Nucleotide sequence analysis and overexpression of the gene encoding a type III chloramphenicol acetyltransferase

Iain A. MURRAY,* Alastair R. HAWKINS,† John W. KEYTE and William V. SHAW

Department of Biochemistry, University of Leicester, Leicester LE1 7RH, U.K.

The gene *cat*_{III}, encoding a type III enterobacterial chloramphenicol acetyltransferase, was cloned from the transmissible plasmid R387 into pBR322 and bacteriophage M13 mp8. Nucleotide sequence analysis of 1160 bp of DNA identified an open reading frame encoding a protein of 213 amino acid residues and a calculated molecular mass of 24965 Da. The predicted *N*-terminal sequence is identical with that determined by Edman degradation of chloramphenicol acetyltransferase purified from *Escherichia coli* harbouring R387. Sequences equivalent to the consensus motifs for initiation and rho-factor-independent termination of transcription in *E. coli* occur 5' and 3' to the *cat*_{III} open reading frame. In contrast with the *cat*_I gene, present on transposon Tn9 and many enterobacterial plasmids, expression of *cat*_{III} is not subject to cyclic AMP-mediated catabolite repression *in vivo* and there is no sequence in the 5' non-coding DNA that resembles that deduced as the consensus for the binding of cyclic AMP receptor protein. Unique restriction-endonuclease cleavage sites were introduced adjacent to the *cat*_{III} reading frame by using oligonucleotide-directed mutagenesis to facilitate insertion into *E. coli* expression vectors. Fully active chloramphenicol acetyltransferase represents 30–50% of the soluble protein component of cell-free extracts of *E. coli* containing the appropriate plasmids.

INTRODUCTION

Enzymic acetylation catalysed by chloramphenicol acetyltransferase (CAT, EC 2.3.1.28) is the commonest mechanism of bacterial resistance to the antibiotic chloramphenicol, an inhibitor of prokaryotic peptidyl-transferase activity (Shaw, 1967). Transfer of the acetyl group of acetyl-CoA to the primary (C-3) hydroxy group of the antibiotic yields 3-acetylchloramphenicol, which fails to bind to bacterial ribosomes (Scheme 1). CAT variants have been isolated from numerous bacterial genera, and in all cases studied the enzyme is a trimer of identical 25 kDa subunits.

Scheme 1. Reactions catalysed by CAT

Isomerization of 3-acetylchloramphenicol and 1-acetylchloramphenicol is non-enzymic and pH-dependent. The mono- and diacetylated adducts of chloramphenicol fail to bind to prokaryotic ribosomes.

Mechanistic studies have in recent times concentrated on the efficient enterobacterial type III variant associated with the transmissible plasmid R387, originally isolated from *Shigella flexneri*. In addition to its high catalytic turnover, this variant yielded crystals suitable for X-ray-diffraction analysis (Leslie *et al.*, 1986), and the structure has been refined to 0.175 nm resolution in the presence and in the absence of bound ligand (Leslie *et al.*, 1988).

Steady-state kinetic analyses of the forward and reverse

Fig. 1. Proposed mechanism of CAT

The imidazole group (N-3) of the catalytic histidine residue acts as a general base, abstracting a proton from the 3-hydroxy group of chloramphenicol (CM), thereby promoting nucleophilic attack at the carbonyl group of the acetyl thioester.

Abbreviations used: CAT, chloramphenicol acetyltransferase; *cat*, the gene encoding CAT (subscripts denote individual variants); Cm⁺, chloramphenicol-resistance phenotype; Sm⁺, streptomycin-resistance phenotype.

* To whom correspondence should be addressed.

† Present address: Department of Genetics, University of Newcastle upon Tyne, Newcastle upon Tyne NE1 7RU, U.K.

These sequence data have been submitted to the EMBL/GenBank Libraries under the accession number Y00723.

reactions suggest a ternary-complex (sequential) catalytic mechanism with independent binding of substrates (Kleanthous & Shaw, 1984). Production of the ternary complex proceeds via a rapid-equilibrium random-order mechanism with a preference for acetyl-CoA as the leading substrate. Inactivation of type III CAT with methyl 4-nitrobenzenesulphonate (Nakagawa & Bender, 1970) and the affinity reagent 3-bromoacetylchloramphenicol both identify an essential histidine residue, which is believed to function as a general base during catalysis (Corney, 1983; Kleanthous *et al.*, 1985). The critical imidazole group is thought to abstract a proton from the primary hydroxy group of chloramphenicol, thus promoting nucleophilic attack at the acetyl-CoA thioester bond (Fig. 1).

The present paper reports the cloning, nucleotide sequence analysis and overexpression in *Escherichia coli*, of the gene *cat_{III}*, encoding type III CAT, to enable further analysis of the catalytic mechanism and ligand-binding properties of CAT via site-directed mutagenesis.

MATERIALS AND METHODS

Materials

Restriction endonucleases, T4 DNA ligase, DNA polymerase I (Klenow fragment and holoenzyme) and T4 polynucleotide kinase were purchased from Bethesda Research Laboratories and used under the manufacturer's recommended conditions. Calf intestinal phosphatase and dideoxynucleotide triphosphates were obtained from Boehringer; radiolabelled nucleotides and deoxynucleotides were from Amersham International. Antibiotics and deoxynucleotide triphosphates were purchased from Sigma Chemical Co. All other reagents were of analytical quality.

Plasmid and M13 RF DNAs were isolated by the method of Birnboim & Doly (1979) and purified by centrifugation in CsCl in the presence of ethidium bromide. *E. coli* cells were transformed by the method of

Kushner (1978). DNA fragments were labelled with [α - 32 P]dATP by nick translation (Rigby *et al.*, 1977).

Bacterial strains, plasmids and bacteriophage

E. coli K12 strains J53/R387 (Shaw *et al.*, 1972) and JM101 (Messing, 1979), plasmids pBR322 (Bolivar *et al.*, 1977) and pUC18 (Yanisch-Perron *et al.*, 1985) and bacteriophages M13 mp8 (Messing & Vieira, 1982) and M13 mp18 (Yanisch-Perron *et al.*, 1985) have been described previously. *E. coli* K12 SK3430 (*thi aroD6 leu hsdR*) was a gift from S. R. Kushner to A. R. H. Strain J53/R387 was tested for carbon-source-mediated catabolite repression of CAT production *in vivo* by the method of Harwood & Smith (1971).

Isolation, purification and assay of CAT

Cell-free extracts of *E. coli* were prepared by sonication or passage through a French pressure cell. CAT was purified by affinity chromatography on chloramphenicol immobilized on Sepharose 4B as previously described (Packman & Shaw, 1981). Protein concentration was determined by the method of Lowry *et al.* (1951) with bovine serum albumin as standard, and purity was determined by polyacrylamide-gel electrophoresis in the presence of SDS (Laemmli, 1970). The spectrophotometric assay method of Kleanthous & Shaw (1984) was used to monitor CAT activity of crude extracts and purified enzyme; 1 unit of activity is defined as 1 μ mol of chloramphenicol acetylated/min.

Oligonucleotide synthesis

Oligonucleotides were synthesized manually by using phosphotriester chemistry with Whatman 3MM paper (Brenner & Shaw, 1985) or controlled-pore glass (Sproat & Bannwarth, 1983) supports and purified by electrophoresis through 50% urea/20% polyacrylamide gels.

Cloning

Restriction-endonuclease digestion products were analysed by horizontal agarose-gel electrophoresis

Fig. 2. Strategy for determination of the nucleotide sequence of *cat_{III}*.

Arrows indicate direction and extent of sequence data derived from individual restriction fragments. Broken arrows indicate extended sequence data determined by the use of specific oligonucleotide primers. Restriction sites are: A, *AluI*; H, *HincII*; Hd, *HindIII*; S, *Sau3A*.

(Maniatis *et al.*, 1982). All ligation reactions included a 2–5-fold molar excess of target DNA fragments over cut and dephosphorylated vector DNA. Recombinant M13 plaques containing *cat*_{III} sequences were identified by DNA–DNA hybridization (Benton & Davis, 1977).

Sequencing

The Sanger chain-terminator protocol (Sanger *et al.*, 1977) with buffer gradient gels and [α -³⁵S]thio]dATP (Biggin *et al.*, 1983) was used throughout. Autoradiography was carried out with Fuji RX film.

Oligonucleotide site-directed mutagenesis

The dual-primer protocol (Zoller & Smith, 1984) was employed with the oligonucleotides 5'-GCGATGGAT-CCAAGT-3' (two mismatches) and 5'-TAGTTCA-TAATCGATCCCTTCTG-3' (three mismatches). The oligonucleotide 5'-AGGCTATCAGGTCATT-3' was synthesized as the secondary primer and is complementary to nucleotides 6915–6930 of the M13 mp18 (+)-strand (6894–6909 of M13 mp8).

RESULTS AND DISCUSSION

Absence of carbon-source-mediated catabolite repression of *cat* expression *in vivo*

E. coli J53 harbouring plasmid R387 (Cm⁺, Sm⁺) was cultured in M9 minimal medium supplemented with L-proline and L-methionine (each 0.02%, w/v), streptomycin sulphate (25 µg/ml) and glucose or glycerol (0.2%, w/v) in the presence and in the absence of either 5 mM-5'-AMP or 5 mM-3',5'-cyclic AMP. In all cases the specific activity of CAT measured in crude cell-free extracts was the same (1.7 ± 0.2 units/mg).

Cloning of *cat*_{III}

Plasmid R387 DNA restricted with *Pst*I was ligated to equivalently cut dephosphorylated pBR322 and the products were used to transform *E. coli* SK3430. Chloramphenicol-resistant transformants were analysed and found to contain recombinant plasmids with a 5.5 kb *Pst*I-digestion fragment derived from R387. The *cat* determinant was localized further via the introduction of *Sau*3A partial-digestion products of one such recombinant plasmid (designated pMH1) into *Bam*HI-cleaved pBR322 and screening SK3430 transformants for chloramphenicol-resistance. Plasmid pJK111, containing approx. 1.5 kb of DNA originally present in R387 including the *cat* gene, served as the source of DNA for subcloning in M13 mp8. The specific activity of CAT purified to apparent homogeneity from *E. coli* harbouring plasmid R387, pMH1 or pJK111 was identical at 800 ± 50 units/mg.

Nucleotide sequence analysis

A single *Hind*III, two *Hinc*II and multiple *Alu*I and *Sau*3A restriction sites were identified within the R387-derived DNA of pJK111 (results not shown). These restriction endonucleases were therefore used to generate DNA fragments for subcloning in M13 mp8 before nucleotide sequence determination. Potential M13 recombinants containing *cat*_{III} DNA were identified, in the first instance, as bacteriophage plaques that gave positive signals when hybridized against a 'nick-translated' pJK111 probe but failed to bind ³²P-labelled pBR322

DNA. Nucleotide sequence data obtained from individual recombinants by use of the M13 'universal' primer were extended, where appropriate, by using synthetic oligonucleotide primers designed according to the derived sequence (Fig. 2).

The nucleotide sequence of 1160 bp of pJK111 DNA was determined from both DNA strands without ambiguities and predicts a single *cat* open reading frame encoding 213 amino acid residues with a calculated molecular mass of 24965 Da (Fig. 3). The deduced N-terminal sequence, Met-Asn-Tyr-Thr-Lys-Phe-Asp, corresponds to that determined by Edman degradation of type III CAT purified from *E. coli* J53 containing R387 (Fitton *et al.*, 1978), and the overall deduced amino acid sequence is in agreement with the partial protein sequence determined by Packman (1978).

The first base of the hexanucleotide AGAAGG, an approximation of the consensus sequence for prokaryotic ribosome binding (AGGAGG; Shine & Dalgarno, 1975), occurs 14 nucleotide residues before the ATG initiation codon. Sequences equivalent to the -10 (TATAAT) and -35 (TTGACA) *E. coli* transcriptional initiation consensus motifs are represented 5' to the *cat* reading frame by TAAAAT and TTGCCT respectively (Pribnow, 1975; Siebenlist, 1979). The 3' non-coding sequence includes an inverted repeat region followed by several thymidine residues characteristic of the proposed secondary structural unit required for rho-factor-independent termination of transcription in *E. coli* (Rosenberg & Court, 1975). The 5' non-coding DNA of *cat*_{III} does not contain sequences equivalent to the consensus for the binding of cyclic AMP receptor protein observed within the *cat*_I determinant present on Tn9, pACYC184 and a number of large enterobacterial plasmids (LeGrice *et al.*, 1982).

The *cat*_{III} open reading frame shows 46% amino acid identity with the type I enzyme encoded by Tn9 (Alton & Vapnek, 1979; Fig. 4). Maximum homology occurs in the region immediately adjacent to the active-site histidine residue (residue 189 in type III CAT). Type III CAT lacks the first five amino acid residues present at the N-terminus of the type I and *Proteus mirabilis* enzymes (Charles *et al.*, 1985), a feature previously believed to be typical of CAT variants isolated from Gram-positive bacteria (Horinouchi & Weisblum, 1982; Harwood *et al.*, 1983; Shaw *et al.*, 1985).

Site-directed mutagenesis and overexpression of *cat*_{III}

Analyses of the effects of site-directed mutations within the *cat*_{III} structural gene (Murray *et al.*, 1986) have been hampered by the low level of gene expression in *E. coli* from plasmids derived from pBR322 and bacteriophage M13 (results not shown). In addition, although *cat*_{III} has a unique *Hind*III restriction site immediately 3' to the proposed transcriptional termination point, there are no unique restriction sites in the 5' non-coding DNA that permit insertion into, and excision from, the multiple cloning site of M13 mp vectors. To overcome these limitations we have introduced novel restriction sites into the 5' non-coding region of *cat*_{III} by using oligonucleotide-directed mutagenesis.

Plasmid pJK111 has a unique *Bgl*II site approx. 100 bp 5' to the region for which sequence data were determined. The *cat* gene was excised by cleavage with *Bgl*II and *Hind*III and ligated to *Bam*HI-*Hind*III-cut and dephosphorylated RF DNA of M13 mp8 yielding the

1 GATCGTATCTGAGAA GTCGTCGCTATGAGC CTCATGTCCAGTCAC GGAAGGATGAATCAG AAGCCATCAAAAACA CGGATTTTAAAGGCC

3'-GTCAA CCTAGGTAGCG-5'

91 ACCGCTGGGTTGTAG AGAGAACACACAGTT GGATGAATCGCTACC GTCGTGTCCTGALTC GTTGGGAGAAAAAGG TCGAAAATTACGAGG

3'-GTCTTCCTAGCTA

101 CAATGCTGCATCTTG CCIGTGCCATCAITG TCTGGAATAAAATCC TTTTGGGATAGGTTT TTAGTAAATAAAAT TCAGAAGGGGATATGT

-35 -10 S.D

A TAC TTG AT-5' 10 20

Met Asn Tyr Thr Lys Phe Asp Val Lys Asn Trp Val Arg Arg Glu His Phe Glu Phe Tyr Arg His Arg

271 T ATG AAC TAT ACA AAA TTT GAT GTA AAA AAT TGG GTT CGC CGT GAG CAT TTT GAG TTT TAT CGG CAT CGT

30 40

Leu Pro Cys Gly Phe Ser Leu Thr Ser Lys Ile Asp Ile Thr Thr Leu Lys Lys Ser Leu Asp Asp Ser Ala

341 TTA CCA TGT GGT TTT AGC TTA ACA AGC AAA ATT GAT ATC ACG ACG TTA AAA AAG TCA TTG GAT GAT TCA GCG

50 60 70

Tyr Lys Phe Tyr Pro Val Met Ile Tyr Leu Ile Ala Gln Ala Val Asn Gln Phe Asp Glu Leu Arg Met Ala

413 TAT AAG TTT TAT CCG GTA ATG ATC TAT CTG ATT GCT CAG GCC GTG AAT CAA TTT GAT GAG TTG AGA ATG GCG

80 90

Ile Lys Asp Asp Glu Leu Ile Val Trp Asp Ser Val Asp Pro Gln Phe Thr Val Phe His Gln Glu Thr Glu

485 ATA AAA GAT GAT GAA TTG ATC GTA TGG GAT TCA GTC GAC CCA CAA TTC ACC GTA TTC CAT CAA GAA ACA GAG

100 110

Thr Phe Ser Ala Leu Ser Cys Pro Tyr Ser Ser Asp Ile Asp Gln Phe Met Val Asn Tyr Leu Ser Val Met

557 ACA TTT TCA GCA CTG AGT TGC CCA TAC TCA TCC GAT ATT GAT CAA TTT ATG GTG AAT TAT TTA TCG GTA ATG

120 130 140

Glu Arg Tyr Lys Ser Asp Thr Lys Leu Phe Pro Gln Gly Val Thr Pro Glu Asn His Leu Asn Ile Ser Ala

629 GAA CGT TAT AAA AGT GAT ACC AAG TTA TTT CCT CAA GGG GTA ACA CCA GAA AAT CAT TTA AAT ATT TCA GCA

150 160

Leu Pro Trp Val Asn Phe Asp Ser Phe Asn Leu Asn Val Ala Asn Phe Thr Asp Tyr Phe Ala Pro Ile Ile

701 TTA CCT TGG GTT AAT TTT GAT AGC TTT AAT TTA AAT GTT GCT AAT TTT ACC GAT TAT TTT GCA CCC ATT ATA

170 180 190

Thr Met Ala Lys Tyr Gln Gln Glu Gly Asp Arg Leu Leu Leu Pro Leu Ser Val Gln Val His His Ala Val

773 ACA ATG GCA AAA TAT CAG CAA GAA GGG GAT AGA CTG TTA TTG CCG CTC TCA GTA CAG GTT CAT CAT GCA GTT

200 210

Cys Asp Gly Phe His Val Ala Arg Phe Ile Asn Arg Leu Gln Glu Leu Cys Asn Ser Lys Leu Lys

845 TGT GAT GGC TTC CAT GTT GCA CGC TTT ATT AAT CGG CTA CAA GAG TTG TGT AAC AGT AAA TTA AAA TAAGCTC

918 AGGTTAAATTAAGG GTTCATATCTAGIG AAGCCCTGTTTTATT ATGTTTTGAGTTAAC TTATCTTCAATGGCG GTCATAGAACCGACT

3' Inverted Repeat

1008 GTTACCCTCAAATA AATGTGTGAGCCGAA TTATTATATCATCCT TCCGCAAAAGGCTAA AAACCATAAACAATA AAATGTTTTTTGTGG

1098 GCTTTTCTTCGCACT TTTTATCTGTTTCTA GGCATTTTATGATAA TTTTATAAGGTAAG CTT

Fig. 3. Nucleotide sequence and predicted open reading frame of *cat_{III}*

Proposed transcriptional initiation/termination and putative ribosome-binding (S.D, Shine-Dalgarno) sequences are underlined. Oligonucleotides used in subsequent site-directed mutagenesis experiments are indicated at appropriate positions above the sequence.

construct mp8:IM2 (Fig. 5). The oligonucleotide 5'-GCGATGGATCCAAGT-3', complementary to nucleotides 116-131 of *cat_{III}*, was used to introduce a two-base substitution, thereby creating a new *Bam*HI recognition sequence (Fig. 6a). The 1034 bp *Bam*HI-*Hind*III-digestion fragment, containing the structural gene and putative promoter/terminator sequences, was

introduced into M13 mp18 RF and plasmid pUC18 to produce mp18:IM3 and pUC18:IM3 (Fig. 5). Constitutive expression of *cat_{III}* in *E. coli* JM101 harbouring pUC18:IM3 is highly efficient, fully active CAT representing 25-40% of the soluble protein in crude cell-free extracts (Fig. 7).

A second cycle of mutagenesis, using the oligo-

TYPE III: **MNYTKFDVKNWVRREHFEFYRHLPCGFSLSKIDITTLKKSLLDSAYKFYPVMI**
 * * * * + * * * * + * + * + * * * * + * * * * *
TYPE I: MEKKITGYTTVDISQWHRKEHFEAFQSVACQCTYNQTVQLDITAFKTVKKNKHKFPYAFI

TYPE III: **YLIAQAVNQFDELRLMAIKDDELIVWDSVDPQFTVFHQETETFSALSCPYSDDIQFMVNY**
 + * * + * * * * * * * * * + * * * * * + * * * * * * * * * + *
TYPE I: HILARLMNAHPEFRMAMKDGELVIWDSVHPCYTVFHEQTETFSLLWSEYHDDFRQFLHIY

TYPE III: **LSVMERYKSDTKLFPQGVTPENHLNISALPHVNFDSFNLNVANFTDYFAPIIITMAKYQQE**
 + * + * * * + * + * * * * * * * * * + * * * + * * *
TYPE I: SQDVACYGDNLAYFPKGF I ENMFFVSANPWVSFTSFDLNVANMDNFFAPVFTMGKYTQ

TYPE III: **GDRLLPLSVQVHHAVCDGFGHVARFINRLQELCNSKLK**
 * * * * * * + * * * * * * * * * * * * + * * * *
TYPE I: GDKVLMPLAIQVHHAVCDGFGHVGRMLNELQQYCDWQGGGA

Fig. 4. Sequence homology between type III and type I variants of CAT

★, Conserved amino acid residues; +, conservative amino acid residue substitutions.

Fig. 5. Construction of plasmids permitting overexpression of *cat*_{III} in *E. coli*

SDM, oligonucleotide site-directed mutagenesis.

Fig. 6. Autoradiograph of dideoxy sequence ladders demonstrating mutations generating new restriction sites

(a) Novel *Bam*HI site (GGATCC). (b) Novel *Cla*I site (ATCGAT). The track order, left to right, is ACGT in each case.

nucleotide 5'-TAGTTCATAATCGATCCCTTCTG-3' and single-stranded mp18:IM3 DNA, was employed to introduce a *Cla*I restriction site adjacent to the ATG initiation codon of *cat*_{III} (Fig. 6b). This mutation facilitates the in-frame insertion of the *cat*_{III} structural gene into the temperature-inducible expression vector pHD, a dual-origin plasmid conferring ampicillin-resistance that contains the promoter, operator and ribosome-binding site of the *trp* attenuator peptide followed by unique *Cla*I and *Eco*RI restriction sites (Thomas, 1987). When present in *E. coli* grown at 30 °C plasmid pHD replicates via a pSC101 origin (three or four copies per chromosome) and transcription from the *trp* promoter is inhibited by the chromosomal *trp* repressor. However, at 37 °C replication is via a ColE1 origin under the control of the P_R promoter of bacteriophage λ, initiated by the inactivation of the temperature-sensitive λ cI857 repressor. The result is a rapid amplification of the plasmid up to several hundred copies per chromosome. At high plasmid copy number the chromosomal *trp* repressor is titrated out, resulting in very efficient transcription from the *trp* promoter and through any appropriately positioned exogenous coding sequence. Expression of *cat*_{III} from the pHD-derivative pIM5 (Fig. 5) in *E. coli* JM101 produces cell-free extracts in which active CAT constitutes 30–50 % of the soluble protein component (Fig. 7). Cultures of *E. coli* harbouring pIM5 can be grown to mid-exponential phase

Fig. 7. Overexpression of *cat*_{III} in *E. coli*

SDS/polyacrylamide-gel electrophoresis of crude cell-free extracts of *E. coli* JM101 harbouring various *cat* plasmids (see the text). Lane A, molecular-mass standards: bovine serum albumin (68 kDa), chymotrypsinogen A (25 kDa) and ribonuclease A (13.7 kDa). Lane B, 10 µg of purified type III CAT. Lane C, pMH1. Lane D, pUC18. Lane E, pUC18:IM3. Lane F, pIM5 uninduced. Lane G, pIM5 induced.

before heat-induction and produce very high yields of CAT within 90–120 min, making this vector the system of choice for the production of mutant CATs that demonstrate enhanced susceptibility to degradation and/or inactivation *in vivo*.

This work was supported by a Science and Engineering Research Council studentship awarded to I. A. M. and a Medical Research Council project grant to W. V. S. We thank Dr. P. C. Engel, Department of Biochemistry, University of Sheffield, Sheffield, U.K., for his assistance during construction of *cat*_{III} expression vectors.

REFERENCES

- Alton, N. E. & Vapnek, D. (1979) *Nature* (London) **282**, 864–869
- Benton, W. D. & Davis, R. W. (1977) *Science* **196**, 180–182
- Biggin, M. D., Gibson, T. S. & Hong, G. F. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 3963–3965
- Birnboim, H. C. & Doly, J. (1979) *Nucleic Acids Res.* **7**, 1513–1523
- Bolivar, F., Rodriguez, R. L., Greene, P. J., Betlach, M. C., Heynecker, H. C., Boyer, H. W., Crosa, J. H. & Falkow, S. (1977) *Gene* **2**, 95–113
- Brenner, D. G. & Shaw, W. V. (1985) *EMBO J.* **4**, 561–568
- Charles, I. G., Keyte, J. W. & Shaw, W. V. (1985) *J. Bacteriol.* **164**, 123–129
- Corney, A. J. (1983) Ph.D. Thesis, University of Leicester

- Fitton, J. E., Packman, L. C., Harford, S., Zaidenzaig, Y. & Shaw, W. V. (1978) in *Microbiology-1978* (Schlessinger, D., ed.), pp. 249-252, American Society for Microbiology, Washington
- Harwood, C. R., Williams, D. M. & Lovett, P. S. (1983) *Gene* **24**, 163-169
- Harwood, J. & Smith, D. H. (1971) *Biochem. Biophys. Res. Commun.* **42**, 57-62
- Horinouchi, S. & Weisblum, B. (1982) *J. Bacteriol.* **150**, 815-825
- Kleanthous, C. & Shaw, W. V. (1984) *Biochem. J.* **223**, 211-220
- Kleanthous, C., Cullis, P. M. & Shaw, W. V. (1985) *Biochemistry* **24**, 5307-5313
- Kushner, S. R. (1978) in *Genetic Engineering* (Boyer, H. W. & Nicosia, S., eds.), pp. 17-23, Elsevier/North-Holland, Amsterdam
- Laemmli, U.K. (1970) *Nature (London)* **227**, 680-685
- LeGrice, S. F. J., Matzura, H., Marcoli, R., Iida, S. & Bickle, T. A. (1982) *J. Bacteriol.* **150**, 312-318
- Leslie, A. G. W., Liddell, J. M. & Shaw, W. V. (1986) *J. Mol. Biol.* **188**, 283-285
- Leslie, A. G. W., Moody, P. C. E. & Shaw, W. V. (1988) *Proc. Natl. Acad. Sci. U.S.A.*, in the press
- Lowry, O. H., Rosebrough, N. H., Farr, A. L. & Randall, R. J. (1951) *J. Biol. Chem.* **193**, 265-275
- Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor
- Messing, J. (1979) *Recomb. DNA Tech. Bull.* **2**, 43-48
- Messing, J. & Vieira, J. (1982) *Gene* **19**, 269-276
- Murray, I. A., Lewendon, A., Kleanthous, C. & Shaw, W. V. (1986) *Biochem. Soc. Trans.* **14**, 1227-1228
- Nakagawa, Y. & Bender, M. L. (1970) *Biochemistry* **9**, 259-267
- Packman, L. C. (1978) Ph.D. Thesis, University of Leicester
- Packman, L. C. & Shaw, W. V. (1981) *Biochem. J.* **193**, 525-539
- Pribnow, D. (1975) *Proc. Natl. Acad. Sci. U.S.A.* **72**, 784-788
- Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. (1977) *J. Mol. Biol.* **113**, 237-251
- Rosenberg, M. & Court, D. (1979) *Annu. Rev. Genet.* **13**, 319-353
- Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463-5467
- Shaw, W. V. (1967) *J. Biol. Chem.* **242**, 687-693
- Shaw, W. V., Sands, L. C. & Datta, N. (1972) *Proc. Natl. Acad. Sci. U.S.A.* **69**, 3049-3053
- Shaw, W. V., Brenner, D. G., LeGrice, S. F. J., Skinner, S. E. & Hawkins, A. R. (1985) *FEBS Lett.* **179**, 101-106
- Shine, J. & Dalgarno, L. (1975) *Nature (London)* **254**, 34-38
- Siebenlist, U. (1979) *Nature (London)* **279**, 651-652
- Sproat, B. S. & Bannwarth, W. (1983) *Tetrahedron Lett.* **24**, 5771-5774
- Thomas, C. D. (1987) Ph.D. Thesis, University of Leicester
- Yanisch-Perron, C., Vieira, J. & Messing, J. (1985) *Gene* **33**, 103-119
- Zoller, M. J. & Smith, M. (1984) *DNA* **3**, 479-487

Received 19 October 1987; accepted 15 January 1988