© 1993 Oxford University Press
Nucleic Acids Research, 1993, Vol. 21, No. 19 4435–4443

A deletion mutant of the type IC restriction endonuclease *Eco*R124I expressing a novel DNA specificity

Agnes Abadjieva, Jay Patel, Michelle Webb, Vitaly Zinkevich and Keith Firman*
Biophysics Laboratories, School of Biological Sciences, The University of Portsmouth, St Michael's Building, Portsmouth PO1 2DT, UK

Received August 10, 1993; Accepted August 20, 1993

ABSTRACT

We have developed a complementation assay which allows us to distinguish between mutations affecting subunit assembly and mutations affecting DNA binding in the DNA recognition subunit (HsdS) of the multimeric restriction endonuclease *Eco*R124I. A number of random point mutations were constructed to test the validity of this assay. Two of the mutants produced were found to be truncated polypeptides that were still capable of complementation with the *Eco*R124I Hsd subunits to give an active restriction enzyme of novel DNA specificity. The N-terminal variable domain (responsible for recognition of GAA from the *Eco*R124I recognition sequence GAAnnnnnnRTCG) and the spacer region (central conserved region) is intact in both of these mutants. One of these mutant genes (*hsdS*(Δ*50*)) has been cloned as an active Mtase. Purification of the Mtase proved to be difficult because the complex is weak. However, Mtase activity was obtained from a soluble cell extract, and this allowed us to determine the DNA recognition sequence of the Mtase to be GAAnnnnnnnTTC. This recognition sequence is an inverted repeat of 5′-end of the *Eco*R124I recognition sequence. This suggests that the mutant Mtase is assembled from two inverted HsdS(D50) subunits, possibly held together by the HsdM subunits.

INTRODUCTION

Type I restriction endonucleases are multimeric proteins comprised of three subunits coded for by the *hsd* genes (for recent reviews see Refs. [1, 2]). The *hsdR* gene codes for the HsdR subunit which is absolutely required for restriction; while the *hsdM* and *hsdS* genes are capable of producing an active DNA methyltransferase (Mtase) independent of the presence of HsdR [3]. HsdS is responsible for DNA recognition while binding of the co-factor required for methylation, *S*-adenosyl methionine (SAM), is provided by HsdM [3–6] (D. Dryden, personal communication). The restriction endonuclease requires all three gene products as well as SAM, ATP and Mg^{2+} as co-factors, and is a multifunctional enzyme capable of both restriction or methylation of DNA dependent upon the methylation status of the recognition sequence [7–11].

The recognition sequences of type I restriction endonucleases are unusual in that they are 'split-sequences' with two specific components separated by a non-specific spacer (see Figure 1).

This suggests that DNA recognition is determined by two domains, separated by a spacer region. In an elegant series of 'domain swapping' experiments it was shown that recognition of the 5′ specific-bases is accomplished by the N-terminal region of the HsdS subunit, and that recognition of the 3′-end of the DNA sequence was accomplished by the C-terminal region of HsdS [13–18].

The type I restriction endonucleases fall into three families based on direction of transcription, protein homology, enzymatic properties and cross-reactivity between antibodies [13, 19–24]. The central conserved region has been shown to encode the spacer that separates the two DNA binding domains [13–15, 25]. The DNA recognition domains lie within the two variable domains, although only specific regions of these domains are absolutely involved in DNA recognition [16].

The recognition sequences of the type IC restriction endonuclease *Eco*R124I and *Eco*R124/3I differ by only one extra nucleotide in the non-specific spacer (Figure 1). When the *hsdS* genes of these two systems were compared following DNA sequence analysis, the only difference was found to be an extra 12-bp repeat within the central conserved region of *Eco*R124/3I. *Eco*R124I was found to have two such repeats and *Eco*R124/3I was found to have three [12]. This strongly suggests that the central conserved region is a spacer separating two DNA binding domains. Extensive mutagenesis of this central conserved region has shown that this section of the protein must be flexible and does indeed define a spacer-region [25]. It has also been proposed that the conserved regions of type I restriction endonucleases may be involved in protein–protein interactions between the three subunits that compose the restriction endonuclease [26].

In this paper we describe a complementation assay which can distinguish DNA-binding mutants from those affected in subunit assembly. A number of point mutants were prepared to test the validity of this assay. Two of these mutants resulted in deletion of the C-terminal conserved region, and part of the C-terminal variable region of the protein and produced a methylase with a novel DNA specificity.

* To whom correspondence should be addressed

MATERIALS AND METHODS

Bacterial strains, bacteriophage and plasmids

Table 1 lists the bacterial strains, bacteriophage and plasmids used in this work. All bacterial cultures were grown in LB supplemented with antibiotics as appropriate (final concentration used were: ampicillin (Ap)—150μg/ml; chloramphenicol (Cm)—25μg/ml; kanamycin (Km)—70μg/ml).

In vivo tests for restriction and modification

High level expression of *hsdS* and *hsdM* was accomplished as previously described [3]. Bacteriophage lambda was grown on lawns of indicator bacteria as previously described [4], and the efficiency of plating (e.o.p.) determined as the fraction of bacteriophage surviving on the restricting host compared to those growing on the permissive host.

ssDNA isolation and site-directed mutagenesis

ssDNA was isolated as described by Amersham International plc in the mutagenesis kit (RPN1523) except that a DNaseI + RNase I digestion of the resuspended phage pellet was performed to ensure no contaminating chromosomal DNA or RNA was present.

Misincorporation mutagenesis

This was carried out as previously described [27]. The misincorporated nucleotide was dCTP, the primer annealed to bases 973–989, the single-stranded template DNA was produced using *E.coli* CJ236 in order to allow incorporation of uracil [28]. This DNA gave a strong selection for mutant DNA over the uracil containing parental DNA when transformed into wild type *E.coli* strains.

Methylation assay

Soluble and pellet fractions, and protein for SDS PAGE were isolated as described in [3]. For the methylation assay 500mg of frozen cell pellet was resuspended in 1ml of distilled water. Cells were disrupted by sonication and insoluble debris removed by centrifugation at 12,000g for 10 min. The presence of the HsdM and HsdS(Δ50) subunits was confirmed by SDS-PAGE. 6μl of this soluble extract was mixed with 10μg of cccDNA and the solution was made 50mM with Tris–HCl (pH 8.0) and 10mM EDTA (pH 8.0) in a final volume of 100μl. Following addition of ^{3}H-*S*-adenosyl methionine (85Ci/mmol–Amersham) to a final concentration of 2.5μM, 15μl aliquots were removed at specific time points, incubated at 65°C for 10 minutes, and then kept on ice for further analysis.

Nucleic acid manipulations

Tritiated DNA produced in the above methylation reaction was loaded onto a 1% low melting agarose gel with 0.1% SDS in the electrophoresis buffer and the loading buffer. The DNA bands were cut from gel, melted at 65°C and mixed with 5 ml of Optiphase (Amersham) scintillation fluid. Incorporation of tritium into the DNA was confirmed by liquid scintillation counting.

To identify the labelled restriction fragments the DNA, from the 4.5 hour aliquot of the previously described methylation assay, was purified by phenol-chloroform extraction, ethanol precipitated, resuspended in a suitable buffer and restricted with the required restriction enzyme following the manufacturer's instructions. The restricted DNA was then separated on either a 6% polyacrylamide gel, or a 0.8% agarose gel as appropriate. The gel was soaked with a fluor (Amplify–Amersham) for 30 min and, after drying, a fluorograph of the gel was prepared [29].

pES491 was constructed from pJS491 [3] by cloning the *Nde*I-*Bam*HI fragment carrying the *hsdS* gene between the *Nde*I and *Bam*HI sites of pET3b [30]. pUS491 was produced from pES491 by cloning the *Xba*I-*Bam*HI fragment of pES491 between the *Xba*I and *Bam*HI sites of pUC119 [31].

RESULTS

Competition assay for identification of DNA binding mutations of HsdS(R124)

The HsdS subunit of the multimeric type I restriction endonucleases is responsible for DNA recognition and can, therefore, be imagined to have two main functions. The first is DNA binding and the other is the ability to interact with the other subunits in the assembly of the active endonuclease. Indeed point mutations within the HsdS subunit of *Eco*K have been associated with interactions between the HsdR and HsdS subunits [32, 33]. In order to be able to isolate DNA binding mutations within the *hsdS* gene of *Eco*R124I it was necessary to be able to distinguish mutants affected in subunit assembly from those affected in DNA binding.

In a complementation experiment in which HsdS(R124) is produced in the presence of *Eco*R124/3I there should be competition between the HsdS(R124) and Hsd(R124/3) subunits for assembly into active endonucleases. Such competition was observed when the *Eco*R124/3I enzyme was produced by pKF650, and pJS491 produced HsdS(R124). Restriction by *Eco*R124/3I was measured as a reduction in the e.o.p. of λ_{vir}.R124, while restriction by *Eco*R124I was measured as a reduction in the e.o.p. of λ_{vir}.R124/3. The competition was reflected in a 10–100 fold reduction in the level of *Eco*R124/3I restriction (Table 2). By analogy, the ability of a mutant (restriction-deficient) HsdS(R124) subunit to also compete with HsdS(R124/3) for production of an active endonuclease would be a strong indicator that restriction-deficiency was due to a failure to bind DNA rather than conformational changes of the mutant HsdS(R124) which prevent subunit assembly.

Therefore, mutant HsdS protein was produced in the presence of pKF650 (*Eco*R124/3I) and the level of restriction of either λ_{vir}.R124 and λ_{vir}.R124/3 compared to that observed for pKF650 (*Eco*R124/3I) alone, and that of pKF650+pJS491 (HsdS(R124)). Mutants showing a reduced level of restriction against λ_{vir}.R124/3 were scored as restriction-deficient; those that show a reduced level of restriction against λ_{vir}.R124 were scored as proteins capable of protein–protein interactions.

Misincorporation mutagenesis within the distal variable region of HsdS(R124)

To investigate the capabilities of the above competition assay a series of point mutations within the C-terminal region of HsdS(R124) were produced. A 17-mer oligonucleotide that hybridises to bases 973–989 of the *hsdS*(R124) gene was used to prime misincorporation mutagenesis of pUS491, covering a 200-bp region, 80-bp from this primer. The limiting nucleotide used was dCTP. 60 colonies were picked and ssDNA isolated for DNA sequencing. Single-track (C-track) sequencing was performed and those isolates showing changes in the DNA sequence were noted.

DNA from the same 60 isolates was transformed into JM109(DE3)[pKF650] and the level of restriction determined

Table 1. Bacterial strains, bacteriophage and plasmids.

Bacterial strain	Genotype or phenotype	Source
JM109(DE3)	F' *traD36, lacI, D(lacZ)M15 proAB/recA1, endA1, gyrA96* (Nal^r), *hsdR17, mcrA, relA1, supE, sbcBC, thi-1, D(lac-proAB)* l(DE3)	Promega
C600	F$^-$ e14$^-$*(mcrA$^-$) hr-1, leuB6, thi-1, lacY1, supE44, rfbD1, fhuA21.*	[35]
CJ236	F'*cat* (=pCJ105;M13^{s}Cm^r) *dut, ung1, thi-1, relA1* /pCJ105(Cm^r)	[36]
Bacteriophage		
λ_{vir}		[37]
KO7	M13 helper phage, Kn^R	[38]
M13mp9		[39]
ϕX174		[40]
Plasmids		
pUS491	*hsdS*(R124) under the control of P_{lac}	This work
pJS491	A derivative of pUC119 and pET3a over-producing HsdS(R124) from the T7 promoter	[3]
pJS4(Δ50)	The equivalent of the above plasmid but carrying *hsdS*(Δ50)	This paper
pJS4M	A derivative of pUC119 and pET3a over-producing M·*Eco*R124I Mtase from T7 promoter(s).	[3]
pKF600	pACYC184 carrying the *BglII-HindIII* fragment from pCP1005 (Firman *et al.*, 1985). Res$^+$Mod$^+$ (R124)	This work
pKF650	pACYC184 carrying the *BglII-HindIII* fragment from pUNG31 (Firman *et al.*, 1985). Res$^+$Mod$^+$ (R124/3)	[3]
pJS4(Δ50)M	*hsdS*(Δ50), hsdM	This paper
pUMI20	*hsdM*(R124) cloned into pUC118 under control of the *lac* promoter	[3]
pCP1005	pUR51 containing the *hsd* genes of *Eco*R124I on a 17kbp *HindIII* fragment	[41]
pUNG31	pBR322 containing the *hsd* genes of *Eco*R124/3I on a 17kbp *HindIII* fragment	[41]
pUC9	*lacZ', lacI,* Ap^R,	[42]
pBR322	Ap^R, Tc^R	[43]
pACYC184	Cm^R, Kn^R	[44]

Table 2. Complementation assay used to identify potential DNA binding mutants in HsdS.

Mutant No.	1	2	3	4	5	6	**7**	8	9	**10**	11	12	13	14	***15***
λ_{vir}.R124/3	10^{-4}	10^{-4}	10^{-4}	10^{-4}	10^{-4}	10^{-3}	**10^{-2}**	5×10^{-4}	10^{-4}	**10^{-2}**	10^{-4}	10^{-4}	10^{-4}	10^{-4}	***1.0***
λ_{vir}.R124	10^{-2}	10^{-2}	10^{-2}	10^{-2}	10^{-2}	10^{-2}	**10^{-2}**	10^{-2}	10^{-2}	**10^{-2}**	10^{-2}	10^{-2}	10^{-2}	10^{-2}	***10^{-4}***
Mutant No.	**16**	**17**	18	19	**20**	**21**	22	23	**24**	25	**26**	27	28	29	30
λ_{vir}.R124/3	**10^{-1}**	**10^{-2}**	10^{-4}	5×10^{-4}	**10^{-3}**	**10^{-2}**	10^{-4}	10^{-4}	**10^{-1}**	5×10^{-4}	**10^{-2}**	10^{-4}	10^{-4}	5×10^{-4}	10^{-4}
λ_{vir}.R124	**10^{-2}**	**10^{-2}**	10^{-2}	10^{-2}	**10^{-2}**	**10^{-2}**	10^{-2}	10^{-2}	**10^{-2}**	10^{-2}	**10^{-2}**	10^{-2}	10^{-2}	10^{-2}	10^{-2}
Mutant No.	31	**32**	**33**	**34**	**35**	36	37	**38**	39	40	41	42	43	44	45
λ_{vir}.R124/3	10^{-4}	**5×10^{-3}**	**10^{-3}**	**10^{-1}**	**10^{-2}**	10^{-4}	10^{-4}	**10^{-2}**	10^{-4}	10^{-4}	10^{-4}	10^{-4}	10^{-4}	5×10^{-4}	10^{-4}
λ_{vir}.R124	10^{-2}	**10^{-2}**	**10^{-2}**	**10^{-2}**	**10^{-2}**	10^{-2}	10^{-2}	**10^{-2}**	10^{-2}	10^{-2}	10^{-2}	10^{-2}	10^{-2}	10^{-2}	10^{-2}
Mutant No.	46	47	**48**	**49**	*50*	51	52	***53***	54	55	*56*	57	58	59	**60**
λ_{vir}.R124/3	10^{-4}	10^{-4}	**5×10^{-3}**	**10^{-2}**	*5×10^{-5}*	10^{-4}	10^{-4}	***1.0***	10^{-4}	10^{-4}	*5×10^{-5}*	10^{-4}	10^{-4}	10^{-4}	**10^{-2}**
λ_{vir}.R124	10^{-2}	10^{-2}	**10^{-2}**	**10^{-2}**	*10^{-2}*	10^{-2}	10^{-2}	***10^{-4}***	10^{-2}	10^{-2}	*10^{-2}*	10^{-2}	10^{-2}	10^{-2}	**10^{-2}**

The e.o.p. of lambda was determined as described materials and methods. All complementation assays were carried out in JM109(DE3) in the absence of IPTG induction. The potential DNA binding mutants are shown in **bold**, unfolded mutants are shown in ***bold+italic***, the 'hyper-restriction' mutants are shown in *italic+underlined*.

against λ_{vir}.R124 or λ_{vir}.R124/3. Table 2 shows that there are 39 wild-type isolates and 17 mutants, showing reduced levels of restriction against λ_{vir}.R124/3, were detected. All 17 restriction-deficient isolates were associated with changes in DNA sequence [28]. This indicates that the competition assay is sufficiently sensitive to detect mutations within the *hsdS* gene. Isolate numbers 15 and 53 were found not to give a positive competition assay and are, therefore, presumed to be altered in their protein configuration and cannot assemble with the other subunits. This is reflected in their total inability to restrict lambda DNA. The point mutations, resulting in alleviation of restriction, were found to be distributed over 63 amino acids between amino acid numbers 233 and 296 (Figure 2).

Mutations exhibiting elevated levels of restriction

The mutations described above were produced within the distal variable region of HsdS(R124) subunit and were expected to comprise mainly DNA binding mutations. However, two isolates (50 and 56) gave unexpectedly high levels of restriction against λ_{vir}.R124/3. ssDNA was isolated from these mutants and the

Figure 1. Regions of amino acid identities between the HsdS subunits of type IC restriction and modification systems. The regions of amino acid identity are shown by the hatched boxes. As with type IA and type IB HsdS subunits there is a large central conserved region and a smaller distal conserved region. The arrows within the central conserved region represent the repeats of 4 amino acids (TAEL), two in *Eco*R124I, and three in *Eco*R124/3I and *Eco*DXX1I [12] (T.Bickle, personal communication).

DNA fully sequenced to determine the precise change(s) that result in this increase in restriction [28]. Mutant 50 was found to be an A to G transition at position 2681 (amino acid 245), resulting in the introduction of a stop codon. Mutant 56 was found

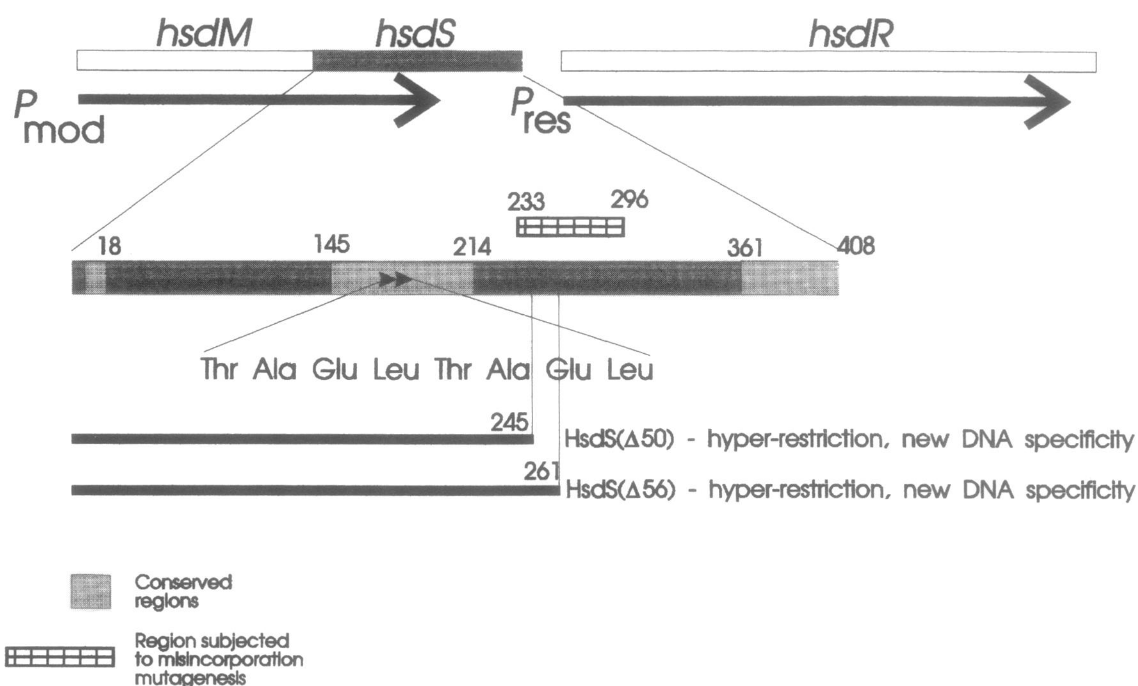

Figure 2. End-points of the deletion derivatives of the *hsdS* gene of *Eco*R124I and location of the point mutations produced by misincorporation mutagenesis. The numbers shown above the *hsdS* gene are the amino acid numbers of HsdS that define the start and end points of the protein, the regions of homology and the sites of mutagenesis.

Table 3. R-M phenotype of the deletion mutants HsdSΔ50 and HsdSΔ56.

Modified phage	e.o.p. of bacteriophage λ on bacterial strains:						
	C600	C600 [pCP1005] (*Eco*R124I)	C600 [pKF650] (*Eco*R124/3I)	JM109 (DE3) [pKF650] [HsdSΔ50]	JM109 (DE3) [pKF650] [HsdSD56]	JM109 (DE3) [pJS491] [pKF650]	JM109 (DE3) [pKF650] [pJS4M]
λ_{vir}.0	1.0	10^{-4}	10^{-4}	10^{-4}	10^{-4}	N.T.	10^{-3}
λ_{vir}.[pCP1005] (*Eco*R124I)	1.0	1.0	10^{-4}	10^{-3}	10^{-3}	10^{-2}	10^{-3}
λ_{vir}.[pKF650] (*Eco*R124/3I)	1.0	10^{-4}	1.0	10^{-4}	10^{-4}	10^{-4}	10^{-2}
λ_{vir}.[pKF650] [HsdSΔ50]	1.0	1.0	1.0	1.0	1.0	N.T.	1.0
λ_{vir}.[pKF650] [HsdSD56]	1.0	1.0	1.0	1.0	1.0	N.T.	1.0
λ_{vir}.[pCP1005] [pKF650]	1.0	1.0	1.0	10^{-4}	10^{-3}	N.T.	1.0

The efficiency of plating were determined as in materials and methods.
N.T.—Not tested.

to be an additional adenine at position 2728 resulting in a frame-shift and stop codon 24 amino acids later (The basepair co-ordinates used are described in Ref.[25]). Both mutants gave a positive competition assay indicating that they can compete for assembly with HsdR and HsdM. It would appear, therefore, that the C-terminal region of HsdS is not an absolute requirement for assembly into an active endonuclease (Figure 2).

Table 3 shows that lambda recovered from strains carrying both *Eco*R124/3I (pKF650) and mutant 50, or 56, is resistant to both *Eco*R124I and *Eco*R124/3I restriction. This suggests that the mutant HsdS subunits are still capable of M·*Eco*R124I methylation. However, when these mutants were screened with lambda DNA which has been 'double-modified' against both *Eco*R124I and *Eco*R124/3I R-M systems both mutants still show restriction (Table 3, line 6). This suggests that the mutant R-M system is both capable of modifying against *Eco*R124I restriction, and restricting with a new DNA specificity. This is the result expected from a mutant that has a degenerate *Eco*R124I DNA specificity.

Cloning *hsdS(Δ50)*

The mutant *hsdS* gene was excised from pUS491(Δ50) on a small *Nde*I-*Bam*HI fragment. This fragment was ligated to the large *Nde*I-*Bam*HI fragment of pJ119 [28] to give pJS(Δ50). This

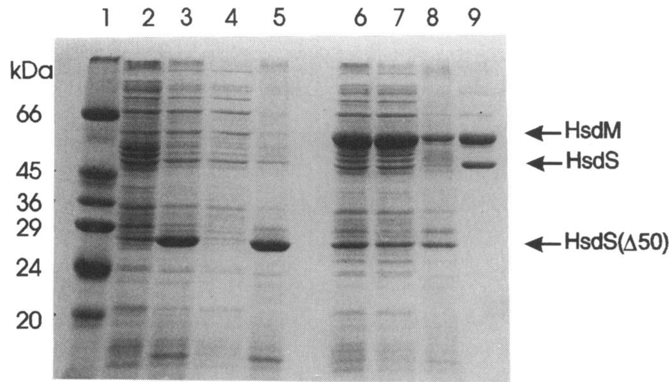

Figure 3. SDS-PAGE showing the HsdS(Δ50) subunit and Mtase(Δ50). **Lane 1,** protein size markers; **lane 2,** JM109(DE3) whole cell extract; **lane 3,** JM109(DE3)[pJS4(Δ50)] whole cell extract; **lane 4,** JM109(DE3)[pJS4(Δ50)] soluble fraction; **lane 5,** JM109(DE3)[pJS4(Δ50)] pellet fraction; **lane 6,** JM109(DE3)[pJS4(Δ50)M] whole cell extract; **lane 7,** JM109(DE3)[pJS4(Δ50)M] soluble fraction; **lane 8,** JM109(DE3)[pJS4(Δ50)M]; **lane 9,** Purified M·*Eco*R124I methylase.

plasmid was used to over-produce HsdS(Δ50) to confirm that the mutant was indeed a deletion of the *hsdS* gene. The HsdS(Δ50) subunit was found to be a protein of approximate molecular weight 27,250Da (Figure 3), this is in agreement with the predicted size of the deleted subunit based on the DNA sequence. The mutant *hsdS* gene was then subcloned from pJS(Δ50) using *Sph*I + *Bam*HI (blunt-ends produced with Klenow + dNTPs), and ligated with the large *Sph*I-*Bcl*I fragment of pJS4M carrying the *hsdM* gene (Figure 4). This ligated DNA was used to transform JM109(DE3) and the presence of the mutant Mtase (pJS4(Δ50)M) was confirmed by SDS-PAGE following IPTG induction and isolation of soluble extract (Figure 3). The presence of soluble HsdS(Δ50) from the Mtase(Δ50) clone suggested that the subunits were assembled into a complex in a manner analogous to that obtained with the wild-type HsdS subunit [3].

Bacteriophage lambda grown on the strain JM109-(DE3)[pJS4(Δ50)M] was screened for modification against restriction by either *Eco*R124I (C600[pCP1005]) or *Eco*R124/3I (C600[pUNG31]) and was found not to be modified against either restriction system (data not shown). This suggests that either the Mtase is non-functional, or that it has a new DNA specificity which is different from that of either *Eco*R124I or *Eco*R124/3I. This data is slightly different from that obtained from complementation between HsdS(Δ50) and the Hsd subunits of *Eco*R124/3I; where the new specificity appeared to overlap the *Eco*R124I specificity (i.e. protection against *Eco*R124I was afforded by the presence of HsdS(Δ50)).

Purification of Mtase(Δ50)

Our first attempt to purify the mutant Mtase(Δ50) involved following the same steps as for the wild-type Mtase [34]. However, it was found that HsdS(Δ50) did not bind to DEAE sephacel, and ran through the column in the dead volume. The result of this was the separation of the HsdS(Δ50) and HsdM subunits (c.f. Ref. [25]). Rather than spend a great deal of time attempting to purify the active Mtase using other techniques, or buffers, it was decided to measure Mtase activity in soluble cell extracts of *E.coli* producing both subunits.

Methylation using crude cell extract containing Mtase(Δ50)

The soluble fraction from JM109(DE3)[pJS4(Δ50)M] was mixed with pUC9 plasmid DNA (which contains no sites for the *Eco*R124I Mtase) and incorporation of tritium from ^{3}H-SAM was measured by scintillation counting following purification of the plasmid DNA by agarose gel electrophoresis. Figure 5 shows that the DNA was labelled by the crude extract, and that this methylation is destroyed by heat which suggests an active enzyme was present. No methylation was observed with JM109(DE3), or with JM109(DE3)[pUMI20]—which produces only HsdM [3]. This indicates that the Mtase clone is functional, and that the DNA sequence methylated is not that recognised by *Eco*R124I or *Eco*R124/3I.

pUC9 DNA from the above experiment was cleaved with a variety of restriction enzymes and the fragments separated on a 6% acrylamide gel. *Hae*III (Figure 6), *Hin*fI, *Nci*I and *Taq*I single and double digests of the labelled DNA showed only two bands were labelled. A computer search using the DNASTAR QUERYDNA program with the parameters of the search set for ABC(N_{5-8})DEFN (where ABCDEF are any specific nucleotide and N is any non-specific nucleotide), revealed two possible sequences with a pattern similar to those recognised by type I restriction enzymes, which were unique to these regions of pUC9−GAAnnnnnnCARG and GAAnnnnnnnTTC. Both sequences were also found to occur in pBR322, but G-AAnnnnnnCARG occurs four times (and would give four labelled bands following *Hae*III digestion), while GAAnnnnnnnTTC occurs only twice (and would give only two bands following *Hae*III digestion). Only two *Hae*III bands were labelled by the Mtase(Δ50) suggesting that the recognition sequence is GAAnnnnnnnTTC (Figure 6). To confirm this recognition sequence further labelling experiments with pACYC184, ϕX174 DNA and with a number of other restriction enzymes (*Taq*I, *Hin*fI & *Nci*I), were carried out. From this set of data the smallest fragment analysed with the DNASTAR QUERYDNA program was the 87-bp *Hae*III fragment of pACYC184 (Figure 6). From the analysis of the DNA sequences of all of the labelled DNA fragments, detected by these labelling experiments, the recognition sequence was identified as **GAA**(C/A/G)-(C/A/G)(C/A)C/G/T)N(T/C)N**TTC** (where the bases shown in bold are always present—Table 4). At this stage the central spacer is not proven to be N_7. To absolutely confirm that the spacer really is N_7 further labelling experiments were required. However, computer prediction showed that bacteriophage lambda has 19 sites for GAA(N_7)TTC, which would make full interpretation of any gel nearly impossible (and this was indeed found to be the case). However, M13mp9 was found to have only four sites, and digestion with *Hin*fI produced three radioactively-labelled bands of 700-bp, 960-bp and ~2050-bp respectively (data not shown). These labelled bands correspond to those predicted from a computer search, using GAA(N_7)TTC as the search sequence, and no other type I recognition sequence was found to occur within these bands.

Table 4 shows we can now predict a recognition sequence of **GAA**N(C/A/G)(C/A/T)(G/C/T)N(T/C/G)N**TTC**. To enable us to unequivocally assign the spacer as N_7 we still require data from another DNA molecule. We were fortunate in that we were able to easily identify two radioactively-labelled fragments from an *Eco*RV digest of lambda DNA on a 0.8% agarose gel. These bands were estimated at ~2700-bp and ~3600-bp respectively (data not shown). Table 4 shows that the only possible recognition

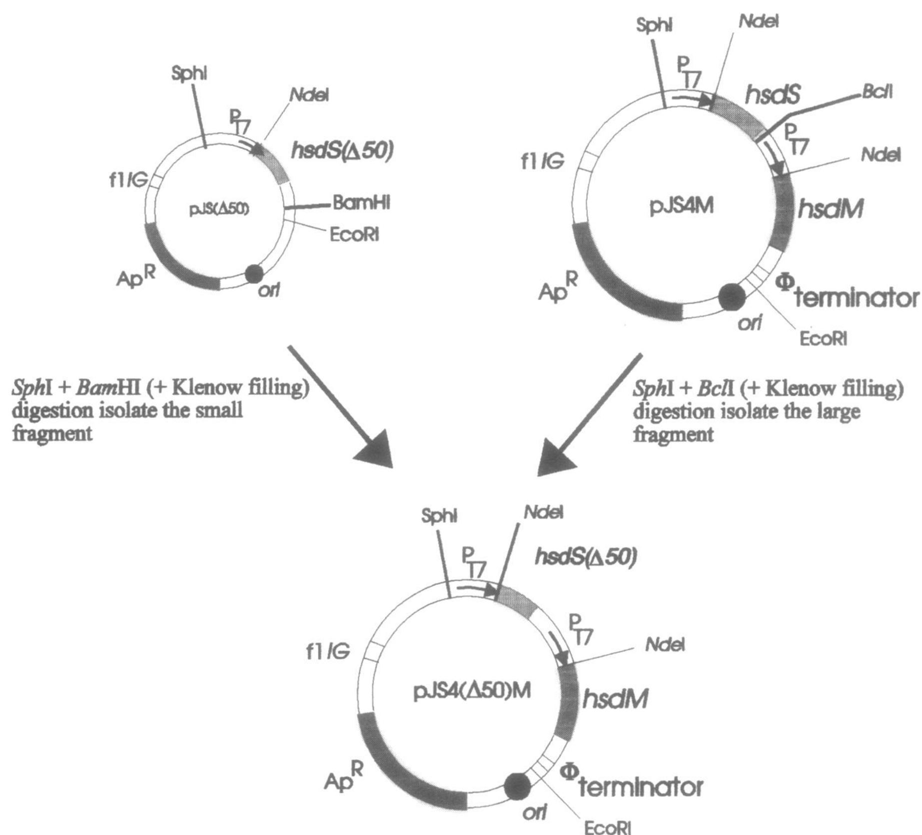

Figure 4. Cloning the mutant Mtase(Δ50).

sequences present on these bands identify the new DNA specificity as GAAnnnnnnnTTC. This sequence is an inverted repeat of the sequence recognised by the N-terminal recognition domain of HsdS(R124).

DISCUSSION

The HsdS subunit of type I restriction endonucleases is responsible for DNA recognition. It has been shown that HsdS(R124) has two domains responsible for DNA recognition, and that the N-terminal domain recognises GAA from the recognition sequence GAAnnnnnnRTCG [15]. It has also been proposed that the conserved regions of the *Eco*K family of HsdS proteins are important for protein-protein interactions [26]. Comparison of the DNA sequences of *Eco*R124I and *Eco*DXX1I (Bickle, personal communication; own unpublished observations) and heteroduplex electron microscopy [15] indicate that the *Eco*R124I HsdS subunit also has these conserved regions (Figure 1).

One can imagine that two classes of mutation in the *hsdS* gene could be isolated. Those that are impaired for DNA binding, and those that are unable to undergo normal subunit assembly. In order to investigate the potential of a complementation assay, to distinguish DNA-binding mutants from mutants affected in protein-protein interactions, a small library of random mutations was constructed. This assay was based on competition between mutant and wild-type HsdS subunits for assembly into an active endonuclease. All of the mutations produced were within the distal variable region of *hsdS*. Isolates were initially screened by C-track DNA sequencing, and then by means of the complementation assay. All restriction-deficient mutants were found to be associated with changes in the DNA sequence;

Table 4. DNA sequences methylated by Mtase(Δ50).

pUC9	1	C	A	C	**G**	**A**	**A**	C	C	C	C	C	C	G	**T**	**T**	**C**	A	G	C
	2	T	T	G	**G**	**A**	**A**	A	A	C	G	T	T	C	**T**	**T**	**C**	G	G	G
pACYC184	1	C	A	C	**G**	**A**	**A**	C	C	C	C	C	C	G	**T**	**T**	**C**	A	G	T
	2	G	A	C	**G**	**A**	**A**	A	A	A	C	A	T	A	**T**	**T**	**C**	T	C	A
φX174	1	T	A	T	**G**	**A**	**A**	G	G	A	T	G	T	T	**T**	**T**	**C**	C	G	T
pBR322	1	C	A	C	**G**	**A**	**A**	C	C	C	C	C	C	G	**T**	**T**	**C**	A	G	C
	2	A	T	G	**G**	**A**	**A**	A	A	C	G	T	T	C	**T**	**T**	**C**	G	G	G
summary from above sequences		N	A T	C G T	G	A	A	C A G	C A G	C A	C G T	N	C T	N	T	T	C	N	G C	N
M13mp9 + *Hin*fI fragment size - ~690-bp	1	G	T	T	**G**	**A**	**A**	T	G	T	G	G	T	A	**T**	**T**	**C**	C	T	A
M13mp9 + *Hin*fI fragment size - ~960-bp	2	A	A	T	**G**	**A**	**A**	T,	A	T	C	C	G	G	**T**	**T**	**C**	T	T	G
M13mp9 + *Hin*fI fragment size - ~ 2000-bp	3*	T	C	A	**G**	**A**	**A**	T	A	A	T	A	G	G	**T**	**T**	**C**	C	G	A
M13mp9 + *Hin*fI fragment size - ~ 2100-bp	4*	C	T	G	**G**	**A**	**A**	C	G	G	T	A	A	A	**T**	**T**	**C**	A	G	A
lambda + *Eco*RV fragment size ~3600-bp	18	A	A	T	**G**	**A**	**A**	C	G	C	A	A	T	A	**T**	**T**	**C**	A	C	A
lambda + *Eco*RV fragment size ~2700-bp	19	G	C	T	**G**	**A**	**A**	A	T	G	T	G	A	T	**T**	**T**	**C**	T	C	T
Final summary		N	T A C	N	**G**	**A**	**A**	N	N	N	N	N	N	N	**T**	**T**	**C**	N	G C T	N

* Bands 3 & 4 were not easily resolved in the H*inf*I digest of M13mp9

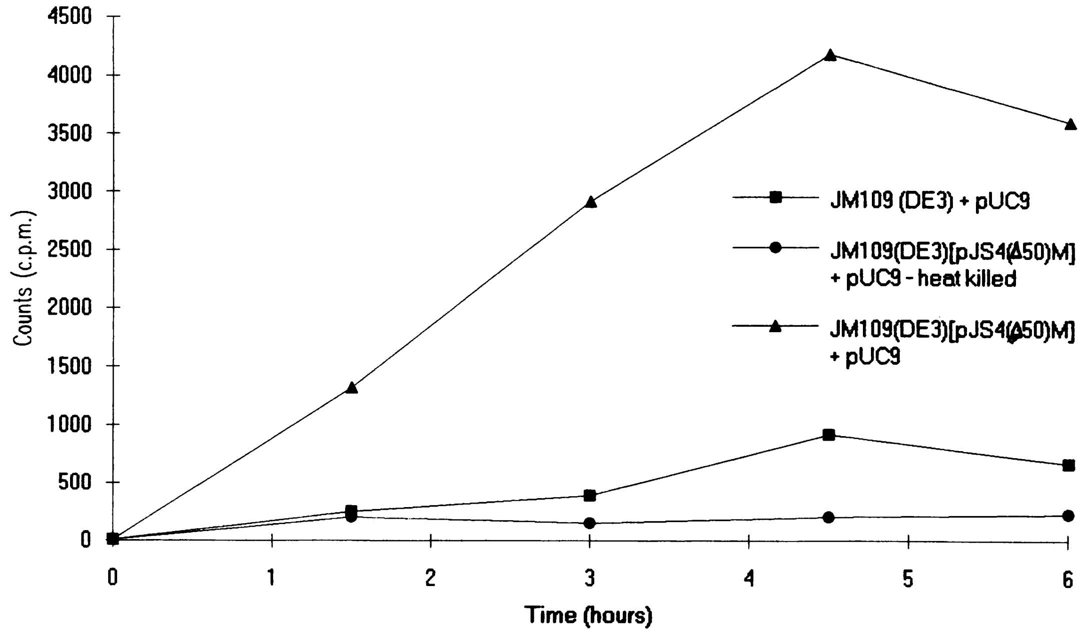

Figure 5. Incorporation of tritiated methyl groups into plasmid pUC9 DNA using soluble cell extracts of *E.coli* producing Mtase(Δ50).

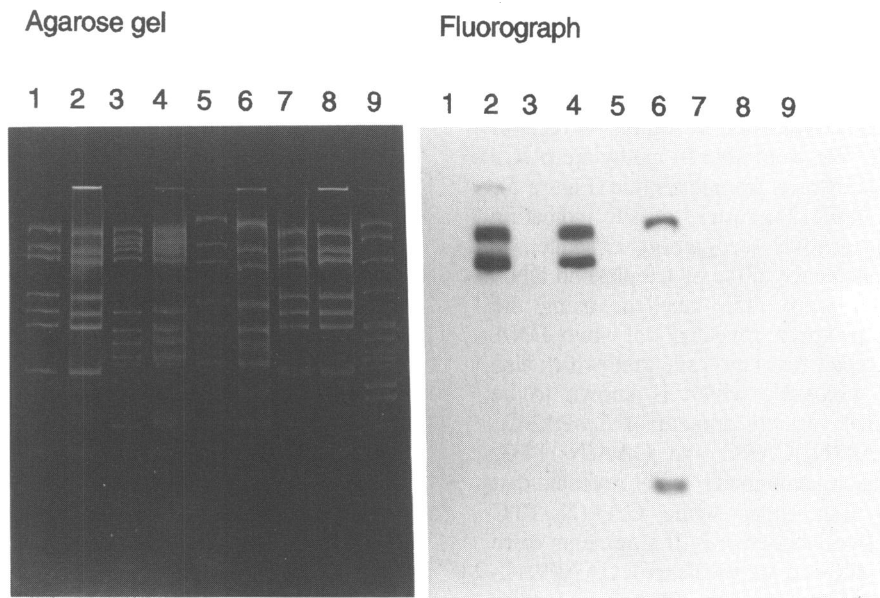

Figure 6. Agarose gel and associated fluorograph showing specific labelling of plasmid DNA in soluble extracts containing Mtase(Δ50). **Lane 1,** pUC9 + *Hae*III; **lane 2,** pUC9 + *Hae*III following 4.5 hours incubation with the soluble extract from JM109(DE3)[pJS4(Δ50)M]; **lane 3,** pBR322 + *Hae*III; **lane 4,** pBR322 + *Hae*III following 4.5 hours incubation with the soluble extract from JM109(DE3)[pJS4(Δ50)M]; **lane 5,** pACYC184 + *Hae*III; **lane 6,** pACYC184 + *Hae*III following 4.5 hours incubation with the soluble extract from JM109(DE3)[pJS4(Δ50)M]; **lane 7,** pUC9 + *Hae*III following 4.5 hours incubation with the soluble extract from JM109(DE3)[pUMI20]; **lane 8,** pUC9 + *Hae*III following 4.5 hours incubation with the soluble extract from JM109(DE3); **lane 9,** pBR322 + *Msp*I marker DNA.

although, a number of sequence changes gave no phenotypic change (as expected from their occurrence at the third base of the codon). This indicates that the complementation assay is sufficiently sensitive, and reliable, as a method for detecting restriction-deficient mutations. Two of the mutations did not show any 'competition' with wild-type HsdS(R124/3), and were totally restriction-deficient (15 and 53). These mutants satisfy the criteria for unfolded, or folding mutations, indicating that the competition assay can also discriminate between the two classes of mutant we predict to occur.

We also isolated two mutations within the distal variable region of HsdS(R124) that resulted in deletion of the entire distal conserved region (and part of the C-terminal variable region). Both mutations result in elevated levels of restriction and

apparently a new DNA specificity. One explanation for the elevated level of restriction produced by these mutants is that they are poor at methylation of DNA and this leads to increased cell lysis following lambda infection (this would be reflected in higher restriction levels). This has not been easy to establish because of our inability to purify the active Mtase to date. The DNA remaining in *hsdS(Δ50)* codes for the variable domain responsible for recognition of 5′-GAA-3′, and the central conserved spacer region. The *hsdS(Δ50)* gene was cloned into the expression plasmid pJS4M, replacing the wild-type *hsdS* gene. JM109(DE3) containing the resultant plasmid (pJS4(Δ50)M) was found to overproduce the HsdM and HsdS(Δ50) subunits, and both were present in the soluble fraction. Since our attempts to purify the active Mtase resulted in separation of HsdS(Δ50) and HsdM on a DEAE sephacel column Mtase(Δ50) may exist as only a weak complex with perhaps only a few percent of the subunits assembled into an active enzyme.

Bacteriophage lambda was grown on JM109(DE3)-[pJS4(Δ50)M] and then screened for modification against the *Eco*R124I and *Eco*R124/3I restriction systems. The bacteriophage was found to be restricted by both systems suggesting that either there was no modification by Mtase(Δ50), or that any modification that takes place was not at the *Eco*R124I or *Eco*R124/3I recognition sequence. However, the initial complementation assay indicated that HsdS(Δ50) modification in a genetic background containing wild-type *Eco*R124/3I afforded protection against *Eco*R124I restriction. That result appears to reflect complex interactions between the Hsd subunits in the presence of wild-type HsdS(R124/3). The Mtase(Δ50) produced by pJS4(Δ50)M does not protect against *Eco*R124I restriction, however, it is an active Mtase in soluble extracts of JM109(DE3)[pJS4(Δ50)M]. We were able to methylate pUC9 DNA in the presence of crude extracts from this strain (Figure 5). pUC9 does not contain an *Eco*R124I restriction site indicating that this methylation was at a novel recognition sequence.

Computer analysis of the sequence of the pUC9 plasmid DNA restriction fragments which were radio-labelled, using the DNASTAR QUERYDNA program revealed only two DNA sequences that resembled a type I restriction site (and which also contained the sequence 5′-GAA-3′, which is known to be recognised by the N-terminal variable domain of *Eco*R124I). These sequences were GAA(N_6)CARG and GAA(N_7)TTC. Analysis of pBR322 DNA in an analogous manner revealed that GAA(N_6)CARG occurred four times while GAA(N_7)TTC occurred only twice. Only two labelled *Hae*III fragments were observed (Figure 6). This allowed us to discard GAA(N_6)C-ARG as a possible recognition sequence. However, further analysis of relatively short DNA molecules did not absolutely show that the internal spacer region (N_7) was totally non-specific. A detailed analysis of the methylation of lambda DNA was hampered by the large number of labelled fragments observed. To overcome this problem data from M13mp9 was used to resolve some of the uncertainty surrounding the nature of the spacer sequence recognised. With this extra information from M13mp9 we were able to identify two radio-labelled bands from an *Eco*RV digest of lambda and use these bands to absolutely show that the spacer is indeed non-specific. Each band contained only one site that fitted the pattern GAA(N_7)TTC (Table 4).

The nature of the recognition sequence is an inverted repeat of the DNA sequence recognised by the N-terminal DNA binding domain of HsdS. Therefore, GAA is present on the top strand and on the bottom strand and the specific regions of the recognition sequence are separated by seven non-specific nucleotides—the preferred separation for type IC spacer sequences [15, 20, 24]. This suggests that recognition is accomplished by means of two copies of the HsdS(Δ50) subunit, in inverted orientation, with one subunit reading one strand of the DNA and the other subunit reading the other strand. This also suggests that there is a contact between the two HsdS(Δ50) subunits, perhaps by means of the spacer region; although it seems equally likely that HsdM contributes to the stability of this structure. It is possible that this inverted dimerisation reflects how the HsdS-HsdM subunits interact in the wild-type enzyme. This would imply a surface on HsdM which interacts with itself and that HsdM-HsdM interactions orientate the HsdS(Δ50) subunits such that they will bind DNA, one (half) subunit reading one strand of the DNA molecule. We are currently investigating the possibility of *in vitro* assembly and stabilisation of the complex by protein-protein cross linking.

ACKNOWLEDGEMENTS

Dr Abadjieva and Dr Patel were supported by research grants from the Wellcome Trust. Dr Zinkevich was supported by a Wellcome Trust European Exchange Fellowship.

REFERENCES

1. Bickle, T.A., (1987) In: *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology* (F.C. Neidhardt, *et al.*, Editors) American Society for Microbiology, Washington DC. p. 692–696.
2. Wilson, G.G. and Murray, N.E., (1991). Ann. Rev. Genet. 25, 585–627.
3. Patel, J., Taylor, I., Dutta, C.F., Kneale, G.G. and Firman, K., (1992). Gene 112, 21–27.
4. Hubacek, J. and Glover, S.W., (1970). J. Mol. Biol. 50, 111–127.
5. Glover, S.W. and Colson, C., (1969). Genet. Res. (Cambs) 13, 227–240.
6. Boyer, H.W. and Roulland-Dussoix, D., (1969). J. Mol. Biol. 41, 459–472.
7. Haberman, A., Heywood, J. and Meselson, M., (1972). Proc. Natl. Acad. Sci. USA 69, 3138–3141.
8. Meselson, M. and Yuan, R., (1968). Nature 217, 1110–1114.
9. Meselson, M., Yuan, R. and Heywood, J., (1972). Ann. Rev. Biochem. 41, 447–462.
10. Yuan, R. and Meselson, M., (1970). Proc. Natl. Acad. Sci. USA 65, 357–362.
11. Yuan, R., Heywood, J. and Meselson, M., (1972). Nature New Biol. 240, 42–43.
12. Price, C., Bickle, T.A., Firman, K. and Glover, S.W., (1989). J. Mol. Biol. 205, 115–125.
13. Cowan, G.M., Gann, A.A.F. and Murray, N.E., (1989). Cell 56, 103–109.
14. Fuller-Pace, F.V. and Murray, N.E., (1986). Proc. Natl. Acad. Sci. USA 83, 9368–9372.
15. Gubler, M., Braguglia, D., Meyer, J., Piekarowicz, A. and Bickle, T.A., (1992). EMBO J. 11, 233–240.
16. Gann, A.F.F., Campbell, A.J.B., Collins, J.F., Coulson, A.F.W. and Murray, N.E., (1987). Mol. Microbiol. 1, 13–22.
17. Nagaraja, V., Shepherd, J.C.W. and Bickle, T.A., (1985). Nature 316, 371–372.
18. Nagaraja, V., Shepherd, J.C.W., Pripfl, T. and Bickle, T.A., (1985). J. Mol. Biol. 182, 579–587.
19. Kannan, P., Cowan, G.M., Daniel, A.S., Gann, A.A.F. and Murray, N.E., (1989). J. Mol. Biol. 209, 335–344.
20. Price, C., Shepherd, J.C.W. and Bickle, T.A., (1987). EMBO J. 6, 1493–1498.
21. Fuller-Pace, F.V., Cowan, G.M. and Murray, N.E., (1985). J. Mol. Biol. 186, 65–75.
22. Murray, N.E., Gough, J.A., Suri, B. and Bickle, T.A., (1982). EMBO J. 1, 535–539.
23. Price, C., Pripfl, T. and Bickle, T.A., (1987). Eur. J. Biochem. 167, 111–115.
24. Skrzypek, E. and Piekarowicz, A., (1989). Plasmid 21, 195–204.
25. Gubler, M. and Bickle, T.A., (1991). EMBO J. 10, 951–957.
26. Gough, J.A. and Murray, N.E., (1983). J. Mol. Biol. 166, 1–19.

27. Lehtovaara, P.M., Koivula, A.K., Bamford, J. and Knowles, J.K.C., (1988). Protein Eng. 2, 63–68.
28. Patel, J., (1992). PhD thesis, University of Portsmouth.
29. Chamberlain, J.P., (1979). Anal. Biochem. 98, 132–135.
30. Rosenberg, A.H., Lade, B.N., Chui, D.-S., Lin, S.-W. and Dunn, J.J., (1987). Gene 56, 125–135.
31. Vieira, J. and Messing, J., (1987) In: *Methods in Enzymology* (R. Wu and L. Grossman, Editor) Academic Press, New York. p. 416–431.
32. Zinkevich, V.E., Weiserova, M., Kryukov, V.M. and Hubacek, J., (1990). Gene 90, 125–128.
33. Zinkevich, V., Heslop, P., Glover, S.W., Weiserova, M., Hubacek, J. and Firman, K., (1992). J. Mol. Biol. 227, 597–601.
34. Taylor, I., Patel, J., Firman, K. and Kneale, G.G., (1992). Nucleic Acids Res. 20, 179–186.
35. Appleyard, R.K., (1954). Genetics 39, 440–452.
36. Raleigh, E.A., Lech, K. and Brent, R., (1989) In: *Current protocols in molecular biology* (F.M. Ausubel, Editor) Publishing Associates and Wiley Interscience, New York.
37. Jacob, F. and Wollman, E.L., (1954). Ann. Inst. Pasteur 87, 653–673.
38. Dente, L., Cesareni, G. and Cortese, R., (1983). Nucleic Acids Res. 11, 1645–1655.
39. Messing, J. In: *Third Cleveland Symposium on Macromolecules*. (1981). Cleveland, Ohio: Elsevier Scientific Publishing Company.
40. Sanger, F., Nickler, S. and Coulson, A.R., (1977). Proc. Natl. Acad. Sci. USA 74, 5463–5467.
41. Firman, K., Price, C. and Glover, S.W., (1985). Plasmid 14, 224–234.
42. Vieira, J. and Messing, J., (1982). Gene 19, 259–268.
43. Bolivar, F., Rodriguez, R.L., Greene, P.J., Betlach, M.F. and Heyneker, H., (1977). Gene 2, 95–113.
44. Chang, A.C.Y. and Cohen, S.N., (1978). J. Bacteriol. 134, 1141–1156.