Available online at www.sciencedirect.com
ScienceDirect
Journal of Biotechnology 134 (2008) 33–45
Journal of BIOTECHNOLOGY
www.elsevier.com/locate/jbiotec

The genome of *Xanthomonas campestris* pv. campestris B100 and its use for the reconstruction of metabolic pathways involved in xanthan biosynthesis

Frank-Jörg Vorhölter[a], Susanne Schneiker[a], Alexander Goesmann[c], Lutz Krause[c], Thomas Bekel[c], Olaf Kaiser[d,1], Burkhard Linke[c], Thomas Patschkowski[d], Christian Rückert[d], Joachim Schmid[e], Vishaldeep Kaur Sidhu[b], Volker Sieber[f], Andreas Tauch[d], Steven Alexander Watt[b], Bernd Weisshaar[d], Anke Becker[d], Karsten Niehaus[b], Alfred Pühler[a,*]

[a] *Universität Bielefeld, Biologie VI (Genetik), Universitätsstr. 25, D-33615 Bielefeld, Germany*
[b] *Universität Bielefeld, Biologie 7 (Proteom und Metabolomforschung), Germany*
[c] *Universität Bielefeld, CeBiTec, Bioinformatics Resource Facility, Germany*
[d] *Universität Bielefeld, CeBiTec, Institut für Genomforschung und Systembiologie, Germany*
[e] *Insilico Biotechnology AG, Nobelstr. 15, D-70569 Stuttgart, Germany*
[f] *Degussa GmbH, Paul-Baumann-Str. 1, D-45772 Marl, Germany*

Received 7 September 2007; received in revised form 4 December 2007; accepted 24 December 2007

Abstract

The complete genome sequence of the *Xanthomonas campestris* pv. campestris strain B100 was established. It consisted of a chromosome of 5,079,003 bp, with 4471 protein-coding genes and 62 RNA genes. Comparative genomics showed that the genes required for the synthesis of xanthan and xanthan precursors were highly conserved among three sequenced *X. campestris* pv. campestris genomes, but differed noticeably when compared to the remaining four *Xanthomonas* genomes available. For the xanthan biosynthesis genes *gumB* and *gumK* earlier translational starts were proposed, while *gumI* and *gumL* turned out to be unique with no homologues beyond the *Xanthomonas* genomes sequenced. From the genomic data the biosynthesis pathways for the production of the exopolysaccharide xanthan could be elucidated. The first step of this process is the uptake of sugars serving as carbon and energy sources wherefore genes for 15 carbohydrate import systems could be identified. Metabolic pathways playing a role for xanthan biosynthesis could be deduced from the annotated genome. These reconstructed pathways concerned the storage and metabolization of the imported sugars. The recognized sugar utilization pathways included the Entner-Doudoroff and the pentose phosphate pathway as well as the Embden-Meyerhof pathway (glycolysis). The reconstruction indicated that the nucleotide sugar precursors for xanthan can be converted from intermediates of the pentose phosphate pathway, some of which are also intermediates of glycolysis or the Entner-Doudoroff pathway. Xanthan biosynthesis requires in particular the nucleotide sugars UDP-glucose, UDP-glucuronate, and GDP-mannose, from which xanthan repeat units are built under the control of the *gum* genes. The updated genome annotation data allowed reconsidering and refining the mechanistic model for xanthan biosynthesis.

Keywords: *X. campestris* pv. campestris genome sequence; Xanthan biosynthesis; Nucleotide sugars synthesis genes; Comparative genomics; Carbohydrate uptake; Carbohydrate metabolism

1. Introduction

Xanthomonas species are members of the γ subdivision of the Gram-negative *Proteobacteria*, which have adopted a plant-associated and usually plant pathogenic lifestyle. Xanthomonads are pathogens of diverse groups of cultivated plants, among

* Corresponding author. Fax: +49 521 1065607.
E-mail address: Puehler@Genetik.Uni-Bielefeld.DE (A. Pühler).
[1] Current address: Roche Diagnostics GmbH, Nonnenwald 2, D-82377 Penzberg, Germany.

0168-1656/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.jbiotec.2007.12.013

them important crops like rice and citrus plants. Based on their specific host ranges, strains of *Xanthomonas campestris* were differentiated into over 140 pathovars (Swings, 1993). The number of pathovars was reduced when *Xanthomonas* strains were reclassified to 20 species (Vauterin et al., 1995). In the course of this reclassification many strains previously accepted as *X. campestris* pathovars became classified as members of distinct species. The largest and most heterogeneous of the new species was *X. axonopodis*, to which amongst others strains of *X. campestris* pv. vesicatoria type A strains were reclassified. *X. campestris* pv. campestris is pathogenic to cruciferous plants, where it causes the "black rot" disease by invading the vascular system of the host plants. The infected crucifers include cultivated *Brassicaceae* like cabbage and cauliflower as well as the model plant *Arabidopsis thaliana*. Besides its importance as a phytopathogen, *X. campestris* pv. campestris is known as the producer of the acid exopolysaccharide xanthan. Xanthan is a heteropolysaccharide with a cellulose-like backbone and trisaccharide side-chains of two mannose and one glucuronate residues that are attached to every second glucose moiety of the main chain (Jansson et al., 1975). There were inconsistent observations concerning the role of xanthan in pathogenicity in the past, and recent data suggesting that xanthan is not required for pathogenicity but contributes to the epiphytic survival of the Xanthomonads (Dunger et al., 2007) have now been challenged for *X. axonopodis* pv. citri (Rigano et al., 2007). Xanthan is commercially produced by fermentation and employed as a thickening agent and emulsifier in the nutritional, pharmaceutical, and oil drilling industries (Becker et al., 1998). In the last years the xanthan production by commercial providers increased significantly, so that it now exceeds 86,000 tons per annum (Seisun, 2006; Sutherland, 1998). Substantial efforts have been put into understanding the xanthan synthesis, the xanthan solution properties and its downstream recovery as well as the kinetics of *X. campestris* pv. campestris growth under xanthan production conditions (Garcia-Ochoa et al., 2000). A model established recently describes the kinetics of xanthan production in a batch reactor (Letisse et al., 2003). Further optimization of the xanthan production process would benefit significantly from utilizing genome data of *X. campestris* pv. campestris model or production strains.

So far genomes of five *Xanthomonas* strains have been sequenced. These include the *X. campestris* pv. campestris strains ATCC 33913 (da Silva et al., 2002) and 8004 (Qian et al., 2005), which cause the "black rot" disease on Brassicaceae. *X. campestris* pv. vesicatoria strain 85-10 (Thieme et al., 2005) is the causative agent of the bacterial spot disease on pepper (*Capsicum* spp.). The *X. oryzae* pv. oryzae strains KACC10331 (Lee et al., 2005) and MAFF 311018 (Ochiai et al., 2005) are both pathogens of rice (*Oryza sativa*) where they cause bacterial blight. The *X. axonopodis* pv. citri strain 306 (da Silva et al., 2002) causes citrus canker, which affects most commercial citrus cultivars. All genomes comprised circular chromosomes of 4,940,217 base pairs (bp) (*X. oryzae* pv. oryzae strain MAFF 311018) to 5,178,466 bp (*X. campestris* pv. vesicatoria strain 85-10), with comparably high G+C contents of 63.7–65.0%. Existence of plasmids was restricted to *X. axonopodis* pv. citri strain 306 and *X. campestris* pv. vesicatoria strain 85-10, while all strains carried multiple insertion sequence (IS) elements. An overview on the general genome features of *X. campestris* strains is given in Table 1. So far the functional analysis of the available genome data was focused on pathogenicity features (da Silva et al., 2002; Lee et al., 2005; Qian et al., 2005; Ochiai et al., 2005; Thieme et al., 2005). Phytopathogenesis was also the focus of genome comparisons related to *Xanthomonas* (da Silva et al., 2002; Moreira et al., 2005; Qian et al., 2005). Although generally mentioned in the genome publications, synthesis of xanthan and related cell-surface carbohydrates was up to now only marginally analyzed at the genomic level.

The experimental analysis of the biosynthesis of xanthan has been carried out mainly with strains that were not subject to genome sequencing so far. The xanthan synthesis is encoded by the *gumBCDEFGHIJKLM* genes, which are located in a single gene cluster of 12 kb that is mainly expressed as an operon from a promoter upstream of the first gene, *gumB* (Katzen et al., 1996; Vojnov et al., 2001). Xanthan synthesis is located at the cell membrane, where defined pentasaccharide repeat units of glucose–glucose–mannose–glucuronate–mannose are built from nucleotide sugars at a polyprenol lipid carrier (Ielpi et al., 1993). This is performed by glycosyltransferases encoded by the *gum* genes D, M, H, K, and I. The outer mannose at the distal position of the side-chain can be pyruvylated; both mannose residues of the repeat units can be acetylated at varying degrees (Stankowski et al., 1993). Completed repeat units are exported and polymerized to form mature xanthan that is subsequently released to the environment. The export and polymerization process is not well understood.

Table 1
Features of *X. campestris* chromosomes

Feature	X. c. pv. campestris B100	X. c. pv. campestris 8004	X. c. pv campestris ATCC 33913	X. c. pv. vesicatoria 8510
Size (bp)	5,079,002	5,148,708	5,076,187	5,178,466
G+C content, %	65.0	64.9	65.0	64.8
CDS				
Predicted no.	4,471	4,273	4,181	4,487
Function assigned	2878	2,671	2,708	2,965
Ribosomal RNA operons	2	2	2	2
Transfer RNAs	54	54	54	54
Insertion sequence elements	59	115	109	58

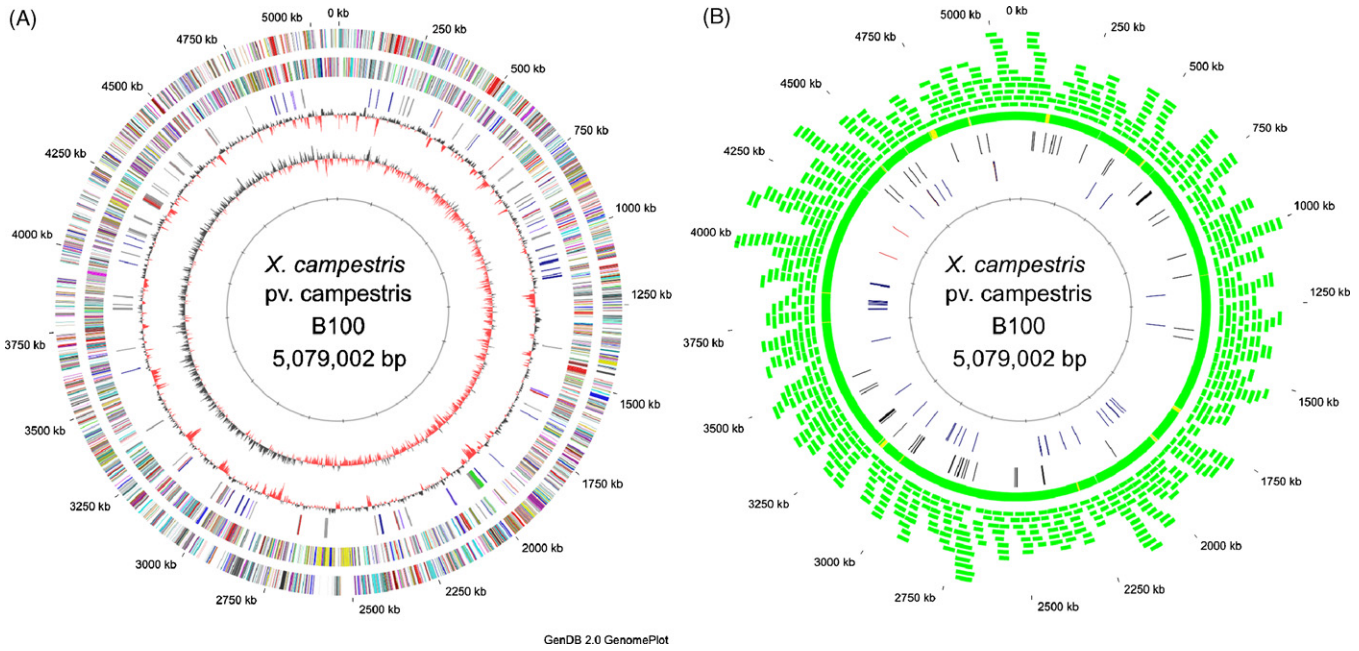

Fig. 1. Circular representations of the *X. campestris* pv. campestris strain B100 genome displaying relevant genome features (A and B) and the validation of the sequence assembly by a fosmid map (B). In the circular genome plot (A) the outermost first circle gives the chromosomal positions of the genome in kilobases (kb), related to the translation start of the *dnaA* gene at the origin of replication. The coding sequences (CDS) of the protein-coding genes are depicted in the second and third circles for the two directions of transcription, respectively. Colors in these circles indicate the functions assigned to the CDS according to the COG categories. CDS related to polysaccharide synthesis and carbohydrate metabolism are shown in the fourth circle. In this circle, the xanthan synthesis genes *gumBCDEFGHIJKLM* are marked in green, while red indicates CDS involved in nucleotide sugars synthesis. CDS related to the synthesis of further cell surface saccharides like LPS or cyclic glucans are given in gray. CDS encoding the carbohydrate metabolism, e.g. glycolysis, pentose phosphate, and Entner-Doudoroff pathways are indicated by dark blue coloring, sugar importer genes are in violet. The fifth circle shows the GC plot. CDS with G+C contents lower than the genome average are marked in red. The GC skew is given in the sixth circle. The innermost seventh circle repeats the scaling ticks of the outermost first circle. In the fosmid map (B) of *X. campestris* pv. campestris B100 the outermost and innermost circles indicate the chromosomal positions similar to the genome plot (A). In the second circle each green arc represents a single fosmid clone mapped to the assembled DNA sequence of the *X. campestris* pv. campestris B100 chromosome. The third circle shows the coverage of the chromosome with fosmid clones. Green symbolizes coverage by more than two fosmids, yellow indicates coverage by one fosmid. The fourth circle gives the positions of the IS elements, while the RNA genes are given in the fifth circle, where the rRNA operons are highlighted in red.

Besides being used to analyze IS elements (Simon et al., 1991) and the TonB-dependent iron uptake (Wiggerich et al., 1997; Wiggerich and Puhler, 2000), *X. campestris* pv. campestris strain B100 has been employed as a model strain to investigate the biosynthesis of the cell surface polysaccharides. Originally, it was isolated in a screening that aimed at identifying a strain suitable for xanthan production. The biosynthesis of xanthan depends on the nucleotide sugar precursors UDP-glucose, GDP-mannose, and UDP-glucuronate, the latter being generated by oxidation of UDP-glucose (Lin et al., 1995). The genes *xanA* and *xanB* encode the bifunctional enzymes phosphoglucomutase/phosphomannomutase and mannose-6-phosphate isomerase/mannose-1-phosphate guanylyltransferase, respectively, which catalyze four reactions and thus most of the steps within the UDP-glucose and GDP-mannose synthesis pathways (Koplin et al., 1992). In addition to the *xanAB* gene products only the glycolysis enzyme phosphoglucose isomerase (Tung and Kuo, 1999) and the UTP-glucose-1-phosphate uridylyltransferase (Wei et al., 1996) are required. At the nucleotide sugar precursors level, synthesis of xanthan is linked to the synthesis of lipopolysaccharides (LPS) (Koplin et al., 1993; Vorholter et al., 2001). The LPS precursor synthesis genes cluster on the *X. campestris* pv. campestris chromosome with the *xanAB* genes (Hotte et al., 1990) and further genes involved in LPS synthesis (Steinmann et al., 1997). All analyses related to this over all 30 kb gene cluster were based on *X. campestris* pv. campestris strain B100. A preliminary analysis of DNA sequence data obtained from sequencing transposon insertion sites of an *X. campestris* pv. campestris strain B100 transposon-tagged shotgun library revealed unexpected differences to the genome of *X. campestris* pv. campestris strain ATCC 33913 (Vorholter et al., 2003). As a consequence it was worthwhile to complete the nucleotide sequence of the *X. campestris* pv. campestris B100, to compare the *Xanthomonas* genes related to xanthan synthesis, and to use this genome data as a basis to reconstruct the metabolic pathways leading to the synthesis of xanthan.

2. Methods

2.1. Whole genome shotgun sequencing

DNA shotgun libraries with insert sizes of 1.5 kb, 2 kb, 3 kb and 4–6 kb were constructed in pUC19 vector (Yanisch-Perron et al., 1985) by Qiagen GmbH. Plasmid clones were end sequenced on ABI3700 capillary sequencing machines (ABI) by Qiagen GmbH and on ABI3730xl DNA analyser (ABI) at Bielefeld University. In addition to this, sequences from a transposon-tagged library (Vorholter et al., 2003) and from another 1.5-kb

shotgun library constructed in pGEM-T Easy by MWG Biotech AG, were incorporated in the genome project. Base calling was carried out using PHRED. High quality reads were defined by a minimal length of 250 bp with an averaging quality value of at least phred 20. Finally, 44,851 high-quality reads, a total of 10,156 (1.36×), 10,191 (1.38×), 6770 (0.87×), 5961 (0.79×) from pUC19 libraries with 1.5-kb, 2-kb, 3-kb and 4–6-kb inserts, respectively and 2385 (0.22×), 9388 (1.12×) from the pGEM-T Easy transposon-tagged and the 1.5-kb shotgun library, respectively, were established (× indicates genome equivalents).

2.2. Sequence assembly and assembly validation

Base calling, quality control and elimination of vector DNA sequences of the shotgun-sequences were performed by using the software package BioMake (Vorholter et al., 2003) as previously described (Kaiser et al., 2003). Sequence assembly was performed by using the PHRAP assembly tool (www.phrap.org) and the CONSED/AUTOFINISH software package (Gordon et al., 1998; Gordon et al., 2001). For gap closure and assembly validation, a fosmid library was constructed and end-sequenced by IIT GmbH using the EpiFOS Fosmid Library Production Kit (Epicentre). Fosmid end sequences were mapped onto the genome sequence by employing the BACCardI software (Bartels et al., 2005). Remaining gaps of the whole genome shotgun assembly were closed by sequencing on shotgun and fosmid clones carried out by IIT GmbH. To obtain a high quality genome sequence, all regions of the consensus sequence were polished to at least phred40 quality by primer walking. Collectively, 576 sequencing reads obtained by high throughput sequencing standard procedures as well as about 400 sequencing reads obtained with a secondary structure dissolving sequencing protocol with a higher melting temperature, were added to the shotgun assembly for finishing and polishing of the genomic sequence. Repetitive elements, i.e. rRNA operons, were sequenced completely by primer walking on fosmid clones.

2.3. Genome annotation, metabolic reconstruction, and xanthan quantification

The finished genome data were annotated by employing the GenDB software (Meyer et al., 2003). Genes were predicted using a combined strategy based on the results of the gene finders GISMO (Krause et al., 2007) and REGANOR (Linke et al., 2006). The functional annotation was performed analogous to the previous *X. campestris* pv. vesicatoria sequencing project (Thieme et al., 2005). In addition relevant information was extracted from the CAZy database (http://www.cazy.org/; Coutinho and Henrissat, 1999). The metabolic pathways related to xanthan synthesis were reconstructed on the basis of the functional annotation by means of the CellDesigner software (Kitano et al., 2005). DNA and protein sequences were compared by BLAST (Altschul et al., 1997). Xanthan was quantified from *X. campestris* pv. campestris cultures grown in Erlenmeyer flasks under identical conditions in TY medium at 30 °C by first precipitating the centrifugation supernatant with cetylpyridinium chloride, then drying the fallout, which was finally weighted.

EMBL accession number: The genome sequence data has been submitted to the EBI EMBL database (accession number AM920689).

3. Results and discussion

3.1. General features of the X. campestris pv. campestris strain B100 genome

The genome of *X. campestris* pv. campestris strain B100 is composed of a circular chromosome of 5,079,002 bp (Fig. 1). A total of 4471 protein-coding sequences (CDS) were predicted within the genome of *X. campestris* pv. campestris strain B100 (Table 1). Further 62 RNA genes were identified, among them 54 tRNAs and 2 rRNA operons (Fig. 1B). Overall 3148 (70%) of the CDS were assigned to functional categories of the COG database (Tatusov et al., 1997). Among these CDS were 270 that encoded conserved hypothetical proteins assigned to the COG functional category S (function unknown), leaving 2878 (64%) CDS with COG-classified known functions assigned. Within the *X. campestris* pv. campestris strain B100 genome several genes related to mobile elements were identified, especially 59 IS elements (Table 1, Fig. 1B).

3.2. Comparison of Xanthomonas genomes

The size of the *X. campestris* pv. *campestris* strain B100 chromosome is roughly equal to those of the *X. campestris* pv. campestris strains ATCC33913 (da Silva et al., 2002) and 8004 (Qian et al., 2005). Like these *X. campestris* pv. campestris strains, and differing from other *Xanthomonas* genomes (Thieme et al., 2005; da Silva et al., 2002), *X. campestris* pv. campestris strain B100 was found to carry no plasmid. Except for an inverted region opposite to the origin of replication the *X. campestris* pv. campestris B100 chromosome is widely collinear to the strain 8004 (Supplemental Fig. 1A). It differs from strain ATCC 33913 by the inversion of a huge chromosomal fragment (Supplemental Fig. 1B). While the CDS of *X. campestris* pv. campestris strains ATCC 33913 and 8004 were highly similar, there were significantly more sequence differences between the CDS of these strains and *X. campestris* pv. campestris strain B100. The genome of *X. campestris* pv. campestris strain B100 contained 496 additional CDS that were not annotated in the other two *X. campestris* pv. campestris strains (Fig. 2), among them 88 CDS also identified in more distant *Xanthomonas* genomes and 408 CDS without representation in any of the sequenced *Xanthomonas* genomes. Most of these new CDS were hypothetical or conserved hypothetical proteins, or matched CDS from plasmids, phages, or IS elements. CDS that had no homologs in the existing annotations of the *X. campestris* pv. campestris strains ATCC 33913 and 8004 were compared to translations of these genomes by means of TBLASTN, which led to 263 hits in the genome of strain ATCC 33913 and to 248 hits in the genome of strain 8004, indicating that about half of the newly identified CDS were also present in the other two *X. campestris* pv. campestris genomes, but have not been annotated so far.

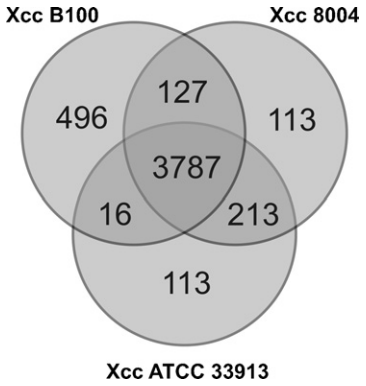

Fig. 2. The distribution of conserved *X. campestris* pv. campestris genes. A Venn diagram illustrates the common gene composition of the sequenced *X. campestris* pv. campestris strains ATCC 33913, 8004 and B100. While the three strains share a common core genome of about 3800 genes, the annotation with the GenDB software, which exploited the new gene prediction tools GISMO and REGANOR, led to the identification of 496 genes not identified in the two other strains.

In the initial stage of this genome project transposon insertion sites in *X. campestris* pv. campestris strain B100 shotgun fragments were mapped to the chromosome of *X. campestris* pv. campestris strain ATCC33913, which revealed 4 genomic regions of *X. campestris* pv. campestris strain ATCC 33913 that were not represented in the *X. campestris* pv. campestris strain B100 transposon-tagged shotgun library (Vorholter et al., 2003). The distinctive region 1 had a size of 37 kb and represented a prophage present in *X. campestris* pv. campestris strain ATCC 33913. This prophage was absent in *X. campestris* pv. campestris strain B100. The distinct region 2 encoded the *X. campestris* pv. campestris strain ATCC 33913 *recDCB* genes. As the finished *X. campestris* pv. campestris strain B100 chromosome sequence revealed homologous *recDCB* genes at an analog location the previous absence of this region may be attributed to cloning bias. Region 3 encoded several metabolic enzymes, which varied in their genomic organization among the two *X. campestris* pv. campestris strains due to differing IS element insertions. Finally, the fourth region not represented was downstream of the *gum* gene cluster that is required for xanthan biosynthesis. All genes from this region were present in all three sequenced *X. campestris* pv. campestris strains, and there was no obvious reason for this region not being covered by transposon-tagged clones.

3.3. Comparative genomics of xanthan production

A comparison of the *Xanthomonas* genomes indicated that the genomic organization of the genes so far characterized to be involved in the synthesis of xanthan (Katzen et al., 1998; Vojnov et al., 2001) and its nucleotide sugar precursors was widely conserved (Fig. 3). In *X. axonopodis* pv. citri and *X. campestris* pv. vesicatoria however an additional gene annotated as *ugd* was located between the *xanAB* and *rmlABCD* nucleotide sugar synthesis genes (Fig. 3A). The function of these additional *ugd* genes remained unclear. Biosynthesis of the exopolysaccharide xanthan is performed by genes of the *gum* cluster. Despite these genes' conservation in size, transcriptional organization, and general similarity, there were differences at the nucleotide sequence level among the *Xanthomonas* strains (Fig. 3). While there were rather few single-nucleotide polymorphisms (SNPs) distinguishing these gene sequences among the *X. campestris* pv. campestris strains, there were more fundamental differences to the sequenced strains of the phylogenetically more distant *Xanthomonas* strains, which probably included small insertion/deletion events. The discrepancies affected in some cases more than 30% of the base sequence. The sequence variation was less manifest but still considerable for the deduced gene products (Supplemental Table 1). Concerning sequence similarity to *X. campestris* pv. campestris the *X. campestris* pv. vesicatoria genes grouped with the genes from more distant *Xanthomonas* strains. It was tempting to find out whether besides variations in the primary structure of the xanthan synthesis genes and gene products, there were also differences in the amounts of xanthan produced by *X. campestris* pv. campestris strains. Xanthan was quantified from cultures of the three sequenced strains ATCC 33913, 8004, and B100, which were grown under identical conditions. The xanthan production in *X. campestris* pv. campestris B100 was raised (Supplementary Fig. 2).

The gene prediction results for *gumK* agree with the recently reported shifting of the translational start of this gene to a position further upstream (Barreras et al., 2004), where in all published *Xanthomonas* genomes a conserved start codon was located (Fig. 3B). The open reading frame of *gumB* was extended when compared to the original GenBank entry U22511, which was obtained from *X. campestris* pv. campestris strain B1459. Upstream of the start codon annotated for strain B1459 this entry included an additional nucleotide that was absent in all the sequenced *Xanthomonas* genomes, where the absence of this reading frame shift caused the first in-frame ATG start codons to be further upstream. For *X. campestris* pv. campestris strain B100 the earlier translational start for *gumB* that was presumed on the basis of predictions by the tools Glimmer, Critica, REGANOR, and GISMO, increased the predicted protein size from 213 to 286 amino acids. The new start was in accordance with the successful knockout of *gumB* by a transposon insertion 15 bp upstream of the previously accepted start codon (Vojnov et al., 1998). The gene prediction results permit earlier translational starts also for the *gumB* genes of the other sequenced Xanthomonads.

3.4. Carbohydrate uptake and interconversion

X. campestris pv. campestris strain B100 expresses a complex set of cell surface carbohydrates. To synthesize these compounds, sugars need to be imported, taken from the glycogen storage or synthesized. Several sugar uptake genes were identified. They encoded fifteen uptake systems for monosaccharides and disaccharides that were classified according to the Transport Classification Database TCDB (Saier, 2000). The uptake systems comprised one ATP-binding cassette (ABC) carbohydrate uptake transporter, two phosphotransferase systems (PTS), seven major facilitator superfamily (MFS) proteins, four glycoside-pentoside-hexuronide:cation (GPH) symporters,

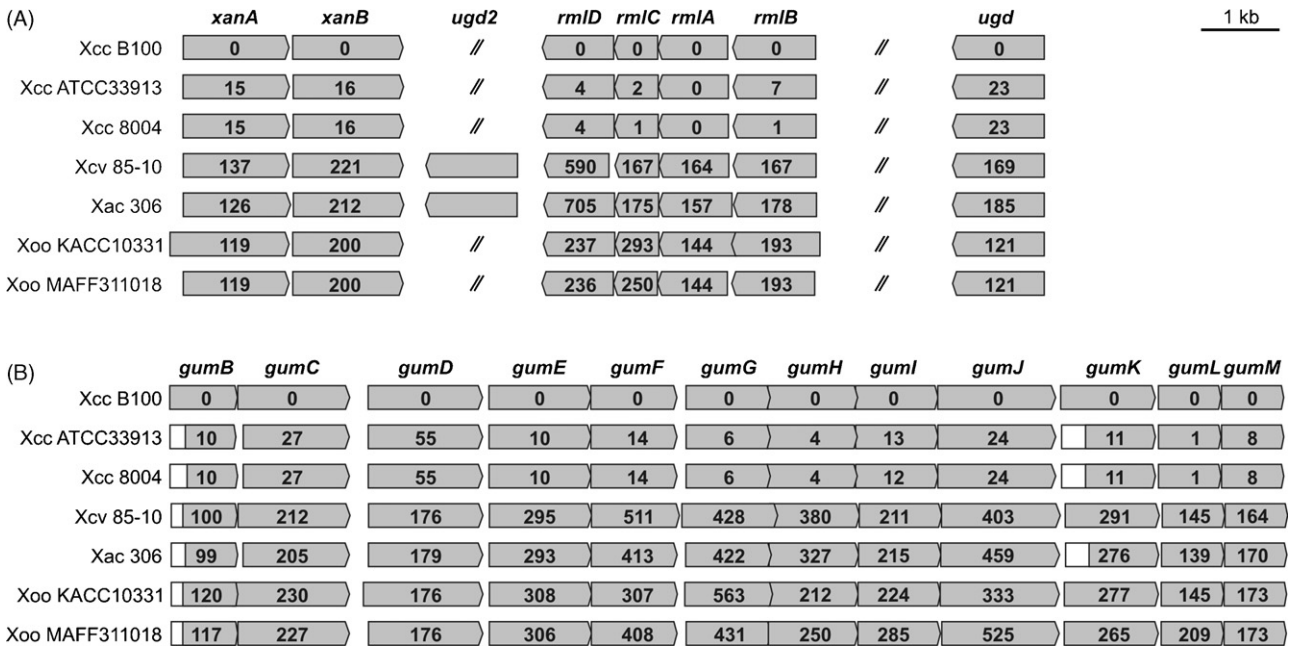

Fig. 3. Comparative representation of the organization of *Xanthomonas* gene clusters directing the synthesis of nucleotide sugars and xanthan. The *Xanthomonas* genes involved in the synthesis of nucleotide sugars (A) and xanthan (B) are symbolized by shaded arrows that illustrate the similarities in size, order, and direction of transcription among these genes. Conventional gene names are given above the arrows of homologous genes. The *xanAB* genes are required for the synthesis of the xanthan precursors UDP-glucose and GDP-mannose. The *rmlABCD* genes encode the synthesis of dTDP-L-rhamnose. In *X. campestris* pv. vesicatoria and *X. axonopodis* pv. citri *ugd* genes are situated in between the otherwise adjacent *xanAB* and *rmlABCD* genes. The functional annotation of the *ugd* genes of *X. axonopodis* pv. citri and *X. campestris* pv. vesicatoria, which are termed *ugd*2 in this figure, appears questionable, as outlined in the text. The genes *gumB* to *gumM* (B) are required for xanthan synthesis. Current gene prediction software indicates for the *gumB* and *gumK* genes earlier translational starts with ATG start codons located further upstream, which are conserved but not annotated in all *X. campestris* pv. campestris strains. Open boxes give the extents of these not annotated upstream regions. The numbers within the arrows give the nucleotides deviating by SNPs or insertion-deletion events from the *X. campestris* pv. campestris B100 gene sequences. While there are few SNPs in the *X. campestris* pv. campestris strains, the differences among the remaining strains are substantial.

among them a recently characterized sucrose importer SuxC (Blanvillain et al., 2007), and one solute:sodium symporter (SSS) protein. The specificity of these uptake systems was predicted on the basis of sequence similarity to known transporters (Table 2). However, because of the diversity of structurally similar but not identical carbohydrates, sequence similarity-based predictions for sugar importers should be generally considered with caution. Abundant glucose not required to provide cellular energy or building blocks can be stored in the form of glycogen. The genome annotation revealed genes that provide the blueprints for the synthesis and reutilization of glycogen.

Table 2
Carbohydrate uptake systems of *X. campestris* pv. campestris strain B100

No.	Description	TC	genes	Predicted specificity[a]
1	ABC superfamily, carbohydrate uptake transporter	3.A.1.1	*malEFG*	Maltose, trehalose, fructooligosaccharide
2	PTS mannose-fructose-sorbose family	4.A.6	*manX*, *ptsHIN*, *hprK*	Mannose, glucose, glucosamine, fructose
3	PTS fructose porter	4.A.2.1.1	*fruABK*	Fructose
4	MFS glucose/fructose:H^+ symporter	2.A.1.1.32	*glcP*	Glucose, fructose
5	MFS glucose uniporter	2.A.1.1.4		Glucose
6	MFS L-fucose/H^+ symporter	2.A.1.7.1	*fucP*	Fucose
7	MFS glucose/galactose porter	2.A.1.7.2	*gluP*1	Glucose, galactose
8	MFS glucose/galactose porter	2.A.1.7.2	*gluP*2	Glucose, galactose
9	MFS glucarate porter	2.A.1.14.1	*gudP*	Glucarate
10	MFS galactarate:H^+ symporter	2.A.1.14.14	*garP*	Galactarate
11	GPH pentoside permease	2.A.2.3.2	*xylP*	Pentoside
12	GPH sucrose:H^+ symporter	2.A.2.4.2	*suxC*	Sucrose
13	GPH maltose/sucrose symporter	2.A.2.6.1	*suc*	Sucrose, maltose
14	GPH isoprimeverose permease	2.A.2.3.3	*gutA*	Glucitol, lactose, α-D-xylopyranosyl-(1,6)-D-glucopyranose,
15	SSS glucose or galactose:Na^+ symporter	2.A.21.3.2	*sglT*	Glucose, galactose

Abbreviations: ABC, ATP-binding cassette; PTS, phosphotransferase system; MFS, major facilitator superfamily; GPH, glycoside-pentoside-hexuronide:cation symporters; SSS, solute:sodium symporter.

[a] In order to likelihood, based on sequence specificity.

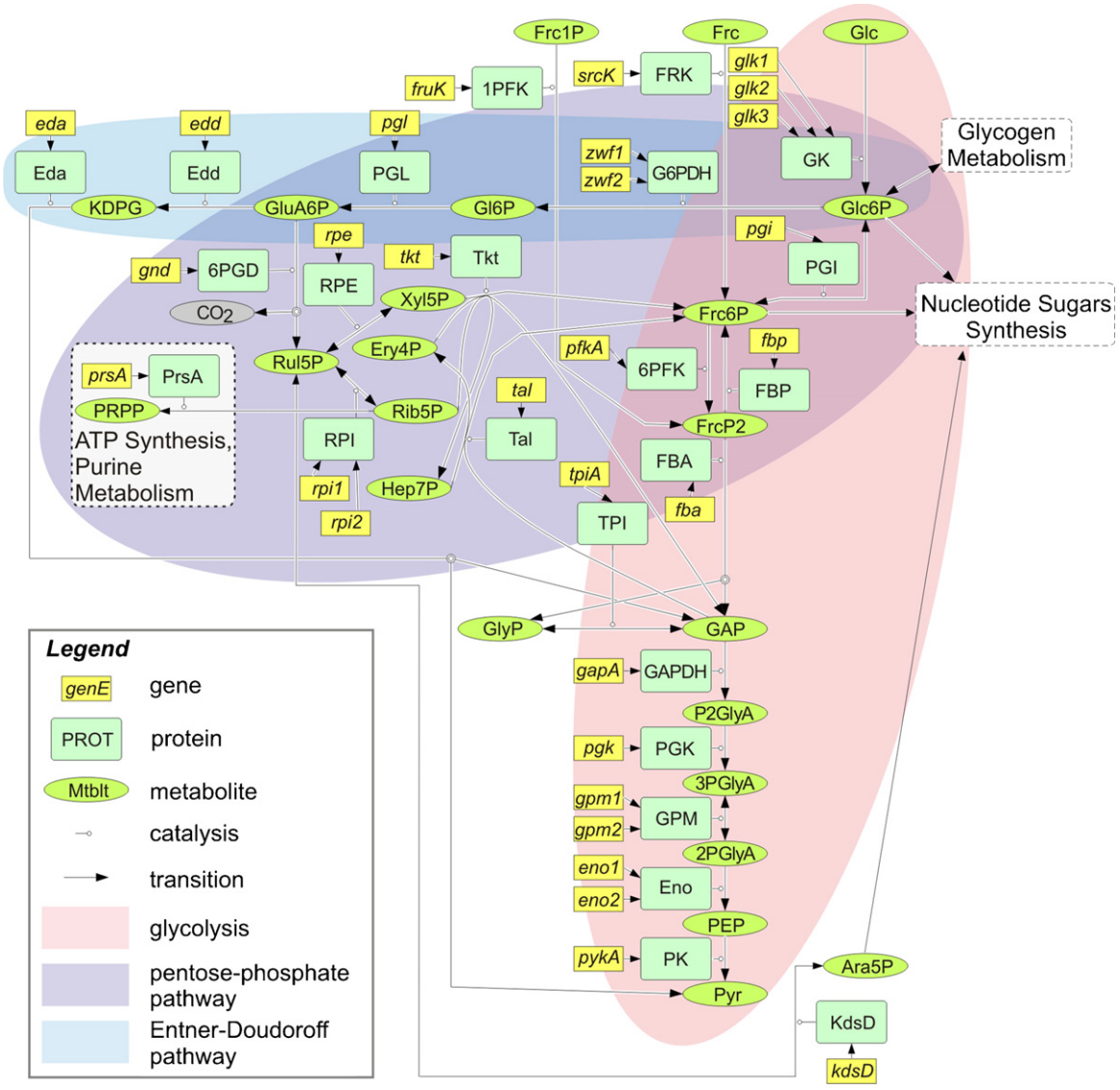

Fig. 4. Reconstruction of the *X. campestris* pv. campestris carbohydrate metabolism. The basic carbohydrate metabolism pathways were reconstructed from the annotated *X. campestris* pv. campestris B100 genome. Metabolic reconstruction was performed with the CellDesigner software and data were stored in the current SBML format. The symbols indicating genes, proteins and metabolites as well as transition and catalysis are given in the legend. The diagram shows the glycolysis, the pentose phosphate cycle, and the Entner-Doudoroff pathway. The final product of all these pathways is pyruvate (Pyr), which can subsequently enter the TCA (not depicted). Some of the enzymes are encoded by two or three isogenes. The intermediate glucose 6-phosphate (Glc6P) is also a substrate for glycogen synthesis, which allows to store imported carbohydrates for later use as carbon or energy source. Phosphorylated intermediates of the carbohydrate metabolism, all of which are intermediates of the pentose phosphate cycle, are the origins of nucleotide sugars biosynthesis, which is illustrated in Fig. 5. Different background colors indicate roughly the extents of the main pathways involved.

Investigations of the glucose flow in *X. campestris* using radiolabeled glucose and enzymological techniques showed that glucose is catabolized mainly through the Entner-Doudoroff pathway and to a small proportion by the pentose-phosphate cycle (Whitfield et al., 1982; Pielken et al., 1988). This genome analysis revealed all genes required for encoding the enzymes of the Entner-Doudoroff and pentose-phosphate pathways (Fig. 4). All essential Entner-Doudoroff genes were combined in a single cluster situated almost exactly vis-à-vis the origin of replication of the circular chromosome. Glycolysis was supposed to be absent in *X. campestris*, as no 6-phosphofructokinase activity had been observed (Whitfield et al., 1982; Pielken et al., 1988). The genome data now allowed identifying a gene (*pfkA*, xccb100_0891) highly similar to 6-phosphofructokinase entries, e.g. of the SwissProt database. Thus genes for all glycolysis and gluconeogenesis enzyme activities were present, implying that *X. campestris* pv. campestris strain B100 can perform glycolysis as well as the reactions of the Entner-Doudoroff and pentose phosphate pathways (Fig. 4). All these pathways shared a common triose-phosphate conversion which started from glyceraldehyde 3-phosphate and which resulted in the synthesis of pyruvate that could finally be utilized in the tricarboxylate (TCA) cycle. *X. campestris* pv. campestris could also perform the reverse reaction of synthesizing hexose sugars from pyruvate

or the TCA. Gluconeogenesis was manifest in *X. campestris* pv. campestris strain B100 by three specific enzymes, which could be identified in the genome. The gluconeogenesis enzymes include the pyruvate, water dikinase, also known as phosphoenolpyruvate synthase (EC 2.7.9.2) encoded by *ppsA*, a gene that was recently shown to be required for gluconeogenesis in *X. campestris* pv. campestris strain 8004 (Tang et al., 2005), the fructose-bisphosphatase (EC 3.1.3.11), encoded by *fbp*, and the malate dehydrogenase (EC 1.1.1.40), encoded by *maeB*. The *maeB* gene product initiates the gluconeogenic pathway from the TCA intermediate malate.

The genome annotation revealed metabolic pathways which extend the capability of *X. campestris* pv. campestris to make use of imported carbohydrates. *X. campestris* pv. campestris strain B100 can route imported gluconate directly to the Entner-Doudoroff pathway for energy production. Further on, *X. campestris* pv. campestris strain B100 was found to have the capacity to metabolize galactose in a separate pathway to glyceraldehyde-3-phosphate and pyruvate that were also the end products of the Entner-Doudoroff pathway, of the pentose phosphate pathway, and of glycolysis. A conversion of imported galactose to glucose or UDP-glucose by means of the Leloir pathway (Frey, 1996) seemed unlikely, as no gene was found coding for a protein with sequence similarity to galactokinase (GalK, EC 2.7.1.6), which performs the first committed step of this conversion. Glucarate and galactarate could be imported and channeled to the TCA by oxidation and decarboxylation to alpha-ketoglutarate. Deviating from the situation in *E. coli* (Monterrubio et al., 2000), the genes for both glucarate and galactarate metabolism and uptake shared in one operon-like gene cluster combined with their respective importers and an outer membrane porin. And finally, imported xylose could, by means of enzymes encoded by *xylAB*, be converted to xylulose-5-phosphate, an intermediate of the pentose-phosphate pathway. The reconstruction of the carbohydrate metabolism is an essential prerequisite for rational strategies to optimize xanthan production. At the same time it sheds light on the utilization of materials originating from pathogenesis.

3.5. Biosynthesis of nucleotide sugars

Most of the cell surface carbohydrates share a common organization of their biosynthesis in successive steps. Initially nucleotide sugars are produced, which provide specifically activated monosaccharides as precursors for the subsequent synthesis steps. The nucleotide sugars' monosaccharide moieties are derived from importation, conversion from other sugars of the primary carbohydrate metabolism, or from breakdown of the cellular glycogen storage. In a second step of biosynthesis monosaccharide moieties taken from the nucleotide sugar precursors are transferred by highly specific glycosyltransferases to undecaprenylphosphate lipid carriers at the inner face of the cell membrane.

The synthesis of all cell surface carbohydrates depends on nucleotide sugars as raw material for the particular synthesis steps. The biosynthesis of xanthan depends on the nucleotide sugars UDP-D-glucose, GDP-D-mannose, and UDP-D-glucuronate. Key factors for the synthesis of these nucleotide sugars are the genes *xanA* and *xanB* (Koplin et al., 1992; Fig. 5). Xanthan precursor biosynthesis was directed besides by *xanAB* by the genes *pgi* (Tung and Kuo, 1999), *galU* (Wei et al., 1996) and *ugd* (Lin et al., 1995, Fig. 5). The products of these genes are enzymes that catalyze the synthesis of UDP-D-glucose, GDP-D-mannose, and UDP-D-glucuronate in a branched pathway. Four of the seven reactions that define this pathway are controlled by *xanAB*, which both encode bifunctional enzymes. The LPS precursor synthesis genes *rmlABCD* (Koplin et al., 1993; Fig. 5) were clustered in the vicinity of *xanAB*. In *X. campestris* pv. vesicatoria and *X. axonopodis* pv. citri *ugd* genes are situated in between the otherwise adjacent *xanAB* and *rmlABCD* genes (Fig. 3). These *ugd* genes of *X. axonopodis* pv. citri and *X. campestris* pv. vesicatoria, which are termed *ugd*2 in the figure, share a homologous primary structure. However the *ugd*2 annotation, which proposes the encoded enzymes to oxidize UPD-glucose to UDP-glucuronate, the third xanthan precursor, and which originates from the *X. axonopodis* pv. citri genome project (da Silva et al., 2002), may be questionable. The *ugd* genes of *X. axonopodis* pv. citri and *X. campestris* pv. vesicatoria, which are located between *xanAB* and *rmlABCD*, are not similar to other *ugd* genes, including the *ugd* gene of *X. campestris* pv. campestris that has been characterized experimentally (Lin et al., 1995). Further more in the *X. axonopodis* pv. citri and *X. campestris* pv. vesicatoria genomes there are additional *ugd* genes at unrelated chromosomal loci that are homologous to the *ugd* genes of the other *Xanthomonas* strains.

Besides these genes, which were known before, further nucleotide sugar synthesis genes were to be expected as they encode precursors of typical surface components of Gram-negative bacteria. Such genes could be identified, resulting in the reconstruction of the synthesis of UDP-*N*-acetylglucosamine, UDP-*N*-acetylmuramate, CMP-3-deoxy-manno-octulosonate (CMP-KDO), and UDP-D-galacturonate, which are precursors of the cell-wall polysaccharide murein, or the LPS (Fig. 5). Genes indicating the synthesis of further nucleotide sugars were identified at different chromosomal loci (Fig. 1A, Fig. 5).

The nucleotide sugar genes identified in the genome of *X. campestris* pv. campestris strain B100 coded in total for 29 enzymes, some of which were bifunctional. These enzymes catalyzed the synthesis of 13 nucleotide sugars as building blocks for cell surface polysaccharides or carbohydrates. All nucleotide sugar synthesis pathways could be routed back to intermediates of the pentose phosphate cycle as their sources. The conversions of glucose-6-phosphate to glucose-1-phosphate and of mannose-6-phosphate to mannose-1-phosphate, which were both encoded by *xanA*, turned out to be required in the synthesis of 10 of the nucleotide sugars.

3.6. Biosynthesis of xanthan

Xanthan synthesis is performed by proteins that are supposed to form a machinery with some similarity to the better characterized EPS or CPS synthesis machineries of *Enterobac-*

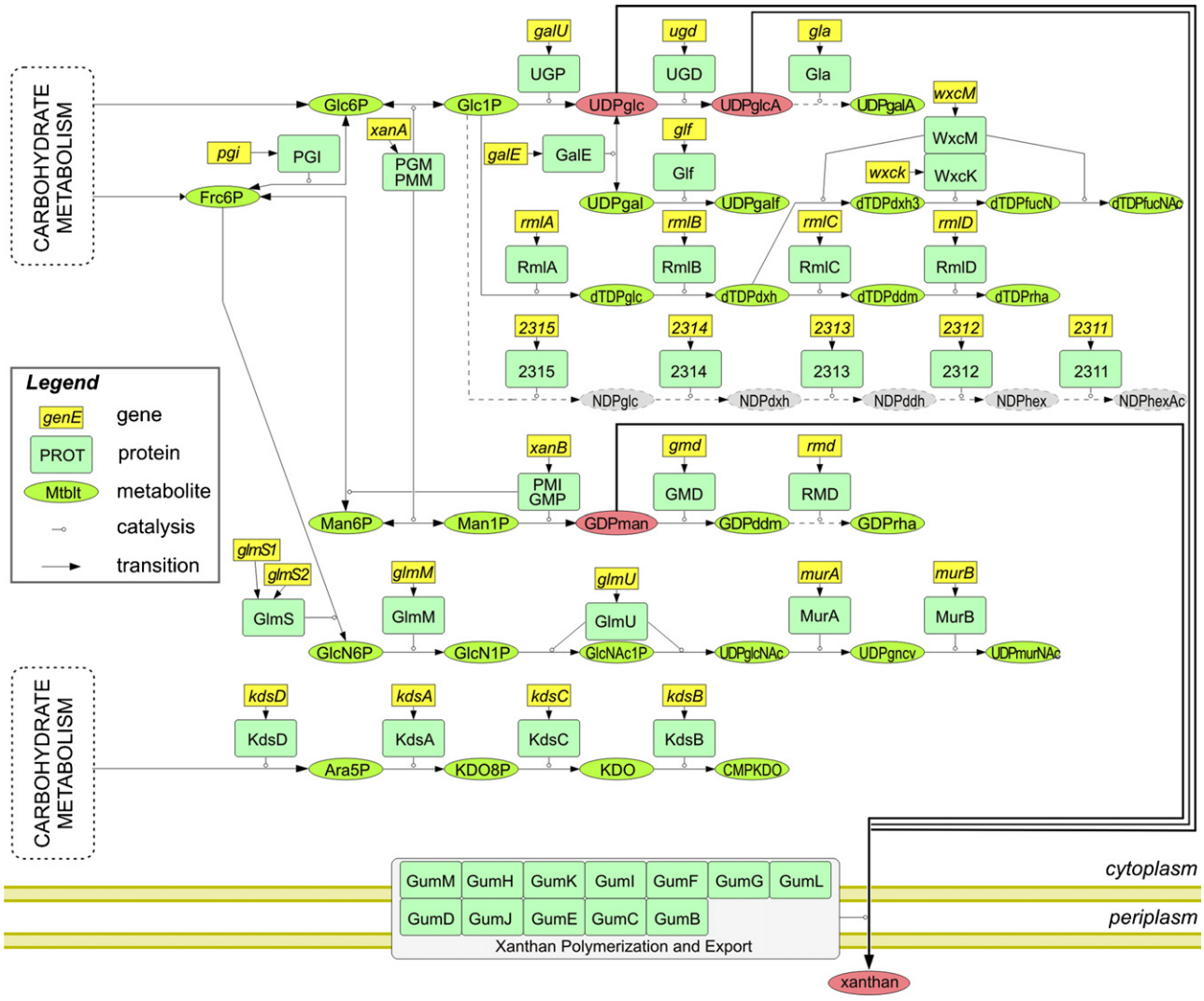

Fig. 5. Reconstruction of the nucleotide sugars synthesis. The biosynthesis of the *X. campestris* pv. campestris B100 nucleotide sugars was reconstructed by means of the CellDesigner. The symbols indicating genes, proteins and metabolites as well as transition and catalysis are given in the legend. Phosphorylated intermediates of the carbohydrate metabolism (Fig. 4) are the origins of nucleotide sugars biosynthesis. Glucose-6-phosphate (Glc6P) and fructose-6-phosphate (Frc6P) are converted to the nucleotide sugars UDP-glucose (UDPglc), UDP-glucuronate (UDPglcA), and GDP-mannose (GDPman), which are highlighted in red. These nucleotide sugars are precursors of xanthan, which is synthesized and exported by the Gum proteins at the cell envelope. The metabolic reconstruction includes the well-characterized synthesis pathway of dTDP-L-rhamnose (dTDPrha, Koplin et al., 1993), and the pathways leading to the LPS and murein precursors CMP-KDO (CMPKDO), UDP-*N*-acetylglucosamine (UDPglcNAc), and UDP-*N*-acetylmuramate. The latter nucleotide sugars are synthesized from fructose-6-phosphate, a substrate of XanB, which links them to the xanthan synthesis. The genome data indicates that further nucleotide sugars are synthesized. The specificity of the probably acetylated nucleotide sugar synthesized by xcc-b100_2311 to xcc-b100_2315 remained unclear, although similarity to the *A. thaliana* UDP-L-rhamnose synthase gene and analogies to the biochemically proposed three-step synthesis (Kamsteeg et al., 1978) point to the UDP-L-rhamnose synthesis from UDP-D-glucose. Enzyme sequence similarities indicated the synthesis of the nucleotide sugars UDP-galactose (UDPgal) and UDP-galactofuranose (UDPgalf).

teriaceae (Whitfield, 2006). The xanthan synthesizing proteins are encoded by genes of the *gum* cluster (Katzen et al., 1998; Vojnov et al., 2001, Fig. 3). Before xanthan can be polymerized from oligosaccharide repeat units and finally exported in the last step of synthesis, the repeat units have to be synthesized. Repeat unit synthesis requires the glycosyltransferases encoded by the *gumDMHKI* genes (Ielpi et al., 1993). The oligosaccharide is built at an undecaprenylphosphate lipid carrier at the interior face of the inner cell membrane. For many but not all *gum* gene products there were counterparts in the non-pathogenic endophyte *Azoarcus* sp., a β-proteobacterium (Supplementary Table 2), which may have acquired the *gum* genes by horizontal gene transfer (Krause et al., 2006).

The initial glycosyltransferase *GumD* is a membrane protein, which transfers the phospho-glucose moiety of UDP-D-glucose to the undecaprenylphosphate lipid carrier to form undecaprenol–diphosphate–glucose (Fig. 6). The protein was conserved in all sequenced strains of *Xanthomonas* and *Xylella*, a phytopathogenic member of the *Xanthomonadaceae* family, which is closely related to *Xanthomonas*. The *GumM* protein is the second glucosyltransferase that transfers a glucose moiety to undecaprenyl–diphosphate–glucose, resulting in a β-

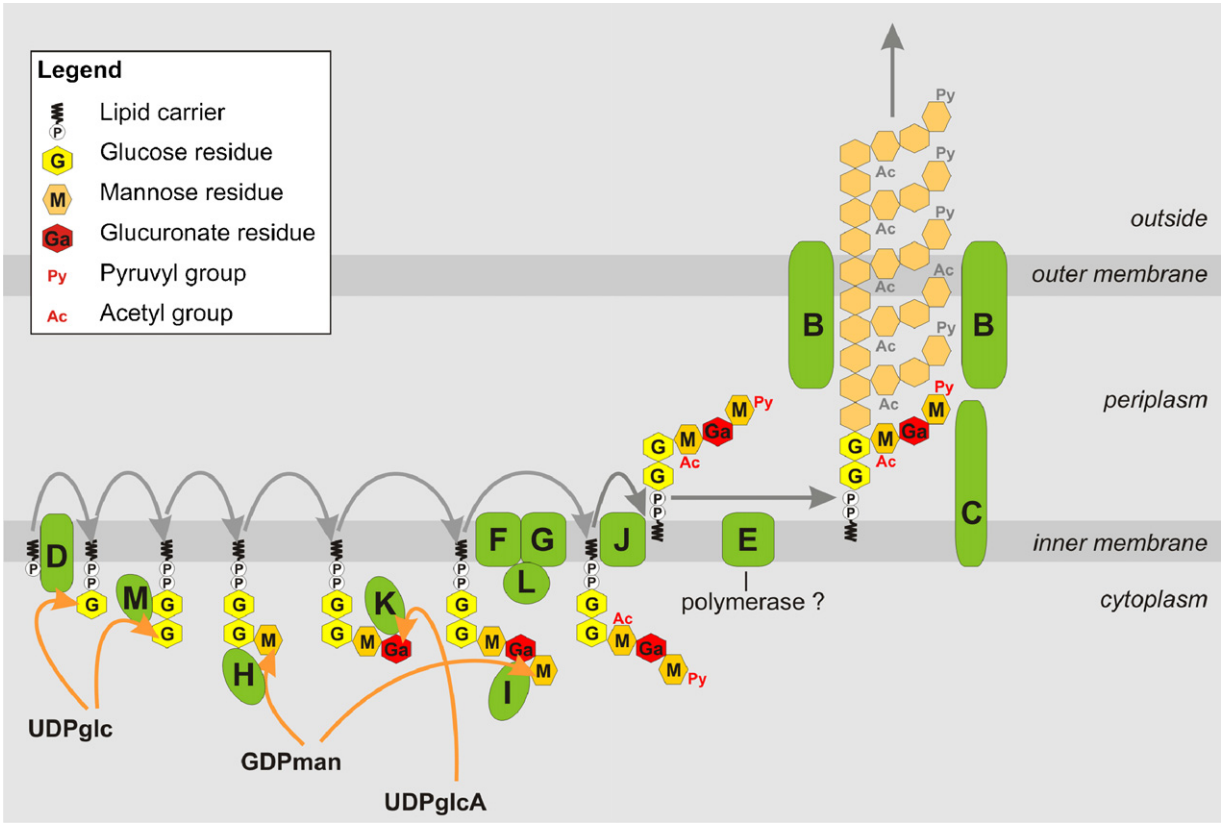

Fig. 6. Mechanistic model for the synthesis of xanthan at the envelope of *X. campestris*. Proteins encoded by the *gum* genes are highlighted in green. In the first reaction the glycosyltransferase GumD transfers a glucose-phosphate residue from UDP-glucose (UDPglc) to an undecaprenylphosphate lipid carrier located at the inner face of the cell membrane. In the four subsequent reaction steps GumM, GumH, GumK and GumI add sequentially a second glucose residue, two mannose residues, and a glucuronate residue from UDPglc, GDP-mannose (GDPman) and UDP-glucuronate (UDPglcA) to form the carbohydrate structure of a xanthan repeat unit. The outer, last-added mannose can be pyruvylated by GumL, both mannose moieties can be specifically acetylated by GumF and GumG, respectively. Finished repeat units are translocated by GumJ to the outer face of the inner membrane. Here, in the periplasm, xanthan could be polymerized by GumE, which transfers immature xanthan polymers to newly translocated repeat units. Finally xanthan is exported. This involves GumC, a protein that is anchored in the inner membrane with a substantial periplasmic domain. When GumC gets into contact with the outer membrane protein GumB, which is assumed to have a large periplasmic domain, too, complexes of both proteins can form open pores which permit exportation of the mature xanthan.

1,4-linked glucose-disaccharide attached to the lipid carrier. GumM is member of the CAZy glycosyltransferase family 26 characterized by an inverting reaction mechanism, and a member of the PFAM WecB/TagA/CpsF glycosyltransferase family. The *GumH* protein transfers a mannose residue from GDP-D-mannose to the lipid-linked disaccharide by establishing an α-1,3-glycosidic link. Thus it adds the first constituent of the prospective xanthan side chain. GumH is member of the CAZy glycosyltransferase family 4, which applies a retaining reaction mechanism, and of the PFAM glycosyltransferase group 1 family. The glucuronyltransferase *GumK* adds a glucuronate moiety from UDP-D-glucuronate to the lipid-linked glucose–glucose–mannose trisaccharide provided by GumH. GumK is a membrane-associated protein and member of the CAZy glycosyltransferase family 70, which creates β-1,4 links to the previously attached mannose residue of the nascent oligosaccharide by an inverting mechanism (Barreras et al., 2004). The last monosaccharide of the xanthan repeat unit is added by the mannosyltransferase *GumI*, which generates a β-1,2 link, and which is not classified in the CAZy database. Significant similarity to the GumI protein was limited to the homologous GumI proteins of the other *Xanthomonas* strains.

Non-sugar decorations are added to the mannose residues of the xanthan repeat unit at varying degrees (Stankowski et al., 1993; Katzen et al., 1998). The *GumL* protein adds a pyruvyl group to the last-added mannose. The acetyltransferase I encoded by *gumF* acetylates the inner mannose, while acetyltransferase II encoded by *gumG* can add acetyl groups to non-pyruvylated outer mannoses. Conserved GumL proteins were restricted to *Xanthomonas* strains. The only relevant exceptions were two *Pseudoalteromonas* proteins with reduced but still significant similarity, as indicated by *E*-values of about 10^{-60} (Supplementary Table 2). The databases kept no further *Pseudoalteromonas* entries similar to Gum proteins, and there were no genes similar to *gumL* in *Xylella* or *Azoarcus* genomes. The two acetyltransferases GumF and GumG were 39% identical. For both proteins 9 transmembrane helices (TMHs) were predicted, respectively, indicating their location in the cell membrane. GumF and GumG are members of the PFAM acyltransferase family 3.

It came as a surprise that in the genome era, with hundreds of bacterial genomes annotated and available, the presence of *gum*-homologous genes seemed so restricted to *Xanthomonas* and *Xylella*, with several homologs clustered in *Azoarcus* sp., which is evolutionary rather distant from Xanthomonas, but shares a plant-associated habitat. Homologous counterparts to the mannosyltransferase GumI and the pyruvyltransferase GumL were even missing in *Xylella*. So the current data gives the impression that xanthan synthesis is limited to *Xanthomonas*. The future will show whether this is a short-lived snapshot or a true biological phenomenon.

While the functions of the proteins involved in repeat unit synthesis has been established on the basis of experimental data (Katzen et al., 1998), there is no evidence for the exact functions of the remaining GumBCEJ proteins. Export and polymerization of xanthan was affected in *gumBCE* mutants. Now the largely extended set of available data from genome sequencing, new structural and functional characterization of model proteins from enteric bacteria, as well as enhanced bioinformatics tools permit to propose a model for the *gumBCEJ* gene products on the basis of homologies to the Wzy-dependent polysaccharide polymerization and export machinery in *Enterobacteriaceae* (Whitfield, 2006, Fig. 6).

A key component for the polymerization process of xanthan is *GumJ*. Mutagenesis of *gumJ* did not affect xanthan polymerization in a cell-free *in vitro* system, but disruption of *gumJ* in the *X. campestris* pv. campestris wild-type was lethal (Katzen et al., 1998). GumJ was conserved among *Xanthomonas* species, and was also similar to the *Escherichia coli* oligosaccharyl exporter Wzx (formerly RfbX). Wzx is involved in the synthesis of capsular polysaccharide (CPS; Whitfield, 2006) and is a member of the PFAM "polysaccharide biosynthesis protein" family. For GumJ 10 TMHs were predicted. The data from the homologous *E. coli* system imply that it will transfer the nascent undecaprenyl diphosphate-linked repeat units across the inner membrane to the periplasm (Fig. 6).

For the export of xanthan through the outer membrane *GumB* was a convincing candidate, due to sequence similarities to the *E. coli* CPS exporter Wza (Dong et al., 2006). Due to a start codon located further upstream the CDS of *gumB* was extended when compared to the original GenBank entry from *X. campestris* pv. campestris strain B1459, see above. This resulted in a predicted CDS of 286 amino acids, compared to 232 amino acids in strain B1459 and 213 amino acids in the sequenced strains 8004 and ATCC 33913. However, regarding the uncertainties of computer-based predictions for translational starts, it will be worth to determine the N-terminus of the gene product experimentally. Like Wza, GumB was a member of the PFAM "polysaccharide biosynthesis/export protein" family. Recent studies revealed that Wza is organized in four domains and located in the outer membrane (Dong et al., 2006).

Proteins homologous to *GumC* beyond the *Xanthomonas* and *Xylella* species were members of the PFAM Wzz-family of "chain-length determinant proteins". However, mutation of *gumC* did not provoke the production of low-molecular-weight xanthan (Katzen et al., 1998). There was also sequence similarity to the *E. coli* protein Wzc, which had a similar characteristic basic topology with two TMHs, a large periplasmic domain and an additional cytoplasmic autokinase domain. Now Wzc was shown to form a protein complex with Wza (Collins et al., 2007). The main function of GumC may be to link the export of xanthan through GumB to the synthesis at the inner membrane in a manner similar to Wza.

Up to now one function in xanthan synthesis is left unassigned, the polymerization of repeat units, and for one protein encoded in the *gum* operon, *GumE*, no function has been assigned. GumE was highly conserved in *Xanthomonas* and *Xylella*, but there were no similar proteins with significant homologies to the GumE amino acid sequence. However, lack of primary structure similarity has been frequently observed for key membrane-components of the Wzy-dependent polysaccharide synthesis (Schnaitman and Klena, 1993). GumE was predicted to be an integral membrane protein with 10–12 TMHs, which resembles the structure of the polymerase Wzy (Whitfield, 2006). Thus GumE is the most suitable candidate for the polymerase function, but experimental data are needed to check this assumption. A detailed analysis of the putative polymerase would be of special value as xanthan production by *X. campestris* is supposed to be very efficient, making GumE a candidate for a highly processive polymerase.

In the light of the new data obtained from the genome annotation a mechanistic model for the biosynthesis of xanthan was established (Fig. 6). It is based on the previous conception for the synthesis of the xanthan repeat unit (Katzen et al., 1998). Now detailed functions could be assigned to the gene products of *gumJ*, *gumC*, *gumD*, and *gumE* in a xanthan polymerization and export machinery. The *gumJ* gene product is assumed to flip the lipid carrier-linked pentasaccharide repeat unit to the outer face of the inner cell membrane. Here GumE may perform the polymerization while upon contact of GumB and GumC, which are anchored in the outer and inner membranes, respectively, a complex periplasm-spanning pore is opened to export the xanthan polymer to the environment.

Acknowledgements

The authors thank Daniela Bartels, Lars Gaigalat, Sascha Mormann, Diana Nakunst, and Jens Plassmeier for checking gene prediction results in the initial phase of the sequencing project. The project was supported by the GenoMik-Plus programme of the German Federal Ministry of Education and Research (BMBF), grant 03138805A, and by the BMBF project BioExPoSys.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jbiotec.2007.12.013.

References

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.

Barreras, M., Abdian, P.L., Ielpi, L., 2004. Functional characterization of GumK, a membrane-associated beta-glucuronosyltransferase from *Xanthomonas campestris* required for xanthan polysaccharide synthesis. Glycobiology 14, 233–241.

Bartels, D., Kespohl, S., Albaum, S., Druke, T., Goesmann, A., Herold, J., Kaiser, O., Puhler, A., Pfeiffer, F., Raddatz, G., Stoye, J., Meyer, F., Schuster, S.C., 2005. BACCardI—a tool for the validation of genomic assemblies, assisting genome finishing and intergenome comparison. Bioinformatics 21, 853–859.

Becker, A., Katzen, F., Puhler, A., Ielpi, L., 1998. Xanthan gum biosynthesis and application: a biochemical/genetic perspective. Appl. Microbiol. Biotechnol. 50, 145–152.

Blanvillain, S., Meyer, D., Boulanger, A., Lautier, M., Guynet, C., Denance, N., Vasse, J., Laubar, E., Arlat, M., 2007. Plant carbohydrate scavenging through tonB-dependent receptors: a feature shared by phytopathogenic and aquatic bacteria. PLoS ONE 2, e224.

Collins, R.F., Beis, K., Dong, C., Botting, C.H., McDonnell, C., Ford, R.C., Clarke, B.R., Whitfield, C., Naismith, J.H., 2007. The 3D structure of a periplasm-spanning platform required for assembly of group 1 capsular polysaccharides in *Escherichia coli*. Proc. Natl. Acad. Sci. U.S.A. 104, 2390–2395.

Coutinho, P.M., Henrissat, B., 1999. Carbohydrate-active enzymes: an integrated database approach. In: Gilbert, H.J., Davies, G., Henrissat, B., Svensson, B. (Eds.), Recent Advances in Carbohydrate Bioengineering. The Royal Society of Chemistry, Cambridge, pp. 3–12.

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., Kitajima, J.P., 2002. Comparison of the genomes of two *Xanthomonas* pathogens with differing host specificities. Nature 417, 459–463.

Dong, C., Beis, K., Nesper, J., Brunkan-Lamontagne, A.L., Clarke, B.R., Whitfield, C., Naismith, J.H., 2006. Wza the translocon for *E. coli* capsular polysaccharides defines a new class of membrane protein. Nature 444, 226–229.

Dunger, G., Relling, V.M., Tondo, M.L., Barreras, M., Ielpi, L., Orellano, E.G., Ottado, J., 2007. Xanthan is not essential for pathogenicity in citrus canker but contributes to Xanthomonas epiphytic survival. Arch. Microbiol. 188, 127–135.

Frey, P.A., 1996. The Leloir pathway: a mechanistic imperative for three enzymes to change the stereochemical configuration of a single carbon in galactose. FASEB J. 10, 461–470.

Garcia-Ochoa, F., Santos, V.E., Casas, J.A., Gomez, E., 2000. Xanthan gum: production, recovery, and properties. Biotechnol. Adv. 18, 549–579.

Gordon, D., Abajian, C., Green, P., 1998. Consed: a graphical tool for sequence finishing. Genome Res. 8, 195–202.

Gordon, D., Desmarais, C., Green, P., 2001. Automated finishing with autofinish. Genome Res. 11, 614–625.

Hotte, B., Rath-Arnold, I., Puhler, A., Simon, R., 1990. Cloning and analysis of a 35.3-kilobase DNA region involved in exopolysaccharide production by *Xanthomonas campestris* pv. campestris. J. Bacteriol. 172, 2804–2807.

Ielpi, L., Couso, R.O., Dankert, M.A., 1993. Sequential assembly and polymerization of the polyprenol-linked pentasaccharide repeating unit of the xanthan polysaccharide in *Xanthomonas campestris*. J. Bacteriol. 175, 2490–2500.

Jansson, P.E., Kenne, L., Lindberg, B., 1975. Structure of extracellular polysaccharide from *Xanthomonas campestris*. Carbohydr. Res. 45, 275–282.

Kaiser, O., Bartels, D., Bekel, T., Goesmann, A., Kespohl, S., Puhler, A., Meyer, F., 2003. Whole genome shotgun sequencing guided by bioinformatics pipelines—an optimized approach for an established technique. J. Biotechnol. 106, 121–133.

Kamsteeg, J., Van Brederode, J., Van Nigtevecht, G., 1978. The formation of UDP-L-rhamnose from UDP-D-glucose by an enzyme preparation of red campion (*Silene dioica* (L) Clairv) leaves. FEBS Lett. 91, 281–284.

Katzen, F., Becker, A., Zorreguieta, A., Puhler, A., Ielpi, L., 1996. Promoter analysis of the *Xanthomonas campestris* pv. campestris *gum* operon directing biosynthesis of the xanthan polysaccharide. J. Bacteriol. 178, 4313–4318.

Katzen, F., Ferreiro, D.U., Oddo, C.G., Ielmini, M.V., Becker, A., Puhler, A., Ielpi, L., 1998. *Xanthomonas campestris* pv. campestris gum mutants: effects on xanthan biosynthesis and plant virulence. J. Bacteriol. 180, 1607–1617.

Koplin, R., Arnold, W., Hotte, B., Simon, R., Wang, G., Puhler, A., 1992. Genetics of xanthan production in *Xanthomonas campestris*: the *xanA* and *xanB* genes are involved in UDP-glucose and GDP-mannose biosynthesis. J. Bacteriol. 174, 191–199.

Koplin, R., Wang, G., Hotte, B., Priefer, U.B., Puhler, A., 1993. A 3.9-kb DNA region of *Xanthomonas campestris* pv. campestris that is necessary for lipopolysaccharide production encodes a set of enzymes involved in the synthesis of dTDP-rhamnose. J. Bacteriol. 175, 7786–7792.

Kitano, H., Funahashi, A., Matsuoka, Y., Oda, K., 2005. Using process diagrams for the graphical representation of biological networks. Nat. Biotechnol. 23, 961–966.

Lee, B.M., Park, Y.J., Park, D.S., Kang, H.W., Kim, J.G., Song, E.S., Park, I.C., Yoon, U.H., Hahn, J.H., Koo, B.S., Lee, G.B., Kim, H., Park, H.S., Yoon, K.O., Kim, J.H., Jung, C.H., Koh, N.H., Seo, J.S., Go, S.J., 2005. The genome sequence of *Xanthomonas oryzae* pathovar *oryzae* KACC10331, the bacterial blight pathogen of rice. Nucleic Acids Res. 33, 577–586.

Krause, A., Ramakumar, A., Bartels, D., Battistoni, F., Bekel, T., Boch, J., Bohm, M., Friedrich, F., Hurek, T., Krause, L., Linke, B., McHardy, A.C., Sarkar, A., Schneiker, S., Syed, A.A., Thauer, R., Vorholter, F.J., Weidner, S., Puhler, A., Reinhold-Hurek, B., Kaiser, O., Goesmann, A., 2006. Complete genome of the mutualistic, N_2-fixing grass endophyte *Azoarcus* sp. strain BH72. Nat. Biotechnol. 24, 1385–1391.

Krause, L., McHardy, A.C., Nattkemper, T.W., Puhler, A., Stoye, J., Meyer, F., 2007. GISMO—gene identification using a support vector machine for ORF classification. Nucleic Acids Res. 35, 540–549.

Letisse, F., Lindley, N.D., Roux, G., 2003. Development of a phenomenological modeling approach for prediction of growth and xanthan gum production using *Xanthomonas campestris*. Biotechnol. Prog. 19, 822–827.

Lin, C.S., Lin, N.T., Yang, B.Y., Weng, S.F., Tseng, Y.H., 1995. Nucleotide sequence and expression of UDP-glucose dehydrogenase gene required for the synthesis of xanthan gum in *Xanthomonas campestris*. Biochem. Biophys. Res. Commun. 207, 223–230.

Linke, B., McHardy, A.C., Neuweger, H., Krause, L., Meyer, F., 2006. REGANOR: a gene prediction server for prokaryotic genomes and a database of high quality gene predictions for prokaryotes. Appl. Bioinformat. 5, 193–198.

Meyer, F., Goesmann, A., McHardy, A.C., Bartels, D., Bekel, T., Clausen, J., Kalinowski, J., Linke, B., Rupp, O., Giegerich, R., Puhler, A., 2003. GenDB—an open source genome annotation system for prokaryote genomes. Nucleic Acids Res. 31, 2187–2195.

Monterrubio, R., Baldoma, L., Obradors, N., Aguilar, J., Badia, J., 2000. A common regulator for the operons encoding the enzymes involved in D-galactarate, D-glucarate, and D-glycerate utilization in *Escherichia coli*. J. Bacteriol. 182, 2672–2674.

Moreira, L.M., De Souza, R.F., Digiampietri, L.A., Da Silva, A.C., Setubal, J.C., 2005. Comparative analyses of *Xanthomonas* and *Xylella* complete genomes. OMICS 9, 43–76.

Ochiai, H., Inoue, Y., Takeya, M., Sasaki, A., Kaku, H., 2005. Genome sequence of *Xanthomonas oryzae* pv. *oryzae* suggests contribution of large numbers of effector genes and insertion sequences to its race diversity. Jpn. Agric. Res. Q. 39, 275–287.

Pielken, P., Schimz, K.L., Eggeling, L., Sahm, H., 1988. Glucose metabolism in *Xanthomonas campestris* and influence of methionine on the carbon flow. Can. J. Microbiol. 34, 1333–1337.

Qian, W., Jia, Y., Ren, S.X., He, Y.Q., Feng, J.X., Lu, L.F., Sun, Q., Ying, G., Tang, D.J., Tang, H., Wu, W., Hao, P., Wang, L., Jiang, B.L., Zeng, S., Gu, W.Y., Lu, G., Rong, L., Tian, Y., Yao, Z., Fu, G., Chen, B., Fang,

R., Qiang, B., Chen, Z., Zhao, G.P., Tang, J.L., He, C., 2005. Comparative and functional genomic analyses of the pathogenicity of phytopathogen *Xanthomonas campestris* pv. campestris. Genome Res. 15, 757–767.

Rigano, L.A., Siciliano, F., Enrique, R., Sendín, L., Filippone, P., Torres, P.S., Qüesta, J., Dow, J.M., Castagnaro, A.P., Vojnov, A.A., Marano, M.R., 2007. Biofilm formation, epiphytic fitness, and canker development in *Xanthomonas axonopodis* pv. citri. Mol. Plant Microbe Interact. 20, 1222–1230.

Saier Jr., M.H., 2000. Families of transmembrane sugar transport proteins. Mol. Microbiol. 35, 699–710.

Schnaitman, C.A., Klena, J.D., 1993. Genetics of lipopolysaccharide biosynthesis in enteric bacteria. Microbiol. Rev. 57, 655–682.

Seisun D., 2006. The market of hydrocolloids and other uses of seaweeds. Aqua Sur–3rd international conference on aquaculture, Puerto Varas, Chile. http://www.aqua-sur.cl/conferencia_pdf/pdf/puertas_lago_manana/DENISS_SEISUN.pdf.

Simon, R., Hotte, B., Klauke, B., Kosier, B., 1991. Isolation and characterization of insertion sequence elements from gram-negative bacteria by using new broad-host-range, positive selection vectors. J. Bacteriol. 173, 1502–1508.

Stankowski, J.D., Mueller, B.E., Zeller, S.G., 1993. Location of a second *O*-acetyl group in xanthan gum by the reductive-cleavage method. Carbohydr. Res. 241, 321–326.

Steinmann, D., Koplin, R., Puhler, A., Niehaus, K., 1997. *Xanthomonas campestris* pv. campestris *lpsI* and *lpsJ* genes encoding putative proteins with sequence similarity to the alpha- and beta-subunits of 3-oxoacid CoA-transferases are involved in LPS biosynthesis. Arch. Microbiol. 168, 441–447.

Sutherland, I.W., 1998. Novel and established applications of microbial polysaccharides. Trends Biotechnol. 16, 41–46.

Swings, J.G. (Ed.), 1993. Xanthomonas. Chapman & Hall, London.

Tang, D.J., He, Y.Q., Feng, J.X., He, B.R., Jiang, B.L., Lu, G.T., Chen, B., Tang, J.L., 2005. *Xanthomonas campestris* pv. campestris possesses a single gluconeogenic pathway that is required for virulence. J. Bacteriol. 187, 6231–6237.

Thieme, F., Koebnik, R., Bekel, T., Berger, C., Boch, J., Buttner, D., Caldana, C., Gaigalat, L., Goesmann, A., Kay, S., Kirchner, O., Lanz, C., Linke, B., McHardy, A.C., Meyer, F., Mittenhuber, G., Nies, D.H., Niesbach-Klosgen, U., Patschkowski, T., Ruckert, C., Rupp, O., Schneiker, S., Schuster, S.C., Vorholter, F.J., Weber, E., Puhler, A., Bonas, U., Bartels, D., Kaiser, O., 2005. Insights into genome plasticity and pathogenicity of the plant pathogenic bacterium *Xanthomonas campestris* pv. vesicatoria revealed by the complete genome sequence. J. Bacteriol. 187, 7254–7266.

Tatusov, R.L., Koonin, E.V., Lipman, D.J., 1997. A genomic perspective on protein families. Science 278, 631–637.

Tung, S.Y., Kuo, T.T., 1999. Requirement for phosphoglucose isomerase of *Xanthomonas campestris* in pathogenesis of citrus canker. Appl. Environ. Microbiol. 65, 5564–5570.

Vauterin, L., Hoste, B., Kersters, K., Swings, J., 1995. Reclassification of *Xanthomonas*. Int. J. Syst. Bacteriol. 45, 472–489.

Vojnov, A.A., Slater, H., Daniels, M.J., Dow, J.M., 2001. Expression of the *gum* operon directing xanthan biosynthesis in *Xanthomonas campestris* and its regulation *in planta*. Mol. Plant Microbe Interact. 14, 768–774.

Vojnov, A.A., Zorreguieta, A., Dow, J.M., Daniels, M.J., Dankert, M.A., 1998. Evidence for a role for the *gumB* and *gumC* gene products in the formation of xanthan from its pentasaccharide repeating unit by *Xanthomonas campestris*. Microbiology 144, 1487–1493.

Vorholter, F.J., Niehaus, K., Puhler, A., 2001. Lipopolysaccharide biosynthesis in *Xanthomonas campestris* pv. *campestris*: a cluster of 15 genes is involved in the biosynthesis of the LPS O-antigen and the LPS core. Mol. Genet. Genomics 266, 79–95.

Vorholter, F.J., Thias, T., Meyer, F., Bekel, T., Kaiser, O., Puhler, A., Niehaus, K., 2003. Comparison of two *Xanthomonas campestris* pathovar campestris genomes revealed differences in their gene composition. J. Biotechnol. 106, 193–202.

Wei, C.L., Lin, N.T., Weng, S.F., Tseng, Y.H., 1996. The gene encoding UDP-glucose pyrophosphorylase is required for the synthesis of xanthan gum in *Xanthomonas campestris*. Biochem. Biophys. Res. Commun. 226, 607–612.

Whitfield, C., Sutherland, I.W., Cripps, R.E., 1982. Glucose metabolism of *Xanthomonas campestris*. J. Gen. Microbiol. 128, 981–985.

Whitfield, C., 2006. Biosynthesis and assembly of capsular polysaccharides in *Escherichia coli*. Annu. Rev. Biochem. 75, 39–68.

Wiggerich, H.G., Klauke, B., Koplin, R., Priefer, U.B., Puhler, A., 1997. Unusual structure of the *tonB-exb* DNA region of *Xanthomonas campestris* pv. *campestris*: *tonB*, *exbB*, and *exbD*1 are essential for ferric iron uptake, but *exbD*2 is not. J. Bacteriol. 179, 7103–7110.

Wiggerich, H.G., Puhler, A., 2000. The *exbD*2 gene as well as the iron-uptake genes *tonB*, *exbB* and *exbD*1 of *Xanthomonas campestris* pv. *campestris* are essential for the induction of a hypersensitive response on pepper (*Capsicum annuum*). Microbiology 146, 1053–1060.

Yanisch-Perron, C., Vieira, J., Messing, J., 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene 33, 103–119.