Reprinted from *J. Mol. Biol.* (1986) **191**, 1–11

Bacillus thuringiensis var. *israelensis* δ-Endotoxin

Nucleotide Sequence and Characterization of the Transcripts in *Bacillus thuringiensis* and *Escherichia coli*

E. S. Ward and D. J. Ellar

J. Mol. Biol. (1986) 191, 1–11

Bacillus thuringiensis var. *israelensis* δ-Endotoxin

Nucleotide Sequence and Characterization of the Transcripts in *Bacillus thuringiensis* and *Escherichia coli*

E. S. Ward and D. J. Ellar†

Department of Biochemistry
University of Cambridge
Tennis Court Road, Cambridge CB2 1QW, England

(Received 20 December 1985, and in revised form 14 April 1986)

The nucleotide sequence of a 1408 base-pair DNA fragment encoding the insecticidal 27,340 M_r δ-endotoxin of *Bacillus thuringiensis* var. *israelensis* has been determined by analysis of a recombinant plasmid from *Escherichia coli*. The hydropathy plot of the protein shows it to be highly hydrophobic, consistent with a postulated cytolytic mechanism of action for the toxin. In addition, the δ-endotoxin transcriptional start points that are used in *B. thuringiensis* and an *E. coli* recombinant have been determined. In *B. thuringiensis* var. *israelensis*, transcription initiates from a single start point, and gene-specific transcripts are not observed before stage II of sporulation. This is the stage at which δ-endotoxin antigen is first detected, indicating that control of expression is primarily at the transcriptional level for this protein. Analysis of gene-specific transcription in *E. coli* indicates that at least three start points are utilized in this organism. Interestingly, the highest level of δ-endotoxin mRNA is seen during mid-exponential growth of *E. coli* and the level appears to decrease as the cells enter the stationary phase of growth.

1. Introduction

Coincidental with sporulation, the Gram-positive spore-forming organism *Bacillus thuringiensis* var. *israelensis* (Goldberg & Margalitt, 1977) produces large amounts of a protein δ-endotoxin that is extremely toxic to the larvae of mosquitoes and blackfly (diptera) (de Barjac, 1978; Thomas & Ellar, 1983*b*). Possibly because of the quantity of δ-endotoxin that is produced (15 to 20% of the cell dry weight), it accumulates in the cytoplasm as a parasporal crystalline inclusion (Somerville, 1978; Bulla *et al.*, 1980) that is irregular in shape and consists of several polypeptides (Huber & Luthy, 1981). Recent reports (Yamamoto *et al.*, 1983; Davidson & Yamamoto, 1984; Thomas, 1984; Armstrong *et al.*, 1985; Sriram *et al.*, 1985) have shown that a 25,000 M_r polypeptide, derived from the 27,000 M_r protein of the *israelensis* crystal, has the insecticidal and haemolytic properties previously associated with purified crystals from this organism (Thomas & Ellar, 1983*b*). In addition, the gene encoding the 27,000 M_r δ-endotoxin has recently been cloned in *Escherichia coli* (Ward *et al.*, 1984; Waalwijk *et al.*, 1985) and in *Bacillus subtilis* (Ward *et al.*, 1986) and sequenced (Waalwijk *et al.*, 1985). Recombinants synthesizing this polypeptide were toxic to mosquito larvae and mosquito cells *in vivo* and *in vitro*, respectively (Ward *et al.*, 1984). Protein extracted from a *B. subtilis* recombinant has also been shown to be haemolytic (Ward *et al.*, 1986).

Despite the considerable scientific and commercial interest in the δ-endotoxins of the strains of *B. thuringiensis*, relatively little is known about the factors controlling their temporal regulation and remarkably high level of expression. Wong, Schnepf and Whiteley have cloned (Schnepf & Whiteley, 1981), sequenced (Schnepf *et al.*, 1985) and identified the transcriptional start points (Wong *et al.*, 1983) of the *B. thuringiensis* var. *kurstaki* δ-endotoxin gene encoding the P_1 protein (Yamamoto & McLaughlin, 1981). This gene encodes a 130,000 M_r lepidopteran toxin (Schnepf & Whiteley, 1981). Interestingly, a tandem promoter for this gene has been found, and the two start points are used differentially during the various stages of sporulation. Tandem promoters have been found for several sporulation-specific genes in *B. subtilis* (Johnson *et al.*, 1983; Wang & Doi, 1984; Wong *et al.*, 1984), and they may be responsible, in part, for the timing and duration of expression of these genes. In addition, Wong & Chang (1985)

† Author to whom all correspondence should be addressed.

0022-2836/86/170001-11 $03.00/0

have recently reported that a 400 bp† restriction fragment containing the transcriptional terminator of the cloned var. *kurstaki* δ-endotoxin gene enhances the expression of cloned genes when placed in either orientation at the distal end. Deletion analysis has indicated that the enhancing activity is located within a 78 bp fragment that contains an inverted repeat sequence, and it has been suggested that transcription of this retro-regulator sequence generates a 3′ stem-and-loop structure, which stabilizes the mRNA against exonucleolytic degradation.

This report describes the nucleotide sequence of an *israelensis* gene encoding a dipteran-specific toxin and the identification of the transcriptional start points used in *B. thuringiensis* var. *israelensis* and *E. coli* by high-resolution S_1 nuclease mapping.

2. Materials and Methods

(a) *Strains and plasmids*

DNA encoding the 27,000 M_r δ-endotoxin gene of *B. thuringiensis* var. *israelensis* was obtained from *E. coli* JM101 containing the plasmid pIP173 (Ward *et al.*, 1984). Plasmid DNA was prepared using a Triton lysis method as described (Ward *et al.*, 1984). Subclones used in this work were derived from pIP173. For DNA sequencing *E. coli* TG1 = K12, Δ(*lac-pro*), *supE, thi, hsdD5*/F′*traD36*, *proA+B+*, *lacI*^{q}, *lacZ*ΔM15 (a kind gift from Dr T. Gibson, MRC Laboratory, Cambridge: Gibson, 1984) and phages M13mp8 and M13mp9 (Messing, 1983) were used.

(b) *Enzymes and reagents*

Restriction enzymes were obtained from either New England Biolabs or Bethesda Research Laboratories. S_1 nuclease was obtained from Bethesda Research Laboratories. Klenow fragment of DNA polymerase and phage T4 DNA ligase were generous gifts from Dr Tim Hunt, University of Cambridge. T4 DNA polymerase was from Pharmacia. Deoxynucleotide and dideoxynucleotide triphosphates were from Boehringer-Mannheim. [α-^{35}S]-dATP and [α-^{32}P]dATP were from Amersham. For restriction enzyme digestions the medium-salt buffer of Maniatis *et al.* was used except with *Eco*RI and *Bam*HI, where the high-salt buffer (Maniatis *et al.*, 1982) was used. For *Sma*I digestions, the buffer recommended by the manufacturer was used. Bovine serum albumin (Boehringer-Mannheim, nucleic acid grade) was added to all restriction enzyme digestions and ligations at a final concentration of 100 μg/ml. Conditions for ligation were as described (Ward *et al.*, 1984).

(c) *DNA sequencing*

The dideoxynucleotide chain termination method (Sanger *et al.*, 1977; Bankier & Barrell, 1983) was used. Restriction fragments encoding the δ-endotoxin gene and flanking sequences were generated using intragenic restriction sites or restriction sites within the polylinker of the cloning vector pUC12 (Messing, 1983). These fragments were usually gel-purified using DEAE-cellulose (Dretzen *et al.*, 1981) before ligating into M13mp8 or M13mp9 cut with the appropriate restriction enzyme(s). Several fragments were end-filled (Maniatis *et al.*, 1982) using T4 DNA polymerase before ligating into one of the blunt-ended cloning sites of M13mp8 or M13mp9.

(d) *RNA preparation*

B. thuringiensis var. *israelensis* cells grown overnight at 30°C in PWYE medium (peptone/water/yeast extract) were diluted into fresh PWYE ($A_{600} = 0{\cdot}1$) and grown to mid-exponential phase before diluting into CCY medium (Stewart *et al.*, 1981). Under these conditions, a high degree of sporulation synchrony was achieved (Ellar & Posgate, 1973), with the proportion of cells in any one sporulation stage rising from 0 to 90% within 1·5 h. Cells were harvested at mid-exponential phase, early stationary phase, and stages I, II, III, IV–V and VI of sporulation, as determined by phase-contrast microscopy. *E. coli* JM101 (pIP173 transformant) cells were grown in L broth (Maniatis *et al.*, 1982) containing 100 μg ampicillin/ml at 37°C overnight and then diluted into fresh L broth. Cells were harvested at mid-exponential phase, early stationary phase and 4 h after the onset of stationary phase. For both *israelensis* and *E. coli*, cells were harvested by centrifugation at the appropriate growth stage and RNA was extracted using the glass bead (B. Braun Melsungen A.G.; 0·1 to 0·11 mm diameter) breakage method described by Nasmyth (1983).

(e) *S_1 nuclease mapping*

To generate probes for S_1 mapping experiments, primer extension using deoxynucleotide triphosphates and [α-^{32}P]dATP followed by restriction enzyme cutting at a primer distal site was used (Farrell *et al.*, 1983; Biggin *et al.*, 1984). The single-stranded DNA probes produced by this method all contain the M13 primer and some of the polylinker sequences. To generate the 393 bp *Taq*I–*Bam*HI probe, an M13mp9 clone with the appropriate insert in the *Acc*I and *Bam*HI site was used, whereas for the 194 bp *Aha*II–*Acc*I probe, M13mp8 with an appropriate end-filled insert in the *Sma*I site was used. Samples of 15 μg of *israelensis* RNA or 30 μg of *E. coli* RNA were mixed with 50 to 200 cts/s probe, the nucleic acids were precipitated with ethanol, evaporated to dryness and carefully resuspended in 25·5 μl of 780 mM-NaCl, 100 mM-PIPES (pH 6·9), 10 mM-EDTA. After incubation at 65°C for 12 to 15 h, 275 μl of ice-cold 4 mM-$ZnSO_4$, 30 mM-sodium acetate, 250 mM-NaCl and 10 μg denatured salmon sperm DNA/ml, containing 10 to 22 units of S_1 nuclease, were added and the mixture was incubated at 37°C for 30 min. The reaction was terminated by addition of EDTA and carrier tRNA to final concentrations of 10 mM and 8 μg/ml, respectively. The S_1-resistant DNA–RNA hybrids were recovered by precipitation with ethanol and analysed using 6% (w/v) acrylamide/8 M-urea sequencing gels.

(f) *Polyacrylamide gel electrophoresis*

Lysates of *israelensis* cells were prepared for protein analysis as follows. Cells were grown as for the RNA preparation, harvested by centrifugation and stored at −80°C. Lysates were made by resuspending cell pellets in 50 mM-Tris–acetate (pH 8·0), 10% (w/v) sucrose followed by addition of lysozyme to a final concentration of 100 μg/ml and incubation at 37°C for 10 min. Cells were

† Abbreviations used: bp, base-pair(s); kb, 10^3 bases or base-pairs. Sequence hyphens are omitted throughout this paper.

then placed on ice and 100 μl of 100 mM-Tris–acetate (pH 8·0), 100 mM-EDTA was added, followed 5 min later by 200 μl of Triton X-100 (2%, v/v). After a further 5 min, DNase and $MgSO_4$ were added to final concentrations of 10 μg/ml and 50 mM, respectively. The volume was then increased to 800 μl using Tris–acetate/10% sucrose. Almost complete lysis of the cells was then achieved by 2 periods of 30 s sonication using a 0·5-inch sonic probe (Dawe Instruments, London) operating at maximum intensity. Ice-cold trichloroacetic acid was added to a final concentration of 12·5% (w/v) to the resulting lysate. The precipitate was collected by centrifugation, washed with 80% (v/v) acetone, and resuspended in double-strength sodium dodecyl sulphate/polyacrylamide gel sample buffer (Stewart *et al.*, 1981). Protein samples were analysed using 13% (w/v) acrylamide gels (Laemmli, 1970) and immunoblotting (Towbin *et al.*, 1979; Hawkes *et al.*, 1982), and using antisera prepared as described (Ward *et al.*, 1984).

3. Results

(a) *Location of the δ-endotoxin gene in pIP173*

The restriction map of a 5·4 kb *Eco*RI restriction fragment (pIPECO 5) derived from the 9·7 kb *Hin*dIII insert of pIP173, with the location of the δ-endotoxin gene, is shown in Figure 1. DNA extracted from a subclone consisting of the 4·6 kb *Pvu*II–*Eco*RI fragment inserted into the cloning vector pUC12 produces a 27,000 M_r δ-endotoxin when used to prime protein synthesis in an *in-vitro* *E. coli* transcription–translation system (Chen & Zubay, 1983; Howe *et al.*, 1982; Ward *et al.*, 1984; data not shown). However, when DNA from a subclone derived from the 3·3 kb *Pvu*II–*Hin*cII fragment is used in the *in-vitro* system (data not shown), a polypeptide product of a slightly lower, or higher, molecular weight than 27,000 is immunoprecipitated, the size of the polypeptide being dependent on the orientation of the insert in the vector. The most likely explanation for these findings is that *Hin*cII cuts at or very close to the carboxy terminus of the δ-endotoxin gene. The restriction fragments shown in Figure 1, derived from the subclone containing the 3·3 kb *Pvu*II–*Hin*cII fragment, or from pIPECO 5, were isolated and cloned into the sequencing vectors M13mp8 and M13mp9 (Messing, 1983). Using the dideoxynucleotide chain termination method (Sanger *et al.*, 1977) the sequence of a 1408 bp segment of DNA that includes the structural gene for the *israelensis* δ-endotoxin has been determined for both the coding and non-coding strands. The nucleotide sequence shown in Figure 2 contains an extended open reading frame (ORF) starting at nucleotide 509 and ending at the termination codon TAA at position 1256. This ORF encodes a polypeptide of 249 residues, molecular weight 27,340. Comparison of the sequence data from nucleotides 369 to 1408 (Fig. 2) with that recently reported by Waalwijk *et al.* (1985) for the var. *israelensis* 27,000 M_r δ-endotoxin shows identity, with the exception of an additional A residue at nucleotide position 370 in our sequence. In addition, 29/30 of the amino acid residues 30 to 59 are in agreement with the

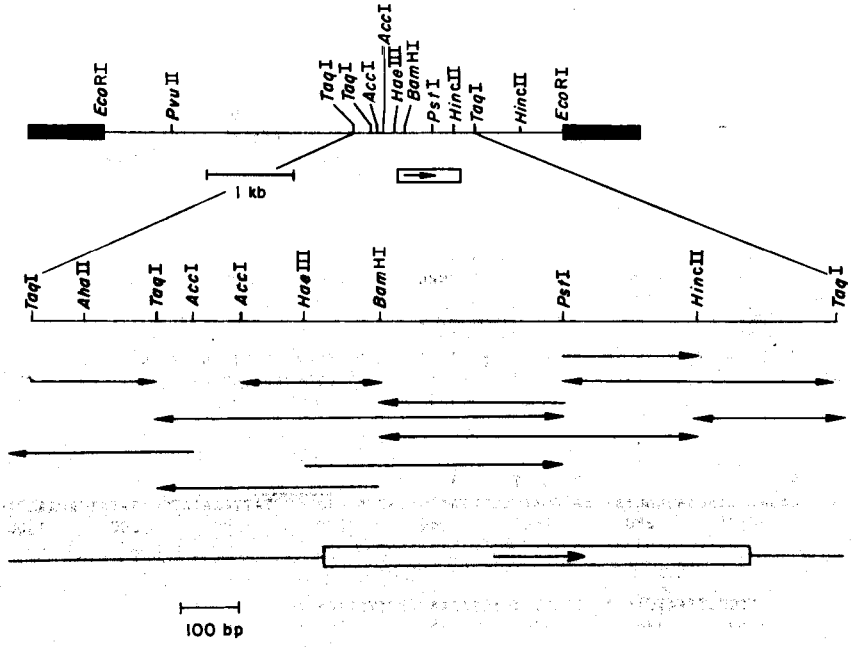

Figure 1. Strategy used for determination of the nucleotide sequence of the δ-endotoxin gene and 5′ and 3′-flanking regions. The restriction endonuclease map of pIPECO 5, with the sequenced region shown enlarged, is shown. The single line represents var. *israelensis* DNA; the filled boxes in the upper diagram represent pUC12 DNA. The location of the 27,000 M_r δ-endotoxin gene, determined using the *E. coli in-vitro* system, is indicated by an open box, with the direction of transcription shown by an arrow. The restriction fragments used in the sequence determination were derived from subclones of pIP173 and cloned in either M13mp8 or M13mp9. The arrows indicate the direction in which the fragments were sequenced.

amino acid sequence data recently reported for the N terminus of a 25,000 M_r polypeptide derived, after proteolysis, from the *israelensis* 27,000 M_r δ-endotoxin (Armstrong *et al.*, 1985).

Examination of the sequences at the 3′ end of the 27,000 M_r δ-endotoxin structural gene has revealed the presence of an IR (inverted repeat) element (shown in Fig. 2), which can potentially form a stem-and-loop structure in the transcript made from this region, with a predicted ΔG value of −26·0 kcal (1 kcal = 4·184 kJ). This is similar in both the 3′ location and free energy to the retroregulator found by Wong & Chang (1985) for the var. *kurstaki* δ-endotoxin gene, and it will be of interest to determine whether this *israelensis* IR affects the level of 27,000 M_r δ-endotoxin expression. In addition, S_1 mapping experiments (data not shown) indicate that although some

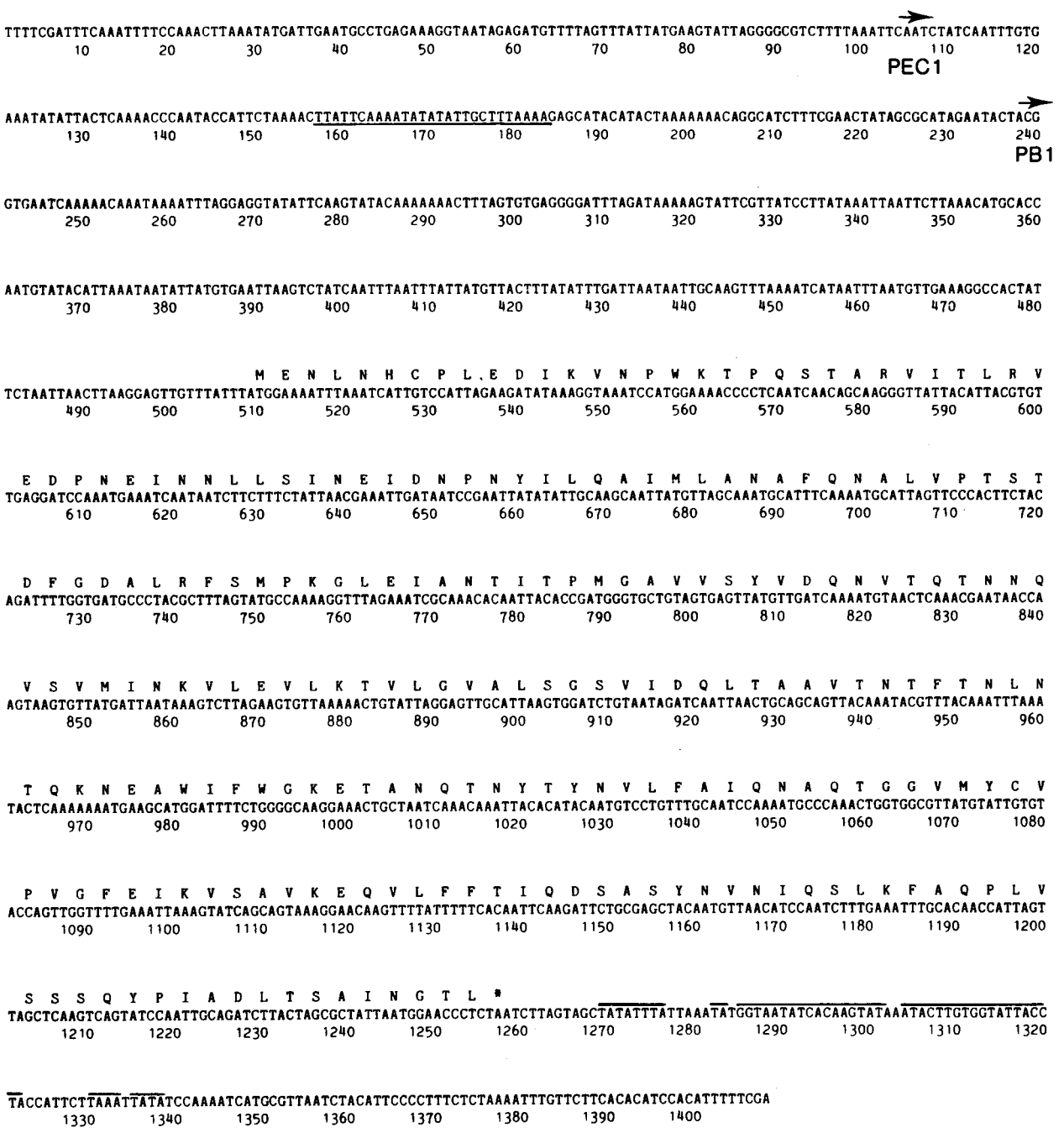

Figure 2. Nucleotide sequence of 1408 bp of var. *israelensis* DNA encoding the 27,000 M_r δ-endotoxin. The amino acid sequence for the open reading frame from nucleotides 509 to 1256 is shown. The initiation sites for transcription used in *B. thuringiensis* var. *israelensis* and *E. coli*, determined by S_1 nuclease mapping are indicated by PB1 (var. *israelensis* and *E. coli*) and PEC1 (*E. coli*). In addition, there appears to be a third initiation site used in *E. coli* upstream from the *Aha*II site at nucleotide 92. This site has not been localized. An inverted repeat sequence beyond the 3′ end of the structural gene is indicated by overlining, and a highly A+T-rich region to the 5′ side of PB1 by underlining. Both these sequences are discussed in the text.

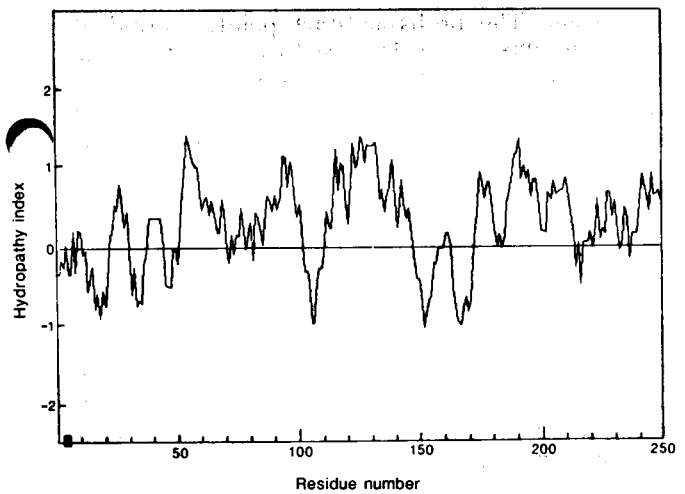

Figure 3. Hydropathy plot of the var. *israelensis* 27,000 M_r δ-endotoxin. The plot was obtained using the computer program by Kyte & Doolittle (1982).

δ-endotoxin transcripts terminate at this potential stem-and-loop structure, a significant proportion extend beyond this sequence. Preliminary results suggest that the length of the 27,000 M_r δ-endotoxin transcript varies with the stage of sporulation at which the mRNA is extracted. Different 3′ termini of the var. *kurstaki* δ-endotoxin gene transcript have been observed in cells harvested during mid and late sporulation (Schnepf *et al.*, 1985).

The hydropathy plot of the δ-endotoxin is shown in Figure 3 (positive values represent hydrophobicity, negative values represent hydrophilicity). As can be seen, there are extensive hydrophobic regions throughout the protein, consistent with the postulated mode of action of the 27,000 M_r δ-endotoxin (Thomas & Ellar, 1983*a*; Ellar *et al.*, 1985). The possible significance of these hydrophobic regions is discussed below.

The amino acid composition (not shown), determined from the nucleotide sequence, shows good agreement with that reported for the 25,000 M_r polypeptide derived from the δ-endotoxin

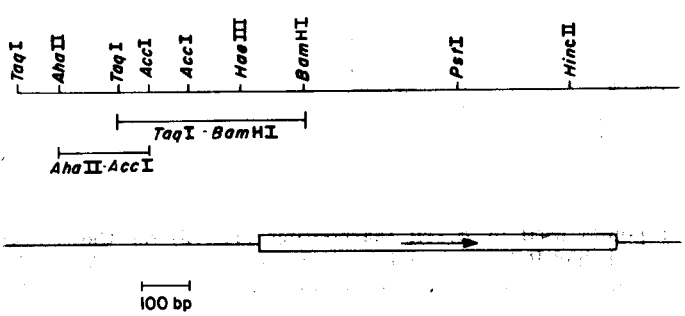

Figure 4. DNA fragments used as probes in the S_1 mapping experiments. The probes were made by the "prime-cut" method (Materials and Methods) using the appropriate M13 clones. Restriction enzyme cleavage sites are shown for each probe.

(Davidson & Yamamoto, 1984; Armstrong *et al.*, 1985).

(b) *Location of the transcriptional start points for the δ-endotoxin gene*

For *israelensis*, RNA was isolated from cultures at the time points indicated in Materials and Methods. The extracted nucleic acids were hybridized to the 393 bp *Taq*I–*Bam*HI probe shown in Figure 4. Figure 5 shows an analysis of the S_1-resistant DNA–RNA hybrids after denaturation using a 6% acrylamide/8 M-urea gel. The T, C, G and A tracks for the dideoxy sequencing reactions of the 393 bp fragment cloned in M13mp9 are

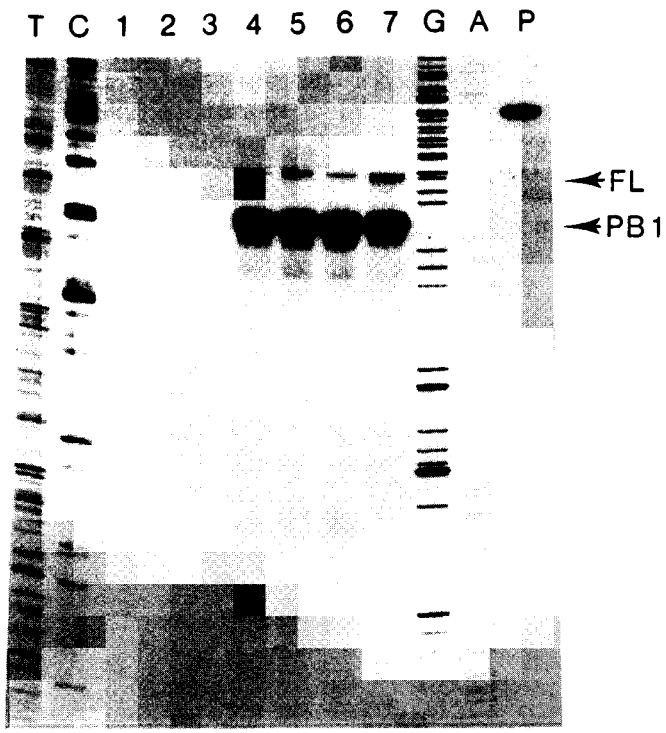

Figure 5. S_1 nuclease mapping of the transcriptional start points used in *B. thuringiensis* var. *israelensis* for the δ-endotoxin gene. DNA fragments derived from the 393 bp *Taq*I–*Bam*HI probe, and protected against S_1 nuclease digestion (see Materials and Methods) by var. *israelensis* mRNA, were analysed on 6% acrylamide/8 M-urea sequencing gels. The T, C, G and A sequencing reactions of the appropriate M13 clone were used to determine the length of the protected probe, taking into account the 5′ primer and polylinker nucleotides present in the dideoxy terminated fragments. RNA samples were extracted from var. *israelensis* cells harvested at the following growth stages: lane 1, mid-exponential; lane 2, early stationary; lane 3, stage I of sporulation; lane 4, stage II; lane 5, stage III; lane 6, stage IV–V; lane 7, stage VI. The lanes labelled P, T, C, G and A correspond to untreated probe (2% of the amount used in each hybridization) and the dideoxy-T, -C, -G and -A sequencing reactions, respectively. The exact size of the protected fragment was determined by running the same samples on a 4% gel for twice as long. FL indicates complete protection of probe, and PB1 corresponds to the start point mapped in var. *israelensis*.

shown (lanes T, C, G and A) as size markers. The start points have been corrected by 27 bp to account for the presence of primer and polylinker sequences at the 5′ end of the single-stranded DNA strands synthesized during the sequencing reactions. Figure 5 shows that mRNA encoding the δ-endotoxin gene was not detected in *israelensis* cells harvested at mid-exponential, early stationary, or stage I of sporulation. However, gene-specific RNA was detected in cells at stages II, III, IV, V and VI of sporulation. The S_1 start point seen in Figure 5, lanes 4, 5, 6 and 7, displays some heterogeneity centred around nucleotide 238, which could either be due to adjacent initiation sites being used *in vivo*, or to an *in-vitro* artefact generated during S_1 nuclease digestion as has been observed by others for both eukaryotic and prokaryotic genes (Hentschel *et al.*, 1980; Wong *et al.*, 1983). The results shown in Figure 5 clearly indicate that the *israelensis* δ-endotoxin gene is only transcribed during stages II to VI of sporulation, and that a start point centred around nucleotide 238 (PB1) is used. The amount of gene-specific RNA appears to increase dramatically during stages I to II and then remain relatively constant at this level throughout sporulation. The probable start point, centred at nucleotide 238, used in *israelensis* is shown in Figure 2 (designated PB1). In addition, three other probes that span this 5′ region from different restriction sites have been used, and yield the same result (data not shown). The transcriptional start point, PB1, differs from that recently reported by Waalwijk *et al.* (1985). Possible reasons for this discrepancy are discussed below.

For *E. coli*, using the 393 bp *Taq*I–*Bam*HI probe, RNA was analysed that had been extracted from cells harvested at the stages of growth indicated in Materials and Methods. Interestingly, as shown in Figure 6(a), gene-specific RNA is produced in *E. coli* during all three stages of growth, but as the cells enter stationary phase, the RNA appears to be less abundant or to undergo degradation, and the start points become less well-defined. In contrast to the findings with var. *israelensis*, at least two start points appear to be used. The major start point is the same as that used by var. *israelensis* (PB1), and the band of higher mobility represents complete protection of the 393 bp probe from S_1 nuclease digestion. This latter species was investigated further using a 194 bp *Aha*II–*Acc*I probe shown in

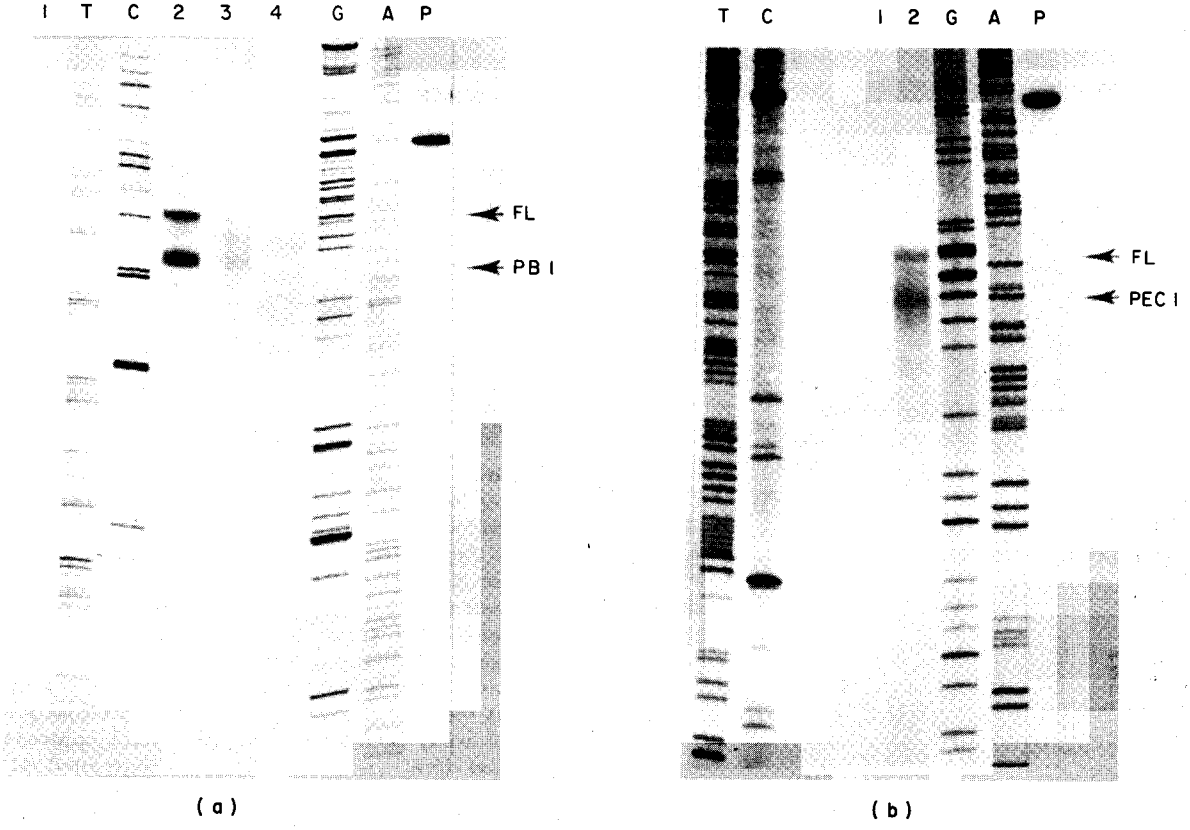

Figure 6. S_1 nuclease mapping of the transcriptional start points used in *E. coli* (harbouring pIP173) for the δ-endotoxin gene. (a) Using the 393 bp *Taq*I–*Bam*HI probe. Lanes labelled P, T, C, G and A are the same as for Fig. 5. RNA samples were harvested from cells at the following growth points: lane 2, mid-exponential; lane 3, early stationary; lane 4, 4 h after onset of stationary phase. Lane 1 is a control using 30 μg of carrier tRNA. (b) Using the 194 bp *Aha*II–*Acc*I probe. The dideoxy-T, -C, -G and -A sequencing restrictions of the M13 clone used to make the probe are indicated, and the lane labelled P corresponds to untreated probe. Lane 1 is a control using 30 μg of carrier tRNA; lane 2 shows the protection of probe by RNA harvested from cells at the mid-exponential stage of growth. In both (a) and (b), FL indicates complete protection of probe, and PB1 and PEC1 correspond to the start points that have been mapped in *E. coli*.

Table 2
Comparison of Bacillus *promoter sequences*

Holoenzyme/ promoter	−35	Spacing	−10
B. subtilis			
$E\sigma^{43}$	TTGACA	17–18	TATAAT
$E\sigma^{37}$	AGG-TT	13–16	GG-ATTG-T
$E\sigma^{32}$	AAATC	14, 15	TA-TG-TT-TA
$E\sigma^{29}$	A-TT-AAAA	14–17	CATATT-T
$E\sigma^{28}$	CTAAA	16	CCGATAT
var. *israelensis*			
PB1	GCATCTTT	12	CATAGAAT
var. *kurstaki*			
Bt1	GCATTTTT	13	CATATGTT

The conserved −10 and −35 regions of *B. subtilis* promoters (Johnson *et al.*, 1983; Tatti & Moran, 1985; Cowing *et al.*, 1985) recognized by different holoenzyme forms of RNA polymerase and the −10 and −35 sequences of PB1 (var. *israelensis*) and Bt1 (var. *kurstaki*) (Wong *et al.*, 1983) are shown. The underlined bases of the var. *israelensis* sequences show identity with the corresponding *kurstaki* nucleotides. The spacing between the 2 conserved regions is indicated in base-pairs.

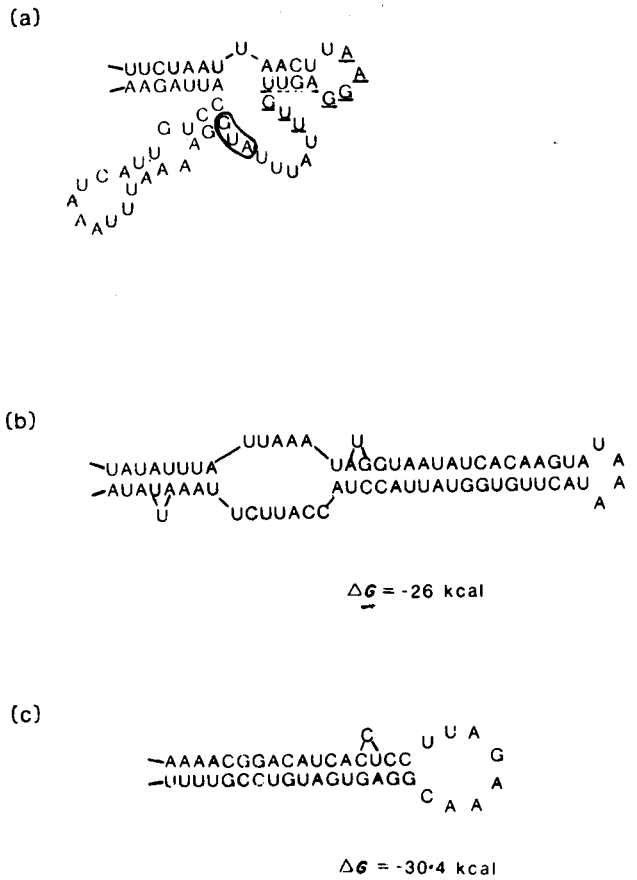

Figure 8. (a), (b) and (c). Potential secondary structures at (a), the 5′ end, nucleotides 480 to 538 (Fig. 2), and (b) the 3′ end, nucleotides 1269 to 1338 (Fig. 2), of the var. *israelensis* 27,000 M_r δ-endotoxin mRNA. The initiation codon of the δ-endotoxin gene in (a) is shown boxed, and the putative Shine–Dalgarno sequence (Shine & Dalgarno, 1975) is underlined. The potential secondary structure at the 3′ end of the var. *kurstaki* P1 δ-endotoxin mRNA (Wong & Chang, 1985) is shown in (c) for comparative purposes.

the respective strains of *B. thuringiensis*. There is in addition an extremely A+T-rich box upstream from PB1, shown underlined in Figure 2, which has also been found in var. *kurstaki* and in the *B. subtilis* spoVG promoter (Banner *et al.*, 1983). For the spoVG promoter, deletion analysis has shown that this A+T-rich region plays an important role in enhancing transcription.

The data obtained in the S_1 mapping and immunoblotting experiments indicate that the temporal control of expression of the *israelensis* gene is at the transcriptional, rather than the translational, level. It will be of interest to elucidate the factors that determine the timing and high efficiency of expression of this gene. Transcriptional initiation at PB1 generates a mRNA with an unusually long 5′ untranslated sequence; several reports have shown that the accessibility of the initiation codon, and to a lesser extent, the ribosome binding site, are important determinants of the level of expression of a gene (Iserentant & Fiers, 1980; Derom *et al.*, 1982; Gheysen *et al.*, 1982; Wood *et al.*, 1984). Analysis of the predicted secondary structure (Tinoco *et al.*, 1973; Jacobson *et al.*, 1984) for the mRNA in the region between nucleotides 401 and 800 indicates that the Shine–Dalgarno sequence and initiation codon may be in exposed regions as shown in Figure 8(a). It is conceivable that this plays a role in enhancing δ-endotoxin gene expression, although further work is required to investigate this possibility.

Interestingly, there is an inverted repeat (IR) element in the sequences 3′ to the structural gene, which could give rise to a stem-and-loop structure in the transcript of this region, with a predicted ΔG value similar to that for the retroregulator found by Wong & Chang (1985) for the var. *kurstaki* δ-endotoxin gene (shown in Fig. 8(b) and (c)). For the var. *kurstaki* P1 δ-endotoxin gene, however, this IR element also appears to act as a transcriptional terminator, whereas, for var. *israelensis*, *S_1 mapping* (data not shown) indicates that some transcription proceeds beyond the IR sequence. As noted by Wong & Chang (1985), termination of transcription immediately beyond the stem-and-loop structure does not appear to be essential for the function of the IR element, since enhancement of expression was observed when the *kurstaki* retroregulator was placed in either orientation at the 3′ end of cloned genes. It will be of significance to determine the possible role of the var. *israelensis* 3′ IR element in enhancing δ-endotoxin expression.

We thank Dr J. Todd, Dr C. Howe and A. Miller for helpful advice, and Drs P. Alefounder and R. Staden for invaluable aid with the computer analysis. We are grateful to L. Jewitt and J. Moss for photographic assistance. This work was supported by the Medical Research Council (E.S.W.) and Agricultural Research Council (D.J.E.).

References

Aoyama, T., Takanami, M., Ohsuka, E., Taniyama, Y., Marumoto, R., Sato, H. & Ikehara, M. (1983). *Nucl. Acids Res.* **11**, 5855–5864.

Armstrong, J. L. Rohrmann, G. F. & Beaudreau, G. S. (1985). *J. Bacteriol.* **161**, 39–46.

Bankier, A. & Barrell, B. (1983). In *Techniques in the Life Sciences,* vol. B508, pp. 1–34, Elsevier, Amsterdam.

Banner, C. D. B., Moran, C. P. Jr & Losick, R. (1983). *J. Mol. Biol.* **168**, 351–365.

Biggin, M., Farrell, P. & Barrell, B. (1984). *EMBO J.* **3**, 1083–1090.

Birdsall, N. J. M., Ellar, D. J., Lee, A. G., Metcalfe, J. C. & Warren, G. B. (1975). *Biochim. Biophys. Acta,* **380**, 344–354.

Briat, J.-F., Gilman, M. Z. & Chamberlin, M. J. (1985). *J. Biol. Chem.* **260**, 2038–2041.

Bulla, L. A. Jr, Bechtel, D. B., Kramer, K. J., Shethna, Y. I., Aronson, A. I. & Fitz-James, P. C. (1980). *Crit. Rev. Microbiol.* **8**, 147–204.

Chen, H.-Z. & Zubay, G. (1983). *Methods Enzymol.* **101**, 674–690.

Cowing, D. W., Bardwell, J. C. A., Craig, E. A., Woolford, C., Hendrix, R. W. & Gross, C. A. (1985). *Proc. Nat. Acad. Sci., U.S.A.* **82**, 2679–2683.

Davidson, E. W. & Yamamoto, T. (1984). *Curr. Microbiol.* **11**, 171–174.

de Barjac, H. (1978). *CR Acad. Sci. Paris, Ser. D,* **286**, 797–800.

Derom, C., Gheysen, D. & Fiers, W. (1982). *Gene,* **17**, 45–54.

Dretzen, G., Bellard, M., Sassone-Corsi, P. & Chambon, P. (1981). *Anal. Biochem.* **112**, 295–298.

Ellar, D. J. & Posgate, J. A. (1973). In *Spore Research 1973* (Barker, A. N., Gould, G. N. & Wolf, J., eds), pp. 21–40, Academic Press, London.

Ellar, D. J., Thomas, W. E., Knowles, B. H., Ward, S., Todd, J., Drobniewski, F., Lewis, J., Sawyer, T., Last, D, & Nichols, C. (1985). In *Molecular Biology of Microbial Differentiation, Proceedings of the Ninth International Spores Conference, Asilomar, California, September, 1984* (Hoch, J. A. & Setlow, P., eds), pp. 230–240, American Soc. Microbiol., Washington, D.C.

Farrell, P. J., Deininger, P. L., Bankier, A. & Barrell, B. (1983). *Proc. Nat. Acad. Sci., U.S.A.* **80**, 1565–1569.

Ganesan, S., Kamdar, H., Jayaraman, K. & Szulmajster, J. (1983). *Mol. Gen. Genet.* **189**, 181–183.

Gheysen, D., Iserentant, D., Derom, C. & Fiers, W. (1982). *Gene,* **17**, 55–63.

Gibson, T. J. (1984). Ph.D. thesis, University of Cambridge.

Gilman, M. Z., Wiggs, J. L. & Chamberlin, M. J. (1981). *Nucl. Acids Res.* **9**, 5991–6000.

Goldberg, L. J. & Margalitt, J. (1977). *Mosquito News,* **37**, 355–358.

Grossman, A. D., Erickson, J. W. & Gross, C. A. (1984). *Cell,* **38**, 383–390.

Hawkes, R., Niday, E. & Gordon, J. (1982). *Anal. Biochem.* **119**, 142–147.

Hentschel, C., Irminger, J.-C., Baucher, P. & Birnstiel, M. L. (1980). *Nature (London),* **285**, 147–151.

Howe, C. J., Bowman, C. M., Dyer, T. & Gray, J. C. (1982). *Mol. Gen. Genet.* **186**, 525–530.

Huber, H. E. & Luthy, P. (1981). In *Pathogenesis of Microbial Diseases* (Davidson, E. W., ed.), pp. 209–234, Allanheld Osmun, New Jersey.

Iserentant, D. & Fiers, W. (1980). *Gene,* **9**, 1–12.

Jacobson, A. B., Good, L., Simonetti, J. & Zuker, M. (1984). *Nucl. Acids Res.* **12**, 45–52.

Jacquet, M.-A. & Ehrlich, R. (1985). *Biochimie,* **67**, 987–997.

Johnson, W. C., Moran, C. P. & Losick, R. (1983). *Nature (London),* **302**, 800–804.

Kyte, J. & Doolittle, R. F. (1982). *J. Mol. Biol.* **157**, 105–132.

Laemmli, U. K. (1970). *Nature (London),* **227**, 680–685.

Landick, R., Vaughn, V., Lau, E. T., VanBogelen, R. A., Erickson, J. W. & Neidhardt, F. C. (1984). *Cell,* **38**, 175–182.

Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982). In *Molecular Cloning: A Laboratory Manual,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

McLaughlin, J. R., Murray, C. L. & Rabinowitz, J. C. (1981). *J. Biol. Chem.* **256**, 11283–11291.

Messing, J. (1983). *Methods Enzymol.* **101**, 20–78.

Moran, C. P., Lang, N., LeGrice, S. F. J., Lee, G., Stephens, M., Sonensheim, A. L., Pero, J. & Losick, R. (1982). *Mol. Gen. Genet.* **186**, 339–346.

Nasmyth, K. (1983). *Nature (London),* **302**, 670–676.

Rosenberg, M. & Court, D. (1979). *Annu. Rev. Genet.* **13**, 319–353.

Sanger, F., Nicklen, S. & Coulson, A. R. (1977). *Proc. Nat. Acad. Sci., U.S.A.* **74**, 5463–5468.

Schnepf, H. E. & Whiteley, H. R. (1981). *Proc. Nat. Acad. Sci., U.S.A.* **78**, 2895–2897.

Schnepf, H. E., Wong, H. C. & Whiteley, H. E. (1985). *J. Biol. Chem.* **260**, 6264–6272.

Shine, J. & Dalgarno, L. (1975). *Nature (London),* **254**, 34–38.

Somerville, H. J. (1978). *Trends Biochem. Sci.* **3**, 108–110.

Sriram, R., Kamdar, H. & Jayaraman, K. (1985). *Biochem. Biophys. Res. Commun.* **132**, 19–27.

Stewart, G. S. A. B., Johnstone, K., Hagelberg, E. & Ellar, D. J. (1981). *Biochem. J.* **196**, 101–106.

Tatti, K. M. & Moran, C. P. Jr (1985). *Nature (London),* **314**, 190–192.

Thomas, W. E. (1984). Ph.D. thesis, University of Cambridge.

Thomas, W. E. & Ellar, D. J. (1983*a*). *FEBS Letters,* **145**, 362–368.

Thomas, W. E. & Ellar, D. J. (1983*b*). *J. Cell Sci.* **60**, 181–197.

Tinoco, I. Jr, Borer, P. N., Dengler, B., Levine, M. D., Uhlenbeck, O. C., Crothers, D. M. & Gralla, J. (1973). *Nature New Biol.* **246**, 40–41.

Towbin, H., Staehlin, T. & Gordon, J. (1979). *Proc. Nat. Acad. Sci., U.S.A.* **76**, 4350–4354.

Waalwijk, C., Dullemans, A. M., van Workum, M. E. S. & Visser, B. (1985). *Nucl. Acids Res.* **13**, 8207–8217.

Wang, P.-Z. & Doi, R. H. (1984). *J. Biol. Chem.* **259**, 8619–8625.

Ward, E. S., Ellar, C. J. & Todd, J. A. (1984). *FEBS Letters,* **175**, 377–381.

Ward, E. S., Ridley, A. R., Ellar, D. J. & Todd, J. A. (1986). *J. Mol. Biol.* **191**, 13–22.

Wong, H. C. & Chang, S. (1985). In *Molecular Biology of Microbial Differentiation (Proceedings of the Ninth International Spores Conference, Asilomar, California, September, 1984* (Hoch, J. A. & Setlow, P., eds), pp. 104–109, American Soc. Microbiol., Washington, D.C.

Wong, H. C., Schnepf, H. E. & Whiteley, H. E. (1983). *J. Biol. Chem.* **258**, 1960–1967.

Wong, S.-L. & Doi, R. H. (1984). *J. Biol. Chem.* **259**, 9762–9767.

Wong, S.-L., Price, C. W., Goldfarb, D. S. & Doi, R. H. (1984). *Proc. Nat. Acad. Sci., U.S.A.* **81**, 1184–1188.

Wood, C. R., Boss, M. A., Patel, T. P. & Emtage, J. S. (1984). *Nucl. Acids Res.* **12**, 3937–3950.
Yamamoto, T. & McLaughlin, R. E. (1981). *Biochem. Biophys. Res. Commun.* **103**, 414–421.
Yamamoto, T., Iizuka, T. & Aronson, J. N. (1983). *Curr. Microbiol.* **9**, 279–284.

Note added in proof. Since this manuscript was prepared, a report has appeared (Thorne, L., Garduno, F., Thompson, T., Decker, D., Zounes, M., Wild, M., Walfield, A. M. & Pollock, T. J. (1986). *J. Bacteriol.* **166**, 801–811) indicating that insertion of transposon Tn*5* between an upstream A+T-rich region and the promoter sequences, or deletion of these upstream A+T-rich sequences, enhances the expression of a cloned *B. thuringiensis* var. *kurstaki* δ-endotoxin gene (the "4·5 gene") in recombinant *E. coli* cells. Similar observations were made by these authors for a second var. *kurstaki* gene (the "5·3 gene").

Edited by J. Miller