Molecular cloning, expression and nucleotide sequence of the *rcsA* gene of *Erwinia amylovora*, encoding a positive regulator of capsule expression: evidence for a family of related capsule activator proteins

MARK COLEMAN,¹ ROWAN PEARCE,¹ ED HITCHIN,² FRANCIS BUSFIELD,¹ JOHN W. MANSFIELD² and IAN S. ROBERTS^{1*}

¹Department of Microbiology, Medical Sciences Building, University of Leicester, Leicester LE1 9HN, UK

²Biochemistry and Biological Sciences Department, Wye College (University of London), Ashford, Kent TN25 5AH, UK

(Received 1 May 1990; accepted 6 June 1990)

A gene encoding a positive activator of the expression of extracellular polysaccharide (EPS) synthesis in the phytopathogen *Erwinia amylovora* has been isolated from a genomic library in *Escherichia coli*. The presence of the cloned gene in *E. coli* stimulated transcription of the genes encoding colanic acid biosynthesis and could complement *rcsA* mutations. Introduction of the gene on a multicopy plasmid into *Er. amylovora* caused a threefold increase in EPS expression. The nucleotide sequence of the gene (designated *rcsA*) was determined. This revealed a single open reading frame encoding an RcsA protein of 23·7 kDa. This was confirmed by minicell analysis in *E. coli*. The predicted amino acid sequence of this RcsA protein showed a high degree of homology to the RcsA protein of *Klebsiella aerogenes*, demonstrating the existence of a family of related RcsA activator proteins capable of stimulating EPS expression. The protein had no significant homology to known DNA-binding activator proteins, indicating, for the first time, that the RcsA family of activator proteins may stimulate expression of EPS synthesis indirectly by acting on other regulatory proteins.

Introduction

The production of bacterial extracellular polysaccharide (EPS) has long been implicated in the virulence of both animal and plant pathogens (Costerton *et al.*, 1981; Ferris & Beveridge, 1985). The precise role of EPS in these systems is still far from clear, although there is considerable evidence that EPS plays a vital role in mediating bacterial–host interactions (Thomashow *et al.*, 1987). The synthesis of EPS is associated with the virulence of phytopathogenic *Erwinia* species (Bradshaw-Rouse *et al.*, 1981), in particular in *Erwinia amylovora*, the causative organism of apple and pear fireblight (Ayers *et al.*, 1979; Bennet, 1980; Steinberger & Beer, 1988). Although EPS from *Er. amylovora* causes wilting in excised cotoneaster and apple shoots the precise part played by the polysaccharide in the fireblight disease process is still unclear.

In contrast to other *Enterobacteriaceae* little is yet

known about the control of EPS production in *Er. amylovora*. In *Escherichia coli* K12 the biosynthesis of extracellular slime, colanic acid, is encoded by an operon of six genes, *cpsA–F* (Trisler & Gottesman, 1984). The control of expression of these genes is complex, being regulated by both the *lon* gene and three other regulatory genes termed *rcsA*, *rcsB* and *rscC* (Gottesman *et al.*, 1985; Brill *et al.*, 1988). The RcsA protein acts as an activator of *cps* expression and is sensitive to the Lon protease. The increased stability of the RcsA protein in a *lon* mutant is likely to account for the overproduction of colanic acid in these cells (Torres-Cabassa & Gottesman, 1987). The *rscB* and *rscC* genes have been proposed to constitute a two-component regulatory system (Stout & Gottesman, 1990) in which the RcsC protein acts as sensor of environmental stimuli with the RcsB protein acting as an effector to stimulate colanic acid biosynthesis. The interplay between this system and the regulation by the RcsA protein has yet to be determined. In *Erwinia stewartii*, the causative organism of leaf blight and vascular wilt in maize (Braun, 1982), the control of EPS expression seems in part to be regulated by a positive-acting factor analogous to the *rscA* gene of *E. coli* (Torres-Cabassa *et al.*, 1987).

Abbreviations: EPS, extracellular polysaccharide; ORF, open reading frame.

The nucleotide sequence reported in this paper has been submitted to GenBank and assigned the accession number M33571.

The *rscA* gene of *Klebsiella aerogenes* K21 has been cloned and sequenced (Allen *et al.*, 1987). The gene encodes a protein of 23 kDa capable of stimulating colanic acid biosynthesis in *E. coli*. Two additional activators of colanic acid biosynthesis in *E. coli* termed *rmpA* and *rmpB* have been cloned from *Klebsiella pneumoniae* (Nassif, *et al.* 1989), but their relationship to the *rscA* gene is as yet unclear.

In this paper we report the cloning, expression and nucleotide sequence of an activator of EPS expression of *Er. amylovora*. The gene was capable of stimulating colanic acid synthesis in *E. coli* and complementing *rscA* mutations. The predicted amino acid sequence of the *Erwinia* RcsA protein revealed a polypeptide highly homologous to the published sequence of the *K. aerogenes* RcsA gene product and indicated that RcsA proteins may stimulate transcription of the *cps* genes indirectly, possibly by acting on other regulatory proteins.

Methods

Bacterial strains and plasmids. The bacterial strains and plasmids used in this study and their relevant properties are shown in Table 1. All bacteria were grown in L broth, supplemented with 100 µg ampicillin ml⁻¹, 25 µg chloramphenicol ml⁻¹ or 20 µg tetracycline ml⁻¹ where appropriate. Triparental filter matings were performed as described by Franklin *et al.* (1981).

DNA procedures. *Er. amylovora* chromosomal DNA was extracted essentially as described by Saito & Miura (1963), except that prior to lysis the cells were washed twice in TE buffer (10 mM-Tris, 1 mM-EDTA, pH 8.0) to remove excess EPS. Plasmid DNA was purified by the method of Clewell & Helinski (1969), or for small-scale rapid purification, by the method of Birnboim & Doly (1979). Restriction endonucleases and DNA ligase were purchased from either BRL or New England Biolabs and used according to the manufacturer's

recommendations. The cosmid genomic library was constructed in the broad-host-range vector pLAFR3 (Staskawicz *et al.*, 1987). Restriction maps were constructed by performing single, double and triple digests directly on the cosmid clones or on subcloned fragments thereof.

DNA-DNA hybridization. Restriction-endonuclease-generated DNA fragments were transferred to Hybond-N (Amersham) as described previously (Roberts *et al.*, 1986). Radiolabelled probes were generated by labelling fragments directly in molten agarose gel slices by extending random hexadeoxynucleotide primers with the Klenow fragment of DNA polymerase I in the presence of [³²P]dCTP (Feinberg & Vogelstein, 1983). Filters were washed in 1 × SSC (150 mM-NaCl, 150 mM-sodium citrate) for 60 min at 65 °C.

Nucleotide sequence analysis. The sequencing strategy involved cloning the 1.0 kb *Bam*HI-*Pst*I fragment of the *Er. amylovora* genome containing the *rscA* gene into both M13mp18 and M13mp19 (Yanisch-Perron *et al.*, 1985). Single-stranded M13 DNA templates were sequenced by the dideoxy-chain-termination method (Sanger *et al.*, 1977), with [³⁵S]thio-dATP and Sequenase (Cambridge Biosciences) and DNA fragments analysed using buffer gradient gels (Biggin *et al.*, 1983). The complete nucleotide sequence on both strands of DNA was determined by using complementary oligonucleotides generated against the existing sequence as sequencing primers. Nucleotide sequences were analysed with the Staden computer programs (Staden, 1984) and a suite of programs assembled by the University of Wisconsin.

Polysaccharide determination. Polysaccharide was extracted from the appropriate *Er. amylovora* cultures as described by Jann (1985). The concentration of carbohydrate in each polysaccharide preparation was determined by the phenol/sulphuric acid method (Dubois *et al.*, 1956) and expressed in terms of galactose equivalents. The mean of triplicate samples was determined and the standard error (SEM) calculated.

β-Galactosidase assays. β-Galactosidase expression in *lac*-fusion strains was visualized on L agar supplemented with the chromogenic dye X-Gal (Sigma). β-Galactosidase was assayed as described by Casaregola *et al.* (1982).

Minicell analysis. Minicells were prepared and labelled as described by Wilkins *et al.* (1982). Labelled products were analysed by sodium dodecyl sulphate-polyacrylamide gel electrophoresis and fluorography (Wilkins *et al.*, 1982).

Table 1. *Bacterial strains and plasmids*

Strain or plasmid	Relevant genotype or properties*	Source or reference
<i>E. coli</i>		
ED8767	<i>supE supF hsdS⁻ M⁺ R⁺ met trpR recA</i>	W. J. Brammar (Leicester University)
DS410	<i>minA minB</i>	Wilkins <i>et al.</i> (1982)
JM101	<i>supE thi Δ(lac-proAB) F'</i> [<i>traD36 proA⁺ proB⁺ lacI^q</i>]	J. Messing (Minnesota University)
SG20799	<i>rscA104 cps::lac</i>	S. Gottesman (National Cancer Institute, Bethesda)
SG20688	<i>rscB15 cps::lac</i>	S. Gottesman
<i>Er. amylovora</i>		
OT1	Prototroph	Bennett & Billing (1978)
Plasmids		
pLAFR3	Tc ^r <i>cos</i>	M. Daniels (John Innes Institute)
pACYC184	Cm ^r Tc ^r p15A replicon	Chang & Cohen (1978)
pUC19	Ap ^r	Yanisch-Perron <i>et al.</i> (1985)
pGSS33	Ap ^r Cm ^r Tc ^r Sm ^r	Sharpe (1984)

* Ap^r, ampicillin resistance; Cm^r, chloramphenicol resistance; Sm^r, streptomycin resistance; Tc^r, tetracycline resistance.

Results

Cloning of the *rscA* gene of *Er. amylovora*

A cosmid library of chromosomal DNA of *Er. amylovora* strain OT1 was constructed in the broad-host-range cosmid vector pLAFR3 (Staskawicz *et al.*, 1987) and introduced into *E. coli* strain ED8767. The library, of 3640 clones, was screened for visibly mucoid clones showing increased colanic acid biosynthesis. This screening procedure identified five mucoid clones. To determine which of these encoded an *rscA*-like function from *Er. amylovora* the cosmid clones were introduced by filter mating into *E. coli* strain SG20799 (*rscA104 cps::lac*). Complementation of the *rscA* mutation in this strain results in an increase in transcription of the *cps::lac* fusion and a Lac⁺ phenotype which can be identified on media supplemented with X-Gal. All five cosmid clones were able to complement the *rscA* mutation in SG20799 and gave rise to blue Lac⁺ colonies on media supplemented with X-Gal. These results were confirmed by β -galactosidase assays (Table 2). All five cosmid clones identified on X-Gal-supplemented agar clearly had increased transcription of the *cps* genes, resulting in a several-hundred-fold increase in the level of β -galactosidase (Table 2). Indeed, the cloned *rscA* gene of *Er. amylovora* increased transcription of the *E. coli cps* genes to a greater extent than the corresponding cloned *rscA* gene from *K. aerogenes* K21 (Table 2). All of the five

cosmid clones failed to complement strain SG20688 (*rscB15 cps::lac*), indicating that none of them encoded a functional RcsB protein.

Subcloning of the *rscA* gene of *Er. amylovora*

Restriction endonuclease analysis of the DNA of the five cosmids that contained the gene revealed a nested set of conserved DNA fragments indicating a common chromosomal origin (data not shown). Indeed two of the cosmids, p29B1 and p29B2 (Table 2) appeared identical and are most likely siblings from a single infection when the library was constructed. To isolate the *rscA* gene on a smaller fragment of DNA, cosmid p26 was cleaved with the restriction endonuclease *Hind*III and ligated to *Hind*III-cleaved vector pACYC184 (Chang & Cohen, 1978). The ligation mixture was used to transform strain SG20799 and *rscA*⁺ transformants were identified on agar supplemented with chloramphenicol and X-Gal. Of the 2000 chloramphenicol-resistant transformants, 20 were clearly Lac⁺, indicating the presence of the cloned *Er. amylovora rscA* gene. Restriction endonuclease analysis of the plasmid DNA indicated the presence of a cloned 3.4 kb *Hind*III fragment in all of the Lac⁺ transformants. One plasmid containing the cloned 3.4 kb *Hind*III fragment alone was studied further and a restriction endonuclease cleavage map constructed (Fig. 1). This plasmid was termed pISR401 and the presence of the subcloned *rscA* gene on the 3.4 kb *Hind*III fragment was confirmed by β -galactosidase assays of strain SG20799(pISR401) (Table 2). Since pISR401 contains two *Bam*HI restriction cleavage sites, one in the cloned 3.4 kb *Hind*III fragment and one in the pACYC-184 vector sequences, the *rscA* gene was localized on the 3.4 kb *Hind*III fragment by digestion of pISR401 with *Bam*HI followed by self-ligation and transformation of strain SG20799. Plasmid DNA was purified from 12 *rscA*⁺ transformants and digested with *Bam*HI. Nine of the transformants had a plasmid of 6.2 kb with a single *Bam*HI cleavage site whilst the remaining transformants contained pISR401. The deletion derivative of pISR401 was termed pISR402. Since the deletion of the 1.2 kb *Bam*HI fragment of pISR401 did not inactivate the *rscA* gene, this locates the *rscA* gene in its entirety on the 2.2 kb *Bam*HI-*Hind*III fragment (Fig. 1). This conclusion was confirmed by β -galactosidase assays of strain SG20799(pISR402) (Table 2). To localize the *rscA* gene further, pISR401 was digested with both *Bam*HI and *Pst*I and ligated to pUC19 (Yanisch-Perron *et al.*, 1985) cleaved with the same restriction endonucleases. The ligation mixture was used to transform *E. coli* JM101. Approximately 1500 ampicillin-resistant transformants were obtained, of which 300 were Lac⁻ and visibly mucoid. Plasmid DNA was purified from 12 such

Table 2. β -Galactosidase expression in strains SG20799 and SG20688 in the presence and absence of cloned *rscA* genes

Strain	β -Galactosidase activity [units (mg dry wt) ⁻¹]
SG20799	2
SG20799(p23)*	313
SG20799(p24)*	260
SG20799(p26)*	205
SG20799(p29b1)*	197
SG20799(p29b2)*	230
SG20799(pLV213)†	56
SG20799(pISR401)	228
SG20799(pISR402)	197
SG20799(pISR403)	223
SG20799(pISR404)	215
SG20799(pISR404::1)	3
SG20799(pISR404::2)	4
SG20688(p23)*	2
SG20688(p24)*	1
SG20688(p26)*	2
SG20688(p29b1)*	3
SG20688(p29b2)*	2
SG20688(pISR404)	3

* These numbers refer to the five original *rscA*⁺ cosmid clones.

† Plasmid pLV213 contains the *rscA* gene of *K. aerogenes* (Allen *et al.*, 1987).

Fig. 1. Restriction endonuclease map of the *rcsA*-encoding 3.4 kb *Hind*III fragment. The restriction enzyme sites are: B, *Bam*HI; H, *Hind*III; Hc, *Hinc*II; P, *Pst*I. The numbered vertical lines denote the sites of *Tn1000* insertion in pISR404::1 and pISR404::2.

Fig. 2. Southern blot using the radiolabelled 3.4 kb *Hind*III fragment as a probe to restriction endonuclease digests of *Er. amylovora* chromosomal DNA. Lane 1, *Pst*I; lane 2, *Bam*HI; lane 3, *Hind*III; lane 4, *Eco*RI.

Fig. 3. Polyacrylamide gel electrophoresis of [35 S]methionine-labelled polypeptides encoded by minicells harbouring: 1, pISR404::1; 2, pISR404::2; 3, pISR404; 4, pUC19.

colonies and digested with *Bam*HI and *Pst*I. All of the transformants had an identical plasmid of 3.7 kb, consisting of pUC19 and the subcloned 1.0 kb *Bam*HI–*Pst*I fragment of pISR401. One of these recombinants, termed pISR404, was used for further study. These experiments locate the *Er. amylovora rcsA* gene on this 1.0 kb fragment of DNA (Fig. 1). This was confirmed by β -galactosidase assays of strain SG20799(pISR404) (Table 2). Introduction of pISR404 into strain SG20688 failed to induce expression of the *cps*::*lac* fusion and gave rise to a Lac[−] phenotype (Table 2). This confirms the results obtained with the initial cosmid clones; notably that an *rcsB* gene has not been cloned and importantly that the cloned *Er. amylovora* RcsA protein requires a functional *E. coli* RcsB protein in order to increase transcription of the *cps* genes. A similar observation has been made for the *E. coli* RcsA protein (Brill *et al.*, 1988).

Southern blot analysis using the cloned *rcsA* gene

To confirm the origin of the cloned *rcsA* gene, Southern blot hybridizations were performed using the radiolabelled 3.4 kb *Hind*III fragment as a probe to *Er. amylovora* chromosomal DNA digested with various restriction endonucleases. This probe hybridized to a single chromosomal *Hind*III fragment of 3.4 kb and a single *Eco*RI fragment of 5.0 kb (Fig. 2). As predicted from the restriction enzyme cleavage map (Fig. 1) the probe hybridized to two *Bam*HI fragments, of 5.0 and 8.0 kb, and to two *Pst*I fragments, of 5.0 and 9.0 kb (Fig. 2). No homology to *E. coli* or *K. aerogenes* chromosomal DNA was detected (data not shown).

Identification of the RcsA protein

Minicells were prepared from *E. coli* strain DS410 carrying pUC19 and pISR404 and the plasmid-encoded polypeptides labelled with [35 S]methionine. The labelled products were analysed by sodium dodecyl sulphate-polyacrylamide gel electrophoresis and fluorography. The autoradiograph obtained from such an experiment (Fig. 3) clearly demonstrates that pISR404 encodes only one non-vector polypeptide of 25 kDa. This suggests that this polypeptide is the *Er. amylovora* RcsA protein. This was confirmed by analysis of minicells prepared from strain DS410 carrying *Tn1000* insertion mutants of pISR404 (see Fig. 1). Both pISR404::1 and pISR404::2 were unable to complement strain SG20799 (Table 2) and neither encoded the 25 kDa polypeptide. This confirms that the 25 kDa polypeptide is indeed the RcsA protein of *Er. amylovora*. Since no truncated polypeptides could be identified (Fig. 3) the direction of transcription of the *rcsA* gene could not be determined.

```

      10              30              50
CGTATTTTGA CTGCTAAATTACACAAACAGGTGTTACGGCGTTTCTTGCCATATCAAATC

      70              90              110
TCAAAC TGTCAGTTGGCAAAAAGCGCATTGTTACCTCGCTGATGATGACGCGTTTTCG

      130             150             170
TAACCTTGGGTAAATTATTACCAGCATGAAAACGGAATTTTGTCTCGTAGCGCAAATGCGT

      190             210             230
AATTCATCGTTGAATTTAACCTATAACAAATTGGTGAGGTCGCTGTTATGCCGACGATT
                                     MetProThrIle

      250             270             290
ATTATGGATTCATGCAATTATACACGGCTGGGTTTAACTGAATATATGACTGTCAAGGGA
IleMetAspSerCysAsnTyrThrArgLeuGlyLeuThrGluTyrMetThrValLysGly

      310             330             350
GTTAAAAAGAAAAATATATCCCTGATCAATGATATTGCACAGTTACAGAATAAATGCCAG
ValLysLysLysAsnIleSerLeuIleAsnAspIleAlaGlnLeuGlnAsnLysCysGln

      370             390             410
CAGCTGAAACCCGGCGTCGTGCTTATTAAACGAAGACTGCTTTATCCATGAATCTGACGCC
GlnLeuLysProGlyValValLeuIleAsnGluAspCysPheIleHisGluSerAspAla

      430             450             470
AGTGAGCGCATCAGGAAAATTATTCTTCAGCATCTGATACACTATTTTCATCTTTATG
SerGluArgIleArgLysIleIleLeuGlnHisProAspThrLeuPhePheIlePheMet

      490             510             530
GCCATCTCTAATATTCACTTTGAAGAATACCTGTATGTGCGTAACAACTGATTATTACC
AlaIleSerAsnIleHisPheGluGluTyrLeuTyrValArgAsnAsnLeuIleIleThr

      550             570             590
TCAAAGCGATCAAAATTTCTACGCTGGACTCCCTGCTCAACGGCTATTTTTCAGAAGAAA
SerLysAlaIleLysIleSerThrLeuAspSerLeuLeuAsnGlyTyrPheGlnLysLys

      610             630             650
TTAAACCTGTCTGTGCGTCATGGCACACATTGAGAGTCCATCCGTTGACGCTTAGCCAG
LeuAsnLeuSerValArgHisGlyThrHisSerGluValHisProLeuThrLeuSerGln

      670             690             710
ACCGAGTCAAACATGCTGAAAATGTGGATGTCTGCCATGATACCATTCAAATATCAGAC
ThrGluSerAsnMetLeuLysMetTrpMetSerGlyHisAspThrIleGlnIleSerAsp

      730             750             770
AAAATGCAAAATCAAAGCCAAAACGGTTTCTTCGCATAAAGGCAATTCAAACGCAAGATC
LysMetGlnIleLysAlaLysThrValSerSerHisLysGlyAsnIleLysArgLysIle

      790             810             830
AAAACCCATAACAAGCAGGTTATCTATCACGTTGTCCGCTGACCGATAATGTGACCAGC
LysThrHisAsnLysGlnValIleTyrHisValValArgLeuThrAspAsnValThrSer

      850             870
GGCATTATGTGTTAACGAAAGATAGGATCC
GlyIleTyrValAsnGluArg

```

Fig. 4. Nucleotide sequence of the *Er. amylovora rcaA* gene. The predicted amino acid sequence of the RcsA protein is indicated in three-letter code below the appropriate nucleotide bases. The putative ribosome-binding site is underlined at -13 bp as are the -10 and -35 regions at -95 bp and -120 bp respectively.

Reintroduction of the cloned *rcaA* gene into *Er. amylovora*

To assay the effect of the cloned *rcaA* gene on expression of EPS by *Er. amylovora* it was necessary to subclone the *rcaA* gene into a broad-host-range plasmid vector. Essentially this involved subcloning the 2.2 kb *Bam*HI-*Hind*III DNA fragment containing *rcaA* (Fig. 1) from

pISR402 into the broad-host-range plasmid pGSS33 (Sharpe, 1984) to yield a recombinant plasmid termed pISR403. The presence of the cloned *rcaA* gene was confirmed by β -galactosidase assays of strain SG20799(pISR403) (Table 2).

Plasmid pISR403 was introduced into *Er. amylovora* strain OT1 by filter mating. Colonies of recombinant *Er.*

10	20	30	40	50	60
MPTIIMDSCNYTRLGLTEYMTVKGVKKKNISLINDIAQLQNKCCQLKPGVVLINEDCFIH					
.....X::					
MSTMIMDLCSYTRLGLTGYLTSRGIKKQEIIEVNSAADLQKHCTSCCPAVVFLNEDCFVH					
10	20	30	40	50	60
70	80	90	100	110	120
ESDASERIRKIILQHPDTLFFIFMAISNIHFEEYLYVRNNLIITSKAIKISTLDSLLNGY					
..... :::::					X ..
DDESNGIIRQIITQNPAITLFVIFMSLANIHFDRLVRKNLLISSKSITPKDLVDVILVNY					
70	80	90	100	110	120
130	140	150	160	170	180
FQKKLNLNLSVRHGHSEVHPLTSLQTESNMLKMWMSGHDTIQISDKMQIKAKTVSSHKGNI					
..... :::::					
----LKYKNTSVGQLTLPLTSLSKTESNMLQMWAGHGTSQLSTQMNIAKAKTVSSHKGNI					
130	140	150	160	170	
190	200	210			
KRKIKTHNKQVIYHVRLTDNVTSGIYVNER					
..... :::::					
KKKIQTTHNKQVIYHIVRLTENITSGIQVNMR					
180	190	200			

Fig. 5. Comparison of the predicted amino acid sequences of the RcsA proteins of *Er. amylovora* and *K. aerogenes* K21. The amino acids are denoted by single-letter code, with the sequence of the *Er. amylovora* RcsA above and that of the *K. aerogenes* RcsA protein below. : denotes identical amino acids, and · denotes similar amino acids of the same family group (Lipman & Pearson, 1985). The dashes denote spaces inserted in the *K. aerogenes* RcsA protein to accommodate the four additional amino acids present in the *Er. amylovora* sequence. X denotes the boundaries of the largest single stretch of unbroken homology.

amylovora harbouring pISR403 were fluidal, producing visibly copious amounts of EPS. The amount of EPS produced by strains OT1 and OT1(pISR403) was determined by extracting the EPS in saline and assaying the total carbohydrate present by the phenol/sulphuric acid method. Since the EPS of *Er. amylovora* is a galactose-rich polymer (T. Smith, personal communication), the total carbohydrate was expressed in terms of galactose equivalents. The results were (μg galactose equivalents per 10^9 cells; mean $\pm$ SEM): OT1, 14.8 ± 6.0 ; OT1(pISR403), 51.1 ± 9.1 ; OT1(pGSS33), 12.6 ± 5.0 . Thus the presence of plasmid pISR403 in strain OT1 induced a greater than threefold increase in the levels of extractable EPS.

Nucleotide sequence of the *rcaA* gene.

The 1.0 kb *Pst*I–*Bam*HI restriction fragment containing the *Er. amylovora* *rcaA* gene was subcloned into the M13 vectors mp18 and mp19 and used as a starting point for the nucleotide sequence analysis. The initial sequence on both DNA strands was extended using oligonucleotide primers derived from the determined sequence. Analysis of the complete 1010 base pairs of sequence revealed a single open reading frame (ORF) of 633 base pairs encoding a putative protein of 211 amino acids and molecular mass 23.7 kDa (Fig. 4). This compares favourably with the molecular mass of the RcsA protein determined from the minicell analysis. Nine bases upstream of the first ATG of the ORF was a putative

ribosome-binding site GAGG, which is in good agreement with the consensus sequence (Shine & Dalgarno, 1974). A further 79 bases upstream is a potential '–10' sequence of TAATTA, whilst 19 bases upstream is the sequence ATGACA, which has very strong homology with the '–35' *E. coli* consensus sequence (Rosenberg & Court, 1979). Gravy score analysis (Kyte & Doolittle, 1982) of the predicted amino acid sequence gave a score of –0.01 and this, together with examination of the hydropathy plots (Kyte & Doolittle, 1982) (data not shown), is consistent with a cytosolic location. Comparison of the predicted amino acid sequences of the *Er. amylovora* and *K. aerogenes* K21 RcsA proteins reveals a high degree of homology, with over 56% identical amino acids (Fig. 5). Computer-aided searches (Lipman & Pearson, 1985) of the NBRF, SWISSPROT and NEW NBRF protein databases failed to identify significant homology to any other protein sequences. In particular, no homology was found to either the LysR family of bacterial activator proteins (Henikoff, *et al.*, 1988; Tao, *et al.*, 1989) or to Cro-like DNA-binding proteins (Dodd & Egan, 1987).

Discussion

The *rcaA* gene, encoding a positive activator of EPS expression in *Er. amylovora*, has been cloned and sequenced. The gene was capable of complementing the corresponding *rcaA* mutation in *E. coli* and stimulating

transcription of the *E. coli cps* genes. The *Er. amylovora rcsA* gene encoded a protein of 23.7 kDa, which is similar in size to the *E. coli* RcsA protein of 27 kDa (Torres-Cabassa & Gottesman, 1987), the RcsA protein of *Er. stewartii* of 25 kDa (Torres-Cabassa *et al.*, 1987) and the RcsA protein of *K. aerogenes* K21 of 23 kDa (Allen *et al.*, 1987). The similar size of these proteins combined with the finding that they are capable of activating EPS expression in heterologous hosts may indicate that they represent a conserved set of functionally related proteins. Indeed comparison of the predicted amino acid sequences of the RcsA proteins of *Er. amylovora* and *K. aerogenes* K21 revealed that these two proteins are highly homologous. It is likely that this homology extends to the RcsA protein of *E. coli* since it has been reported that there is extensive homology between the RcsA proteins of *E. coli* and *K. aerogenes* K21 (A. Torres-Cabassa, unpublished results). These results indicate the presence of conserved systems for the regulation of EPS synthesis in these micro-organisms. This in part may reflect a common mechanism for the management and utilization of surplus carbohydrates in the medium, since it is known that slime/EPS synthesis increases when the level of available carbohydrate is high (Linton *et al.*, 1988).

The presence of the cloned *rcsA* on a multi-copy plasmid in *Er. amylovora* induced a greater than threefold increase in the expression of EPS, clearly indicating a role for this gene in the control of EPS expression in *Er. amylovora*. The effect of this RcsA protein is likely to be to increase the transcription of the genes encoding enzymes involved in the biosynthesis of EPS in a manner analogous to that of the RcsA protein of *Er. stewartii* and *E. coli* (Torres-Cabassa *et al.*, 1987; Gottesman *et al.*, 1985), a conclusion supported by the observation that the cloned *Er. amylovora rcsA* gene increased transcription of the *cps* genes in *E. coli*. To determine the possible role of the RcsA protein in acting as a transcriptional activator, the predicted amino acid sequence of the *Er. amylovora* RcsA protein was compared to the consensus sequence of the LysR family of bacterial activator proteins. The LysR family consists of at least 10 different bacterial proteins known to act as transcriptional activators (Henikoff *et al.*, 1988; Tao *et al.*, 1989), with each protein having a conserved helix-turn-helix DNA binding motif near to its N-terminus. The observations that the *Er. amylovora* RcsA protein has no homology to this group of functionally related activator proteins and that no DNA-binding domains could be identified in the RcsA protein using the parameters of Dodd & Egan (1987) might indicate that the *Er. amylovora* RcsA protein stimulates transcription of the *cps* genes indirectly. Indeed, it is possible that the RcsA protein may mediate transcription of the *cps* genes

by modification of the RcsB protein, since the cloned *Er. amylovora rcsA* gene failed to induce transcription of the *cps* genes in an *E. coli rcsB* mutant (Table 2). Such a model has been proposed for the mode of action of the *E. coli* RcsA protein, albeit in the absence of any published amino acid sequence for this protein (Brill *et al.*, 1988). Clearly the determination of the predicted amino acid sequence of the functionally related *Er. amylovora* RcsA protein is the first evidence for such a role. In *Er. amylovora* little is yet known concerning the regulation of EPS expression, for example whether proteins analogous to the *E. coli* RcsB and RcsC exist and the possible mechanisms by which the *Er. amylovora* RcsA protein may mediate its activity.

The observation that *in planta* the expression of EPS by *Er. amylovora* is greatly increased (Bennet & Billing, 1978) may indicate a role for the RcsA protein in mediating interactions between *Er. amylovora* and the host plant and in the subsequent rapid turn-on of EPS expression. Indeed it is possible that the regulation of *rcsA* expression may be a key event in the initial stages of an *Er. amylovora* infection. Experiments are in progress to generate the necessary gene fusions in the *Er. amylovora rcsA* gene to test this hypothesis.

The authors gratefully acknowledge the kind gift of strains and advice from Dr E. Billing, Dr S. Gottesman, Dr J. R. Saunders and Professor C. A. Hart. This work was supported by a grant to I.S.R. from the Science and Engineering Research Council of the United Kingdom.

References

- ALLEN, P., HART, C. A. & SAUNDERS, J. R. (1987). Isolation from *Klebsiella* and characterization of two *rsc* genes that activate colanic acid capsule biosynthesis in *Escherichia coli*. *Journal of General Microbiology* **133**, 331–340.
- AYERS, A. F., AYERS, S. B. & GOODMAN, R. N. (1979). Extracellular polysaccharide of *Erwinia amylovora*: a correlation with virulence. *Applied and Environmental Microbiology* **38**, 659–666.
- BENNET, R. A. (1980). Evidence for two virulence determinants in the fireblight pathogen *Erwinia amylovora*. *Journal of General Microbiology* **116**, 351–356.
- BENNET, R. A. & BILLING, E. (1978). Capsulation and virulence in *Erwinia amylovora*. *Annals of Applied Biology* **89**, 41–45.
- BIGGIN, M. D., GIBSON, T. J. & HONG, G. F. (1983). Buffer gradient gels and ³⁵S label as an aid to rapid DNA sequence determination. *Proceedings of the National Academy of Sciences of the United States of America* **80**, 3963–3965.
- BIRNBOIM, H. C. & DOLY, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Research* **1**, 1513–1523.
- BRADSHAW-ROUSE, J. J., WHATLEY, M. A., COPLIN, D. L., WOODS, A., SEQUEIRA, L. & KELMAN, A. (1981). Agglutination of strains of *Erwinia stewartii* with corn agglutinin: correlation with extracellular polysaccharide production and pathogenicity. *Applied and Environmental Microbiology* **42**, 344–350.
- BRAUN, E. J. (1982). Ultrastructural investigation of resistant and susceptible maize inbreds infected with *Erwinia stewartii*. *Phytopathology* **72**, 159–166.
- BRILL, J. A., QUINLAN-WALSHE, C. & GOTTESMAN, S. (1988). Fine-structure mapping and identification of two regulators of capsule synthesis in *Escherichia coli* K-12. *Journal of Bacteriology* **170**, 2599–2611.

- CASAREGOLA, S., D'ARI, R. & HUISMAN, O. (1982). Quantitative evaluation of *recA* gene expression in *Escherichia coli*. *Molecular and General Genetics* **185**, 430–439.
- CHANG, A. C. Y. & COHEN, S. N. (1978). Construction and characterisation of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. *Journal of Bacteriology* **134**, 1141–1156.
- CLEWELL, D. B. & HELINSKI, D. R. (1969). Supercoiled circular DNA-protein complex in *Escherichia coli*: purification and induced conversion to an open circular DNA form. *Proceedings of the National Academy of Sciences of the United States of America* **62**, 1159–1166.
- COSTERTON, J. W., IRVIN, R. T. & CHENG, K. J. (1981). The role of bacterial surface structures in pathogenesis. *Critical Reviews in Microbiology* **8**, 303–338.
- DODD, I. B. & EGAN, J. B. (1987). Systematic method for the detection of potential λ Cro-like DNA-binding regions in proteins. *Journal of Molecular Biology* **194**, 557–564.
- DUBOIS, M., GILLES, K. A., HAMILTON, J. K., REBERS, P. A. & SMITH, F. (1956). Colorimetric method for the determination of sugars and related substances. *Analytical Chemistry* **28**, 350–356.
- FEINBERG, A. P. & VOGELSTEIN, B. (1983). A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Analytical Biochemistry* **132**, 6–13.
- FERRIS, F. G. & BEVERIDGE, T. J. (1985). Functions of bacterial cell surface structures. *Bioscience* **35**, 172–177.
- FRANKLIN, F. C. H., BAGDASARIAN, M., BAGDASARIAN, M. M. & TIMMIS, K. N. (1981). Molecular and functional analysis of the Tol plasmid pWW0 from *Pseudomonas putida* and cloning of the genes for the entire regulated aromatic ring meta cleavage pathway. *Proceedings of the National Academy of Sciences of the United States of America* **78**, 7458–7462.
- HENIKOFF, S., HAUGHN, G. W., CALVO, J. M. & WALLACE, J. C. (1988). A large family of bacterial activator proteins. *Proceedings of the National Academy of Sciences of the United States of America* **85**, 6602–6606.
- GOTTESMAN, S., TRISLER, P. & TORRES-CABASSA, A. (1985). Regulation of capsular polysaccharide synthesis in *Escherichia coli* K-12: characterisation of three regulatory genes. *Journal of Bacteriology* **162**, 1111–1119.
- JANN, K. (1985). Isolation and characterisation of capsular polysaccharides (K antigens) of *Escherichia coli*. In *The Virulence of E. coli*, pp. 373–379. Edited by M. Sussman. London: Academic Press.
- KYTE, J. & DOOLITTLE, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology* **157**, 105–132.
- LIPMAN, D. J. & PEARSON, W. R. (1985). Rapid and sensitive protein similarity searches. *Science* **227**, 1435–1441.
- LINTON, J. D., GOULDNEY, D. & WOODARD, S. (1988). Efficiency and stability of exopolysaccharide production from different carbon sources by *Erwinia herbicola*. *Journal of General Microbiology* **134**, 1913–1921.
- NASSIF, X., HONORE, N., VASELON, T., COLE, S. T. & SANSONETTI, P. J. (1989). Positive control of colanic acid synthesis in *Escherichia coli* by *rmpA* and *rmpB*, two virulence-plasmid genes of *Klebsiella pneumoniae*. *Molecular Microbiology* **3**, 1349–1359.
- ROBERTS, I. S., MOUNTFORD, R., HIGH, N., BITTER-SUERMAN, D., JANN, K., TIMMIS, K. N. & BOULNOIS, G. J. (1986). Molecular cloning and analysis of genes for production of the K5, K7, K12 and K92 capsular polysaccharides in *Escherichia coli*. *Journal of Bacteriology* **168**, 1228–1233.
- ROSENBERG, M. & COURT, D. (1979). Regulatory sequences involved in the promotion and termination of RNA transcription. *Annual Review of Genetics* **13**, 319–353.
- SAITO, H. & MIURA, K. I. (1963). Preparation of transforming deoxyribonucleic acid by phenol treatment. *Biochimica et Biophysica Acta* **72**, 619–629.
- SANGER, F., NICKLEN, S. & COULSON, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America* **74**, 5463–5467.
- SHARPE, G. S. (1984). Broad host range cloning vectors for Gram-negative bacteria. *Gene* **29**, 93–102.
- SHINE, J. & DALGARNO, L. (1974). The 3' terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. *Proceedings of the National Academy of Sciences of the United States of America* **71**, 1342–1346.
- STADEN, R. (1984). Graphic methods to determine the function of nucleic acid sequences. A summary of ANALYSEQ options. *Nucleic Acids Research* **12**, 505–519.
- STASKAWICZ, B., DAHLBECK, D., KEEN, N. & NAPOLI, N. (1987). Molecular characterisation of cloned avirulence genes from race 0 and race 1 of *Pseudomonas syringae* pv. *glycinea*. *Journal of Bacteriology* **169**, 5789–5794.
- STEINBERGER, E. M. & BEER, S. V. (1988). Creation and complementation of pathogenicity mutants of *Erwinia amylovora*. *Molecular Plant-Microbe Interactions* **1**, 135–144.
- STOUT, V. & GOTTESMAN, S. (1990). RcsB and RcsC: a two component regulator of capsule synthesis in *Escherichia coli*. *Journal of Bacteriology* **172**, 659–669.
- TAO, K., MAKINO, K., YONEI, S., NAKATA, A. & SHINAGAWA, H. (1989). Molecular cloning and nucleotide sequencing of the *oxyR*, a positive regulatory gene of a regulon for an adaptive response to oxidative stress in *Escherichia coli*: homologies between OxyR and a family of bacterial activator proteins. *Molecular and General Genetics* **218**, 371–376.
- THOMASHOW, M. F., KARLINSEY, J. E., MARKS, J. R. & HURLBERT, R. E. (1987). Identification of a new virulence locus in *Agrobacterium tumefaciens* that affects polysaccharide composition and plant cell attachment. *Journal of Bacteriology* **169**, 3209–3216.
- TORRES-CABASSA, A. & GOTTESMAN, S. (1987). Capsule synthesis in *Escherichia coli* K-12 is regulated by proteolysis. *Journal of Bacteriology* **169**, 981–989.
- TORRES-CABASSA, A., GOTTESMAN, S., FREDERICK, R. D., DOLPH, P. J. & COPLIN, D. L. (1987). Control of extracellular polysaccharide synthesis in *Erwinia stewartii* and *Escherichia coli* K-12: a common regulatory function. *Journal of Bacteriology* **169**, 4525–4531.
- TRISLER, P. & GOTTESMAN, S. (1984). *lon* transcriptional regulation of genes necessary for capsular polysaccharide synthesis in *Escherichia coli* K-12. *Journal of Bacteriology* **160**, 184–191.
- WILKINS, B. M., BOULNOIS, G. J. & LANKA, E. (1982). A plasmid DNA primase active in discontinuous bacterial DNA replication. *Nature, London* **290**, 217–221.
- YANISCH-PERRON, C., VIEIRA, J. & MESSING, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of M13mp18 and pUC19 vectors. *Gene* **33**, 103–119.