The yjbEFGH locus in Escherichia coli K-12 is an operon encoding proteins involved in exopolysaccharide production

Lionel Ferrières, Shazia N. Aslam, Richard M. Cooper and David J. Clarke

INTRODUCTION

Exopolysaccharides (EPSs) are major components of the bacterial cell envelope and play an important role in the interaction between the bacterium and its environment. Several types of EPS have been characterized in Escherichia coli including the serotype-specific lipopolysaccharide (LPS) O-antigen, capsular K-antigen, colanic acid and the recently identified (1→4)-β-glucan bacterial cellulose and (1→6)-β-N-acetylglucosamine polymer (PGA), which are both involved in cell attachment and biofilm formation (Wang et al., 2004; Whitfield, 2006; Zogaj et al., 2001).

Capsules are polymers of high-molecular-mass polysaccharides that are firmly attached to the cell surface (for review see Whitfield, 2006). Capsules are well-established virulence factors and they have been classified into four groups, according to genetic and biochemical criteria (Whitfield & Roberts, 1999). In E. coli, group 1 and 4 capsules share a common assembly system and are found essentially in isolates that cause intestinal infections. On the other hand, group 2 and 3 capsules are found in E. coli isolates and this capsule is believed to play a role in the lifestyle outside the host (Majdalani & Gottesman, 2005).

Colanic acid is a water-soluble acidic polysaccharide produced by many members of the family Enterobacteriaceae, particularly E. coli. The production and secretion of colanic acid requires proteins that are encoded by the cps/wca operon. The expression of the cps/wca operon is regulated by a complex signal transduction pathway called the Rcs phosphorelay (Stout & Gottesman, 1990). The Rcs phosphorelay shares many features with members of the family of signalling proteins called two-component pathways (Majdalani & Gottesman, 2005). The Rcs phosphorelay is composed of two membrane-anchored proteins, RscC and RscD, and a cytoplasmic protein, RscB. In response to a signal, the sensor kinase RscC autophosphorylates on a His residue and the phosphoryl group is transferred to an Asp residue in RscB, via the HPt domain of RscD (Clarke et al., 2002; Takeda et al., 2001). Phosphorylated RscB is a DNA-binding protein that acts as a transcriptional regulator. At certain promoters, the activity of phospho-RscB is modulated by the availability of an auxiliary protein, RscA (Stout et al., 1991). Signal perception by RscC requires the presence of an outer-membrane lipoprotein, RscF (Castanié-Cornet et al., 2006; Majdalani et al., 2005). While the environmental signal leading to activation of the Rcs phosphorelay remains elusive, the phosphorelay has been shown to respond to conditions that affect the cell envelope and to growth on a solid surface (Ferrières & Clarke, 2003; Huang et al., 2006; Majdalani & Gottesman, 2005).

The RcsCDB phosphorelay was originally identified as the main regulator of colanic acid biosynthesis in Escherichia coli K-12. However, recent transcriptomic analyses have identified more than 150 genes belonging to the Rcs regulon, including yjbE, yjbF, yjbG and yjbH. These genes are clustered on the genome and oriented in the same direction but their function remains unknown. In this work it is shown that yjbE, yjbF, yjbG and yjbH are transcribed as a single operon and it is confirmed that the expression of this operon is controlled by the Rcs phosphorelay, in a manner that is dependent on the auxiliary regulatory protein RcsA. Interestingly, Northern blot analysis revealed that the amount of yjbE transcripts in the cell is higher than the amount of yjbEFGH transcripts and it is proposed that this differential expression is mediated by the presence of a strong stem–loop structure in the yjbE-yjbF intergenic region. Finally, evidence is provided that the overexpression of yjbEFGH affects colony morphology and leads to the production of an extracellular polysaccharide that binds Congo red and toluidine blue-O.

Abbreviations: EPS, exopolysaccharide; PGA, (1→6)-β-N-acetylglucosamine polymer.
In addition to the cps/wca operon, the Rcs phosphorelay controls the expression of genes involved in cell division (fitS, Carballes et al., 1999), the response to osmotic shock (osmB, Boulanger et al., 2003; osmC, Davalos-Garcia et al., 2001), flagella biosynthesis (fliDC, Francez-Charlot et al., 2003) and the post-transcriptional regulation of the stress-response sigma factor σE (rprA, Majdalani et al., 2002). Recent transcriptomic analyses revealed that more than 150 genes belong to the Rcs regulon (Ferrieres & Clarke, 2003; Hagiwara et al., 2003). It is predicted that at least 50 % of the Rcs-regulated genes encode proteins whose function is related to the cell envelope, suggesting that the Rcs phosphorelay plays an important role in remodelling the E. coli cell surface in response to environmental signals (Ferrieres & Clarke, 2003).

Previous studies had revealed four genes of unknown function, yjbE, yjbF, yjbG and yjbH, that were located at the same locus on the E. coli genome and were all highly induced by the Rcs phosphorelay (Ferrieres & Clarke, 2003). In this paper, we show that the yjbEFGH genes of E. coli K-12 form a single operon which is regulated by the Rcs phosphorelay in a RcsA-dependent manner, and we provide evidence that this operon is involved in EPS production.

METHODS

Bacterial strains, plasmids, phages and growth conditions. All E. coli strains, plasmids and phages used in this study are described in Table 1. Bacteria were routinely grown at 30 °C in Luria–Bertani (LB) broth. When required, antibiotics were added to the following concentrations: ampicillin (Ap), 100 μg ml⁻¹; chloramphenicol (Cm), 20 μg ml⁻¹; kanamycin (Km), 30 μg ml⁻¹; tetracycline (Tet), 15 μg ml⁻¹; spectinomycin (Sp), 100 μg ml⁻¹. The mutations were transferred from strain to strain by transduction with either P1vir or P1cml. The plasmids were introduced into the different backgrounds by chemical transformation with CaCl₂ (Sambrook et al., 1989).

Colony morphology and EPS staining. The morphology of colonies grown for 36 h on LB agar plates was observed with a Nikon TMS microscope. EPS production was detected by growing bacteria on LB agar plates containing different polysaccharide-staining dyes at the following concentrations: calcofluor (Fluo brightener 28, Sigma), 80 μg ml⁻¹; ruthenium red, 10–40 μg ml⁻¹; Congo red, 150 μg ml⁻¹; toluidine blue-O, 40 μg ml⁻¹ (Weiner et al., 1999). In the case of calcofluor, dye binding was assessed by illuminating the plates with UV light.

Cloning of the yjbEFGH region, the individual yjb genes and rcsA. The yjbEFGH region, the individual yjb genes and rcsA were cloned under the control of the tac promoter in the plasmid pTrc99a (Amann et al., 1988). Briefly, the genes were amplified from purified genomic DNA by PCR using the Expand High Fidelity PCR system (Roche) and pairs of primers designed to add Ncol and XbaI restriction sites at the 5′ end and 3′ end of the PCR product, respectively (Table 2). PCR fragments corresponding to yjbE, yjbF, yjbG, yjbH, rcsA and yjbEFGH were purified by agarose gel electrophoresis, digested and ligated to the Ncol/XbaI-digested pTrc99a plasmid to give pBMM500, pBMM501, pBMM502, pBMM503, pBMM505 and pBMM506, respectively. The integrity of each plasmid insert was checked by sequencing.

Deletion of the yjbE, yjbF, yjbG and rcsA genes. The yjbE, yjbF, yjbG and rcsA genes were deleted from the chromosome and replaced with a chloramphenicol-resistance cassette using Red recombinase, as described previously (Datsenko & Wanner, 2000; Yu et al., 2000). Briefly, linear DNA cassettes carrying the chloramphenicol-resistance gene were engineered by PCR using pKD3 as a template (Datsenko & Wanner, 2000). The primers used were ~60 bp long and designed so that the 3′ end hybridized with the chloramphenicol cassette of pKD3 and the 5′ end (~40 bp) was homologous to the beginning (forward primer) or the end (reverse primer) of the gene to be deleted (Table 2). After purification on agarose gel, digestion with DpnI to remove traces of the plasmid template and subsequent column purification (StrataPrep PCR Purification Kit, Stratagene), 100 ng of PCR product was electroporated into E. coli strains expressing the red recombinase genes from the chromosome (strain DY330, Yu et al., 2000) or from the pKD46 plasmid (Datsenko & Wanner, 2000). Separation of electrocompetent cells expressing the red recombinase has been described previously (Thomason et al., 2003). The transformants were selected on LB agar plates supplemented with 10 μg chloramphenicol ml⁻¹ at 30 °C and the proper integration of the cassette was confirmed by colony PCR, using two primers that flanked the point of insertion. Finally, the mutations were transferred into the required backgrounds by transduction with P1vir.

RNA analysis. Total RNA was extracted from exponentially growing cultures of PSG1031/pPSC961-31 and PSC1038/pPSC961-31 after induction of the Rcs phosphorelay by DjlA overproduction, as described previously (Ferrieres & Clarke, 2003). RNA integrity was checked by electrophoresis on an agarose gel and the absence of contaminating DNA was confirmed by trying to amplify the cat gene carried by pPSC961-31 by PCR. Semi-quantitative analysis of the yjb transcript was performed by RT-PCR as follows. First-strand cDNA synthesis was performed on 100 ng of total RNA with the SuperScript II reverse transcriptase (Roche) according to the manufacturer’s instructions. Subsequently, newly synthesized cDNA was amplified by PCR and the resulting products were analysed by electrophoresis on 1 % (w/v) or 2 % (w/v) agarose gels. The primer pairs used for the amplification of the yjb intergenic regions are listed in Table 2.

Northern blot analysis of total RNA was performed using probes complimentary to either the individual yjbE, yjbF, yjbG or yjbH genes or the entire yjbEFGH locus. The probes were amplified from genomic DNA by PCR, purified on agarose gels and radiolabelled by the incorporation of [α-³²P]dATP using the DNA polymerase Klenow fragment and a mix of random hexamers (Promega). After synthesis, unincorporated radionucleotides were removed by purification with a Sephadex G-50 gel filtration Nick column (Pharmacia Biotech). Northern blot analysis was performed as described previously with a few modifications (Sambrook et al., 1989). Briefly, 20 μg RNA was separated on a 2 % (w/v) formaldehyde, 1.5 % (w/v) agarose gel, transferred by Hybond-N+ membrane (Amersham) and fixed to the membrane by UV cross-linking (Stratagene Stratalink). The membrane was pre-hybridized at 60 °C for 1 h in Church buffer [0.6 M sodium phosphate, pH 7.0, 7 % (w/v) SDS, 1 mM EDTA, 1 % (w/v) BSA] before the addition of the radiolabelled yjb probe. Hybridization was performed by incubation at 60 °C for 1 h followed by overnight incubation at 55 °C with gentle rocking. The membrane was washed twice by immersion in 2 x SSC solution (1 x SSC solution: 150 mM NaCl, 15 mM sodium citrate, pH 7.0), containing 0.5 % (w/v) SDS, for 20 min at 60 °C, with gentle rocking, and exposed to a phosphoimager screen (Fujiﬁlm imaging plate BAS-MS). In order to assess the quality of the RNA loading, the membrane was stripped in a boiling 0.5 % (w/v) SDS solution for 20 min and hybridized with a second oligonucleotide probe that was complementary to the 55 rRNA (5′- ACTACACATGGCCGCTACGGC-3′), as described previously (Joyce & Clarke, 2003).
Bacteria grown on LB agar plates were scraped from the surface, washed and resuspended in 0.9% (w/v) NaCl. Cell suspensions were agitated vigorously on a magnetic stirrer for 2 h at room temperature and centrifuged at 100,000 g for 24 h to pellet cellular debris and contaminating lipopolysaccharides. The supernatant was recovered, concentrated 10-fold by freeze-drying and dialysed against water for 48 h to remove any low-molecular-mass compounds. The resulting crude extract was freeze-dried and the resultant compound mixture was dissolved, to a concentration of 20 mg ml⁻¹. The subsequent step was then transferred to a sterile nitrocellulose membrane (0.45 μm, WCN type, Whatman) by vacuum filtration. The membrane was washed twice with 10 ml PBS and placed on a pre-warmed LB agar plate. After incubation at 30 °C for 150 min, the cells were recovered from the membrane by vigorous washing in PBS, lysed with chloroform/SDS and gene expression was assayed by measuring β-galactosidase activity (Miller, 1972).

### EPS extraction.

The expression of yjhH–lacZ transcriptional fusion in response to a solid surface. The yjhH–lacZ transcriptional fusion was isolated during the screening of a library of λlacMu53 insertion mutants for genes regulated by the Rcs phosphorelay (Ferrie`res & Clarke, 2003). The expression of yjhH–lacZ was followed after induction of the Rcs phosphorelay by transferring the cells from liquid broth to a solid surface, as described previously (Ferrie`res & Clarke, 2003). Briefly, cells grown to mid-exponential phase (OD₆₀₀ 0.5) were harvested by centrifugation, resuspended to OD 600 0.1 in PBS and 10 ml of the suspension was then transferred to a sterile nitrocellulose membrane (0.45 μm, WCN type, Whatman) by vacuum filtration. The membrane was washed twice with 10 ml PBS and placed on a pre-warmed LB agar plate. After incubation at 30 °C for 150 min, the cells were recovered from the membrane by vigorous washing in PBS, lysed with chloroform/SDS and gene expression was assayed by measuring β-galactosidase activity (Miller, 1972).
Chemical analysis of EPS. The total amount of carbohydrates in the purified EPS was determined by the phenol/H₂SO₄ method, using glucose as a standard (Dubois et al., 1956). Uronic acids were measured by the m-hydroxydiphenyl/sulphuric acid assay, using glucuronic acid as a standard (Blumenkrantz & Asboe-Hansen, 1973). The different sugars were quantified after hydrolysis with 4 M HCl (100 °C, 4 h) as follows: the content of reducing sugars was determined by the Nelson-Somogyi method, using glucose as a standard (Nelson, 1944); amino sugars were quantified by the method of Rondle & Morgan (1955), using glucosamine as a standard. The amount of glucose was determined with the glucose assay kit (glucose oxidase/peroxidase-o-anisidine, Sigma), according to the manufacturer’s instructions. Galactose was estimated with a similar galactose oxidase/peroxidase system. Protein contamination was determined by the method of Bradford (1976) (Bradford Reagent, Sigma), using bovine serum albumin (BSA) as a standard.

Table 2. Primers used in this study

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence*</th>
<th>Gene</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gene cloning</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LF94</td>
<td>5’-CATGCCATGGGGTGCAACGATTATATTGATTTG-3’</td>
<td>rcsA</td>
</tr>
<tr>
<td>LF95</td>
<td>5’-CATCCGACGCTGCTGACTGTTGGGAAC-3’</td>
<td>rcsA</td>
</tr>
<tr>
<td>LF96</td>
<td>5’-CATGCCATGGGGGAAAAAGTCTCTAGATGTTGGATTTTGG-3’</td>
<td>yjbE/yjbEFGH</td>
</tr>
<tr>
<td>LF97</td>
<td>5’-CATCTAGACATATTATACCTGACTGTTGGGAAC-3’</td>
<td>yjbE</td>
</tr>
<tr>
<td>LF98</td>
<td>5’-CATGCCATGGGGGAAAAAGTCTCTATTAGCTGTTGGATTTTGG-3’</td>
<td>yjbF</td>
</tr>
<tr>
<td>LF99</td>
<td>5’-CATCTAGACATATTATACCTGACTGTTGGGAAC-3’</td>
<td>yjbF</td>
</tr>
<tr>
<td>LF100</td>
<td>5’-CATGCCATGGGGGAAAAAGTCTCTATTAGCTGTTGGATTTTGG-3’</td>
<td>yjbG</td>
</tr>
<tr>
<td>LF101</td>
<td>5’-CATCTAGACATATTATACCTGACTGTTGGGAAC-3’</td>
<td>yjbG</td>
</tr>
<tr>
<td>LF102</td>
<td>5’-CATCTAGACATATTATACCTGACTGTTGGGAAC-3’</td>
<td>yjbH</td>
</tr>
<tr>
<td>LF103</td>
<td>5’-CATCTAGACATATTATACCTGACTGTTGGGAAC-3’</td>
<td>yjbH/yjbEFGH</td>
</tr>
<tr>
<td><strong>Gene deletion</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LF37</td>
<td>5’-AAGGGAAATTTGTAAGAAAAAGTTCTGTATGGCATTTCGTTATGGGAACGCTGCTTACTGCTTC-3’</td>
<td>yjbE::Cm</td>
</tr>
<tr>
<td>LF38</td>
<td>5’-TTTGGGAGATACATCTTTTATTACTGGGTACTGGG ATTTGGAACGCTGCTTACTGCTTC-3’</td>
<td>yjbE::Cm</td>
</tr>
<tr>
<td>LF39</td>
<td>5’-CCCTCTGAGAAGAGTCGCTGAAGCGACCTGACTCATGTTGGGAACGCTGCTTACTGCTTC-3’</td>
<td>yjbF::Cm</td>
</tr>
<tr>
<td>LF40</td>
<td>5’-GTTGTTTTTATCATGTGGTTTTGCAGCTGTTATGAGACTCATGGAAATATTCCCTCTCTTACTGCTTC-3’</td>
<td>yjbF::Cm</td>
</tr>
<tr>
<td>LF41</td>
<td>5’-TCAAGCGGCGAAAATGATTATAACAAACTATTGTCGTCGTGAAGCGACCTGACTCATGTTGGGAACGCTGCTTACTGCTTC-3’</td>
<td>yjbG::Cm</td>
</tr>
<tr>
<td>LF42</td>
<td>5’-AAACAGATGTTCTTTTTTCTTATTATGGGATATCCTGACGTGGTCGTGAAGCGACCTGACTCATGTTGGGAACGCTGCTTACTGCTTC-3’</td>
<td>yjbG::Cm</td>
</tr>
<tr>
<td>LF58</td>
<td>5’-ATTGGATGGAAGATGTTTACCTGTTGCAACGCTGACTCATGTTGGGAACGCTGCTTACTGCTTC-3’</td>
<td>rcAS::Cm</td>
</tr>
<tr>
<td>LF59</td>
<td>5’-ACAGCTGAAATGTGTTTACGCGATTTGTAACAACTATTGTCGTCGTGAAGCGACCTGACTCATGTTGGGAACGCTGCTTACTGCTTC-3’</td>
<td>rcAS::Cm</td>
</tr>
<tr>
<td><strong>RT-PCR on yjbEFGH intergenic regions</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LF72</td>
<td>5’-TAAGCTCAGCGCGGTGG-3’</td>
<td>3’ end of yjbE</td>
</tr>
<tr>
<td>LF73</td>
<td>5’-TAGCTCGTGACGGAGCAG-3’</td>
<td>5’ end of yjbF</td>
</tr>
<tr>
<td>LF74</td>
<td>5’-CGGAAGATGTTACCTCGG-3’</td>
<td>3’ end of yjbF</td>
</tr>
<tr>
<td>LF75</td>
<td>5’-GCTCTGCGATTACCC-3’</td>
<td>5’ end of yjbG</td>
</tr>
<tr>
<td>LF76</td>
<td>5’-CGCCGAAAGTACCTCAG-3’</td>
<td>3’ end of yjbG</td>
</tr>
<tr>
<td>LF77</td>
<td>5’-GGATATGCGAGCTCATCT-3’</td>
<td>5’ end of yjbH</td>
</tr>
</tbody>
</table>

* NcoI and XbaI restriction sites are underlined. Italicized sequences are complementary to the Cm-resistance cassette of plasmid pKD3.

RESULTS

The yjbEFGH genes are transcribed as an operon

The yjbE, yjbF, yjbG and yjbH genes are found together on the E. coli genome, between the phosphoglucone-isomerase-encoding gene, pgi, and the phosphate-starvation-inducible ptsE gene (formerly yjbA), and the four genes are orientated in the same direction (Fig. 1a). The yjbE gene is predicted to encode a small (80 amino acids) protein with a stretch of threonine residues at the C-terminal end and yjbG is predicted to encode a 245 amino acid protein. Both YjbE and YjbG are predicted to contain N-terminal signal
peptides, suggesting that these proteins are probably localized in the periplasm. On the other hand, yjbF and yjbH are predicted to encode β-barrel, outer-membrane lipoproteins, 212 amino acids and 698 amino acids in length, respectively (Zhai & Saier, 2002). The ORFs of yjbF, yjbG and yjbH overlap: yjbF and yjbG share four nucleotides while yjbG and yjbH have a single base in common. In contrast, 114 bp separates the predicted stop codon of yjbE and the start codon of yjbF. This organization, in addition to the fact that these genes are all regulated by the Rcs phosphorelay (Ferrières & Clarke, 2003), strongly suggested that these genes might be transcribed as a single operon.

In order to confirm this, we followed the expression of a reporter gene in response to the activation of the Rcs phosphorelay by DjlA overproduction, as described in Methods. Amplification of the mRNA corresponding to the constitutive chloramphenicol-resistance gene (cat) encoded by pPSG961-31 was used as a control of cDNA synthesis. The expected size of each PCR product is as follows: yjbE-yjbF, 269 bp; yjbF-yjbG, 260 bp; yjbG-yjbH, 1755 bp; and yjbE-yjbH, 976 bp, strongly suggesting that all of the genes are transcribed on the same mRNA transcript (Fig. 1c). Moreover, when RNA extracted from the rcsC mutant strain was used as a template, no PCR product was detected for the yjbE-yjbF, yjbF-yjbG and yjbE-yjbG intergenic regions while the amplification of the yjbG-yjbH and yjbF-yjbH intergenic regions resulted in bands of the expected size but of very poor intensity compared to the respective bands obtained with RNA extracted from the wild-type. This confirmed that the expression of the yjbEFGH operon is strongly dependent on the Rcs phosphorelay and that there is no detectable expression of the yjbEFGH operon in the absence of Rcs activation.

We further examined the nature of the yjbEFGH transcript by Northern blotting. As in the previous RT-PCR experiments, we prepared total RNA from wild-type and rcsC mutant cells, following induction of the Rcs phosphorelay.
by DjlA overproduction. We observed a single band when the RNA was hybridized to the yjbEFGH probe (Fig. 2a).
This band was present in the wild-type strain only, confirming the Rcs-dependent regulation of this locus. Interestingly, the band observed displayed a molecular size of ~300 nucleotides (as estimated by comparison with the migration of 23S, 16S and 5S rRNA), which was much smaller than the size expected for the full yjbEFGH mRNA (~4000 nucleotides) but was consistent with the size of a transcript carrying yjbE only. Further hybridizations with probes specific for each gene of the yjbEFGH region revealed that this band was specific for yjbE. Moreover, despite the clear identification of a significant level of yjbE transcript, no band was obtained with probes corresponding to yjbF, yjbG or yjbH (Fig. 2a and data not shown). Given that yjbE is clearly separated from the other genes of the yjb operon, we analysed the yjbE-yjbF intergenic sequence in silico and found that a strong stem-loop secondary structure was predicted in this region ($\Delta G = -20.61 \text{ kcal mol}^{-1}$; $-86.2 \text{ kJ mol}^{-1}$) (Fig. 2b). Altogether, the data suggest that, although part of the same operon, yjbE and yjbFGH are differentially expressed, probably due to the presence of a strong secondary structure between yjbE and yjbF.

### The expression of yjbEFGH is RcsA-dependent

It has been previously shown that RcsA is involved in the regulation of a subset of Rcs-regulated genes. To investigate the role of RcsA in the regulation of the yjbEFGH operon, we tested for the rcsA-dependence of yjbH-lacZ expression following activation of the Rcs phosphorelay by transfer of wild-type, rcsA and rcsB cells from liquid medium to a solid surface (Ferrières & Clarke, 2003). Under these conditions the expression of yjbH-lacZ was induced 50-fold in the wild-type strain and this induction was dependent on the rcsB gene, confirming the importance of the Rcs phosphorelay in the expression of the yjbEFGH operon (data not shown). In the rcsA mutant strain the level of yjbH-lacZ induction was only 14-fold, suggesting that RcsA is required for full yjbEFGH expression. RcsA is thought to act as a heterodimer with RcsB in binding DNA targets and a consensus RcsAB box has been defined (Wehland & Bernhard, 2000). In silico analysis identified a potential RcsAB binding site located 251 to 264 bp upstream of the yjbE start codon (Fig. 3). The potential RcsAB box is identical to the consensus RcsAB box in 10 out of 14 nucleotides, including the six most conserved positions (Wehland & Bernhard, 2000). Furthermore, the location of the potential RcsAB box is consistent with the positions reported for other RcsAB binding sites i.e. 100 to 500 bp upstream of the translation start site.

#### yjbEFGH overexpression alters colony morphology

To determine the role of the yjbEFGH operon we constructed deletion mutations in yjbE, yjbF, yjbG and yjbH but the mutants obtained did not have any apparent phenotype when compared to the wild-type strain. In particular, although the Rcs phosphorelay is required for normal biofilm formation in E. coli, we did not observe any defect in the ability of the yjb mutants to form biofilms on PVC plastic (Ferrières & Clarke, 2003, and data not shown). Therefore, we decided to look at any effect associated with the overexpression of yjbEFGH. In order to do that, we cloned the full-length yjbEFGH operon under the control of the ptrc promoter, in the vector pTrc99a, resulting in pBMM506, and the plasmid was introduced into ZK2686. The induction of yjbEFGH expression by IPTG appeared to affect growth on plates, suggesting that it might be toxic for the cells (data not shown). Therefore, for future experiments, we used the basal activity of the ptrc promoter to express yjbEFGH. Northern blot analysis confirmed that the

![Fig. 2. Differential expression of yjbE and yjbFGH. (a) Northern blot analysis of total RNA isolated from PSG1031/pPSG961-31 (wt) and the isogenic mutant PSG1038/pPSG961-31 (rcsC) after activation of the Rcs phosphorelay by DjlA overproduction. The yjbE, yjbF and yjbEFGH probes were used as indicated. The positions of the 23S rRNA, 16S rRNA and the expected yjbEFGH mRNA are indicated. Hybridization with the yjbG and yjbH probes gave the same result as hybridization with the yjbF probe (data not shown). (b) Schematic representation of the secondary structure ($\Delta G = -20.61 \text{ kcal mol}^{-1}$; $-86.2 \text{ kJ mol}^{-1}$) of the yjbE-yjbF intergenic region as predicted by DNAsite (Lasergene). The beginning and end of the yjbE and yjbF genes, respectively, are highlighted in grey. The position of the truncated ERIC sequence predicted by Wilson & Sharp (2006) is indicated in bold type.](http://mic.sgmjournals.org)
yjbEFGH operon was overexpressed in ZK2686/pBMM506 compared to ZK2686/pTrc99a (data not shown). Under these conditions, colonies of the strain containing the pBMM506 plasmid displayed an altered morphology on LB agar plates (Fig. 4). Strains overexpressing yjbEFGH produced translucent, crater-like colonies that contrasted with the opaque, regular, cone-shaped colonies formed by the vector-carrying strain. Changes in colony morphology have often been associated with modifications of the expression status of cell-surface components, such as capsule and adhesins (Hasman et al., 2000; Schembri et al., 2004). In order to better understand the nature of the changes occurring in the presence of pBMM506, we observed the colony morphology of mutants overexpressing yjbEFGH but deficient in the production of different surface components such as colanic acid (gmd::λp lacMu53), curli (csgD::uidA-Km) and type I fimbriae (fimB::Cm). Although the disruption of curli and fimbrial production did affect colony morphology, the introduction of pBMM506 led to the formation of crater-like colonies in all the backgrounds tested, suggesting that the change in colony morphology conferred by yjbEFGH overexpression is independent on the production of colanic acid, curli and type I fimbriae (Fig. 4b). As YjbE, F, G and H are all predicted to be secreted proteins, it was also possible that the overproduction of these proteins would perturb the cell envelope. To exclude this possibility the activation status of two systems known to respond to alterations in the cell envelope, RcsCDB and CpxAR, was monitored, using gmd–lacZ and cpxR–lacZ reporter fusions, respectively, after overexpression of the yjbEFGH operon (Majdalani & Gottesman, 2005; Raivio et al., 1999). There was no change in the level of either gmd–lacZ or cpxR–lacZ expression, suggesting that the overexpression of yjbEFGH had no significant effect on the integrity of the cell envelope (data not shown).

**yjbEFGH overexpression leads to increased binding of toluidine blue-O and Congo red**

*E. coli* contains paralogues of the yjbEFGH genes, namely the ymcDCBA genes. The ymcDCBA locus is an operon that also includes, at its 3’ end, the yccZ, etp and etk genes. The ymcDCBA-yccZ-etyl etk operon is involved in the production of a group 4 capsule in EPEC O127 but the expression of

---

**Fig. 3.** Alignment of the yjbEFGH upstream region and the consensus RcsAB box (Wehland & Bernhard, 2000). The first base pair of yjbE was used as the starting point for base pair numeration.

---

**Fig. 4.** Colony morphology of wild-type *E. coli* and mutants overexpressing yjbEFGH. Bacteria were grown on LB-Ap agar plates for 36 h at 30 °C. (a) Overview of wild-type *E. coli* colonies (ZK2686) carrying pTrc99a (pvector) or pBMM506 (pyjbEFGH), respectively. (b) Photographs of colonies formed by ZK2686 (wt) and isogenic mutants defective in colanic acid expression (BMM526, cps mutant), curli expression (BMM568, csg mutant) or fimbriae expression (ZK2688, fim mutant), and carrying pTrc99a (pvector) or pBMM506 (pyjbEFGH), as indicated.
this operon is blocked in K-12 strains by the insertion of an IS1 element upstream of ymcD (Ilan et al., 1999; Peleg et al., 2005; Vincent et al., 2000). However, the homology with ymcDCBA does suggest that the yjbEFGH locus may encode proteins involved in EPS production. In order to check EPS production in strains overexpressing the yjbEFGH operon, both ZK2686/pTrc99a and ZK2686/pBMM506 were streaked on LB agar plates containing dyes known to recognize different types of EPS i.e. calcofluor (CF), ruthenium red (RR), toluidine blue-O (TB) and Congo red (CR) (Weiner et al., 1999). CF is a fluorochrome that binds to (1→3)-β- and (1→4)-β-D-glucopyranosides, such as cellulose. Consistent with observations reported for other E. coli K-12 strains (Da Re & Ghigo, 2006), ZK2686 did not bind CF, and we found that the introduction of pBMM506 did not reverse this phenotype, suggesting that the overexpression of yjbEFGH does not mediate the production of a cellulose-like compound (data not shown). RR is known to stain acidic EPS but we did not observe any difference between the RR-binding ability of ZK2686/pBMM506 and ZK2686/pTrc99a (data not shown). In contrast, both TB and CR appeared to stain ZK2686/pBMM506 specifically (Fig. 5). TB recognizes negatively charged or helical polysaccharide while CR binds (1→4)-α-D-glucopyranosides and basic or neutral EPS. CR has also been shown to interact with various proteinaceous compounds, including curli (Hammar et al., 1995). However, CR and TB binding was not affected in the csgD mutant background, suggesting that the enhanced dye-binding capability of the yjbEFGH-overexpressing strain was not the result of an increase in curli production (data not shown). Moreover, we did not observe any significant staining when the genes of the yjbEFGH operon were individually expressed in ZK2686, excluding the possibility that CR and TB may be directly binding one of the overproduced proteins (Fig. 5). Therefore, based on dye staining properties, we can conclude that the overexpression of the yjbEFGH operon likely results in the production of a polysaccharide that is not cellulose (does not bind CF), does not contain any (1→3)-β- and (1→4)-β-D-glucopyranosides (does not bind CF), could be helical (binds TB) and could contain (1→4)-α-D-glucopyranoside units (binds CR).

**Overexpression of yjbEFGH leads to increased production of EPS**

In order to firmly establish that yjbEFGH overexpression leads to the production of EPS, we compared the total EPS content of cells carrying pBMM506 with cells carrying the vector only. To avoid any contamination by colanic acid during this analysis we used strain BMM526, containing an insertion in gmd and therefore deficient in colanic acid production. Chemical analysis of the EPS purified from BMM526/pTrc99a and BMM526/pBMM506 revealed a 1.8-fold increase and a 2.2-fold increase in total carbohydrates and reducing sugars, respectively, in the yjbEFGH-overexpressing strain, confirming that the yjbEFGH operon is involved in the production of EPS (Table 3). The EPS purified from BMM526/pBMM506 contained [% (w/w) total carbohydrate] reducing sugars (29.7), amino sugars (10.6) and uronic acids (15.7) (Table 3). Interestingly, galactose and glucose, two major components of colanic acid, were either absent or present at a very low level (4.9 %, in the case of glucose), confirming the absence of contamination by colanic acid. Therefore the yjbEFGH operon appears to be involved in the production of an EPS that is distinct from colanic acid.

**DISCUSSION**

The genes yjbE, yjbF, yjbG and yjbH are located at the same locus on the chromosome of E. coli and have all been identified as members of the Rcs regulon (Ferrières & Clarke, 2003). In this work we used polar cassette insertion and RT-PCR to show that these genes are transcribed as a single unit. Surprisingly, Northern blot analysis of the mRNA corresponding to this region revealed the existence of a short transcript corresponding to yjbE alone. The yjbE gene is separated from the downstream genes by 114 bp that are predicted to fold into a strong secondary structure. In a recent study, the longest loop of this 114 bp region (Fig. 2b)
was predicted to be a truncated ERIC (enterobacterial repetitive intergenic consensus) sequence (Wilson & Sharp, 2006). ERIC sequences are 127 bp imperfect palindromes that occur in multiple copies in the genome of enterobacteria and vibrio. Their function is still unknown but, because of their preferential location in the intergenic regions of co-transcribed genes, a role in regulation of gene expression has been postulated (Sharples & Lloyd, 1990). Analysis of the yjbEFGH mRNA transcript by Northern blotting strongly supports this hypothesis and suggests that ERIC sequences participate in the differential expression of co-transcribed genes, either by increasing the stability of the upstream part of the mRNA or by decreasing the transcription efficiency of the downstream part. As a consequence, although yjbEFGH is transcribed as a single mRNA transcript, yjbE transcripts are present at a higher level in the cell than yjbEFGH transcripts. The reason for this differential expression remains to be elucidated.

The genes of the Rcs regulon can be divided into two classes, depending on whether their expression is dependent on RcsA. RcsA-dependent genes, e.g. the cps/wca operon, flhDC and rcsA, possess a consensus RcsAB box located 50–100 bp upstream of the promoter elements (Majdalani & Gottesman, 2005; Wehland & Bernhard, 2000). The observation that the Rcs-mediated activation of yjbEFGH is partially dependent on RcsA, together with the identification of a potential RcsAB box upstream of yjbE, strongly suggests that yjbEFGH is regulated by the RcsAB complex.

Together with the cps/wca operon, yjbEFGH now becomes the second Rcs-regulated locus that is involved in EPS production. However, the yjbEFGH-encoded proteins are not likely to be involved in the synthesis of EPS per se since none of them is predicted to have any enzymic activity. In fact, all the proteins are predicted to be secreted and, moreover, two of them, YjbF and YjbH, are predicted to be \( \beta \)-barrel lipoproteins (Zhai & Saier, 2002). \( \beta \)-Barrel proteins are found in the outer membrane of Gram-negative bacteria, where they form pore-like structures which facilitate the transport of a broad range of molecules, e.g. proteins and sugars, across the membrane (Wimley, 2003). It is therefore highly possible that the yjbEFGH operon encodes a system involved in EPS secretion. The genes required for yjbEFGH-related EPS biosynthesis remain to be identified but preliminary experiments suggest that they are distinct from the colanic acid biosynthesis genes since a mutation in gmd does not abolish the production of this EPS, while it impairs colanic acid production (see Results).

E. coli K-12 is known to produce only two types of EPS: colanic acid and the \((1\rightarrow6)\)-\( \beta \)-N-acetylglucosamine polymer, also called PGA (Wang et al., 2004; Whitfield, 2006). Chemical analyses revealed that the two major components of colanic acid, galactose and glucose, are absent or present at a low level, respectively, in the EPS extracted from the yjbEFGH-overexpressing strain, indicating that the yjbEFGH-related EPS is distinct from colanic acid. Moreover, the amount of glucosamine residues is similar in the strain overproducing YjbEFGH to that in the non-overproducing strain, suggesting that the yjbEFGH-related EPS is different from PGA. Therefore, it is likely that yjbEFGH may contribute to the production of a new kind of EPS in E. coli, the structure of which remains to be determined. Paralogues of the yjbEFGH genes, ymcDCBA, are found on the chromosome of E. coli K-12, where they belong to a seven-gene operon that is required for the production of group 4 capsule, formerly ‘O-antigen capsule’, in EPEC O127 (Peleg et al., 2005). Vibrio cholerae O139 also contains yjbFGH homologues, called wbfDCB (formerly otnEFG). The function of these genes remains unknown but they are located in the vicinity of wzm and wzz, which are involved in the formation of the O139 O-antigen capsule (Bik et al., 1996). It is therefore tempting to speculate that the yjbEFGH-related EPS might be a group 4 capsule. In EPEC O127, secretion of the group 4 capsule requires all of the genes in the ymcDCBA-yczc-egt-ek operon (Peleg et al., 2005). The last three genes of this operon are not found in the yjbEFGH operon but they are very similar to the first three genes of the cps/wca operon (wzc, wzb and wza, respectively). As the expression of these genes is not affected in the gmd-\( \Delta \)placMu53 insertion mutant used in this study, we cannot exclude the possibility that the production of EPS by yjbEFGH overexpression requires wza-wzb-wzc. Indeed, this would explain why the expression of the yjbEFGH and cps operons is linked by the Rcs phosphorelay.

Like most bacteria in the environment, E. coli has the ability to form sessile communities on solid surfaces that are enclosed in a polysaccharide-rich matrix, called biofilms (Stoodley et al., 2002). To date, colanic acid, PGA and, in some cases, cellulose have been identified as components of the E. coli biofilm matrix and mutations in genes encoding these EPSs result in decreased biofilm formation (Danese et al., 2000a; Wang et al., 2004; Zogaj et al., 2001). Although we did not observe any effect on biofilm formation with any of the yjb deletion mutants produced during this work, a recent study has reported that the deletion of yjbE does reduce biofilm formation in another strain of E. coli K-12 (Herzberg et al., 2006). This suggests a possible role for

### Table 3. Composition of EPS produced by overexpression of yjbEFGH

The values are expressed as \( \mu \)g dry weight \( \text{mg}^{-1} \) and represent the mean ± SD of three determinations.

<table>
<thead>
<tr>
<th></th>
<th>Vector</th>
<th>yjbEFGH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total carbohydrates</td>
<td>254 ± 27</td>
<td>451 ± 46</td>
</tr>
<tr>
<td>Reducing sugars</td>
<td>62 ± 2</td>
<td>134 ± 8</td>
</tr>
<tr>
<td>Amino sugars</td>
<td>45 ± 10</td>
<td>48 ± 5</td>
</tr>
<tr>
<td>Uronic acids</td>
<td>53 ± 1</td>
<td>71 ± 4</td>
</tr>
<tr>
<td>Glucose</td>
<td>15 ± 2</td>
<td>22 ± 9</td>
</tr>
<tr>
<td>Protein</td>
<td>49 ± 3</td>
<td>43 ± 3</td>
</tr>
</tbody>
</table>

...
yjbEFGH in biofilm formation and, perhaps, in matrix production. Genome analysis reveals that the yjbEFGH operon is present in all sequenced strains of *E. coli* (data not shown) and further studies will be aimed at characterizing the role of this operon in both capsule production and biofilm formation.

ACKNOWLEDGEMENTS

We would like to thank Donald Court, Corinne Dorel, Roberto Kolter, Tom Silhavy and Barry Wanner for providing strains and plasmids. We would like also to thank Robert Jackson and all members of the D. J. C. laboratory for useful discussion, especially Susan Joyce for help with Northern blots. This work was supported by a grant from the BBSRC to D. J. C. and a grant from the Leverhulme Trust to R. M. C.

REFERENCES


Edited by: I. R. Henderson