Volume 229, number 2, 283–288 FEB 05658 March 1988

The bacteriophage Mu transposase protein can form high-affinity protein-DNA complexes with the ends of transposable elements of the Tn *3* family

Robin K. Cameron, Allan M. Jarjour, Peter P. Tolias and Michael S. DuBow

McGill University, Department of Microbiology and Immunology, 3775 University Street, Montreal, Quebec H3A 2B4, Canada

Received 29 December 1987

The 37 kb transposable bacteriophage Mu genome encodes a transposase protein which can recognize and bind to a consensus sequence repeated three times at each extremity of its genome. A subset of this consensus sequence (5'-PuCGAAA(A)-3') is found in the ends of many class II prokaryotic transposable elements. These elements, like phage Mu, cause 5 bp duplications at the site of element insertion, and transpose by a cointegrate mechanism. Using the band retardation assay, we have found that crude protein extracts containing overexpressed Mu transposase can form high-affinity protein-DNA complexes with Mu *att R* and the ends of the class II elements Tn *3* (right) and IS*101*. No significant protein-DNA complex formation was observed with DNA fragments containing the right end of the element IS*102*, or a non-specific pBR322 fragment of similar size. These results suggest that the Mu transposase protein can specifically recognize the ends of other class II transposable elements and that these elements may be evolutionarily related.

Transposase protein; Class II prokaryotic transposon; Protein-DNA recognition; Phage Mu

1. INTRODUCTION

Mobile genetic elements are ubiquitous in nature and share a number of common characteristics [1]. Each element is a discrete, self-contained unit that is usually bounded by inverted or directly repeated DNA sequences. Moreover, the elements are often flanked by short duplications of host DNA at the site of element insertion, the size of which is characteristic of each class of element. These elements move by a concerted nicking-ligation reaction mediated by enzymes called 'transposases' usually encoded by the elements themselves [1]. The transposase enzymes recognize specific DNA sequences in the inverted or direct repeats of the element [1,2].

The transposase protein of the transposable coliphage Mu recognizes and binds to three consensus sites in each end of the phage DNA [2,3] in order to catalyze the transposition reactions required for the propagation of the viral genome [4]. Within this consensus sequence is the sequence 5'-PuCGAAA(A)-3', which is also found in the ends of non-viral class II bacterial transposons, such as Tn *3*, $\gamma\delta$, Tn *951* and IS*101* [5,6]. Like phage Mu, these class II elements cause 5 bp duplications at the site of element insertion [7]. By using the band retardation assay [8] we demonstrate that an *Escherichia coli* crude protein extract containing overexpressed Mu transposase can specifically recognize and retard the migration of DNA fragments containing the ends of the class II elements Tn *3* and IS*101*, but not other DNA fragments including that of the non-class II element IS*102*.

Correspondence address: M.S. DuBow, McGill University, Department of Microbiology and Immunology, 3775 University Street, Montreal, Quebec H3A 2B4, Canada

Abbreviations: bp, base pair; CT, calf thymus; Tn, transposable element; IS, insertion sequence; Pu, purine; kb, kilobase pair

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/88/$3.50

2. MATERIALS AND METHODS

The Mu *A* and *B* genes were cloned into the expression vector pOP95-15K as a *BalI* fragment from Mu*cts*-13-4 [3]. The Mu *A* gene was subcloned as a 2.1 kb *Eco*RI-*Dra*I fragment from the recombinant plasmid pPZ341 which contains both the Mu *A* and *B* genes cloned in pOP95-15K [9] (fig.1). The Mu *A* gene fragment was made 'blunt-ended' by treatment with the Klenow fragment of DNA polymerase I (Pharmacia) plus dATP and dTTP in nick-translation buffer [10]. The *A* gene fragment was then ligated with the *lac* UV5 expression vector pOP95-15K [9] previously hydrolyzed with *Sma*I and treated with calf intestinal alkaline phosphatase (Dupont) [11]. Plasmid-encoded proteins were visualized by the chloramphenicol release procedure as described [12].

Cells [strain LF1900 containing plasmid pAJ11 in *E. coli*

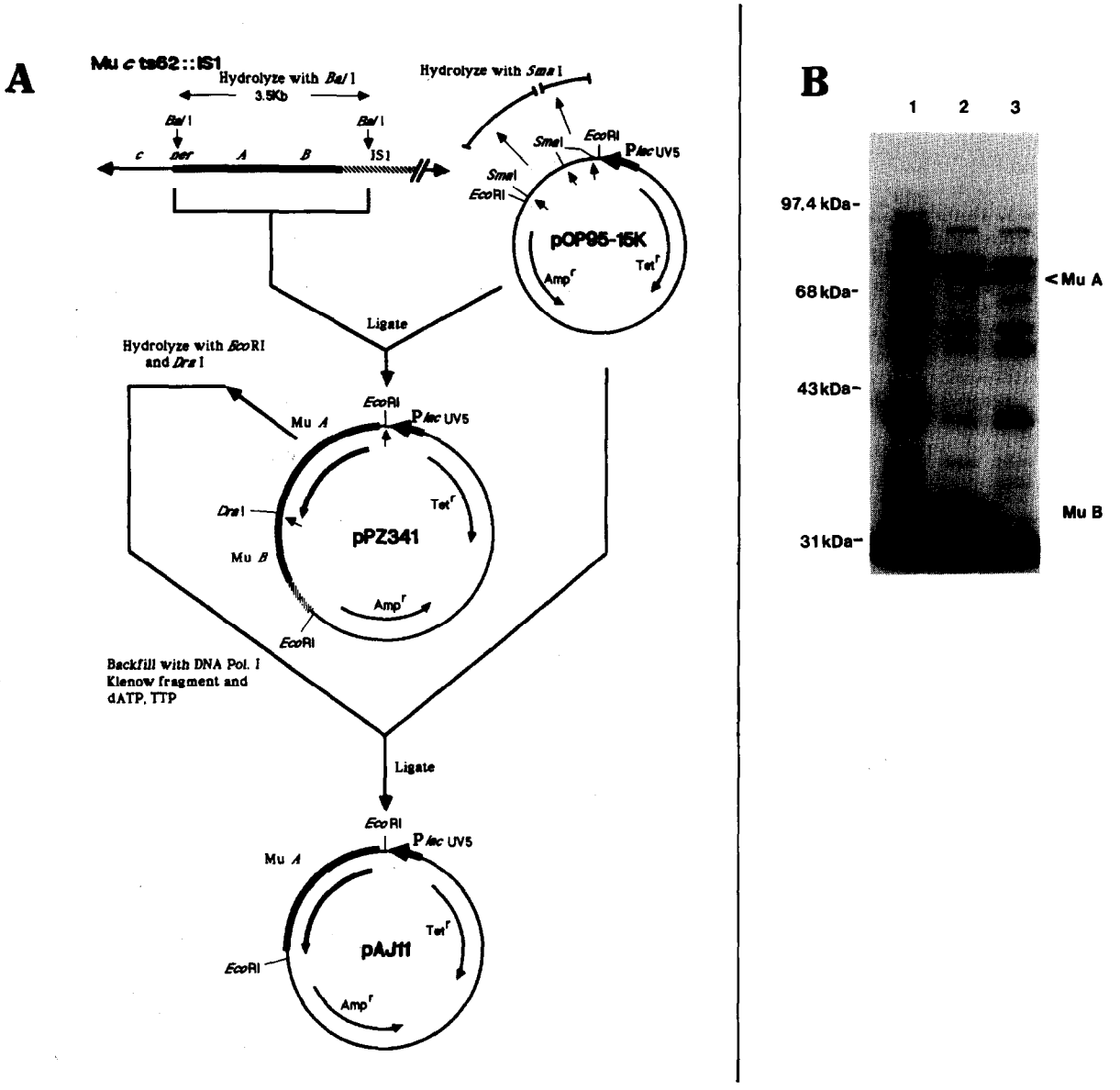

Fig.1. Cloning of the Mu *A* gene and visualization of its product. (A) The Mu *A* gene was subcloned from plasmid pPZ341 (see section 2) as a 2.1 kb *Eco*RI-*Dra*I fragment into pOP95-15K as described in section 2. (B) 10% SDS-polyacrylamide electrophoretic analysis of the plasmid-encoded gene products by means of the chloramphenicol release procedure as described in section 2. Lanes: 1, pOP95-15K; 2, pPZ341; 3, pAJ11. Arrows indicate the Mu *A* (70 kDa) and Mu *B* (33 kDa) gene products [2,27].

NM522 (*Δpro lac, rps, thi, hsd R⁻M⁺,sup E44/F',tra D36,pro AB,lac Iᵠ,* zΔM15) [13], strain LF1906 containing plasmid pPZ344 in *E. coli* NM522, and strain LF1907 containing plasmid pOP95-15K in *E. coli* NM522] were grown in LB broth [14] containing 40 μg/ml ampicillin and 1.0 mM IPTG to induce plasmid-encoded gene expression. Crude protein extracts were prepared by the method of Tolias and DuBow [15].

The DNA restriction fragments (see fig.2) used in these experiments were extracted from 5% polyacrylamide gels [10] by the 'crush and soak' method of Maxam and Gilbert [16]. These were made radioactive with [α-^{32}P]dNTPs by 'backfilling' the restriction sites at the extremities of each fragment [15]. The fragments used in these assays contained the Tn *3*-right extremity [17] as a *Hpa*II 404 bp pBR322 fragment; Mu *att R* [18] as an *Eco*RI/*Bam*HI 425 bp pJK5 fragment; the entire IS*101* element [19] as an *Ava*II/*Bgl*I 403 bp pSC101 fragment; IS*102*-right extremity [19] as a *Pvu*II/*Xho*I 396 bp pSC101 fragment; and a non-specific control DNA as a *Sau96A* 350 bp pBR322 fragment [20].

Specific protein-DNA interactions were detected using a modification [9] of the band retardation assay of Strauss and Varshavsky [8]. We determined specific binding of the various fragments (see fig.2) by a crude protein extract containing overexpressed Mu *A* protein using the following conditions: 20 μg of crude protein extract, 5 ng of an (α-^{32}P)-labelled fragment, and increasing concentrations of unlabelled competitor DNA (sonicated CT DNA). The reactions and subsequent gel electrophoresis conditions were carried out according to Tolias and DuBow [15] in a reaction volume of 20 μl.

3. RESULTS

The plasmids pAJ11 and pPZ344 (fig.1A) overproduce the Mu transposase protein (70 kDa) under *lac* UV5 transcriptional control [21]. The predicted plasmid-encoded proteins were visualized by the chloramphenicol release procedure (fig.1B). A crude protein extract made from *E. coli* strain LF1900 harbouring plasmid pAJ11 and overexpressing the Mu transposase (*A* gene product) was used in the band retardation assay with the various class II and control DNAs (fig.2). Specific DNA-protein interactions can be detected in this assay because the mobility of the protein-DNA complexes is reduced in low ionic strength polyacrylamide gels. Moreover, the relative affinity of the transposase protein in the crude extract for the DNA restriction fragments can be measured by its ability to recognize a labelled DNA fragment and cause the fragment's retardation in reactions containing increasing amounts of unlabelled competitor calf thymus (CT) DNA.

In fig.3A, the Mu *att R* DNA substrate was greatly retarded in its migration by the

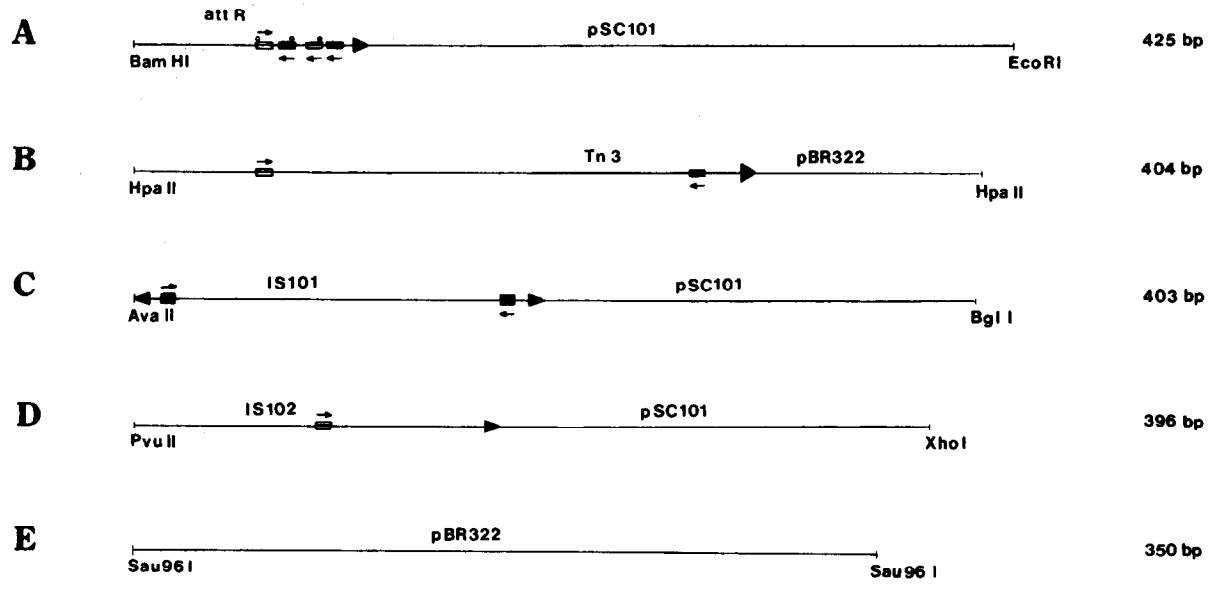

Fig.2. Fragments used in the band retardation assay. All fragments were labelled at their extremities as described in section 2. (A) Mu *att R* DNA substrate; (B) Tn *3*-right DNA substrate; (C) IS*101* DNA substrate; (D) IS*102*-right DNA substrate; (E) non-specific DNA substrate. Closed boxes indicate the location of the consensus sequence 5'-ACGAAAA-3', open boxes, 5'-ACGAAA-3', and boxes with a circle indicate that the purine at the beginning of the sequence is a G. Small arrows above or below the boxes indicate the direction of the sequence. Large arrows indicate the element end.

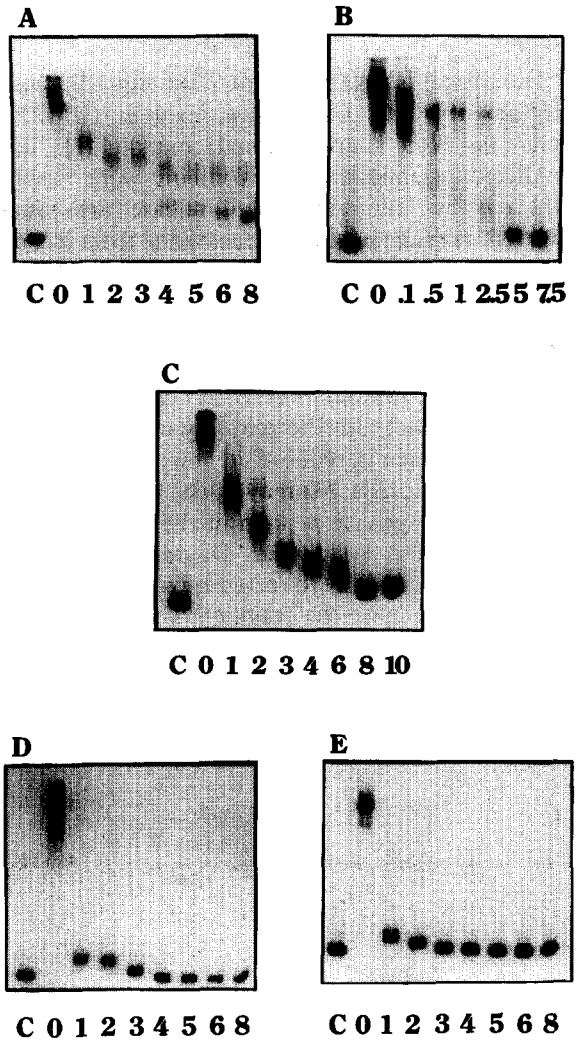

Fig.3. Detection of relative affinity of Mu transposase protein DNA-binding activity by the band retardation assay. In A–E, the numbers along the bottom represent the quantity of sonicated CT DNA (in μg/reaction) present in the reaction (run as described in section 2). Lanes marked C contain labelled DNA substrate incubated without crude protein extract. 5 ng of each DNA substrate were added per reaction along with 20 μg crude protein extract containing the overexpressed Mu transposase protein. (A) Mu *att*R; (B) IS*101*; (C) Tn*3*-right; (D) IS*102*; (E) the non-specific pBR322 derived fragment.

transposase-containing extracts, even at concentrations up to 8 μg (0.4 mg/ml) of competitor CT DNA per reaction. However, the non-specific DNA substrate (fig.3E) was retarded in its migration only when competitor DNA was not added to the reaction. By approx. 2–3 μg of competitor DNA per reaction (0.1–0.15 mg/ml), no significant protein-DNA complexes were observed (fig.3E).

The DNA from the ends of the two class II elements, IS*101* and Tn *3* were retarded in tneir migration with the addition of Mu transposase overexpressing crude extracts (figs 3B and C). The IS*101* substrate (fig.2B) was still retarded in its migration at 7.5 μg of competitor DNA per reaction (0.38 mg/ml). The Tn *3*-right DNA fragment was retarded in its migration even at concentrations of CT DNA up to 10 μg/reaction (0.5 mg/ml), but to a lesser degree than the IS*101* fragment.

The IS*102*-right substrate was slightly retarded in its migration (fig.3D) when compared to the non-specific DNA fragment (fig.3E), but the extent of retardation was much less than that seen with the ends of the class II elements and Mu *att R*. Moreover, no protein-DNA complexes were observed at CT competitor concentrations of greater than 3–4 μg per reaction (0.15–0.2 mg/ml).

4. DISCUSSION

The results presented here report evidence that phage Mu transposase protein (present in crude extracts) possesses the capacity to form specific protein-DNA complexes with the extremities of non-viral class II transposable elements. Mu transposase can bind to and retard the migration of DNA fragments containing the ends of Tn *3* and IS*101* with a much higher affinity than was observed with the non-class II element end IS*102*-right. It is interesting to note, however, that the ends of Tn *3* and IS*101* were retarded in their migration to a lesser extent than was the Mu *att R* substrate.

As expected, Mu transposase containing crude extracts could not form high-affinity protein-DNA complexes with the non-class II element IS*102*. The IS*102* restriction fragment fortuitously contained a presumed transposase recognition subsequence 5′-ACGAAA-3′ located 200 bp from the IS*102* right-end. The slightly greater retardation of this IS*102*-right substrate, compared to the non-specific substrate, may be attributed to this sequence. This result plus the additional fact that Mu transposase binds to three consensus sites (which

contain the sequence 5′-PuCGAAA(A)-3′) in its ends [2], suggest that other DNA sequences must be present and correctly organized in the ends of a transposable element in order to obtain significant transposase protein-DNA complex formation.

When the ends of class II transposable elements (Mu, Tn *3*, *γδ*, Tn *951* and IS*101*) are aligned, it can be seen that the sequence 5′-PuCGAAA-(A)-3′, is also aligned and begins 21 bp from the element ends. However, the class II DNA sequences surrounding the sequence 5′-ACGAA-AA-3′ in the ends of these elements show no homology to this same region in the Mu extremities. When compared to each other, these class II elements (Tn *3*, *γδ*, Tn*951*, and IS*101*) do have very homologous ends. The 5′-PuCGAAA-(A)-3′ sequence and its location in the element ends appear to have been conserved and both are probably important in transposase recognition of the element extremities. For those class II elements (Tn *21*, Tn*501*, Tn *1721*) which do not contain 5′-PuCGAAA(A)-3′, a very A-rich sequence begins at bp 21 [3]. It has been shown that insertions or deletions of an adenine residue in the run of As present in 5′-ACGAAAA-3′ in the right-end of Mu results in mutant phages which are incapable of transposing [22]. Perhaps these 3–4 A-T base pairs, present in many class II elements and in Mu and the related coliphage D106 [23], are crucial residues which interact with the transposase protein for protein-DNA recognition.

The number of binding sites located in the ends of the elements may also play a role in the strength of transposase binding. IS*101* was cloned as a complete element and therefore contained two 5′-ACGAAAA-3′ sites in the correct orientation, whereas the Tn *3*-right substrate contained only the right end, plus a fortuitous 5′-ACGAAA-3′ located 194 bp from the right-end that is not required for transposition [17]. This may explain why IS*101* was retarded by transposase containing crude extracts to a greater extent than the Tn *3*-right fragment.

Tn *501* and Tn *21* are class II transposable elements which do not contain the sequence 5′-PuCGAAA(A)-3′ in their ends [5]. However when the protein sequence of the Tn *501* and Tn *21* transposases were compared with that of Tn *3*, their amino-terminal regions were found to be 53% homologous [24], suggesting that these class II transposase proteins once shared a common ancestor. Moreover, IS*30* transposase (a non-class II element) is significantly homologous to phage Mu repressor protein in both the amino (DNA-binding region)- and the carboxyl-terminal regions [25]. Additionally, the amino-termini of the Mu [26,27] and D108 [15] transposase proteins contain the sequence specific DNA-binding domains that are homologous to their respective repressor proteins [12,27,28] and are capable of recognizing each others extremities, even though they differ in amino acid sequence. It is interesting to note, however, that Mu transposase and IS*30* transposase exhibited little homology to one another [25]. Because we have shown that Mu transposase can recognize the ends of the class II elements Tn *3* and IS*101*, we can speculate that Mu and these class II transposases, Tn *3*, Tn *501*, Tn *21*, along with that of IS*30* and Mu repressor, may have evolved from a distant ancestral DNA-binding protein (or domain) encoded by a mobile genetic element, and have maintained a strong selection for the conservation of their mechanism of DNA substrate recognition.

Acknowledgements: We would like to thank Paul Zanga for the construction of plasmid pPZ344 and Piet van de Putte for helpful discussions. M.S.D. is a Chercheur Boursier of the Fonds de la Recherche en Santé du Québec and R.K.C. is supported by an Agriculture Canada Studentship (86027). This work was supported by the Medical Research Council of Canada grant MA6751.

REFERENCES

[1] Shapiro, J.A. (1983) Mobile Genetic Elements, Academic Press, New York.
[2] Craigie, R., Mizuuchi, M. and Mizuuchi, K. (1984) Cell 39, 387–394.
[3] Groenen, M.A.M., Timmers, E. and Van de Putte, P. (1985) Proc. Natl. Acad. Sci. USA 82, 2087–2091.
[4] Howe, M.M. (1987) in: Phage Mu (Symonds, N. et al. eds) pp.25–40, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
[5] Heffron, F. (1983) in: Mobile Genetic Elements (Shapiro, J.A. ed.) pp.223–260, Academic Press, New York.
[6] Reed, R., Young, R., Steitz, J.A., Grindley, N. and Guyer, M. (1979) Proc. Natl. Acad. Sci. USA 76, 4882–4886.
[7] Calos, M.P. and Miller, J.H. (1980) Cell 20, 579–595.
[8] Strauss, F. and Varshavsky, A. (1984) Cell 37, 889–901.
[9] Tolias, P.P. and DuBow, M.S. (1985) EMBO J. 4, 3031–3037.

[10] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
[11] Ullrich, A., Shine, J., Chirgivin, J., Pictet, R., Tischer, R., Rutter, W.J. and Goodwin, H.M. (1977) Science 196, 1313–1319.
[12] Levin, D.B. and DuBow, M.S. (1987) FEBS Lett. 222, 199–203.
[13] Gough, J.A. and Murray, N.E. (1983) J. Mol. Biol. 166, 1–19.
[14] Miller, J.H. (1972) Experiments in Molecular Genetics, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
[15] Tolias, P.P. and DuBow, M.S. (1987) Virology 157, 117–126.
[16] Maxam, A.M. and Gilbert, W. (1980) Methods Enzymol. 65, 499–560.
[17] Heffron, F. and McCarthy, B.J. (1979) Cell 18, 1153–1163.
[18] Harel, J. (1986) PhD Thesis, McGill University.
[19] Bernardi, G. and Bernardi, F. (1984) Nucleic Acids Res. 12, 9415–9426.
[20] Sutcliffe, J.G. (1978) Cold Spring Harbor Symp. Quant. Biol. 43, 77–90.
[21] Fuller, F. (1982) Gene 19, 43–54.
[22] Burlingame, R.P., Obukowicz, M.G., Lynn, D.L. and Howe, M.M. (1986) Proc. Natl. Acad. Sci. USA 83, 6012–6016.
[23] DuBow, M.S. (1987) in: Phage Mu (Symonds, N. et al. eds) pp.201–214, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
[24] Evans, L.R. and Brown, N.L. (1987) EMBO J. 6, 2849–2853.
[25] Dalrymple, B. (1986) FEBS Lett. 208, 7–10.
[26] Nakayama, C., Teplow, B. and Harshey, R. (1987) Proc. Natl. Acad. Sci. USA 84, 1809–1813.
[27] Harshey, R.M., Getzoff, E.D., Baldwin, D.L., Miller, J.L. and Chaconas, G. (1985) Proc. Natl. Acad. Sci. USA 82, 7676–7680.
[28] Toussaint, A., Faelen, M., Desmet, L. and Allet, E. (1983) Mol. Gen. Genet. 190, 70–79.