The EMBO Journal vol.4 no.13B pp.3887–3893, 1985

Transcriptional activation of a Pap pilus virulence operon from uropathogenic *Escherichia coli*

Monica Båga, Mikael Göransson, Staffan Normark and Bernt Eric Uhlin

Department of Microbiology, University of Umeå, S-901 87 Umeå, Sweden

Communicated by K.Nordström

A gene cluster mediating production of pili in uropathogenic *Escherichia coli* was analysed with respect to regulation of pili synthesis. Two cistrons, *papB* and *papI*, were localized upstream of the major pilus subunit gene, *papA*. The *papI-papB-papA* region was characterized by nucleotide sequencing and by transcriptional analysis. The *papA* gene was primarily represented by an 800 nucleotide long transcript but was also co-transcribed with *papB* as a less abundant 1300 nucleotide long mRNA. Both transcripts presumably terminated at the same site downstream of the *papA* coding sequence. The weakly expressed *papI* gene was transcribed in the opposite direction to that of *papB* and *papA*. Studies with *lacZ* operon fusions showed that the *papB* gene encoded a *trans*-active effector required for *papA* transcription. Similarly, the *papI* gene stimulated *papB* transcription in *trans*. Furthermore, full expression of *papA* was *cis* dependent upon the *papI-papB* region. Transcription of the *papB* gene was shown to be dependent upon cAMP and its receptor protein. A binding site for the cAMP-CRP complex was postulated in the DNA sequence upstream of the *papB* promoter.

Key words: bacterial virulence/cAMP-CRP/nucleotide sequence/pilus gene regulation/*trans*-active gene products

Introduction

The ability of the pathogenic bacteria to adhere to certain host tissues is of primary importance in diseases such as diarrhoea, gonorrhoea and urinary tract infections (Beachey, 1981). The uropathogenic *Escherichia coli* frequently express filamentous appendages called P-fimbriae or Pap pili that mediate binding to an α-D-Gal*p*-(1-4)-β-D-Gal*p* moiety present in the globoseries of glycolipids (Källenius *et al.*, 1980; Leffler and Svanborg Edén, 1980). Eight closely linked *pap* genes involved in the biogenesis of Pap pili have been identified in a gene cluster from the uropathogenic *E. coli* isolate J96 (Norgren *et al.*, 1984; Normark *et al.*, 1985; Uhlin *et al.*, 1985a, 1985b). One of these genes, *papB*, was suggested to have a regulatory role in Pap pilus expression (Norgren *et al.*, 1984).

The expression of pili is affected by growth conditions such as temperature and composition of the growth media (Gaastra and de Graaf, 1982; de Graaf *et al.*, 1980). In the case of Pap pili transcription of the structural gene for the pilus subunit (*papA*) is regulated in response to the growth temperature (Göransson and Uhlin, 1984). The presence of glucose in the growth medium in general seems to reduce expression of pili and adhesive properties of strains carrying type 1 pili (Old and Duguid, 1970), P-fimbriae (Svanborg Edén and Hansson, 1978) or K99 pili (Isaacson, 1980). The effect resembles the catabolite repression which regulates the bacterial metabolic activity in response to several sugars. Evidence for involvement of adenosine 3′,5′-monophosphate (cAMP) and its receptor protein (CRP) has been obtained in the case of type 1 pili of *Salmonella typhimurium* (Saier *et al.*, 1978), and K99 pili of *E. coli* (Isaacson, 1980) when pili production was monitored in *cya*$^-$ or *crp*$^-$ strains. For type 1 pili of *E. coli* it has been suggested that the glucose effect is a manifestation of selective outgrowth of non-piliated cells in the population and not a result of regulatory features in pili synthesis (Eisenstein and Dodd, 1982).

To elucidate the molecular genetics of regulatory functions involved in pili synthesis we have made a detailed characterization of the genes mediating Pap pili synthesis. By nucleotide sequencing and transcriptional analysis we have defined the region around the *papA* pilin gene and obtained evidence for positive control of its transcription.

Results

Gene organization and nucleotide sequence of the pap control region

Several *pap* cistrons involved in the biogenesis of Pap pilin adhesin were previously localized in the plasmid pRHU845 (Norgren *et al.*, 1984). It was also shown that transposon Tn*5* insertions in this plasmid upstream from the *papA* pilin gene decrease the expression of pilin antigen. Two such Tn*5* insertions, pSN025 and pSN020, are associated with the loss of an ~13 000-dalton polypeptide in a minicell expression system (Norgren *et al.*, 1984). The corresponding gene, *papB*, was, therefore, tentatively localized 1.1–1.8 kb from the left *Eco*RI site of pRHU845 (Figure 1a). One Tn*5* insertion, pSN022, located 0.64 kb from the *Eco*RI site, expressed the *papB* gene product but resulted in a slightly lower expression of the PapA pilin protein. This insertion abolished the expression of an ~12 000-dalton polypeptide (data not shown). The corresponding gene is hereby designated *papI*.

We anticipated that either *papI*, *papB* or regulatory sequences in this region are directly involved in the control of *papA* expression. To define this region more precisely, we determined the DNA sequence of both strands from the *Eco*RI site to 200 bp downstream from the *papA* gene. The 1483 nucleotides upstream of the *papA* start codon and the sequence of the *papA-papH* intercistronic region are shown in Figure 2. The DNA sequence of *papA* has been published earlier (Båga *et al.*, 1984).

The sequence contains two sets of anti-parallel reading frames. Based on Tn*5* insertion mapping, computer prediction of ribosomal binding sites and Northern blotting data (see below) the *papB* gene was shown to have the same polarity as *papA* and to code for an ~11 700-dalton polypeptide. The *papI* gene was, by similar analyses, determined to be transcribed in the opposite direction, giving rise to a 77 or 73 amino acid polypeptide depending on which AUG is used for initiation of translation.

Transcriptional analysis of the papA regulatory region

To localize transcriptional products in the *papA* regulatory region,

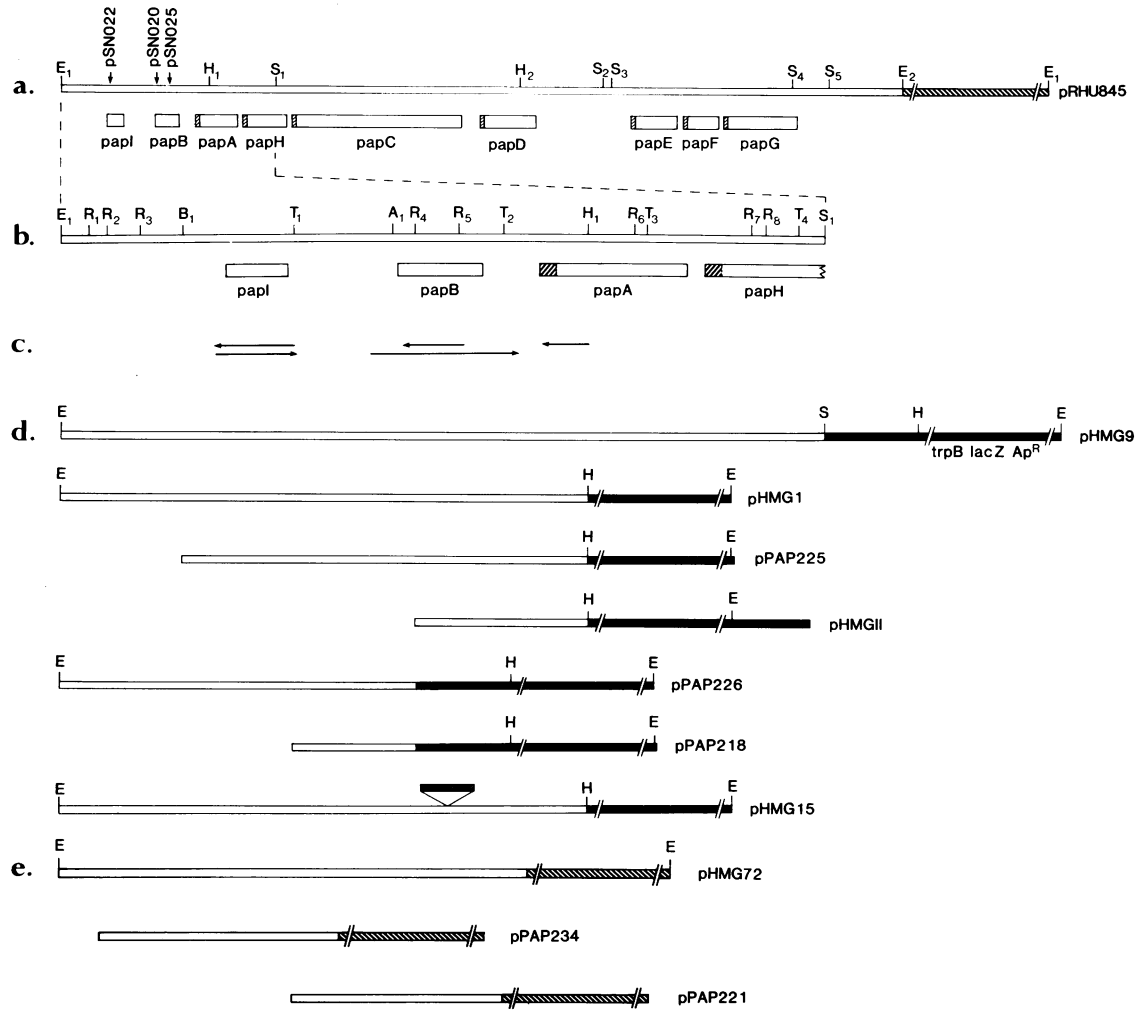

Fig. 1. (**a**) Genetic and physical map of the Pap plasmid pRHU845. Vertical arrows indicate the location of Tn*5* insertions in the derivatives pSN020, pSN022 and pSN025. The open bar shows *pap* DNA, whereas the vector pACYC184 is represented by a hatched bar. Hatched areas in the boxes representing *pap* gene products show the signal peptide of the respective protein. (**b**) Physical map of the 2.8-kb *Eco*RI$_1$-*Sma*I$_1$ region of the plasmid pRHU845. The recognition sites for the endonucleases *Eco*RI (E), *Bal*I (B), *Rsa*I (R), *Taq*I (T), *Apa*I (A), *Hind*III (H) and *Sma*I (S) are indicated. (**c**) Arrows indicate *pap* DNA cloned into M13 vectors used for isolation and labelling of strand-specific probes in Northern hybridization experiments. (**d**) Constructs of *pap-lacZ* transcriptional fusions. The vector pRZ5202 DNA is represented by filled bars. The position of Tn*5* DNA in the *papB* gene after construction of pHMG73 (see Materials and methods) is indicated by a bar above the map of the *lac* fusion derivative pHMG15. (**e**) Plasmids used in *trans*-complementation experiments.

Northern blot hybridizations were performed using *papB, papI* and *papA* single strand sequences as hybridization probes (Figure 1c). No distinct hybridization was detected in control experiments with RNA extracted from cells carrying the vector plasmid pBR322 (Figure 3; lanes A2 – A4, B2 – B4 and C2). An internal *papB* probe complementary to *papB* mRNA hybridized to a weakly expressed 1300 base long transcript (Figure 3, lane B7). In addition, a transcript of a similar size was detected by the *papA* probe (Figure 3, lane A7). This latter probe also hybridized to a heavily expressed 800 base long transcript (lanes A5 – A7). In hybridization experiments with mRNA extracted from the *E. coli* J96 clinical isolate, the *papA* and the *papB* probes gave the same pattern (data not shown) as shown here for the cloned *pap* genes. Neither of these transcripts hybridized to probes proximal or distal to the *papB, papA* region nor to probes complementary to the sense strand (data not shown). These data indicate that a promoter is located between *pap*I and *papB* and directs transcription of both the *papB* and *papA* genes. Furthermore, the 5′ end of the 800 base long *papA* transcript must be positioned between *papB* and *papA*. Immediately distal to the 3′ end of the *papA* pilin gene the sequence contains a dyad symmetry at base pairs 2339 – 2365 (Figure 2) that at the RNA level can form a stem and loop structure with a ΔG value of −20.0 kcal/mol, as calculated according to the rules of Tinoco *et al.* (1973). This dyad symmetry is followed by a stretch of T residues. Hence this region has all the features of a rho-independent transcriptional terminator (Rosenberg and Court, 1979). Defining this uridine stretch as the end point for both transcriptions, the start point for the 1300 base long transcript must be around position 1070 and the start for the short 800-base transcript around position 1570 (Figure 2). In a computerized search for possible promoters three potential sequences were found upstream from the *papB* initiation codon, as indicated in Figure 2. In the intercistronic region between *papB* and *papA* no obvious promoter sequence was found, despite the fact that the *papA* gene is the most efficiently transcribed gene.

A weakly expressed 600 base long transcript was found with a *papI* probe showing complementarity to the reading frame which is opposite to the polarity of *papB*. No detectable hybridization was found with a *papI* probe exhibiting the opposite polari-

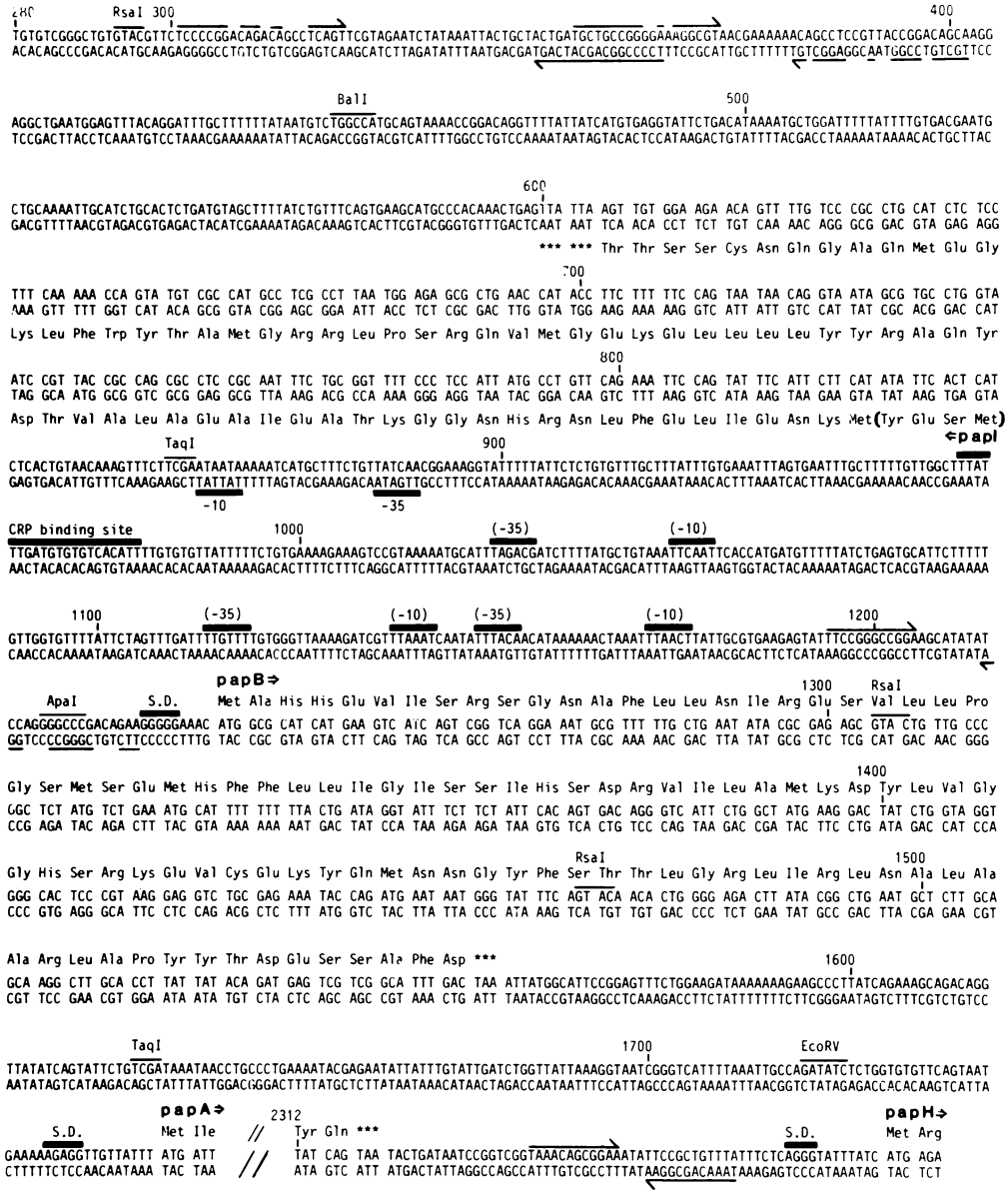

Fig. 2. Nucleotide sequence of the *papA* control region and the region between *papA* and *papH*. The nucleotides are numbered in relation to the *Eco*RI$_1$ site (cf. Figure 1). The amino acid sequences of PapI and PapB were deduced from the DNA sequence. Three potential promoter regions (postulated −35 and −10 sequences) for *papB* transcription and one for *papI* transcription are indicated by thick horizontal lines. Similarly, the positions for postulated ribosomal binding sites and a cAMP-CRP binding site are shown above the sequence. Horizontal arrows show regions with dyad symmetries. Recognition sequences for some of the restriction enzymes mentioned in the text are indicated by thin lines.

ty. These data clearly support our hypothesis that *papI* and *papB* are transcribed in opposite directions. Two potential terminator structures with predicted ΔG values of −26.4 and −34.6 kcal/mol, respectively, were found in the sequence from bases 301 to 402. It is therefore likely that the 600 base long *papI* transcript terminates at one of these sites and that the *papI* promoter is the one indicated in Figure 2.

The transcriptional activities were monitored also by a set of *lacZ* fusion constructs (Figure 1d). A quantitative determination of the *papH* transcription using the *papH-lacZ* fusion pHMG9 showed a roughly 5-fold lower activity in comparison with the *papA-lacZ* fusion pHMG1 (Table I). This result supports the evidence from DNA sequencing that a transcriptional terminator exists in the intercistronic region between *papA* and *papH*. The *papA-lacZ* fusion pHMG1 showed a 4-fold higher β-galactosidase expression than did the *papB-lacZ* fusion pPAP226 (Table I).

Table I. Analysis of *pap* gene transcription by *lacZ* operon fusions

Plasmid	*pap* genotype	β-Galactosidase specific activity (units)	
		−Glucose	+Glucose[a]
pHMG9	*papI*$^+$,*B*$^+$,*A*$^+$,*H-lacZ*	95	45
pHMG1	*papI*$^+$,*B*$^+$,*A-lacZ*	450	340
pPAP225	*papI*$^+$,*B*$^+$,*A-lacZ*	490	370
pHMG11	*papA-lacZ*	8	8
pPAP226	*papI*$^+$,*B-lacZ*	120	34
pPAP218	*papB-lacZ*	50	12
pHMG15	*papI*$^+$,*B1*,*A-lacZ*	100	29
pRZ5202	Vector control	5	3

[a]MC1029 cells harbouring *pap-lacZ* fusion plasmids or the vector pRZ5202 were grown in Casa medium with or without 1% glucose to analyze the effect of a continuously high concentration of glucose.

Thus, as was evident from the Northern blot hybridizations, the *papA* gene is more efficiently transcribed than the *papB* gene.

Transcription of the papA pilin gene is influenced by papB and papI

The role of the *papI, B* region in *papA* transcription was studied with *papA-lacZ* constructs that have different 5′ end points (Figure 1d; Table I). There was no difference in expression between pHMG1 and pPAP225, indicating that the region from *Eco*RI$_1$ to *Bal*I$_1$ does not affect *papA* transcription. Construct pHMG11 carries even less *pap* DNA and lacks the *papB* promoter region. It has its 5′ end within the coding sequence of *papB*, thus containing the intercistronic region between *papB* and *papA* from which the 800-base transcript is presumably initiated *in vivo*. This construct had nearly the same low β-galactosidase activity as the fusion vector pRZ5202. Based on these *papA-lacZ* constructions it appears that *papA* transcription is positively affected by the *papB-papI* region.

To find out whether or not the *papB* gene is responsible for this activation, a *papB* mutant, pHMG73, was constructed. A *papA-lacZ* fusion derivative, pHMG15, of this mutant showed a 5-fold reduction of β-galactosidase activity as compared with pHMG1 (Table I). The activity of this *papB* mutant was of the same level as in the *papB-lacZ* construct pPAP226. The data suggest that the *papB* promoter is solely responsible for the transcriptional activity through *papA* in this *papB* mutant. Furthermore, it was evident that an intact *papB* gene is needed for full activation of *papA* transcription.

The construct pPAP218 defines transcriptional activity from the *papI-papB* intercistronic region in the absence of expression of these two proteins. The expression was low for this construct but was significantly higher than for the vector (Table I). When comparing pPAP218 with the *papI,B-lacZ* fusion pPAP226 the latter expressed a 2-fold higher β-galactosidase activity. Thus transcription of *papB* seemed positively affected by the *papI* gene.

The above analyses indicate that transcription of the *papA* pilin gene may be dependent on several regulatory components. To test for *trans*-active regulatory functions a series of complementation experiments were performed. From analysis of proteins expressed in minicells the plasmid pPAP221 was found to produce the PapB protein. Construct pHMG72 produced both the PapI and PapB proteins, and plasmid pPAP234 expressed only the *papI* gene product (data not shown). These constructs (Figure 1e) were introduced into the *E. coli* strain MC1029 harbouring the *papA-lacZ* or *papB-lacZ* operon fusions (Table II). None of the complementing plasmids altered significantly the expression of β-galactosidase from the *papA-lacZ* fusion pHMG11 which completely lacks the *papI* gene and any promoter in the *papI-papB* intercistronic region. However, the *papB*$^+$ plasmids did restore full *papA-lacZ* β-galactosidase expression in the case of the *papB1* mutant construct pHMG15 (Figure 1d; Table II). The fully active *papA* gene transcription therefore seems to require that the *papI-papB* region is present in *cis*. The complementation tests also provided evidence that *papB* transcription is stimulated in *trans* by the *papI* region as well as by the *papB* gene itself. Introduction of the plasmid pPAP234 or plasmid pPAP221, into cells habouring the *papB-lacZ* fusion pPAP218 resulted in a 2-fold increase in β-galactosidase expression (Table II).

Catabolite repression of pilin synthesis operates via papB

Having defined the *papA* regulatory region we investigated its possible role in glucose repression. Preliminary determinations of the glucose effect on β-galactosidase expression from *papB-* and *papA-lacZ* fusions indicated that the *pap* regulatory region may contain the target for the glucose effect. Addition of 1% glucose to the Casa medium markedly decreased the expression of β-galactosidase from the two *papB-lacZ* fusions, pPAP218 and pPAP226 (Table I). A significant, but less dramatic, decrease of activity was also found for the *papA* fusions carrying the *papB* region (pHMG1 and pPAP225; Table I). Furthermore, Northern blot analyses showed that both the 1300-base *papB-papA* transcript and the 800-base *papA* transcript were significantly less abundant in *E. coli* cells growing in LB medium or Casa medium containing 1% glucose as compared with cells growing in Casa medium (Figure 3, lanes A5 – A7, B5 – B7). Hence, the glucose effect on Pap piliation must operate at the level of transcription.

Since the effect of glucose resembled what could be expected if catabolite repression was occurring we decided to analyse if cAMP and its receptor protein (CRP) might be involved. This can be done genetically be employing strains devoid of the adenylate cyclase (*cya*$^-$ mutants) or the receptor protein (*crp*$^-$ mutants) (Ullmann and Danchin, 1983). Therefore, *papA*- and *papB-lacZ* fusions were analysed in isogenic pairs of *crp*$^+$, *crp*$^-$ (Table III) and *cya*$^+$, *cya*$^-$ strains (Table IV). In both the *crp*$^-$ and *cya*$^-$ strains all constructs showed a low β-galactosidase activity. The activity from pHMG1 was reduced 20- to 30-fold in the *crp*$^-$ and *cya*$^-$ mutants relative to the respective wild-type and also *papB* transcription was significantly decreased (cf. pPAP218 and pPAP226). Addition of 1 mM cAMP to the *cya*$^-$ strains restored the expression of β-galactosidase activity (Table IV).

A computerized search for potential cAMP-CRP binding sites in the *pap* regulatory region identified one sequence (5′-TT-TTTGA--T---TCA-ATT-3′) (Figure 2) in the *papI-papB* intercistronic region (bases 959 – 979) that showed strong homology to the consensus sequence (5′-AA-TGTGA--T---TCA-AT$^A/_T$-3′) of cAMP-CRP binding sites (Ebright *et al.*, 1984). Taken together, the data strongly suggest that cAMP-CRP activates the *papB* promoter and that this leads to stimulation of *papA* transcription.

Discussion

The formation of pili is presumably regulated both at the level of gene expression and at the level of protein-protein interaction in the secretion-assembly process. Here we have focused our attention on the transcriptional organization behind the *papA* pilin gene expression.

Our data showed that *in vivo* the *papA* pilin gene was preferentially transcribed as an 800 base long mRNA. This transcription was dependent on a regulatory region upstream from the start of the 800-base transcript. This region contained two cistrons denoted *papB* and *papI*. The *papB* transcription proceeded through the entire *papA* gene and presumably terminated at the same site as the *papA* transcript, whereas the poorly expressed *papI* gene was transcribed in the opposite direction.

The transcriptional activity of *papA* as measured by *pap-lacZ* fusions was 50 times greater in the presence of the *papI, papB* control region than in its absence. This activation required an intact *papB* gene and could be due to the PapB protein functioning as an activator for *papA* transcription. That *papB* mediates a *trans*-active product was evident from the complementation of the *papB* structural gene mutant (pHMG15) by co-resident plasmids expressing the PapB protein (Table II). However, when the *papI* gene and the entire *papB* promoter region were deleted (pHMG11), we were unable to demonstrate *trans*-activation of

Table II. Complementation test of *papB, papI* genes using *pap-lacZ* fusions

Plasmid	*pap* genotype	β-Galactosidase specific activity (units)			
		pHMG1 *papI*$^+$,*B*$^+$,*A-lacZ*	pHMG15 *papI*$^+$,*B1*,*A-lacZ*	pPAP218 *papB-lacZ*	pHMG11 *papA-lacZ*
pPAP234	*papI*$^+$	570	140	120	17
pPAP221	*papB*$^+$	400	450	110	14
pHMG72	*papI*$^+$,*papB*$^+$	920	740	190	16
pACYC184	Vector control	650	83	47	17

Cells were grown in Casa medium.

papA transcription. Several possible explanations for this apparent *cis*-dependent position effect of the *papB* gene on *papA* transcription could be considered. An important question is whether or not the two mRNA molecules detected with the *papA* and *papB* probes really were completely separate transcripts. If the 1300 nucleotide long *papB* transcript was endonucleolytically cleaved one could envisage how an 800 nucleotide long mRNA encoding only *papA* might be generated. In that case, transcription of *papA* would be *cis*-dependent upon the *papB* gene and promoter. Processing and differential degradation of specific parts of polycistronic mRNA molecules seem to play a regulatory role in some bacterial operons (Saito and Richardson, 1981; Burton *et al.*, 1983; Schmeissner *et al.*, 1984; Belasco *et al.*, 1985). Since the *papA* transcript, which was encoded more distally to the *papB* promoter, was also more abundant (Figure 3), a processing mechanism would have to involve preferential degradation of the 5′ portion of the mRNA. The *papB* gene product would then act as a positive effector by stabilizing the *papA*-encoding mRNA product.

An alternative explanation for the *cis* dependence could be that low level transcription from the *papB* promoter alters the DNA conformation in the region upstream of *papA*. If the *papB* protein can only interact with DNA of such an altered conformation the activation of a putative *papA* promoter would depend upon the *papB* gene in *cis*. Another possibility, for which there may be precedence, is that the PapB protein can interact with two separate DNA regions and that an interaction between these sites is critical for activation of *papA* transcription. It was recently proposed that repression of the *araBAD* promoter by the AraC repressor depends on binding to two sites separated by >170 bp (Dunn *et al.*, 1984), and a similar mechanism has been proposed for the *gal* operon (Irani *et al.*, 1983; Majumdar and Adhya, 1984).

The *lacZ* fusion data indicated that both the *papI* and *papB* genes have a positive effect on *papB* transcription. It was possible to increase transcription of a *papB-lacZ* fusion by expressing the PapI and/or PapB proteins in *trans*. This suggests that these proteins might bind to the intercistronic region between *papI* and *papB*, and act as positive regulators. However, the exact role of the two proteins in terms of *papB* gene transcription remains to be elucidated. It is also evident from the observed stimulatory differences between the *papI*$^+$,B$^+$ plasmid and the plasmids carrying either *papB*$^+$ or *papI*$^+$ (Table II) that we need to know more about the sites of action of the two proteins.

The presence of glucose in the medium reduced the expression of both transcripts through *papB* and *papA* (Figure 3). The experiments with *crp* and *cya* mutant strains (Tables III and IV) showed that the *papB* promoter is activated by cAMP-CRP. In addition, *papA* transcription is activated by cAMP-CRP as a consequence of the above discussed *papB* dependence. This

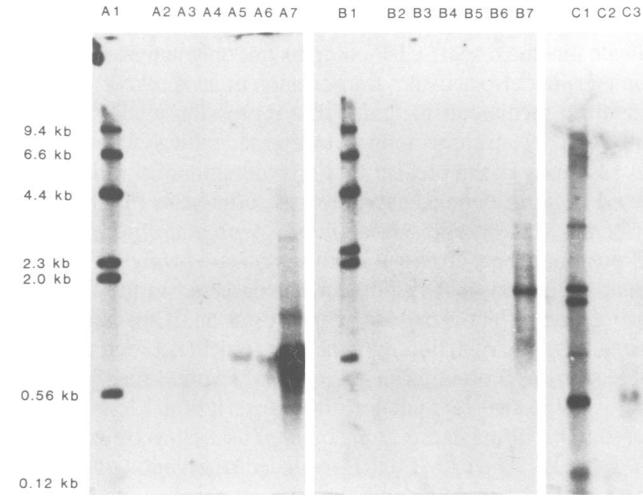

Fig. 3. Northern blot hybridizations using *papA* (**lanes A2 – A7**), *papB* (**lanes B2 – B7**) or *papI* (**lanes C2 – C3**) probes complementary to the sense strand. **Lanes A2 – A4, B2 – B4** and **C2** are hybridizations to RNA prepared from strain HB101 harbouring pBR322, whereas **A5 – A7, B5 – B7** and **C3** are hybridizations to RNA extracted from cells carrying pPAP5. In **lanes A2, A5, B2** and **B5** cells were grown in LB medium; **lanes A3, A6, B3** and **B6** represent cells grown in Casa medium in the presence of 1% glucose, whereas RNA from cells grown in glucose-free Casa medium was used in **lanes A4, A7, B4, B7** and **C3**. The mobility of ^{32}P-labelled single-stranded *Hin*dIII λDNA is shown in **lanes A1, B1** and **C1** and the molecular size of each fragment is shown to the left of **lane A1.**

Table III. Expression of β-galactosidase specific activity in *crp*$^{+/-}$ isogenic pairs of strains harbouring *pap-lacZ* transcriptional fusions

Plasmid	*pap* genotype	β-Galactosidase specific activity (units)	
		M182 *crp*$^+$	M182* *crp*$^-$
pHMG1	*papI*$^+$,*B*$^+$,*A-lacZ*	620	17
pPAP218	*papB-lacZ*	49	15
pRZ5202	Vector control	10	10

Samples were grown in Casa medium supplemented with 0.2% glucose.

regulatory circuit therefore provides a molecular explanation for the effect of glucose on pili formation. The nucleotide sequencing provided evidence for a binding site, although it remains to be directly demonstrated that a cAMP-CRP complex does interact with DNA in the promoter proximal region of *papB*. The orientation of such binding sites is independent of the direction of transcription and the distance to the transcription initiation site varies from 36 to 106 nucleotides in the different systems (de Crombrugghe *et al.*, 1984). Interestingly, the postulated *papI* promoter is located ~110 nucleotides away in the other direction

Table IV. Expression of β-galactosidase-specific activity in *cya*$^{+/-}$ isogenic pairs of strains harbouring *pap-lacZ* transcriptional fusions

Plasmid	*pap* genotype	β-Galactosidase specific activity (units)		
		CSH50 *cya*$^+$ −cAMP	VL391 *cya*$^-$ −cAMP	VL391 *cya*$^-$ +cAMP
pHMG1	*papI*$^+$,*B*$^+$,*A-lacZ*	870	40	490
pPAP226	*papI*$^+$,*B-lacZ*	140	41	135
pPAP218	*papB-lacZ*	67	34	71
pRZ5202	Vector control	7	23	20

Cells were grown in Casa medium containing 0.01% glucose (required for growth of VL391) without or with 1 mM cAMP.

from this site (Figure 3). Hence, it is possible that cAMP-CRP acts as an activator also for the *papI* promoter. If so, this would indicate that the cAMP-CRP complex not only activates the *papB* promoter, but also activates transcription of an activator for *papB*. Growth in media causing catabolite repression would therefore multiply affect transcription determining pilin synthesis.

A summary of the present findings on regulatory functions involved in *papA* gene transcription is outlined in Figure 4. The horizontal arrows representing *papA*, *papB* and *papI* transcripts are positioned according to the postulated promoter and terminator sequences (cf. Figure 2). As depicted in the figure the *papB* gene product is activating transcription of the *papA* gene. The exact mode of action remains to be elucidated but, as shown above, the PapB protein can act in *trans*. Expression of the *papB* gene itself is also regulated at the transcriptional level and appeared to be stimulated in *trans* by *papI* as well as by *papB* itself (cf. Table II). The cAMP-CRP-mediated stimulation of *papB* expression is presumably acting through the postulated binding site located upstream of the *papB* promoter. The present identification of a positive regulation of Pap pili synthesis will promote a characterization of the mode of action at the molecular level.

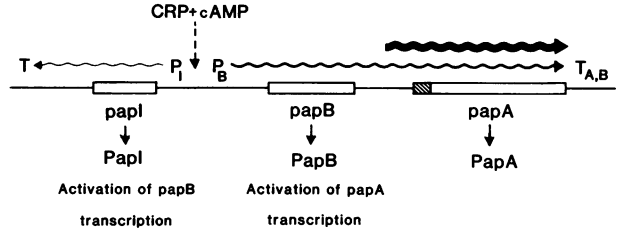

Fig. 4. Summary of transcriptional organization and regulatory features in the *papA* regulatory region. The horizontal wavy arrows indicate length and origins of the transcripts detected (see Results). The letters P and T show proposed promoters and terminators. The dashed vertical arrow indicates the postulated binding site for the cAMP-CRP complex.

Materials and methods

Bacterial strains and culture conditions

The *E. coli* strains used in this study were: HB101 (Boyer and Roulland-Dussoix, 1969); JM103 (Messing *et al.*, 1981); MC1029 (Casadaban and Cohen, 1980); M182 (Casadaban and Cohen, 1980) and an otherwise isogenic *crp*$^-$ derivative M182* (Busby *et al.*, 1983); CSH50 (Miller, 1972) and an isogenic *cya*$^-$ derivative VL391 (Eisenstein and Dodd, 1982). Cells were grown aerobically at 37°C. L-broth was prepared as described by Bertani (1951), supplemented with medium E (Vogel and Bonner, 1956). Casa medium consisted of medium E supplemented with 1 μg/ml thiamine, and 1.5% casamino acids (Difco Laboratories). Where indicated, glucose was added. For growth on plates, all media were solidified with 1.5% Bacto-Agar (Difco). The Lac phenotype was monitored on MacConkey lactose indicator agar plates (Miller, 1972) or on glycerol minimal plates containing casamino acids and 40 μg/ml 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal) (Sigma). Concentrations of antibiotics used for selection were: 100 μg/ml for carbenicillin, 10 μg/ml for tetracycline, 30 μg/ml for chloramphenicol and 100 μg/ml for streptomycin.

DNA techniques

Standard procedures were used for plasmid isolation, agarose gel electrophoresis, isolation of DNA fragments from polyacrylamide gels and transformation of cells (Maniatis *et al.*, 1982). The use of restriction endonucleases, *Bal*31 nuclease, Klenow fragment of DNA polymerase I and T4 DNA ligase were as recommended by the manufacturers (New England Biolabs and Boehringer Mannheim).

Construction of recombinant plasmids

Plasmid pRHU845 and its Tn*5* derivatives, and pPAP5 have been characterized before (Norgren *et al.*, 1984; Normark *et al.*, 1983; Lindberg *et al.*, 1984). Both pRHU845 and pPAP5 carry the same complete *pap* gene cluster. A derivative of the Tn*5* insertion mutant pSN025 was obtained by restriction with *Xho*I which cleaves only within the transposable element, followed by *Bal*31 nuclease treatment, ligation and transformation into strain HB101. In the resulting plasmid, pHMG73, ~200 bp of the Tn*5* DNA remains within the *papB* gene and the mutant allele was denoted *papB1*. Except for the *papB1* mutation, plasmid pHMG73 has the same physical structure as pRHU845. The 1.95-kb *Eco*RI$_1$-*Hind*III$_1$ fragment from plasmid pPAP5 was ligated with cleaved pBR322 (Bolivar *et al.*, 1977), and the resulting recombinant plasmid was named pPAP211. Construction pHMG72 carries the 1.7-kb *Eco*RI$_1$-*Eco*RV$_1$ fragment of pPAP5 inserted into the *Eco*RI, *Pvu*II sites of the plasmid vector pACYC184 (Chang and Cohen, 1978). Plasmid pPAP221 was obtained by cloning the 0.78-kb *Taq*I$_1$-*Taq*I$_2$ fragment of pPAP211 into the *Cla*I-site of pACYC184. The *papB* gene is transcribed in the opposite direction from the tetracycline genes in this construct. After isolation of the 0.88-kb *Sau*3A$_2$-*Sau*3A$_3$ fragment (base pairs 147 − 1030 from the *Eco*RI$_1$ site) of pPAP211, the fragment was cloned into the *Bam*HI site of the vector pACYC184. In the resulting plasmid, pPAP234, the *papI* gene is inserted in the anti-sense orientation with respect to the tetracycline promoter.

All transcriptional fusions between *pap* DNA and a promoterless *lacZ* gene were made by cloning into the vector pRZ5202 (L.Munson and W.S.Reznikoff, unpublished data) followed by transformation into the Lac$^-$ strain MC1029. The Lac fusion pHMG15 was constructed by cloning the 2.15-kb *Eco*RI$_1$-*Hind*III$_1$ fragment of pHMG73 into pRZ5202. The constructs pHMG9 and pHMG1 were obtained by cloning the 2.8-kb *Eco*RI$_1$-*Sma*I$_1$ and the 1.95-kb *Eco*RI$_1$-*Hind*III$_1$ fragments, respectively, from pPAP5. Plasmid pPAP225 has the 1.5-kb *Bal*I$_1$-*Hind*III$_1$ fragment from pPAP211 inserted into the *Sma*I, *Hind*III sites of the vector. For construction of the fusion plasmid pHMG11, a plasmid lacking the central 4.1-kb *Hind*III *pap* DNA fragment (pPAP208, not shown) was restricted with *Apa*I followed by *Bal*31 nuclease treatment and insertion of an *Xho*I linker, and transformation. One deletion derivative obtained was digested with *Xho*I and *Hind*III and the resulting 0.64-kb fragment with one end in *papB* and the other end in *papA* was cloned into the *Sal*I, *Hind*III sites of pRZ5202. To obtain the plasmid pPAP218, the 0.45-kb *Taq*I$_1$-*Rsa*I$_4$ fragment of pPAP211 was first cloned into the *Acc*I, *Sma*I sites of the M13mp9 vector. After restriction with the endonucleases *Pst*I and *Eco*RI, which have recognition sites on both sides of the inserted *Taq*I-*Rsa*I fragment, the isolated 0.45-kb *Pst*I-*Eco*RI fragment was made blunt ended by using the Klenow fragment of DNA polymerase I and thereafter ligated with *Sma*I-restricted pRZ5202. The resulting plasmid pPAP218 contains a single *Apa*I site within the inserted *Taq*I$_1$-*Rsa*I$_4$ fragment. This made it possible to construct pPAP226 by ligating the 1.22-kb *Eco*RI$_1$-*Apa*I$_1$ fragment from pPAP211 into pPAP218 digested with *Eco*RI and *Apa*I.

DNA sequence analysis

Appropriate restriction fragments were ligated into M13mp8 and M13mp9 (Messing and Vieira, 1982) vectors and transformed into the host strain JM103. Single-stranded DNA was prepared as described by Sanger *et al.* (1980) and sequenced by using the dideoxy chain-termination method of Sanger *et al.* (1977).

β-Galactosidase assay

The specific activity of β-galactosidase was assayed as described by Miller (1972).

Each value given represents the average activity obtained from at least three independent experiments. *o*-Nitrophenyl-β-D-galactoside (ONPG) and cAMP were from Sigma.

Labelling of hybridization probes and λHindIII DNA

In Figure 1c we have outlined the DNA regions cloned into M13mp8 and M13mp9 vectors, respectively, and subsequently used for making strand-specific probes used in Northern hybridization experiments. Partially double-stranded, labelled probes were prepared (Hu and Messing, 1982; Ricca *et al.*, 1982) using [α-^{32}P]dGTP and a hybridization probe primer from New England Biolabs. The specific activity of the probes was 1×10^8 c.p.m./μg DNA. λ*Hind*III DNA was labelled with [α-^{32}P]dGTP by filling in recessed 3′ ends using the Klenow fragment of DNA polymerase I.

Isolation of mRNA and Northern hybridization analysis

E. coli HB101 cells harbouring the plasmid pPAP5 or the plasmid pBR322 were grown at 37°C in LB-medium or Casa medium in the presence or absence of 1% glucose to a cell density of $\sim 2 \times 10^8$ cells/ml. RNA was extracted from the cells according to von Gabain *et al.* (1983).

The RNA was dissolved in sample buffer consisting of 50% formamide, 2.2 M formaldehyde and MOPS buffer (20 mM morpholinopropanesulfonic acid, 1 mM EDTA and 5 mM sodium acetate), heated at 60°C for 5 min, followed by a quick cooling and addition of 1/6 volume of dye mix (0.1% xylenecyanol, 0.1% bromphenol blue, 10 mM EDTA and 95% formamide). ^{32}P-Labelled λ*Hind*III DNA was denatured by heating in sample buffer at 95°C for 3 min, quick cooling on ice followed by addition of dye mix. 10 μg (*papA* and *papB* analyses) or 40 μg (*papI* analysis) of RNA samples were loaded per slot on a 1% agarose, 2.2 M formaldehyde slab gel. After electrophoresis, the gel was washed in water for 30 min and subsequently the RNA was transferred to nitrocellulose paper (Schleicher and Schüll, BA85) as described by Thomas (1980). The filters were washed in 3 × standard saline citrate (1 × SSC = 150 mM NaCl, 15 mM Na-citrate pH 7.0), dried and baked *in vacuo* for 2 h at 80°C. Pre-hybridization was performed for 5 – 10 h at 42°C in 50% formamide, 2 × Denhardt's (Denhardt, 1966), 5 × SSC, 0.2% SDS and 500 μg/ml denatured calf thymus DNA. Hybridization was carried out for 20 – 30 h under the same conditions, except that 1 × Denhardt's was used in the presence of 10^6 c.p.m./ml probe. Filters were washed at room temperature for 20 min with four changes of 2 × SSC, 0.5% SDS and then at 45°C with two changes of 0.1 × SSC, 0.5% SDS. Dried fiters were exposed to X-ray films at −80°C for 4 – 14 days using intensifying screens.

Acknowledgements

We would like to thank Arne Olsén and Eva Skogman for excellent technical assistance, and Drs. S.Busby, M.J.Casadaban, B.I.Eisenstein and W.S.Reznikoff for providing strains and plasmids. This work was supported by grants from the Swedish Medical Research Council (Projects 5428 and B85-16P-6893-02B), the Swedish Natural Science Research Council (Projects BU 3373-109 and BU 1670-100) and the Board for Technological Development (Project 81-3384).

References

Båga,M., Normark,S., Hardy,J., O'Hanley,P., Lark,D., Olsson,O., Schoolnik,G. and Falkow,S. (1984) *J. Bacteriol.*, **157**, 330-333.
Beachey,E.H. (1981) *J. Infect. Dis.*, **143**, 325-345.
Belasco,J.G., Beatty,J.T., Adams,C.W., von Gabain,A. and Cohen,S.N. (1985) *Cell*, **40**, 171-181.
Bertani,G. (1951) *J. Bacteriol.*, **62**, 293-300.
Bolivar,F., Rodriques,R.L., Greene,P.J., Betlach,M.C., Heyneker,H.L. and Boyer,H.W. (1977) *Gene*, **2**, 95-113.
Boyer,H.W. and Roulland-Dussoix,D. (1969) *J. Mol. Biol.*, **41**, 459-472.
Burton,Z.F., Gross,C.A., Watanabe,K.K. and Burgess,R.R. (1983) *Cell*, **32**, 335-349.
Busby,S., Kotlarz,D. and Buc,H. (1983) *J. Mol. Biol.*, **167**, 259-274.
Casadaban,M.J. and Cohen,S.N. (1980) *J. Mol. Biol.*, **138**, 179-207.
Chang,A.C.Y. and Cohen,S.N. (1978) *J. Bacteriol.*, **134**, 1141-1156.
de Crombrugghe,B., Busby,S. and Buc,H. (1984) *Science*, **224**, 831-838.
de Graaf,F.K., Wientjes,F.B. and Klaasen-Boor,P. (1980) *Infect. Immun.*, **27**, 216-221.
Denhardt,D.T. (1966) *Biochem. Biophys. Res. Commun.*, **23**, 641-652.
Dunn,T.M., Hahn,S., Ogden,S. and Schleif,R.F. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 5017-5020.
Ebright,R.H., Cossart,P., Gicquel-Sanzey,B. and Beckwith,J. (1984) *Nature*, **311**, 232-235.
Eisenstein,B.I. and Dodd,D.C. (1982) *J. Bacteriol.*, **151**, 1560-1567.
Gaastra,W. and de Graaf,F.K. (1982) *Microbiol. Rev.*, **46**, 129-161.
Göransson,M. and Uhlin,B.E. (1984) *EMBO J.*, **3**, 2885-2888.
Hu,N. and Messing,J. (1982) *Gene*, **17**, 271-277.
Irani,M.H., Orosz,L. and Adhya,S. (1983) *Cell*, **32**, 783-788.
Isaacson,R.E. (1980) *Infect. Immun.*, **28**, 190-194.
Källenius,G., Möllby,R., Svenson,S.B., Winberg,J., Lundblad,A., Svensson,S. and Cedergren,B. (1980) *FEMS Microbiol. Lett.*, **7**, 297-302.
Leffler,H. and Svanborg Edén,C. (1980) *FEMS Microbiol. Lett.*, **8**, 127-134.
Lindberg,F.P., Lund,B. and Normark,S. (1984) *EMBO J.*, **3**, 1167-1173.
Majumdar,A. and Adhya,S. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 6100-6104.
Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning. A Laboratory Manual*, published by Cold Spring Harbor Laboratory Press, NY.
Messing,J. and Vieira,J. (1982) *Gene*, **19**, 269-276.
Messing,J., Crea,R. and Seeburg,P.H. (1981) *Nucleic Acids Res.*, **9**, 309-321.
Miller,J.H. (1972) *Experiments in Molecular Genetics*, published by Cold Spring Harbor Laboratory Press, NY.
Norgren,M., Normark,S., Lark,D., O'Hanley,P., Schoolnik,G., Falkow,S., Svanborg-Edén,C., Båga,M. and Uhlin,B.E. (1984) *EMBO J.*, **3**, 1159-1165.
Normark,S., Lark,D., Hull,R., Norgren,M., Båga,M., O'Hanley,P., Schoolnik,G. and Falkow,S. (1983) *Infect. Immun.*, **41**, 942-949.
Normark,S., Båga,M., Göransson,M., Lindberg,F.P., Lund,B., Norgren,M. and Uhlin,B.E. (1985) in Mirelman,D. (ed.), *Microbial Lectins and Agglutinins: Properties and Biological Activity*, John Wiley and Sons, Inc., NY, in press.
Old,D.C. and Duguid,J.P. (1970) *J. Bacteriol.*, **103**, 447-456.
Ricca,G.A., Taylor,J.M. and Kalinyak,J.E. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 724-728.
Rosenberg,M. and Court,D. (1979) *Annu. Rev. Genet.*, **13**, 319-353.
Saier,M.H., Schmidt,M.R. and Leibowitz,M. (1978) *J. Bacteriol.*, **134**, 356-358.
Saito,H. and Richardson,C.C. (1981) *Cell*, **27**, 533-542.
Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463-5467.
Sanger,F., Coulson,A.R., Barrell,B.G., Smith,A.J.H. and Roe,B.A. (1980) *J. Mol. Biol.*, **143**, 161-178.
Schmeissner,U., McKenney,K., Rosenberg,M. and Court,D. (1984) *J. Mol. Biol.*, **176**, 39-53.
Svanborg Edén,C. and Hansson,H.A. (1978) *Infect. Immun.*, **21**, 229-237.
Thomas,P.S. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 5201-5205.
Tinoco,I.,Jr., Borer,P.N., Dengler,B., Levine,M.D., Uhlenbeck,O.C., Crothers,D.M. and Gralla,J. (1973) *Nature New Biol.*, **246**, 40-41.
Uhlin,B.E., Båga,M., Göransson,M., Lindberg,F.P., Lund,B., Norgren,M. and Normark,S. (1985a) *Curr. Top. Microbiol. Immunol.*, **118**, 163-178.
Uhlin,B.E., Norgren,M., Båga,M. and Normark,S. (1985b) *Proc. Natl. Acad. Sci. USA*, **82**, 1800-1804.
Ullmann,A. and Danchin,A. (1983) *Adv. Cyclic Nucleotide Res.*, **15**, 1-53.
Vogel,H.J. and Bonner,D.M. (1956) *J. Biol. Chem.*, **218**, 97-106.
von Gabain,A., Belasco,J.G., Schottel,J.L., Chang,A.C.Y. and Cohen,S.N. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 653-657.

Received on 26 July 1985; revised on 22 October 1985