODS

Thermoplasma acidophilum: Strain 122-1B3 (ATCC 25905) was originally isolated from a burning refuse pile at the Friar Tuck coal mine in Indiana (4). Cells were cultured with aeration in a 115 litre container at 59°C at pH 1.7 in the medium originally described (5) supplemented with 0.05M sucrose. In late log phase the cultures were adjusted to pH 3-4 with 5M NH_4OH, cooled to 20°C and harvested by centrifugation. The cell pellets were stored at -70°C.

Assay for restriction enzyme activity: Assays were carried out in 10 µl of 0.01 M tris HCl pH 7.4, 0.01 M $MgCl_2$, 1 mM dithiothreitol containing 3 µg of λplac 5 DNA. Either 1 or 2 µl of enzyme sample were added and the reactants incubated at 60°C or 37°C for 60 minutes. The reactions were stopped by adding 1 µl of 0.5% sodium dodecylsulphate 20% sucrose 0.2% xylene cyanol FF 0.2% bromophenol blue. Aliquots of 5 µl were electrophoresed on 5% acrylamide slab gels (20 x 20 x 0.2 cm). The electrophoresis buffer was 0.1 M tris borate pH 8.3, 2 mM EDTA. Electrophoresis was at 100-250 V for 2 - 4 hours. DNA was detected by staining in ethidium bromide as described (6).

DNA sequencing: This was carried out by the method of Maxam and Gilbert (3).

DNA: DNA was purified from T7 and λplac 5 phage as described by McConnell and Bonner (7). Thermoplasma acidophilum DNA was prepared as described by Searcy and Doyle (8). ØX174 viral DNA and RF I were gifts from K. Koths and J. Sims. DNA from the plasmid p BR322 (9) was prepared according to Tanaka and Weisblum (10).

Restriction enzymes: Hae III, Hpa I and II and Taq I were prepared according

to the general method of Humphries et al (11). Hind II, Uba II and Hinf I were gifts from J. Sims, M. Zabeau and D. Hourcade respectively. Hha I was purchased from Bethesda Research Laboratories, Inc. Rockville, Md. 20850.

RESULTS

Isolation of a restriction enzyme Tha I: Frozen cells (3 g) of *Thermoplasma acidophilum* were thawed in 10 ml of 0.075 M NaCl, 0.025 M sodium EDTA pH 8.0. Tris base (1 M) was added to bring the pH to 6.3 lysing the cells. The viscous suspension was sheared in a VirTis '45' homogeniser for 10 minutes at maximum power. The homogenate was centrifuged at 30,000 g for 20 mins. and the supernatant chromatographed on a 200 ml Sepharose 2B column in 0.2 M NaCl, 1 mM EDTA pH 6.25. Nucleoprotein elutes just after the excluded volume (5) well separated from the bulk of the lower molecular weight proteins. The fractions of low molecular weight protein were combined, concentrated by dialysis against equal volume of glycerol and stored at -20°C. Restriction enzyme was detectable in this crude extract.

P-11 chromatography: One half of the extract described above was dialysed against Buffer A (5 mM tris pH 7.4, 0.1 mM dithiothreitol, 5% glycerol) and applied to a phosphocellulose (Whatman P-11) column (20 x 7 cm) equilibrated at 4°C with buffer A. The column was developed with a 300 ml gradient (0.05 M - 0.8 M KCl in buffer A). Enzyme activity eluted as a single peak at about 0.5 M KCl.

DE52 chromatography: The phosphocellulose peak fractions were combined, dialysed against buffer A containing 0.005 M KCl and applied to a DEAE cellulose (Whatman DE52) column (1.2 x 20 cm) equilibrated at 4°C with Buffer A 0.005 M KCl. A 200 ml gradient (0.005 M - 1.0 M KCl in Buffer A) was applied to the column. The enzyme eluted in the early fractions at about 0.02 M KCl.

Yield of enzyme activity: Fractions from the DE52 column containing high enzyme activity were combined and concentrated by dialysis against buffer A containing 50% glycerol and stored at -20°C. The total volume was about 10 ml. At 60°C, 1 ul of this enzyme preparation cleaved 12 ug of λplac 5 DNA in 1 hour giving a total yield from 1.5 g of cells of 120,000 units (1 unit of restriction enzyme cuts 1 μg of λ DNA in 1 hour). At 37°C the activity is about six times lower (Fig. 1). The enzyme has been completely stable at -20°C for 2 months.

Contaminant enzymes: The enzyme purified by DEAE chromatography has some 3' exonuclease but has no detectable 5' exonuclease or phosphotase activity.

Specificity of Tha I: The enzyme cuts λplac 5 and T7 DNA (Fig. 2) producing more than 50 fragments in each case. More than ten fragments are produced

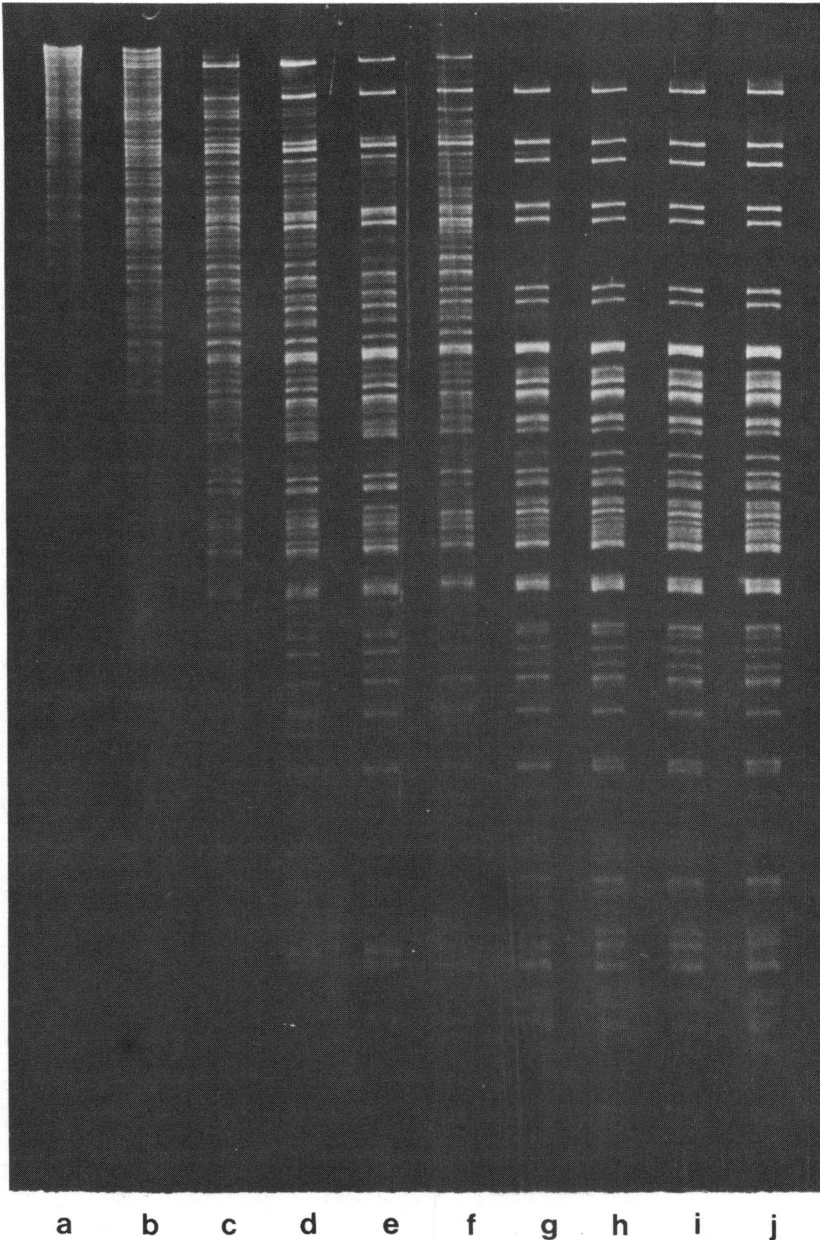

Figure 1. Tha I enzyme activity at 37°C (a - e) and 60°C (f - j). Two reactions each contained λplac 5 DNA (12 μg) and Tha I (1 μl of DE52 purified enzyme). Aliquots were taken from the reaction after 15, 30, 60, 90 and 120 minutes incubation at each temperature and electrophoresed on a 5% polyacrylamide gel. The reaction at 60°C was complete by 60 minutes. The reactions at 37°C for 90 minutes and at 60°C for 15 minutes are approximately equivalent providing a measure of comparison.

from ØX174 RF I or pBR322 DNA (Fig. 2). These results suggest either a four or five base recognition sequence or one with some flexibility such as Hind II or more than one activity. Time courses of the reaction at 37°C or 60°C show that some sites are cleaved more slowly than others (Fig. 1). This discrimination does not necessarily imply more than one enzyme activity as similar results are commonly observed for other purified enzymes and are presumably caused by differences in sequences adjacent to the recognition sequences (e.g. ref. 12).

The Tha I fragment pattern for T7 DNA differed from those produced by Hind II, Hpa I and II, Alu I, Taq I, Hha I, Hae III, Hinf I, all of which produce more than 50 fragments. Some of the comparisons are shown in Fig.2.

T. acidophilum DNA was not cleaved by Tha I, so that the recognition site must be absent or more likely modified. Traces of a base resembling 7-methylguanine have been detected in this DNA (8). Single-stranded ØX174 viral DNA was not cleaved by Tha I either at 37°C or 60°C.

Fragment pattern of ØX174 DNA: The ØX174 RF I fragment pattern for a restriction enzyme of given specificity can be determined from the complete base sequence of the DNA (13). The Tha I digest pattern corresponded closely to that predicted for cleavage at CGCG. There were three differences between the predicted and observed patterns. A large fragment of 1562 base pairs predicted by the ØX174 sequence was absent, and two smaller ones of about 700 and 870 base pairs were present though not predicted. These differences have been resolved by a correction to the published sequence which introduces another CGCG sequence about position 3548 splitting the 1562 base pair fragment (14). The ØX174 RF I fragment patterns produced by Tha I and Uba II are indistinguishable (Fig. 3), which is expected since Uba II is believed to recognise CGCG (15).

The Tha I cutting sequence: In principle the recognition sequence for a restriction enzyme may differ from the cutting sequence. This is for example true for the enzyme Hph I which recognises GGTGA and cuts the DNA 8 base pairs away in the 3' direction (16). Restriction enzymes may cut sequences symmetrically or assymmetrically. We have taken advantage of the DNA sequencing technique of Maxam and Gilbert (3) to determine directly the sequence cut by Tha I and whether the cut is symmetric or assymetric.

Consider a fragment known to contain a Tha I site and labelled on one 5' end. If this fragment is digested with the enzyme the labelled product will be a molecule which ends with the nucleotide to the 5' side of the cleavage site. If this product is electrophoresed in parallel with the products of

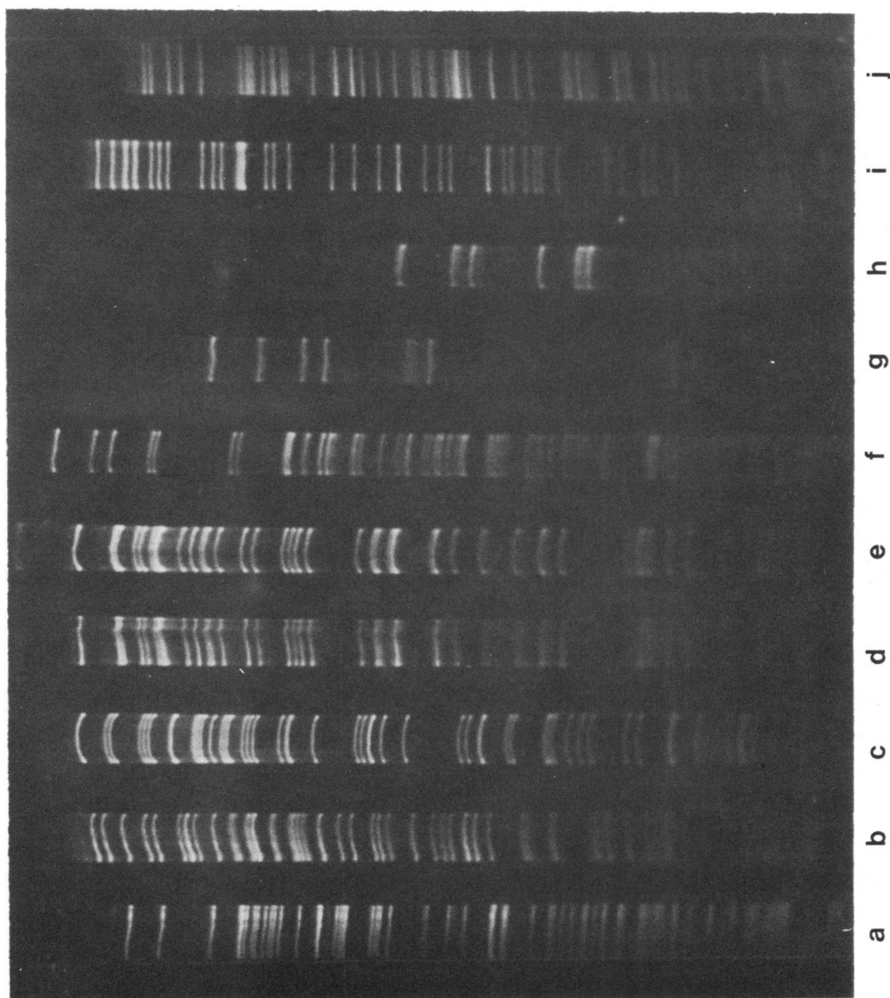

Figure 2. Digests of T7 DNA by Alu I (a), Hpa I+II (b), Hae III (c), Tha I (d,e), Hind II (i) or Hinf I (j); and of λplac 5 (f), ØX174 RF I (g) and pBR322 (h) DNA by Tha I.

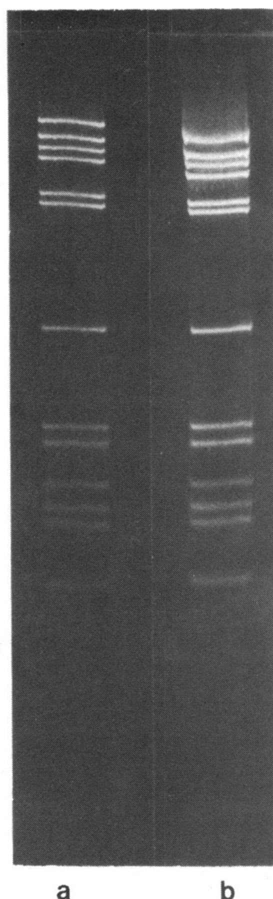

Figure 3. Digests of ØX174 RF I DNA by Uba II (a) and Tha I (b).

the four chemical reactions used for sequencing it will migrate to a point which is equivalent to the distance moved by one of the chemical products, and hence the enzymatic cleavage site can be directly ascertained. The results of such an experiment are shown in Figure 4. A fragment of pBR322 DNA labelled on one 5' end was cut by Tha I. It was also subjected to the chemical cleavages as for sequencing. The Tha I product has nearly the same mobility as a fragment cleaved to the 5' side of a C which lies in the sequence 5'-CG<u>C</u>G-3'. A fragment which appears as a result of chemical cleavage at C has lost that C and ends with the nucleotide to the 5' side of that C, in this case a G. Therefore Tha I cut this strand in the centre of the sequence CGCG. The Tha I product moves slightly more slowly than the corresponding chemical product, presumably because it lacks the 3' terminal phosphate which remains after chemical cleavage. Tha I therefore cleaves

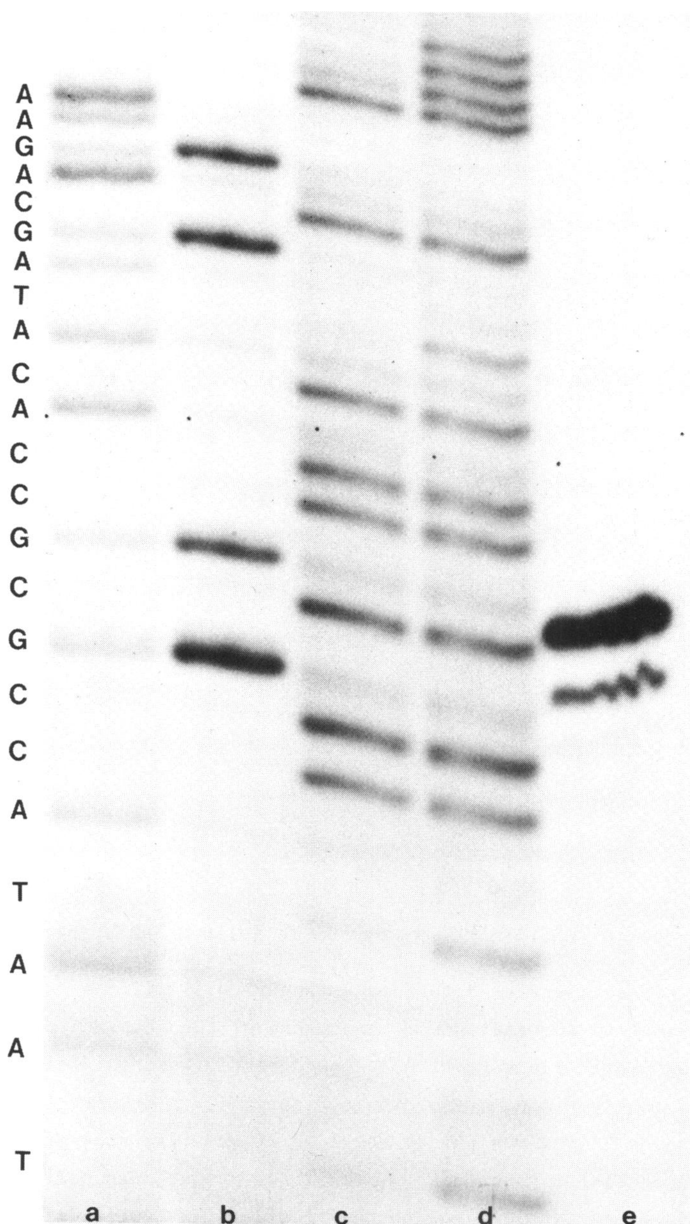

Figure 4. The sequence at a Tha I cleavage site in a fragment of pBR322 DNA. Fragments in (a)(b)(c) and (d) were produced by chemical cleavage before A>G, G>A, C and C+T respectively. The major fragment in (e) was produced by Tha I. The minor fragments are caused by a 3' exonuclease contaminant.

the 3' sugar phosphate bond. The same experiment was repeated on a different fragment of pBR322 with identical results.

Taken with the ØX174 cleavage data these observations establish that Tha I recognises the sequence CGCG and cuts the same sequence symmetrically.

The sequence CGCG is expected a priori to overlap the sequence GCGC with a frequency of 0.5, which is the probability that a G will precede or a C follow CGCG. Hence Tha I will often cut in the same DNA sequence as Hha I which recognises GCGC (17). We have observed that some Hha sites are adjacent to Tha I sites while others are not. Because Hha I cuts assymmetrically, GCGC, the single strands produced by HhaI cleavage will sometimes be the same length as those produced by Tha I when Hha I and Tha I sites coincide.

DISCUSSION

A type II restriction enzyme Tha I which recognises the tetranucleotide CGCG has been partially purified from Thermoplasma acidophilum. The yield is about 10^5 units of purified enzyme per g of cell when enzyme is assayed at 60^oC, and about six times less when assayed at 37^oC. For most practical purposes the enzyme can be used at 60^oC, and indeed this may make the enzyme especially valuable for some experiments. The yield as measured at that temperature is very high, comparable to the amount of Bsp I (2.6×10^5 units per gram) produced from Bacillus sphaericus (18), and much higher for example than the yield of Hae III from Haemophilus aegyptius (5×10^3 units per gram) (19). T. acidophilum is easy to grow and because of its remarkable growth conditions there are few problems of contamination. The purification procedure is straightforward and reproducible and the enzyme is apparently stable. The enzyme preparation is essentially free of nuclease and phosphatase activity after the DE52 column, and has been used in sequencing studies (20, 21). The Tha I digest pattern of ØX174 is indistinguishable from that predicted for any enzyme recognising CGCG.

Tha I cuts in the centre of the sequence CGCG. This was established directly by comparing the size of ^{32}P end-labelled fragments generated by Tha I with those produced by chemical cleavage during sequencing of molecules known to be cut by Tha I. Two fragments cut by Tha I were shown to be cut at CGCG. The Uba II and Tha I digest patterns of ØX174 are apparently identical. Uba II is also thought to cut at CGCG (15).

T. acidophilum was first isolated from a burning coal refuse pile (4). It was found to have pH and temperature optima of 1.7 and 60^oC respectively. It does not grow at 37^oC and lyses spontaneously at neutral pH. On the other hand it is not lysed at 100^oC(22). It has been classified as a myco-

plasma according to its morphology, pattern of sensitivity to antibiotics and DNA content (4,22). Its genome is about 30% as large as that of *E. coli* (8).

It was surprising to find a restriction enzyme in *T. acidophilum*. The discovery of Tha I in such large quantities suggests that *T. acidophilum* is susceptible to viral infection. No viruses have been described for *T. acidophilum* but viruses and biological evidence for modification and restriction have been described for other mycoplasma (23-26). Tha I does not cleave DNA from *T. acidophilum* although the DNA is 46% GC and should therefore contain many CGCG sequences. This leads to the suggestion that there is a modification system in *T. acidophilum*. This is supported by the observation of an unusual base with the same properties as N-7-methylguanine in *T. acidophilum* DNA (8).

The enzyme has been called Tha I a slight variation from the convention (1), to distinguish it easily from Taq I.

ACKNOWLEDGEMENTS

We are grateful to Dr. Jack Maniloff for helpful discussions and Dr. Wally Gilbert for the hospitality and support of his laboratory. D. McC. was supported by an Eleanor Roosevelt Fellowship of the International Union Against Cancer.

REFERENCES

1. Roberts, R.J. (1976) C.R.C. Critical Reviews in Biochemistry 4, 123.
2. Sanger, F. and Coulson, A.R. (1975) J. Mol. Biol. 94, 441.
3. Maxam, A.M. and Gilbert, W. (1977) Proc.Nat. Acad.Sci.U.S.A. 74, 506.
4. Darland, G., Brock, T.D., Samsonoff, W. and Conti, S.F. (1970) Science, 170, 1416.
5. Searcy, D.G. (1975) Biochim. Biophys. Acta 395, 535.
6. Sharp, P.A., Sugden, B. and Sambrook, J. (1973) Biochemistry, 12, 3055.
7. McConnell, D.J. and Bonner, J. (1972) Biochemistry, 11, 23.
8. Searcy, D.G. and Doyle, E.K. (1975) Int. J. Syst. Bacteriol. 25, 286.
9. Bolivar, F., Rodriguez, R.L., Boyer, H.W. and Goodman, H.M. (1977) Gene (in press).
10. Tanaka, J. and Weisblum, B. (1975) J. Bacteriol. 121, 354.
11. Humphries, P., Gordon, R.L., McConnell, D.J. and Connolly, P. (1974) Virology, 58, 25.
12. Thomas, M. and Davis, R.W. (1975) J. Mol. Biol. 91, 315.
13. Sanger, F., Air, G.M., Barrell, B.G., Brown, N.L., Coulson, A.R., Fiddes, J.C., Hutchison, C.A. III, Slocombe, P.M. and Smith, M. (1977) Nature, 265, 687.
14. Sanger, F. (1977) personal communication.
15. Zabeau, M. and Roberts, R.J. (1977) personal communication.
16. Kleid, D., Humayun, Z., Jeffrey, A. and Ptashne, M. (1976) Proc. Nat. Acad. Sci. U.S.A. 73, 293.
17. Roberts, R.J., Myers, P.A., Morrison, A. and Murray, K. (1976) J. Mol. Biol. 103, 199.

18. Kiss, A., Sain, B., Csordas-Toth, E. and Venetianer, P. (1977) Gene 1, 323.
19. Roberts, R.J., Breitmeyer, J.B., Tabachnik, N.F. and Myers, P.A. (1976) J. Mol. Biol. 91, 121.
20. McConnell, D.J. (1977) manuscript in preparation.
21. Sutcliffe, J.G. (1977) manuscript in preparation.
22. Belly, R.T. and Brock, T.D. (1972) J. Gen. Microbiol. 67, 9.
23. Gourlay, R.N. (1970) Nature 225, 1165.
24. Liss, A. and Maniloff, J. (1971) Science 173, 725.
25. Gourlay, R.N. (1974) C.R.C. Crit. Rev. Microbiol. 3, 315.
26. Maniloff, J. and Das, J. (1975) In Goulian, M., Hanawalt, P. and Fox, C.F. (ed.), DNA synthesis and its regulation. W.A. Benjamin, Inc., Reading.