CHAPTER 7

CELLULOSE BIOSYNTHESIS IN ENTEROBACTERIACEAE

UTE RÖMLING*

Department of Microbiology, Tumor and Cell Biology (MTC), Box 280, Karolinska Institutet, SE-17177 Stockholm, Sweden

Abstract

Distinct bacterial species belonging to the family of Enterobacteriaceae harbor a characteristic cellulose biosynthesis operon (*bcs*). A regulatory network for cellulose biosynthesis has been identified in ***Salmonella typhimurium***. Transcription of the bcs operon is constitutive, while cellulose biosynthesis is activated on the post-transcriptional level by AdrA, a GGDEF domain containing protein. AdrA is under the tight positive control of the transcriptional regulator CsgD, which itself is regulated by a wide variety of environmental stimuli and global regulatory proteins. However, regulation of cellulose biosynthesis varies widely among species and even within a species. In ***S. typhimurium*** in cellulose is commonly coexpressed with curli fimbriae, a proteinaceous component whereby the two extracellular matrix components interact with each other fulfilling distinct roles in cell-cell interactions and biofilm formation.

Keywords

Curli fimbriae, electron microscopy, environmental conditions, *Escherichia coli*, GGDEF domain, *ompR*, regulation, *rpoS*, *Salmonella typhimurium*, thin aggregative fimbriae.

Abbrevations

bp, base pair; ORF, open reading frame.

1 INTRODUCTION

Recently, enzymatic and chemical analysis in combination with genetic studies revealed that *Salmonella enterica* serotype Typhimurium (*S. typhimurium*) is capable to produce cellulose as an exopolysaccharide (Zogaj et al. 2001). The

* For correspondence: Tel: +46-8-524-87319; Fax: +46-8-330744; e-mail: ute.romling@mtc.ki.se

R.M. Brown, Jr. and I.M. Saxena (eds.), Cellulose: Molecular and Structural Biology, 107–122.

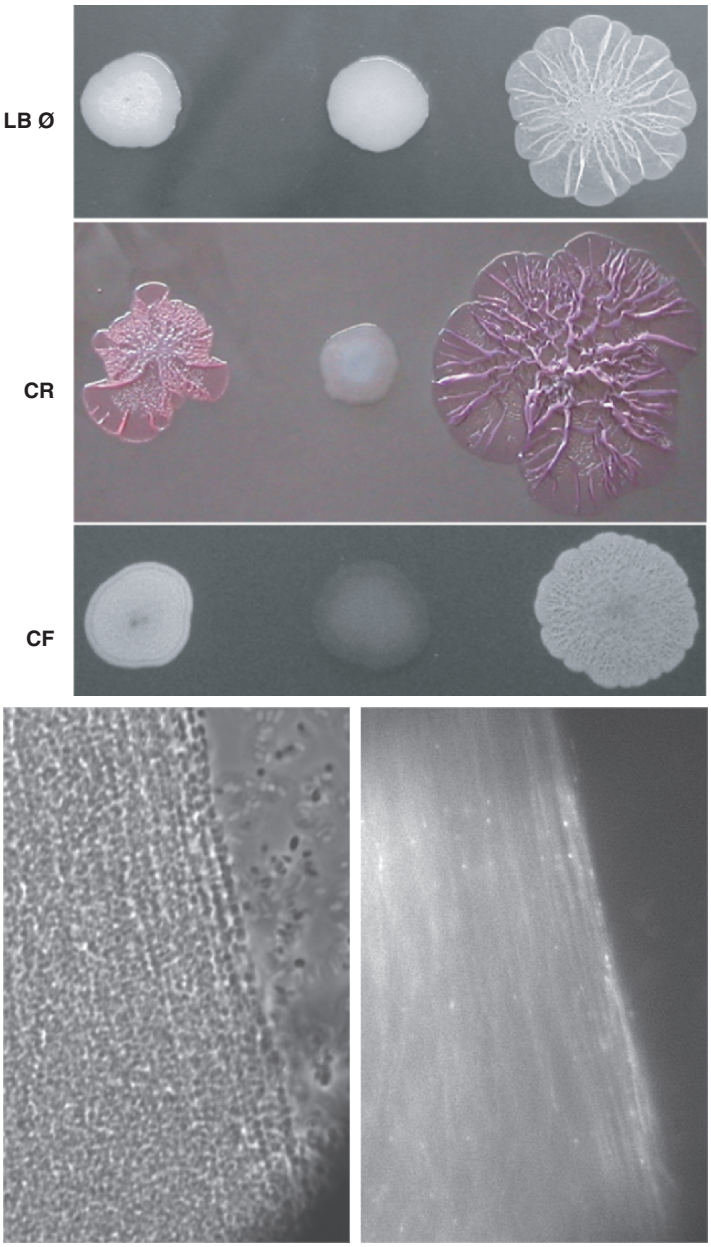


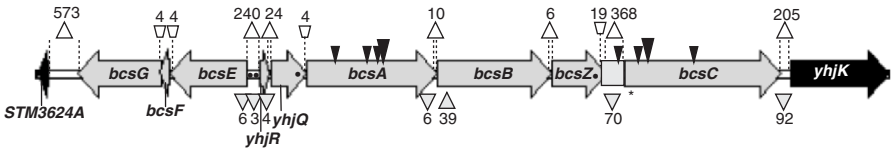
Figure 7-1. Salmonella typhimurium expressing cellulose. (**a**) *S. typhimurium* colonies grown for 48 h on a regular Luria Bertani agar plate without salt (LBØ), a Congo Red plate (CR) and a Calcofluor plate (CF). *Left*: strain that expresses cellulose; *middle*: strain that does not express cellulose. *Right*: strain that expresses cellulose and curli fimbriae. (**b**) Microscopy of cellulose expressing *S. typhimurium* after treatment with Calcofluor. *Left*: Phase contrast; *right*: fluorescence microscopy. Magnification ×600 (*See Color Plate of this figure beginning on page 355*)

bacteria that express cellulose display a characteristic colony morphology when grown on an agar plate and show fibrous cellulose fibers by fluorescence microscopy (Figure 7-1).

Cellulose production confers bacterial cell–cell interactions, adhesion to abiotic surfaces (biofilm formation) and chlorine resistance to the organism (Römling et al. 2000; Zogaj et al. 2001; Solano et al. 2002). Beginning studies shed some light on the molecular mechanisms of cellulose biosynthesis and regulation in *S. typhimurium* and the epidemiology of cellulose biosynthesis in Enterobacteriaceae.

2 THE CELLULOSE BIOSYNTHESIS OPERON IN *Salmonella typhimurium* AND *Escherichia coli*

In *S. typhimurium* and *Salmonella enteritidis*, the two divergently transcribed operons, *yhjRQbcsABZC-bcsEFG* are required for cellulose biosynthesis (Zogaj et al. 2001, Figure 7-2; Solano et al. 2002). The two operons had been identified by random transposon mutagenesis while selecting for mutants with altered capacities to bind the dyes Congo Red and/or Calcofluor when grown on agar plates.



Gene	Position in LT2 genome	ORF (bp)	size of protein (kDa)	Shine–Dalgarno sequence	TAAGGAGGT
bcsG	3806420 - 3808099	1680	62. 2	3806400 – 3806408	A**AAG**TCAAG
bcsF	3806232 - 3806423	192	7. 2	3806217 - 3806225	CGC**GGAG**CG
bcsE	3804664 - 3806235	1572	59. 3	3804644 – 3804652	**TAA**AC**AG**T**T**
yhjR	3804216 - 3804419	204	7. 6	3804426 – 3804434	A**AAG**G**AG**CA
yhjQ	3803463 - 3804191	729	26. 7	3804208 – 3804216	A**A**T**GG**C**GA**T
bcsA	3800842 - 3803466	2625	100. 4	3803480 – 3803488	A**AA**C**G**TCCG
bcsB	3798531 - 3800831	2301	84. 3	3800839 – 3800847	C**AA**T**GA**T**G**A
bcsZ	3797418 - 3798527	1107	41. 6	3798535 – 3798543	TG**A**CC**A**T**G**A
*bcsC**	3793894 - 3797055	3162	115.9	3797067 – 3797076	**TA**TCA**A**TC**T**
bcsC	3793894 - 3797436	3543	129.8	3797446 – 3797455	**TA**CCTGAC**T**

Figure 7-2. Structure of the cellulose biosynthesis operon *bcs* in *Salmonella typhimurium* and *Escherichia coli*. *Arrowheads* represent the open reading frames (ORFs). Symbols above the ORFs show overlap of (▽) or distance between (Δ) ORFs in bps. Symbols below the ORFs show insertions (▲) or deletions (▼), which occur in *S. typhimurium* LT2 as compared to *Escherichia coli* K-12. *Closed arrows* just above the ORFs indicate transposon insertions in *bcsA* and *bcsC*. The *larger arrow* indicates the position of the transposon used to study transcriptional regulation of the respective gene. The table summarizes the features of *bcs* genes using the positioning in the genome of the sequenced LT2 strain. The start codon proposed for bcsC in E. Coli K-12 leads to a shorter ORF (bcsC*).

Highly homologous operons are present in *E. coli*, and deletion of *bcsA*, the catalytic subunit of cellulose synthase in a recently isolated fecal strain, has proven that it also encodes for cellulose biosynthesis in this species ((Zogaj et al. 2001, our unpublished data). Besides minor changes, annotation of open reading frames in *S. typhimurium* LT2 and *E. coli* K-12 is virtually identical with the exception of *bcsC*. The proposed start codon for *bcsC* in *E. coli* has no equivalent in *S. typhimurium*. Alternative start codons common to both species are found up- and downstream of this site. The proposed upstream start codon leads to an ORF which overlaps with *bcsZ* by 19 bp. Strikingly, the use of this sites would lead to an preterminated out-of-frame product in the laboratory strain *E. coli* K-12, but not in all other up to now sequenced natural *E. coli* strains, which include two enterohemorrhagic, one uropathogenic, one enteroaggregative and one enteropathogenic strain. This fact would explain why *E. coli* K-12 does not produce cellulose.

When compared to the classical type 1 cellulose biosynthesis operon of *Gluconacetobacter xylinus* that produces the cellulose I allomorph under laboratory conditions the cellulose biosynthesis operons of *Salmonella* spp. and *E. coli* have both, homologous and unique components (Figure 7-3). As in *G. xylinus bcsA,* which encodes for the catalytic subunit of the cellulose synthase, and *bcsB,* which

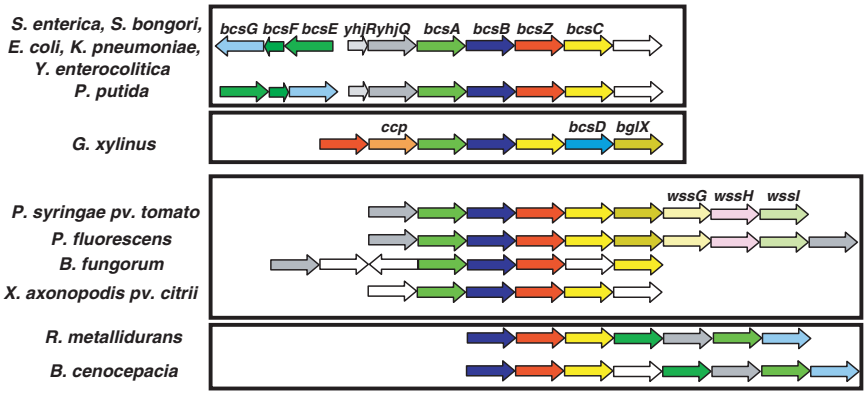


Figure 7-3. Comparison of the cellulose biosynthesis operons *bcs* of Enterobacteriaceae with organisatorically closely related *bcs* operons. ORFs that encode homologous genes involved in cellulose biosynthesis have the same color. *White arrows* indicate genes with no apparent association with cellulose production. Sequences: *Salmonella typhimurium* (AJ315148); *Escherichia coli* (NC 000913); *Pseudomonas putida* KT2440 (NC 002947); *Gluconacetobacter xylinus* (AB015802); *Pseudomonas fluorescens* SJW25 (AY074776); *Xanthomonas axonopodis* pv. *citrii* (NC_003919). Preliminary sequence data for *Pseudomonas syringae* pv. *tomato* were obtained from The Institute for Genomic Research at http://www.tigr.org, for *Burkholderia fungorum* LB400 and *Ralstonia metallidurans* CH34 from the DOE Joint Genome Institute at http://www.jgi.doe.gov/JGI_microbial/html/index.html, for *Burkholderia cenocepacia* J2315, *Salmonella bongori* and *Yersinia enterocolitica* 0:8 from the Sanger center at http://www.sanger.ac.uk/Projects/Microbes/ and for *Klebsiella pneumoniae* from the University of Washington at http://genome.wustl.edu/projects/bacterial/ (*See Color Plate of this figure beginning on page 355*)

encodes a c-di-GMP binding protein, are subsequently arranged in the operon. A cellulase family D member, *bcsZ*, is located within the cellulose biosynthesis operon in *Salmonella* spp. and *E. coli* downstream of *bcsAB*. Although the laboratory strain *E. coli* K-12 is not capable of producing cellulose, a functional cellulase gene could be cloned from the *bcs* operon (Park and Yun 1999). *BcsZ* is required for cellulose biosynthesis in *G. xylinus* (Koo et al. 1998), but in this species the cellulose biosynthesis operon does not contain the gene. However, it is frequently found in close vincinity upstream of the operon.

BcsC, which is required for cellulose biosynthesis *in vivo*, is part of the cellulose biosynthesis operon in Enterobacteriaceae and *G. xylinus*. *BcsC* contains a N-terminal membrane domain and several tetratrico peptide repeats (TRPs) motifs, indicating that it might participate in protein–protein interactions.

Other genes are unique to *G. xylinus* and the Enterobacteriaceae. There is no evidence that Enterobacteriaceae have a homologue of *bcsD*, the last gene in the *bcs* operon in *G. xylinus*, on their chromosome. Also the ccp gene (alternatively called ORF2), located just upstream of *bcsA* is unique to *G. xylinus*. These two genes have been shown to be required for optimal cellulose production and are involved in the control of crystallization by assembling the glucan chains into cellulose I allomorphs (Saxena et al. 1994; Nakai et al. 2002). The lack of those genes in Enterobacteriaceae suggests that the crystallization structure of the glucan chains might be different in those bacteria.

There are also several genes that seem to be unique for cellulose biosynthesis in Enterobacteriaceae. The *yhjQ* gene located upstream of *bcsA* might actually be part of the *bcs* operon as its open reading frame overlaps with the one of *bcsA* by 4 bps (Figure 7-2). Sequence homology search identified *yhjQ* to encode a homologue belonging to the Soj-family, chromosomally encoded ATPases involved in chromosome partitioning and cell division. Consistent with this *in silico* analysis, insertional inactivation of *yhjQ* in *E. coli* K-12 caused abnormal cell division which resulted in incomplete partitioning of the chromosome and filamentous cells at 42°C (Kim et al. 2002). How a *yhjQ* mutation affects cellulose biosynthesis was not studied, since *E. coli* K-12 does not produce cellulose. However, one can envisage that *yhjQ* coordinates cellulose biosynthesis with DNA replication and cell division as the Soj protein coordinates basic cellular processes in *Bacillus subtilis* (Sullivan and Maddock 2000). Why there is a tight coupling of cellulose biosynthesis, and a possible precise positioning of the cellulose synthase complex in the membrane in Enterobacteriaceae, but not in *G. xylinus* remains to be elucidated. Little is known about the role of the divergently transcribed *bcsEFG* operon in cellulose biosynthesis. *BcsE* is predicted to encode for a cytoplasmatically located protease, while *bcsG* encodes for an inner membrane protein of unknown function. Whether *yhjR* (ORF that encodes for a 67 aa long polypeptide) and *yhjT* (63 aa) encode for functional polypeptides involved in cellulose biosynthesis, remains to be shown. Both genes are conserved, since they are also found in the *Pseudomonas putida* cellulose biosynthesis operon (Figure 7-3), although not annotated in the original sequence information.

3 REGULATION OF THE EXPRESSION OF THE *bcsABZC* OPERON

The organization of the *yhjRQbcsABZC-bcsEFG* operons is depicted in Figure 7-2. The two groups of convergently transcribed ORFs are either separated by only a few bps or overlap. Computional analysis in *E. coli* predicts four transcriptional units with one sigma 70-like promoter each*; yhjRQ*, *bcsABZ*, *bcsC* and *bcsEFG*. However, the transcriptional regulatory pattern must be more complex, since, for example, complementation of polar mutations in *bcsZ* does not readily restore cellulose biosynthesis (our unpublished results).

In *S. typhimurium* the transcriptional regulation of the *bcsABZC* operon by environmental conditions was studied with *lacZ*-fusions located in *bcsA* and *bcsC* (Zogaj et al. 2001). Transcription of both genes is growth phase dependent with approximately threefold higher expression in the stationary than logarithmic growth phase. Expression of both genes was highest under aerobic conditions in liquid culture (Luria Broth (LB) without the salt component) and decreased more than twofold under all other conditions such as microaerophilic or anaerobic growth conditions, plate-growth, iron depletion, on minimal medium, under high salt and in rich medium supplemented with glucose as carbon source. However, under all growth conditions substantial transcriptional activity has been found, so that transcription of the *bcsABZC* operon can be considered constitutive. Most surprisingly, transcription of neither *bcsA* nor *bcsC* was dependent on positive regulators of cellulose biosynthesis (see below), *rpoS*, the starvation sigma factor in stationary phase, and *csgD*, a transcriptional response regulator, suggesting that the activation of cellulose biosynthesis takes place at a post-transcriptional level. Those data indicate a situation similar as in *G. xylinus* where membrane fractions showed cellulose biosynthesis activity despite no obvious cellulose production of corresponding whole cells (Saxena and Brown, Jr. 1995) suggesting that synthesis of the cellulose synthase and its activation are separated events. Constitutive expression of the structural genes, but a missing factor for post-transcriptional activation of cellulose biosynthesis could be an explanation for this phenomenon.

4 REGULATION OF CELLULOSE BIOSYNTHESIS

A regulatory cascade leading to the activation of cellulose biosynthesis has been established in *S. typhimurium* (Figure 7-4a). In plate-grown cells, which harbor an intact cellulose biosynthesis operon, expression of AdrA from a low copy number plasmid is sufficient to initiate temperature independent cellulose biosynthesis (Zogaj et al. 2001).

In the natural situation, the chromosomally encoded *adrA* itself is tightly regulated by CsgD (formerly called AgfD), a response regulator of the UhpA (FixJ) family on the transcriptional level (Römling et al. 2000). Under all environmental conditions examined, throughout the bacterial growth phase in liquid culture, plate-growth at 28°C and 37°C, on minimal medium, under iron

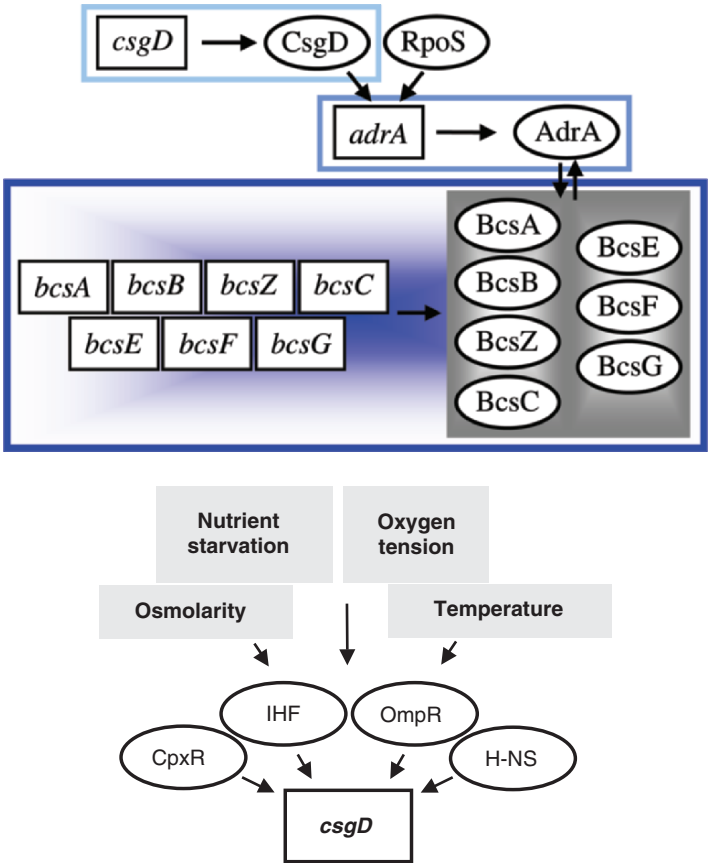


Figure 7-4. (**a**) The cellulose biosynthesis module with regulatory units. Cellulose biosynthesis is activated by the direct or indirect interaction of the AdrA protein with gene products from the *bcs* operons. Transcription of *adrA* is conducted by CsgD together with the second principal sigma factor in stationary phase, RpoS. (**b**) Environmental conditions and global regulatory proteins that influence expression of *csgD* (*See Color Plate of this figure beginning on page 355*)

depletion, under anaerobic conditions and under aerobic conditions in liquid medium, stationary-phase expression of *adrA* was strictly dependent on *csgD* expression. At high salt concentrations, *csgD* is not expressed and hence, *adrA* expression does not take place. However, CsgD is only required for transcription of *adrA*, but not for further steps downstream in the regulatory cascade leading to cellulose biosynthesis (Zogaj et al. 2001).

Another gene, which is required for cellulose biosynthesis is *rpoS* (Römling et al. 1998b, 2000). *RpoS* encodes for the second principal sigma factor in stationary phase, conferring survival properties to various stresses (Hengge-Aronis 1999). As in the case of *csgD*, *rpoS* is solely required for the transcriptional activation of *adrA*,

but not for further steps downstream in the regulatory cascade leading to cellulose biosynthesis. Consequently, the transcription of *bcsA* and *bcsC* was not dependent on *rpoS* (Zogaj et al. 2001).

However, activation of cellulose biosynthesis via the *csgD-adrA* regulatory pathway is not absolute. We have described one *S. enteritidis* strain, where cellulose biosynthesis is at least partially independent of *csgD*, *adrA* and *rpoS* (Römling et al. 2003). In *S. enteritidis*, specific environmental conditions have been reported where cellulose biosynthesis is at least independent of *csgD* and *rpoS* (Solano et al. 2002). In adherence test medium (ATM), which contains sufficient amount of carbon source, but does not support the growth of bacteria, since it is lacking phosphate and the divalent cation magnesium, cellulose was produced after less than 40 min of incubation. Addition of inorganic sources of phosphate, nitrogen, and sulfur or iron, magnesium, or calcium ions abolished cellulose production (Solano et al. 1998). Again, those circumstances resemble activation of cellulose biosynthesis in *G. xylinus* where cellulose biosynthesis was observed in resting cells (Hestrin 1954).

5 REGULATION OF *csgD* EXPRESSION

Under most environmental conditions, cellulose biosynthesis is activated through the transcriptional regulator *csgD*. CsgD itself is regulated on the transcriptional and presumably also post-transcriptional level by a wide variety of environmental stimuli (Gerstel and Römling 2001; Römling et al. 1998b). Nutrient depletion, oxygen tension, osmolarity, and temperature are major factors that influence the expression of *csgD* (Figure 7-4b). *CsgD* is expressed when the bacterial cells reach a certain density in the culture, approximately 3×10^8 cells/ml. Under those conditions, starvation by various nutrients like phosphorus and nitrogen is the trigger to increase expression of *csgD*. Oxygen tension regulates *csgD* expression in a complex interplay with the nutrient source. Reduced oxygen tension (microaerophilic conditions) provided an optimum of *csgD* expression in rich medium, while aerobic conditions are optimal for *csgD* expression in a medium limited for nutrients (minimal medium). High osmolarity abolishes the expression of *csgD*. Cellulose biosynthesis seems to follow this expression pattern via *adrA* expression as judged from the phenotype on plates, since at the moment, no quantitative assay for cellulose production in *S. typhimurium* is available.

Another level of regulation of *csgD* expression is mutations on the chromosome. As the well studied isolates *S. typhimurium* ATCC14028 and SR-11, virtually all *S. typhimurium* and *S. enteritidis* strains isolated from human infections, animal and food express the *csgD* gene in a temperature regulated way, whereby transcription is observed at temperatures below 30°C, but not at 37°C (Römling et al. 1998a, 2003). However, independently isolated mutants of ATCC14028 and SR-11 showed a temperature deregulated expression of *csgD*. The two mutants had individual point mutations in the *csgD* promoter region, which conferred the derepressed expression (Römling et al. 1998b). Consequently, cellulose

biosynthesis was temperature independent in the mutants, but expressed only at temperatures below 30°C in the wild type.

On the molecular level, several globally acting DNA-binding protein bind to the *csgD* promoter region and build up a three-dimensional nucleoprotein complex to ensure tight regulation of expression (Prigent-Combaret et al. 2001; Gerstel et al. 2003). As a key regulator, the transcriptional regulator *ompR* is absolutely required, yet not sufficient to confer *csgD* expression. The alternative sigma factor *rpoS* is also required, but only in strains that express *csgD* temperature regulated.

6 FUNCTION OF AdrA

The function of CsgD in cellulose biosynthesis is solely the activation of *adrA*. But what are the specific mechanisms with which AdrA regulates cellulose biosynthesis? *AdrA,* in a monocistronic operon, encodes for a 371 aa long protein. It contains a highly hydrophobic N-terminal integral membrane domain, named MASE2 (membrane associated sensor; (Nikolskaya et al. 2003)), and a C-terminal GGDEF domain, also called DUF1 (domain of unknown function). The GGDEF domain is considered to be the effector domain. Also in other bacteria, *Rhizobiium leguminosarum* bv. *trifolii* and *Pseudomonas fluorescens*, GGDEF domain containing proteins activate cellulose biosynthesis (Ausmees et al. 1999, 2001; Spiers et al. 2002). In *G. xylinus*, two highly homologous proteins with the domain structure <sensory domain-GGDEF-EAL> confer either cyclization of two GTP molecules or cleavage of c-di-GMP (Tal et al. 1998).

By sequence similarity in combination with molecular modeling, the GGDEF domain has been suggested to confer nucleotide cyclization activity (Pei and Grishin 2001), but this function has to be experimentally proven. However, it is not very far fetched to speculate that the GGDEF domain might synthesize the cyclic nucleotide c-di-GMP, which has been identified as the allosteric activator for cellulose biosynthesis in *G. xylinus* (Ross et al. 1991).

The GGDEF/DUF1 domain is highly abundant, over 750 proteins of the presently sequenced 256 microbial genomes contain this domain, although a phenotype or function has been reported for only a handful of those genes mainly in the context of studies concerning bacterial development (Hecht and Newton 1995; Jones et al. 1999; Gronewold and Kaiser 2001; Boles and McCarter 2002). If the function of the GGDEF domain is in fact the production of c-di-GMP, those data mean that c-di-GMP is an important, yet unidentified global second messenger in bacteria. However, GGDEF domain proteins are not equally distributed among the microorganisms (Galperin et al. 2001). GGDEF domain containing proteins are highly abundant in Gram-negative free-living bacteria, which also can have pathogenic potential. In Gram-positive bacteria this domain occurs in much lower numbers, if at all; obligate parasites have only one or no copy of the GGDEF domain; and archaea miss this domain at all.

S. typhimurium has 12 copies of this domain. That all 12 proteins with this domain are involved in the regulation of cellulose biosynthesis, but under different

environmental conditions, seems unlikely, but cannot be completely excluded. Alternatively, other cellular processes might be regulated by GGDEF domain containing proteins.

7 OCCURRENCE OF THE CELLULOSE BIOSYNTHESIS OPERON AMONG ENTEROBACTERIAL SPECIES

The cellulose biosynthesis operon is not present in all species within the family of Enterobacteriaceae, nor is cellulose constitutively expressed by those species that harbor the genetic information. Whole genome sequence analysis revealed that the cellulose biosynthesis operon is present in *S. enterica* serovars and *S. bongori*, in *E. coli, Shigella* spp., which are actually subspecies of *E. coli*, *Klebsiella pneumoniae*, *Erwinia chrysantemii* strain 3937, *Erwinia carotovora* subsp. *atroseptica* SCRI 1043.06 and *Yersinia enterocolitica* type O:8. In addition, work in our group has detected *bcsA*, the catalytic subunit of cellulose synthase, in *Citrobacter* spp., *Citrobacter freundii*, *Citrobacter koserilfarmeri*, *Enterobacter aerogenes*, *Enterobacter cloacae, Enterobacter sakazakii*, *Klebsiella oxytoca*, and *Raoultella ornithinolytica* (Zogaj et al. 2003).

But where had *S. typhimurium* or, more accurate, an enterobacterial common ancestor acquired this operon? The G+C content of the *yhjRQbcsABZC* operon in *S. typhimurium* is 58%, whereby the average G+C content of the *S. typhimurium* genome is 53%. This fact suggested that the cellulose biosynthesis operon comes from a bacterial species with higher G+C content. Actually, highly homologous sequences, not only on the protein, but also on the nucleotide level are found in the unrelated saprophytic soil bacterium *Pseudomonas putida* KT2440 (Figure 7-3). However, already in *P. putida* only the core genes *yhjRQbcsABZC* are in the same order as in Enterobacteriaceae. The sequence of the *bcsEFG* operon is rearranged through reversion. In the plant pathogen *Pseudomonas syringae* pv. *tomato* and the rhizosphere isolate *P. fluorescens* SWB25, there are additional variants of the cellulose biosynthesis operon (Spiers et al. 2002). Downstream of the respective *bcsC* homologue there is a gene cluster called *wssGHI*, which is paralogous to a gene cluster carrying out acetylation of another exopolysaccharide, namely alginate, in *P. aeruginosa*. It has been suggested that the *wssGHI* gene cluster confers acetylation of the glucan chain, whereby a second copy of the Soj-homologue *yhjQ*, *wssJ*, located downstream of the *wssGHI* genes in *P. fluorescens*, positions the enzyme complex nearby the cellulose synthase.

Also other soil bacteria like species from the medically, agriculturally and environmentally important Burkholderia complex, *Burkholderia fungorum* LB400 and *Burkholderia cenocepacia* J2315 (previously called *Burkholderia* (*Pseudomonas*) *cepacia*) and the heavy metal resistant *Ralstonia metallidurans* CH34 (previously called *Ralstonia eutropha* and *Alcaligenes eutrophus*) and the plant pathogen *Xanthomonas axonopodis* pv. *citrii* (da Silva et al. 2002) harbor variations of the cellulose biosynthesis operon. The similarities in organization of the *bcs* operon are reflected by the distances of the respective catalytic subunits of the cellulose synthase BcsA in the phylogram (compare Figure 7-2 and Figure 7-5).

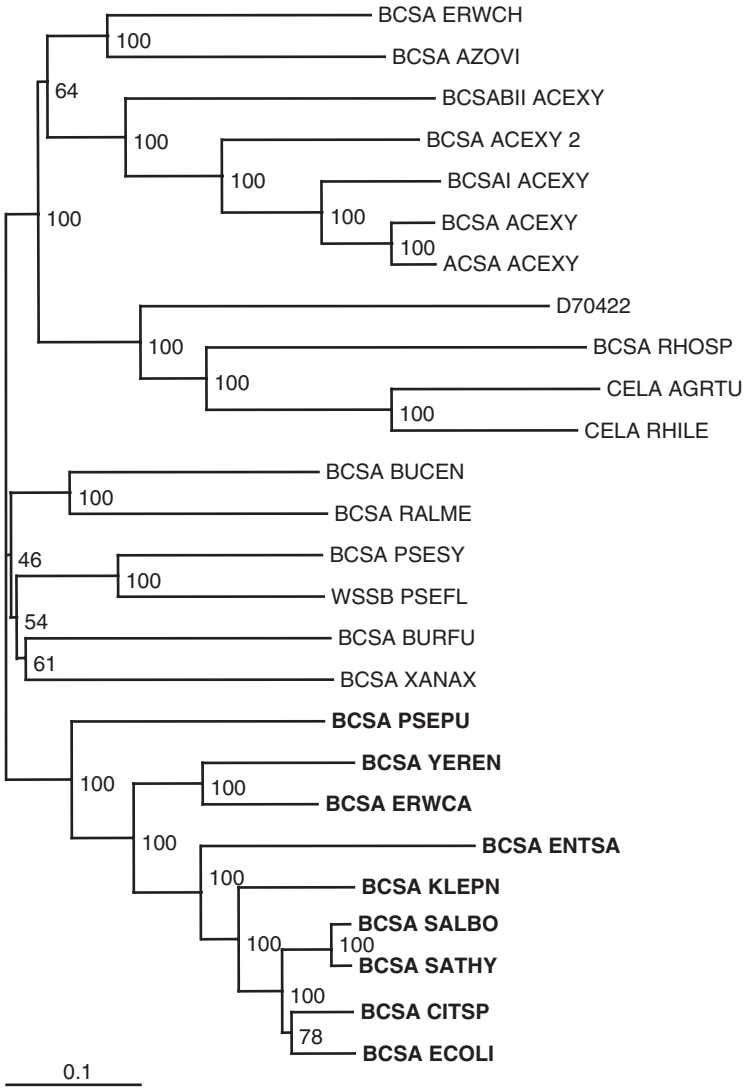


Figure 7-5. Phylogram of relationships of bacterial cellulose synthase BcsA. All enterobacterial BcsA proteins besides *Erwinia chrysantemii* cluster with BcsA from *Pseudomonas putida* KT2440. Sequences were aligned in Clustalx using the default mode and processed manually afterwards. The tree was constructed using the neighborhood joining (NJ) method of Saitou and Nei and subjected to 10,000 bootstrap trials. Numbers on nodes indicate percentages of bootstrap values. The tree was drawn with TreeView. The scale indicates amino acid substitutions per site. Protein sequences used in Figure 7-5: D70422 (*Aquifex aeolicus* BcsA); BCSA_ACEXY (BAA31463.1); ACSA_ACEXY (P19449); BCSAI_ACEXY (BAA77585.1); BCSA_ACEXY_2 (P21877); BCSABII_ACEXY (BAA77593.1); CELA_AGRTU (NP_357298.1); CELA_RHILE (AAC41436.1); BCSA_PSEFL (AAL71842); BCSA_PSEPU (NP_744780); BCSABCSA_SATHY (CAC44015.1); BCSA_ECOLI (P37653); BCSA_ENTSA (CAD56669); BCSA_CITSP (CAD56668); BCSA_XANAX (AAM38358). BCSA_PSESY and BCSA_ERWCH were from The Institute for Genomic Research at http://www.tigr.org, BCSA_AZOVI, BCSA_BURFU, BCSA_RALME and BCSA_RHOSP from the DOE Joint Genome Institute at http://www.jgi.doe.gov/JGI_microbial/html/index.html and BCSA_YEREN, BCSA_ERWCA, BCSA_SALBO and BCSA_BUCEN from the Sanger center at http://www.sanger.ac.uk/Projects/Microbes/ and BCSA_KLEPN from the University of Washington at http://genome.wustl.edu/projects/bacterial/

8 DIFFERENTIAL EXPRESSION OF CELLULOSE AMONG *Enterobacteriaceae*

Study on the expression of cellulose in over 800 *S. enterica* isolates from human infections, food and animals revealed serovar specific expression patterns, which could be correlated with disease severity in the respective hosts (Römling et al. 2003). *S. typhimurium* and *S. enteritidis* isolates consistently expressed cellulose at 28°C on agar plates, while isolates of the serovars *Salmonella typhi*, *Salmonella choleraesuis* and of the variant *S. typhimurium* var. Copenhagen did not express cellulose. However, when expressed by strains of serovars Typhimurium and Enteritidis, we observed that cellulose is always coexpressed with proteinaceous appendaces, the curli fimbriae (see below).

Cellulose expression showed a more variable pattern in *E. coli*, an important inhabitant of the human gastrointestinal tract as well as a pathogen. A substantial proportion of commensal isolates of *E. coli* expressed cellulose at 28°C and/or 37°C, although always together with curli fimbriae (our unpublished data). In uropathogenic isolates, however, cellulose expression could occur without concomitant curli expression (our unpublished data). Otherwise, no systematic investigations about the expression of cellulose have been carried out in *E. coli* strains. Bacterial species isolated from the gastrointestinal tract showed a variable, genus specific expression of cellulose (Zogaj et al. 2003). While *Citrobacter* isolates showed temperature dependent expression of cellulose, *Enterobacter* isolates displayed temperature independent or preferential expression at 37°C. *Klebsiella* isolates did not express cellulose when plate-grown. However, we recently discovered few cellulose-positive colonies derived from *K. pneumoniae* strain DSM12082, which was isolated from a pond (Zogaj et al. 2001).

9 COEXPRESSION OF CELLULOSE WITH CURLI FIMBRIAE

In *S. typhimurium* and *S. enteritidis* serovars and in commensal *E. coli* strains cellulose is usually coexpressed with curli fimbriae, a proteinaceous component. As cellulose biosynthesis, expression of curli fimbriae is regulated by CsgD on the transcriptional level. In fact, the *csgD* gene is part of the *csgDEFG-csgBA(C)* biosynthesis operon and presumably directly activates the *csgBA(C)* operon (Römling et al. 1998a).

The major characteristic of curli fimbriae is their binding capacity to diverse substrates, ranging from proteins present in the human host to hydrophilic and hydrophobic abiotic surfaces such as glass and polystyrene (Ben Nasr et al. 1996; Olsen et al. 1989; Austin et al. 1998; Herwald et al. 1998; Olsen et al. 1998; Römling et al. 1998b). Not unexpectedly, curli fimbriae also interact with the glucan chains of cellulose coexpressed with curli fimbriae on the bacterial surface (Zogaj et al. 2001; White et al. 2003). The interaction is evident in that the otherwise free-floating glucan bundles of cellulose are tightly wrapped around

the bacterial cells when curli fimbriae were coexpressed. Through those, presumable noncovalent, interactions the properties of the bacterial colonies were change dramatically (Figure 7-1a). The bacteria are firmly interconnected in a rigid network and their surface is highly hydrophobic. Actually, cellulose from those bacteria was not digestible with cellulase even after extended hours using high concentrations of enzyme, although the digestion of cellulose alone was no problem (Zogaj et al. 2001).

The interaction and distinct roles of cellulose and curli fimbriae can be seen in various assays. Electron microscopy studies revealed that fine cellulose fibers were produced peritrichously by *S. typhimurium* (Römling and Lunsdorf 2003). On the other hand, curli fimbriae appear as 2–3 nm wide, curled appendages. When expressed together, the material appears diffuse (White et al. 2003). Distinct biofilms are formed by cellulose and curli fimbriae (Römling et al. 2001). Cellulose mainly provided loose adherence at the air–liquid interface, while curli fimbriae mediated tight interactions below the surface of the liquid. The two extracellular matrix components also fullfill different roles in bacterial cell–cell interaction. Cellulose fibers provide elastic, long-range interconnections, while curli fimbriae mediate rigid, but easily breakable connections between individual cells. It can be concluded that cellulose fibers have a structural function, while curli fimbriae provide stabilization. Because of the functional similarities of cellulose and accessory components to the plant system, we have referred to the phenotype of the microbial colonies as "bacterial wood" (Zogaj et al. 2001).

10 CONCLUSIONS

Recently, the molecular basis of cellulose biosynthesis has been detected in *S. typhimurium* and other Enterobacteriaceae. With this discovery, however, new questions did arise concerning various aspects such as the mode of cellulose biosynthesis, its regulation, function, epidemiology, structure and interaction of cellulose with other components. At present, answers are only partially available, if at all. The availability of well characterized and fully sequenced strains together with efficient tools for genetic manipulation, however, gives hope that fairly soon light will be shed at least to some aspects of cellulose biosynthesis in Enterobacteriaceae.

Acknowledgments

Research of U.R. is supported by the Karolinska Institutet ("Elitforskartjänst"), and Vetenskapsradet grants B5107-20005701/2000 and K2002-06X-14232-01A. I thank Xhavit Zogaj for experimental support and preparation of Figure 7.1b and 7-2. Release of sequence data prior to publication from the DOE Joint Genome Institute, the Sanger Center, TIGR and University of Washington is gratefully acknowledged.

REFERENCES

Ausmees N., Jonsson H., Hoglund S., Ljunggren H., and Lindberg M. 1999. Structural and putative regulatory genes involved in cellulose synthesis in *Rhizobium leguminosarum* bv. trifolii. Microbiology 145:1253–1262.

Ausmees N., Mayer R., Weinhouse H., Volman G., Amikam D., Benziman M., and Lindberg M. 2001. Genetic data indicate that proteins containing the GGDEF domain possess diguanylate cyclase activity. FEMS Microbiol Lett 204:163–167.

Austin J.W., Sanders G., Kay W.W., and Collinson S.K. 1998. Thin aggregative fimbriae enhance *Salmonella enteritidis* biofilm formation. FEMS Microbiol Lett 162:295–301.

Ben Nasr A., Olsen A., Sjobring U., Muller-Esterl W., and Bjorck L. 1996. Assembly of human contact phase proteins and release of bradykinin at the surface of curli-expressing *Escherichia coli*. Mol Microbiol 20:927–935.

Boles B.R. and McCarter L.L. 2002. *Vibrio parahaemolyticus* ***scrABC***, a novel operon affecting swarming and capsular polysaccharide regulation. J Bacteriol 184:5946–5954.

da Silva A.C., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R., Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F., Alves L.M., do Amaral A.M., Bertolini M.C., Camargo L.E., Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P., Cicarelli R.M., Coutinho L.L., Cursino-Santos J.R., El-Dorry H., Faria J.B., Ferreira A.J., Ferreira R.C., Ferro M.I., Formighieri E.F., Franco M.C., Greggio C.C., Gruber A., Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G., Lemos M.V., Locali E.C., Machado M.A., Madeira A.M., Martinez-Rossi N.M., Martins E.C., Meidanis J., Menck C.F., Miyaki C.Y., Moon D.H., Moreira L.M., Novo M.T., Okura V.K., Oliveira M.C., Oliveira V.R., Pereira H.A., Rossi A., Sena J.A., Silva C., de Souza R.F., Spinola L.A., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I., Trindade dos Santos M., Truffi D., Tsai S.M., White F.F., Setubal J.C., and Kitajima J.P. 2002. Comparison of the genomes of two ***Xanthomonas*** pathogens with differing host specificities. Nature 417:459–463.

Galperin M.Y., Nikolskaya A.N., and Koonin E.V. 2001. Novel domains of the prokaryotic two-component signal transduction systems. FEMS Microbiol Lett 203:11–21.

Gerstel U. and Römling U. 2001. Oxygen tension and nutrient starvation are major signals that regulate ***agfD*** promoter activity and expression of the multicellular morphotype in *Salmonella typhimurium*. Environ Microbiol 3:638–648.

Gerstel U., Park C., and Romling U. 2003. Complex regulation of ***csgD*** promoter activity by global regulatory proteins. Mol Microbiol 49:639–654.

Gronewold T.M. and Kaiser D. 2001. The ***act*** operon controls the level and time of C-signal production for *Myxococcus xanthus* development. Mol Microbiol 40: 744–756.

Hecht G.B. and Newton A. 1995. Identification of a novel response regulator required for the swarmer-to-stalked-cell transition in *Caulobacter crescentus*. J Bacteriol 177:6223–6229.

Hengge-Aronis R. 1999. Interplay of global regulators and cell physiology in the general stress response of *Escherichia coli*. Curr Opin Microbiol 2:148–152.

Herwald H., Morgelin M., Olsen A., Rhen M., Dahlback B., Muller-Esterl W., and Bjorck L. 1998. Activation of the contact-phase system on bacterial surfaces–a clue to serious complications in infectious diseases. Nat Med 4:298–302.

Hestrin S. and Schramm, M. 1954. Synthesis of cellulose by *Acetobacter xylinum*. Biochem J 58:345–352.

Jones H.A., Lillard J.W., Jr., and Perry R.D. 1999. HmsT, a protein essential for expression of the haemin storage (Hms+) phenotype of *Yersinia pestis*. Microbiology 145 (Pt 8):2117–2128.

Kim M.K., Park S.R., Cho S.J., Lim W.J., Ryu S.K., An C.L., Hong S.Y., Park Y.W., Kahng G.G., Kim J.H., Kim H., and Yun H.D. 2002. The effect of a disrupted ***yhjQ*** gene on cellular morphology and cell growth in *Escherichia coli*. Appl Microbiol Biotechnol 60:134–138.

Koo H.M., Song S.H., Pyun Y.R., and Kim Y.S. 1998. Evidence that a beta-1,4-endoglucanase secreted by *Acetobacter xylinum* plays an essential role for the formation of cellulose fiber. Biosci Biotechnol Biochem 62:2257–2259.

Nakai T., Nishiyama Y., Kuga S., Sugano Y., and Shoda M. 2002. ORF2 gene involves in the construction of high-order structure of bacterial cellulose. Biochem Biophys Res Commun 295:458–462.

Nikolskaya A.N., Mulkidjanian A.Y., Beech I.B., and Galperin M.Y. 2003. MASE1 and MASE2: two novel integral membrane sensory domains. J Mol Microbiol Biotechnol 5:11–16.

Olsen A., Jonsson A., and Normark S. 1989. Fibronectin binding mediated by a novel class of surface organelles on *Escherichia coli*. Nature 338:652–655.

Olsen A., Wick M.J., Morgelin M., and Bjorck L. 1998. Curli, fibrous surface proteins of *Escherichia coli*, interact with major histocompatibility complex class I molecules. Infect Immun 66:944–949.

Park Y.W. and Yun H.D. 1999. Cloning of the *Escherichia coli* endo-1,4-D-glucanase gene and identification of its product. Mol Gen Genet 261:236–241.

Pei J. and Grishin N.V. 2001. GGDEF domain is homologous to adenylyl cyclase. Proteins 42:210–216.

Prigent-Combaret C., Brombacher E., Vidal O., Ambert A., Lejeune P., Landini P., and Dorel C. 2001. Complex regulatory network controls initial adhesion and biofilm formation in *Escherichia coli* via regulation of the ***csgD*** gene. J Bacteriol 183:7213–7223.

Ross P., Mayer R., and Benziman M. 1991. Cellulose biosynthesis and function in bacteria. Microbiol Rev 55:35–58.

Römling U. and Lunsdorf H. 2003. Cellulose biosynthesis in *Enterobacteriaceae*. Cellulose 11:413–418.

Römling U., Bian Z., Hammar M., Sierralta W.D., and Normark S. 1998a. Curli fibers are highly conserved between *Salmonella typhimurium* and *Escherichia coli* with respect to operon structure and regulation. J Bacteriol 180:722–731.

Römling U., Sierralta W.D., Eriksson K., and Normark S. 1998b. Multicellular and aggregative behaviour of *Salmonella typhimurium* strains is controlled by mutations in the ***agfD*** promoter. Mol Microbiol 28:249–264.

Römling U., Rohde M., Olsen A., Normark S., and Reinköster J. 2000. AgfD, the checkpoint of multicellular and aggregative behaviour in *Salmonella typhimurium* regulates at least two independent pathways. Mol Microbiol 36:10–23.

Römling U., Bokranz W., Rabsch W., Zogaj X., Nimtz M., and Tschäpe, H. 2003. Occurrence and regulation of the multicellular morphotype in *Salmonella serovars* important in human disease. Int J Med Microbiol 293:273–285.

Römling U., Rohde M., Olsen A., Normark S., and Reinkoster J. The multicellular morphotypes of *Salmonella typhimurium* and *Escherichia coli* produce cellulose as the second component of the extracellular matrix. Mol Microbiol 39:1452–1463.

Saxena I.M., Kudlicka K., Okuda K., and Brown, Jr. R.M., 1994. Characterization of genes in the cellulose-synthesizing operon (***acs*** operon) of *Acetobacter xylinum*: implications for cellulose crystallization. J Bacteriol 176:5735–5752.

Saxena I.M. and Brown, Jr. R.M., 1995. Identification of a second cellulose synthase gene (*acsAII*) in *Acetobacter xylinum*. J Bacteriol 177:5276–5283.

Solano C., Sesma B., Alvarez M., Humphrey T.J., Thorns C.J., and Gamazo C. 1998. Discrimination of strains of *Salmonella enteritidis* with differing levels of virulence by an *in vitro* glass adherence test. J Clin Microbiol 36:674–678.

Solano C., Garcia B., Valle J., Berasain C., Ghigo J.M., Gamazo C., and Lasa I. 2002. Genetic analysis of *Salmonella enteritidis* biofilm formation: critical role of cellulose. Mol Microbiol 43:793–808.

Spiers A.J., Kahn S.G., Bohannon J., Travisano M., and Rainey P.B. 2002. Adaptive divergence in experimental populations of *Pseudomonas fluorescens*. I. Genetic and phenotypic bases of wrinkly spreader fitness. Genetics 161:33–46.

Sullivan S.M. and Maddock J.R. 2000. Bacterial sporulation: pole-to-pole protein oscillation. Curr Biol 10: R159–161.

Tal R., Wong H.C., Calhoon R., Gelfand D., Fear A.L., Volman G., Mayer R., Ross P., Amikam D., Weinhouse H., Cohen A., Sapir S., Ohana P., and Benziman M. 1998. Three ***cdg*** operons control cellular turnover of cyclic di-GMP in *Acetobacter xylinum*: genetic organization and occurrence of conserved domains in isoenzymes. J Bacteriol 180:4416–4425.

White A.P., Gibson D.L., Collinson S.K., Banser P.A., and Kay W.W. 2003. Extracellular polysaccharides associated with thin aggregative fimbriae of *Salmonella enterica* and *Serovar enteritidis*. J Bacteriol 185:5398–5407.

Zogaj X., Nimtz M., Rohde M., Bokranz W., and Römling U. 2001. The multicellular morphotypes of *Salmonella typhimurium* and *Escherichia coli* produce cellulose as the second component of the extracellular matrix. Mol Microbiol 39:1452–1463.

Zogaj X., Bokranz W., Nimtz M., and Römling U. 2003. Production of cellulose and curli fimbriae by members of the family *Enterobacteriaceae* isolated from the human gastrointestinal tract. Infect Immun 71:4151–4158.