J. Biochem. **131**, 183–191 (2002)

Molecular Characterization of a Novel β1,3-Galactosyltransferase for Capsular Polysaccharide Synthesis by *Streptococcus agalactiae* Type Ib[1]

Masaki Watanabe,* Katsuhide Miyake,*,†,2 Kouji Yanae,* Yohei Kataoka,* Satoshi Koizumi,‡ Tetsuo Endo,‡ Akio Ozaki,‡ and Shinji Iijima*

**Department of Biotechnology, Graduate School of Engineering, and †Research Center for Advanced Waste and Emission Management, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8603; and ‡Tokyo Research Laboratories, Kyowa Hakko Kogyo Co., Ltd., 3-6-6, Asahi-machi, Machida, Tokyo 194-8533*

Received October 14, 2001; accepted November 20, 2001

A group B streptococcus, *Streptococcus agalactiae* type Ib, produces a high-molecular-weight polysaccharide consisting of the following pentasaccharide repeating unit: →4)-[α-D-Neu*p*NAc-(2→3)-β-D-Gal*p*-(1→3)-β-D-Glc*p*NAc-(1→3)]-β-D-Gal*p*-(1→4)-β-D-Glc*p*-(1→. The type-specific capsular polysaccharide (CP) synthesis (*cps*) genes of this strain were cloned and analyzed. A cloned 10-kb DNA fragment contained *cpsIbE* to *L* and *neu* (neuraminic acid synthesis gene) *B*. Comparison of the gene products with those of *S. agalactiae* type Ia, which has a similar but distinct CP, showed that the translation products of *cpsIa* and *cpsIb* genes exhibited very high homology except for those of *cpsJ* and *K*. In the type Ia strain, *cpsIaJ* encodes β1,4-galactosyltransferase, which catalyzes the transfer of galactose as the fourth monosaccharide of the sugar repeating unit. In the type Ib CP, this galactose forms a β1,3-linkage to GlcNAc. The low homology between the type Ia and Ib CpsJs seems to reflect this difference. By enzymatic activity measurement, the *cpsIbJ* product was found to display β1,3-galactosyltransferase activity. Furthermore, hydrophobic cluster analysis clarified the similarities and differences of the structures in N-terminal regions, including the DXD motif, between the galactosyltransferases.

Key words: β1,3-galactosyltransferase, capsular polysaccharide, *cps* gene cluster, *Streptococcus agalactiae*.

Group B streptococci, *Streptococcus agalactiae*, are human pathogens causing invasive diseases such as sepsis, meningitis, and pneumonia in infants (*1*). These bacteria produce type-specific capsular polysaccharides (CPs) that distinguish *S. agalactiae* into nine serotypes. The CPs are assumed to be virulence factors as well as targets for protective immunity. The chemical structures of these polysaccharides have already been determined (*2–5*).

Recently, genes involved in CP synthesis (*cps*) and the mechanisms of biosynthesis have been elucidated in many bacteria (*6, 7*). *cps* gene clusters have also been analyzed in many *S. pneumoniae* strains (*8, 9*) and several *S. agalactiae* strains (*10–12*). We have previously analyzed *cps* genes of *S. agalactiae* type Ia (*13*). The type-specific CP of *S. agalactiae* type Ia has a linear backbone of →4)-β-D-Gal*p*-(1→4)-β-D-Glc*p*-(1→ repeating units with trisaccharide side chains of α-D-Neu*p*NAc-(2→3)-β-D-Gal*p*-(1→4)-β-D-Glc*p*NAc-(1→ linked to the C_3 of each β-D-galactose residue of the backbone (*2*) (Fig. 1). The biosynthesis of CPs involves a complex enzymatic pathway starting with the uptake or synthesis of monosaccharides and their activation to nucleotide derivatives. Membrane-bound transferase complexes then catalyze the coupling of the monosaccharides to a membrane-bound lipid carrier, followed by polymerization of the sugar subunits and subsequent export and attachment of the complete CP to the cell surface (*6, 7*).

Since bacterial CPs are diverse in structure and bacteria have a variety of sugar transferases responsible for the synthesis of CP, bacterial *cps* genes have attracted much interest as a source of glycosyltransferases useful for glycoengineering. In a previous study, we detected the activity of four glycosyltransferases in *Escherichia coli* expressing *cpsIa* genes of *S. agalactiae* (*13*). Although the type-specific CP of *S. agalactiae* type Ib is very similar to that of type Ia, the type of sugar linkage of galactose to GlcNAc in the sugar side chain is distinct (Fig. 1). Therefore, comparison of these *cps* gene clusters and their products, especially galactosyltransferases, may improve our understanding of the mechanism of CP synthesis and the structure-function relationship of sugar transferases.

In this study, the structure of the *cps* gene cluster of *S.*

[1] The nucleotide sequence from *cpsIbE* to *neuB* reported in this paper has been submitted to the GenBank database through DDBJ with accession number AB050723.

[2] To whom correspondence should be addressed. Tel: +81-52-789-4278, Fax: +81-52-789-3221, E-mail: miyake@proc.nubio.nagoya-u.ac.jp

Abbreviations: FCHASE, 6-(5-fluorescein-carboxamido)-hexanoic acid succimidyl ester; Gal, galactose; Glc, glucose; GlcNAc, *N*-acetylglucosamine; HPAEC-PAD, high-performance anion-exchange chromatography with pulsed amperometric detection; LNnT, lacto-*N*-neotetraose; LNT, lacto-*N*-tetraose; LNT2, lacto-*N*-triose; NeuNAc, *N*-acetylneuraminic acid.

agalactiae type Ib was examined and compared with that of *S. agalactiae* type Ia. The *cpsIbJ* gene product was identified as a novel β1,3-galactosyltransferase that can synthesize LNT.

MATERIALS AND METHODS

Bacterial Strains, Media, and Materials—*S. agalactiae* type Ib strain OI2 was kindly supplied by Dr. Michio Ohta (Nagoya University). This strain was an occasional isolate from a vaginal swab from a patient with no symptoms of infection. The strain was confirmed to express the type Ib capsule using type Ib–specific antiserum (Denka Seiken, Tokyo), which was prepared with CP of a type strain from the WHO collaborate center, the National Collection of Type Cultures at the Institute of Hygiene, the Czech Republic.

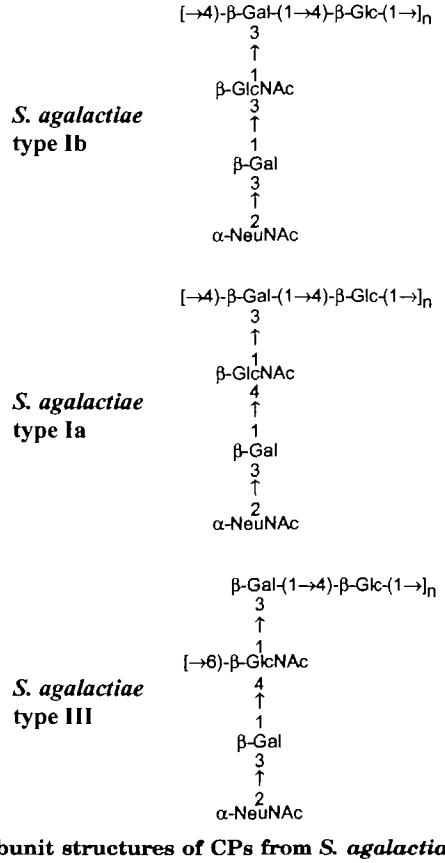

Fig. 1. **Subunit structures of CPs from *S. agalactiae* type Ia, Ib, and III.**

S. agalactiae type Ib strain OI2 was cultured in Todd-Hewitt broth (Becton Dickinson and Company, Sparks, MD) supplemented with 2% glucose and 1.5% Na_2HPO_4 at 37°C (*14*). *E. coli* DH5α was used as the host for the genomic DNA library. *E. coli* JM109 was used as the host for the expression plasmids. All *E. coli* clones were routinely grown in Luria-Bertani broth containing ampicillin.

FCHASE was purchased from Funakoshi (Tokyo). FCHASE-aminophenyl-GlcNAcβ1-3Galβ1-4Glc was synthesized enzymatically from FCHASE-aminophenyl-Galβ1-4Glc and UDP-GlcNAc, with a *N*-acetylglucosaminyltransferase (LgtA). This enzyme was produced by a recombinant *E. coli* overexpressing the *lgtA* gene of *Neisseria gonorrhoeae* (*15*).

DNA Manipulations—Most DNA manipulations were performed according to standard procedures (*16*). Chromosomal DNA was isolated as reported previously (*17*). ^{32}P-labeled probes were prepared with a *Bca*BEST™ labeling kit (Takara, Kyoto). PCR was performed with KOD plus DNA polymerase (Toyobo, Osaka) according to the manufacturer's instructions.

DNA Sequencing—DNA sequences of both strands were determined using an ABI 373S automated DNA sequencer (Applied Biosystems, Foster City, CA). The sequencing data were compared with those in the DDBJ, EMBL and GenBank databases using the BLAST network service at the National Center for Biotechnology Information, National Institutes of Health, Bethesda, MD. The multiple sequence alignments were performed using GENETYX (Software Development, Tokyo). The HCA plots were obtained from the drawhca server on the Internet (http://www.lmcp.jussieu.fr/~soyer/www-hca/hca-form.html). The prediction of α-helices and β-strands was based on the observed association of specific hydrophobic cluster shapes with secondary structures and results obtained using the Jpred2 secondary structure prediction program on the Internet (http://jura. ebi.ac.uk:8888/) (*18*). The three dimensional structures of the CpsIbJ and CpsIaJ proteins were predicted by an automated homology modeling method, which is available on the Internet (http://phychem.pharm.kitasato-u.ac.jp) (*19*).

Construction of Expression Plasmids—For construction of an expression plasmid, the *cpsIbJ* gene was directly amplified with chromosomal DNA from *S. agalactiae* type Ib using cpsIbJ-5'Eco, 5'-CCGGAATTCGAAAAGGTAAAGTGTCTCCGAAA-3' (the *Eco*RI site is underlined) and cpsIbJ-3'Xho, 5'-CGGCTCGAGTCCCAATAGGCGTTGCATC-

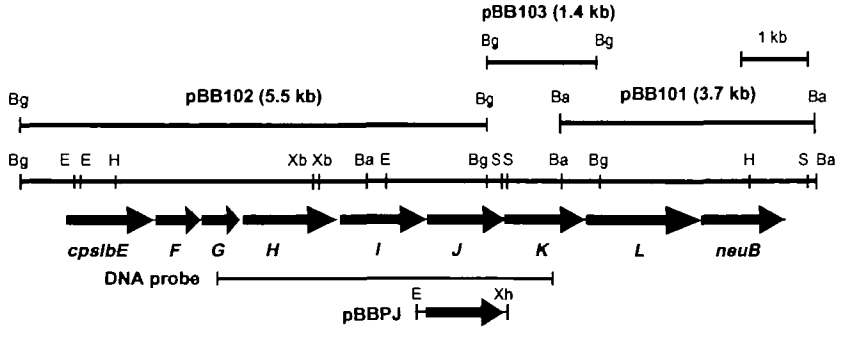

Fig. 2. **Restriction map of the *cps* locus of *S. agalactiae* type Ib.** The locations of ORFs and the direction of transcription are shown by arrows. Gene designations are indicated below as arrows. The DNA probe used for colony hybridization and the expression plasmid used for galactosyltransferase assays are shown below the map. Restriction sites are: Ba, *Bam*HI; Bg, *Bgl*II; E, *Eco*RI; H, *Hin*dIII; S, *Sac*I; Xb, *Xba*I; Xh, *Xho*I.

3′ (the *Xho*I site is underlined). The PCR product was digested with *Eco*RI and *Xho*I to cleave the recognition sites within each primer sequence and ligated into pBluescript II SK+ (Stratagene, La Jolla, CA). The expression plasmid containing the complete coding region of *cpsIbJ* (from 268 bp upstream of the *cpsIbJ* start codon to 274 bp downstream of the stop codon) was designated pBBPJ (Fig. 2). This plasmid was sequenced to confirm that no mutation was introduced. pBAPJ constructed previously was used for the expression of CpsIaJ (*13*). *E. coli* JM109 were transformed with these plasmids. *cpsJ* genes were under the control of the *lac* promoter of pBluescript II SK+. Membranes of recombinant *E. coli* cells were isolated 2 h after induction with 1 mM IPTG.

Galactosyltransferase Assays with Fluorescein-Labeled Oligosaccharides—The recombinant *E. coli* membrane fractions were prepared essentially as described by Kolkman *et al.* (*8*). *E. coli* harboring the plasmid pBluescript II SK+ were used as a negative control. For the galactosyltransferase reaction, 3 μl of membrane fraction (approximately 2.5 μg of protein) was incubated at 37°C for 48 h with LNT2-FCHASE, 5 mM UDP-Gal, 50 mM Tris-acetate, pH 8.3, and 5 mM $MnCl_2$ in a final volume of 15 μl. One-fifth of each these reaction solutions was digested by β1,3- and β1,4-galactosidase (Sigma, St. Louis, MO), respectively. One unit of β1,3-galactosidase was incubated overnight at 37°C with the galactosyltransferase reaction product in 50 mM sodium citrate, pH 4.5. For digestion of β1,4-galactosidase, 0.1 unit of the enzyme was incubated at 37°C 1 h with the sample in 50 mM sodium phosphate, pH 6.0. These reactions were stopped by heating at 70°C for 20 min, and the samples were frozen at –80°C for 30 min. Samples were analyzed by TLC on an HPTLC silica gel plate (Merck, Darmstadt, Germany), developed in ethyl acetate/methanol/water/acetate (7:2:1:0.1), and detected by UV.

Galactosyltransferase Assays with Non-Labeled Oligosaccharides—Acceptor specificity was determined with the membrane fraction of *E. coli* harboring expression plasmids pBAPJ or pBBPJ. *E. coli* harboring the plasmid pBluescript II SK+ were used as a negative control. The reaction was carried out at 37°C for 48 h in 20 μl of 50 mM Tris-acetate, pH 8.3, 5 mM $MnCl_2$, 5 mM acceptor, 5 mM UDP-galactose, and 0.2 mg/ml membrane fraction. The sugar acceptors used in these reactions were LNT2, NOS-B (GlcNAcβ1-3Galβ1-4GlcNAcβ1-3Galβ1-4Glc), chitotriose (GlcNAcβ1-4GlcNAcβ1-4GlcNAc), and GlcNAc. The reactions were stopped by boiling for 2 min, and samples were analyzed by HPAEC-PAD using a DX-500 system with a CarboPac PA-10 column (Dionex, Sunnyvale, CA) (*20*).

RESULTS

Cloning and Sequencing of the cps Locus from S. agalactiae Type Ib—The DNA sequence of the *cps* locus (*cpsA* to *ung*) of *S. agalactiae* type Ia has been reported previously (*13*). Since the DNA sequence of type Ib *cps* genes was expected to be very similar to that of *S. agalactiae* type Ia as judged by the similarity of CPs, we designed primers based on the *cpsIaG* and *cpsIaK* sequences to amplify *cps* genes of the type Ib strain. A 4.4-kb DNA fragment was amplified from chromosomal DNA of *S. agalactiae* type Ib strain OI2 using these primers. This fragment was confirmed to contain *cps* genes by DNA sequencing. To obtain the *cpsE-L* locus from the type Ib strain, two genomic DNA libraries (*Bgl*II and *Bam*HI) of strain OI2 were constructed with pBluescript II SK+ as a vector, and screened by colony hybridization with the PCR product. The location of this probe is shown in Fig. 2. One positive clone and two positive clones were chosen from the *Bam*HI and *Bgl*II libraries, respectively. The positive clone from the *Bam*HI library was designated pBB101 and contained a 3.7-kb *Bam*HI fragment. The two clones from the *Bgl*II library were designated pBB102 and pBB103. These plasmids contained a 5.5-kb *Bgl*II and a 1.4-kb *Bgl*II fragment, respectively. Sequence analysis of these clones revealed that a 10-kb region containing the *cpsE-neuB* genes was cloned.

The DNA sequence of 9,987 nucleotides was determined and nine complete open reading frames (ORFs), designated *cpsIbE* to *cpsIbL*, and *neuB* were identified (Fig. 2). The names of the *cps* genes were assigned according to a previous study (*13*). All ORFs were in the same orientation, and spaced one behind the other at short distances. A possible Shine-Dalgarno sequence was identified just upstream of the potential initiation codon of each ORF. All ORFs except *cpsIbK* were preceded by ATG codons. Only *cpsIbK* was preceded by GTG. The usage of terminator codons is consistent with the usual *E. coli* preference. TAA was used eight times, and TGA was used once.

These gene products show high homology to *cps* genes of several other streptococci. Notably, the *cpsIb* gene products show very high sequence identity with the corresponding *cpsIa* gene products (92%<) except for CpsH, CpsJ, and CpsK (Fig. 3). The previous study revealed that the *cpsIaE*, *cpsIaG*, *cpsIaI*, and *cpsIaJ* gene products display glucosyltransferase, β1,4-galactosyltransferase, β1,3-*N*-acetylglucosaminyltransferase, and β1,4-galactosyltransferase activity, respectively. Therefore, CpsIbE, CpsIbG, and CpsIbI seem to have the same glycosyltransferase activity as those enzymes of *S. agalactiae* type Ia. On the other hand, CpsIbJ shows very low sequence identity with CpsIaJ (15.7%). CpsIbK, which is assumed to be a sialyltransferase (Watanabe, M., unpublished results), also shows relatively low

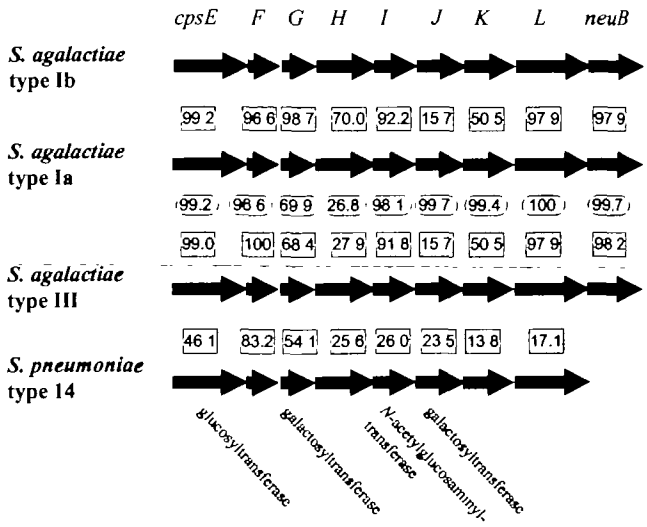

Fig. 3. **Similarity of *cps* gene products between *S. agalactiae* type Ia, Ib, III, and *S. pneumoniae* type 14.** Percentages shown in rectangles are homology to CpsIb proteins, and those shown in orals are homology between *cps* products of type Ia and III strains.

homology to CpsIaK (50.5%) (Fig. 3). Despite the low homology between type Ia and Ib, CpsJ, and K of type Ia and III have almost identical amino acid sequences. Among CPs of these three strains, only type Ib CP has a β1,3-linked galactose as the fourth saccharide (Fig. 1). These differences in homology between *S. agalactiae* strains for the respective genes may be related to the CP structure of each strain. CpsH is unique in the type III strain, although type Ia and Ib CpsH share high homology. The unique structure of CpsIIIH seems to reflect a difference in sugar subunit polymerization, since CpsH has been reported to be a CP polymerase (*11*).

The cpsIbJ Gene Product Shows β1,3-Galactosyltransferase Activity—Our previous study showed that CpsIaJ of *S. agalactiae* type Ia is a β1,4-galactosyltransferase that transfers galactose and produces lacto-*N*-neotetraose (LNnT) (*13*). Although CpsIbJ shows no similarity to the CpsIaJ protein, it was expected to have galactosyltransferase activity from the arrangement of the *cpsIb* gene cluster. To examine the enzymatic activity of CpsIbJ, an expression plasmid was constructed (pBBPJ) and introduced into *E. coli*. The membrane fraction of *E. coli* harboring the expression plasmid was used as a source of enzyme as previously described (*13*) and FCHASE-labeled LNT2 as an artificial acceptor (*15*). As shown in Fig. 4A, the membrane fraction of *E. coli* harboring pBBPJ, which contains *cpsIbJ* downstream of the *lac* promoter, showed clear galactosyltransferase activity. As a positive control, *E. coli* harboring pBAPJ containing *cpsIaJ* was used. These results suggest that the CpsIbJ protein has galactosyltransferase activity as observed for CpsIaJ. To investigate the type of linkage formed by the enzyme reaction, the products were digested with β1,4- and β1,3-specific galactosidases. The reaction product of CpsIbJ was not digested by β1,4-specific galactosidase, while that of *E. coli* containing pBAPJ was cut by this enzyme (Fig. 4A). On the other hand, β1,3-specific galactosidase released galactose from the reaction product of pBBPJ. Together, these results suggest that CpsIbJ has β1,3-galactosyltransferase activity.

The natural acceptor of CpsJ is LNT2 linked to a lipid carrier. To study whether CpsIbJ transfers galactose to the oligosaccharide acceptor without a lipid carrier, non-labeled

TABLE I. **Acceptor specificity of CpsIaJ and CpsIbJ.**

Acceptors	CpsIaJ	CpsIbJ
GlcNAcβ1-3Galβ1-4Glc (LNT2)	1.00	1.00
GlcNAcβ1-3Galβ1-4GlcNAcβ1-3Galβ1-4Glc	5.09×10^{-2}	8.05×10^{-3}
GlcNAcβ1-4GlcNAcβ1-4GlcNAc	5.14×10^{-2}	N.D.
GlcNAc	7.86×10^{-2}	N.D.

The activity for the natural sugar acceptor (LNT2) is set at 1.00. N.D., Not detected.

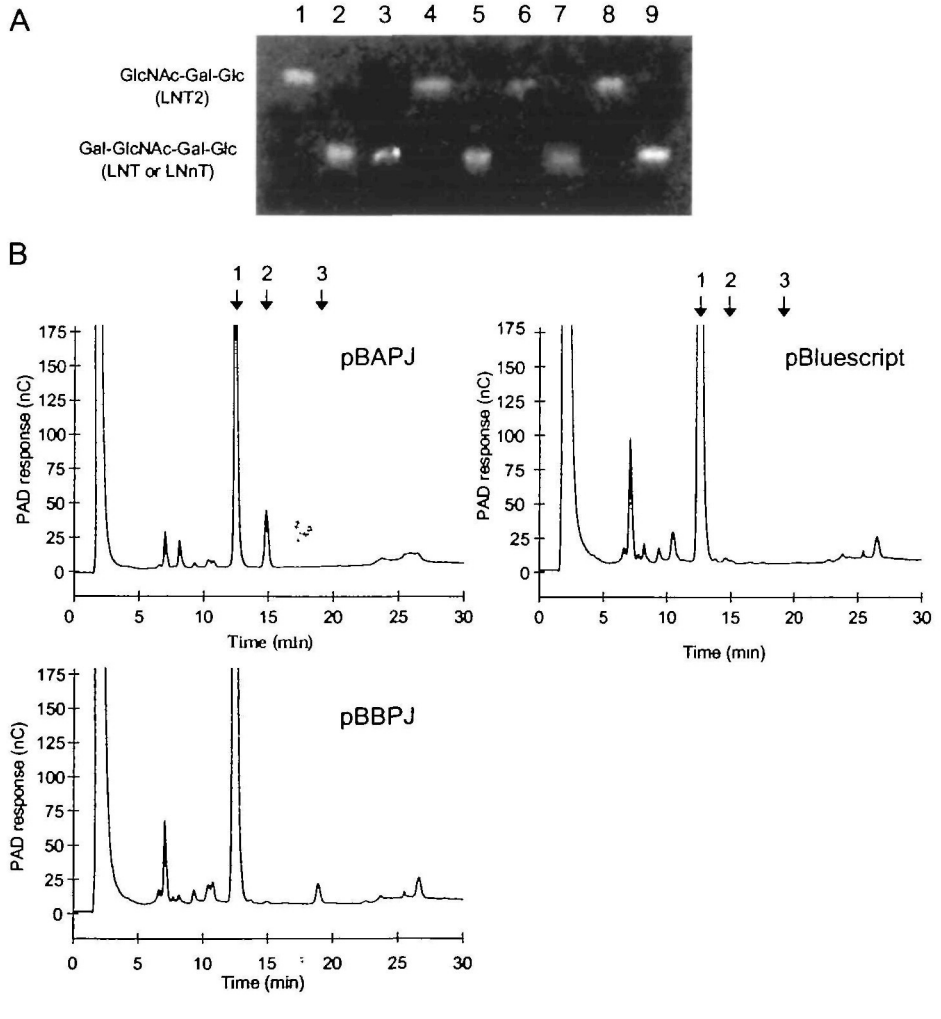

Fig. 4. **Detection of β1,3-galactosyltransferase activity of the *cpsIbJ* gene product.** (A) Thin-layer chromatogram of FCHASE-labeled oligosaccharide produced by CpsIaJ or CpsIbJ enzyme and the degradation products produced by galactosidases. Lane 1, FCHASE-LNT2. The reaction product of the membrane fraction of *E. coli* harboring pBAPJ (lane 2), that digested with β1,3-galactosidase (lane 3), and that digested with β1,4-galactosidase (lane 4). The reaction product of the membrane fraction of *E. coli* harboring pBBPJ (lane 5), that digested with β1,3-galactosidase (lane 6), and that digested with β1,4-galactosidase (lane 7). The reaction product of the membrane fraction of *E. coli* harboring pBluescript II SK+ alone (lane 8); FCHASE-LNnT (lane 9). (B) HPAEC-PAD chromatograms of products of the galactosyltransferase reaction with the membrane fraction of *E. coli* harboring pBAPJ or pBBPJ. The membrane fraction of *E. coli* harboring pBluescript II SK+ alone was used as a negative control. Arrows 1, 2, and 3 indicate the elution positions of sugar standards: 1, LNT2; 2, LNnT; 3, LNT.

LNT2 was reacted with UDP-galactose and membrane fractions of *E. coli* harboring pBBPJ or pBAPJ, and the reaction products were analyzed by HPAEC-PAD. As shown in Fig. 4B, membrane fractions from *E. coli* harboring pBBPJ and pBAPJ produced LNT and LNnT, respectively. This result indicates again that the *cpsIbJ* gene codes for β1,3-galactosyltransferase, and that both CpsIaJ and CpsIbJ can transfer galactose to the sugar acceptor without a lipid carrier.

Acceptor Specificity of CpsIbJ—Since LNT2 contains GlcNAc at the non-reducing end, several GlcNAc-containing oligosaccharides were studied to determine the charac-

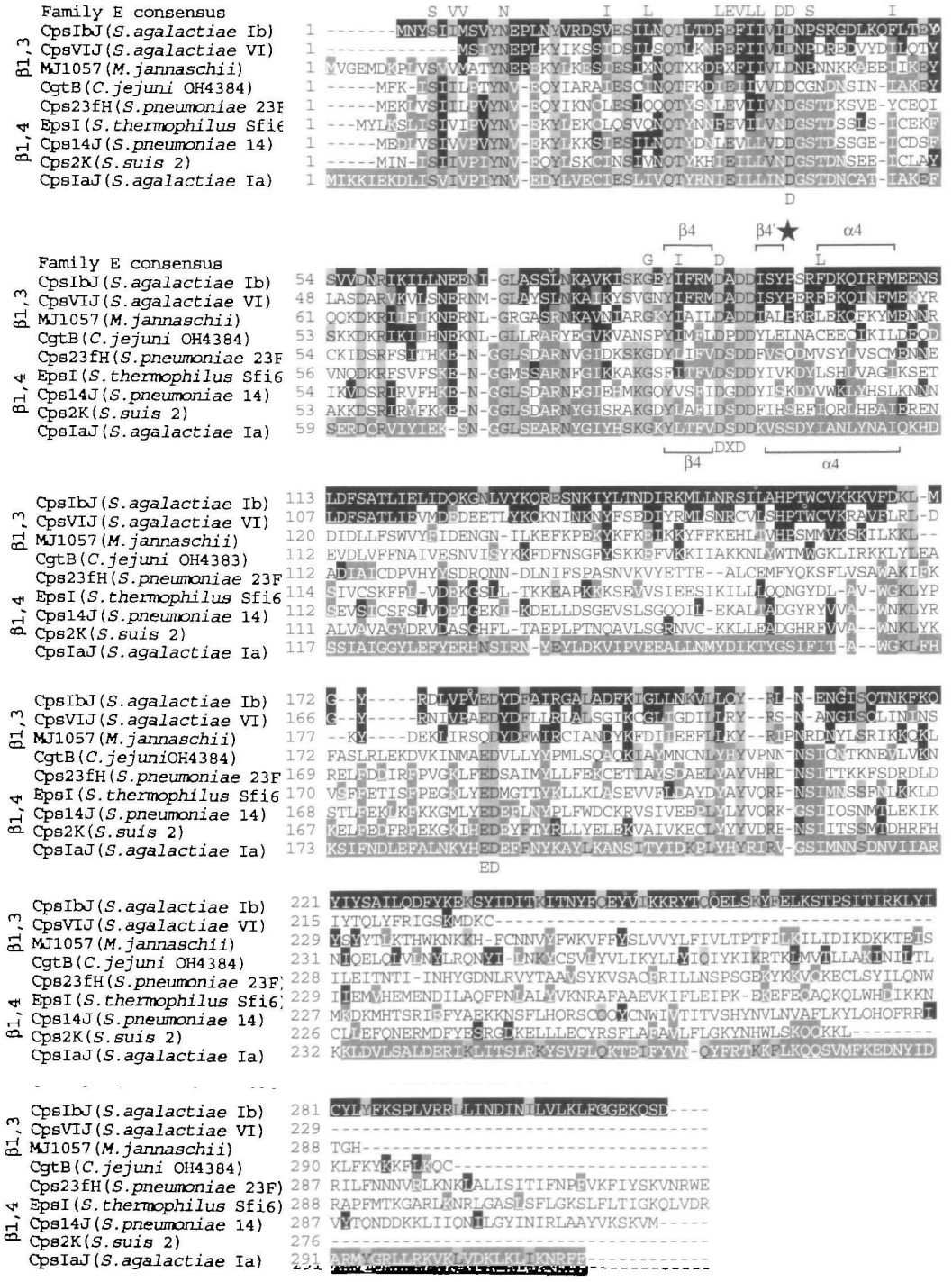

Fig. 5. **Multiple alignment of β1,3- and β1,4-galactosyltransferases derived from *S. agalactiae* and other bacteria.** Amino acids identical to CpsIbJ are given in white letters on a black background, and those identical to CpsIaJ are indicated by white letters on a gray background. Amino acids identical to both CpsIaJ and CpsIbJ are indicated by black letters on a gray background. Family E consensus residues are indicated at the top of the alignment. The DXD motif, ED motif, and the aspartic acid residue following the β2 region are described below the alignment. Predicted secondary structures surrounding the DXD motif of CpsIaJ and CpsIbJ are described in the figure. The proline residue near the DXD motif is indicated by a star.

teristics of the acceptor. As acceptor substrates, GlcNAc, chitotriose, and NOS-B were used. As shown in Table I, CpsIaJ, which has β1,4-galactosyltransferase activity, can transfer galactose to all GlcNAc-containing acceptors, although the enzyme activity for these oligosaccharides was 5–10% of that for LNT2. In contrast, CpsIbJ showed only negligible activity with these saccharides. These results suggest that the CpsIaJ enzyme transfers galactose to sugar acceptors containing just GlcNAc at the non-reducing end, and that the β1,3-galactosyltransferase from the type Ib strain requires the GlcNAcβ1-3Galβ1-4Glc structure for its reaction.

Structural Similarities and Differences between CpsIbJ and CpsIaJ—The amino acid sequences of CpsIbJ and other bacterial β1,3- and β1,4-galactosyltransferases were compared (Fig. 5). CpsIaJ and Cps14J of *S. pneumoniae* type 14 (*8*) are β1,4-galactosyltransferases that transfer galactose to LNT2 and yield LNnT. Cps2K of *S. suis* type 2 (*21*), Cps23fH of *S. pneumoniae* type 23F (*9*), and EpsI of *S. thermophilus* Sfi6 (*22*) are assumed to be β1,4-galactosyltransferases based on the CP structures of the original strains. CgtB of *Campylobacter jejuni* is a β1,3-galactosyltransferase, which has only recently been reported, and is the first example of a β1,3-specific enzyme of bacterial origin. CgtB catalyzes the transfer of galactose to GalNAc (*23*). CpsVIJ of *S. agalactiae* type VI, whose sequence was recently submitted to the GenBank database (GenBank accession number AF337958) by McKinnon *et al.*, is also assumed to be a β1,3-galactosyltransferase based on the CP structure of the bacterium. MJ1057 of *Methanococcus jannaschii* seems to be a β1,3-galactosyltransferase because it shows a high degree of similarity to CpsIbJ (35.5%) (*24*). Analysis of the protein alignment reveals that the N-terminal one-third of the protein sequences are similar among β1,3-galactosyltransferases. The corresponding region shows high similarity among β1,4-galactosyltransferases (Fig. 5). Especially, the DXD motif, which is known to be essential for binding the UDP-galactose donor (*25*), is conserved in all galactosyltransferases including CpsIbJ. Furthermore, the ED motif, which is supposed to play an important role in acceptor recognition (*26*), is also observed in all galactosyltransferases (Fig. 5).

To analyze the structural similarities and differences between these galactosyltransferases in detail, we performed hydrophobic cluster analysis (HCA). HCA is a powerful tool for making sequence comparisons, and can detect similarity in secondary structures among proteins (*27, 28*). This method plots the two-dimensional patterns of protein sequences and allows visual comparisons and detection of conserved structural features. HCA has been successfully applied to glycosyltransferases for the grouping of proteins exhibiting low sequence identity, and for predicting of catalytic residues (*26, 29*).

Using HCA, we compared several bacterial β1,3- and β1,4-galactosyltransferases. The secondary structures were also assigned by Jpred2 servers (*18*). The predicted secondary structures of these enzymes are similar, especially in the N-terminal one-third of the proteins (Fig. 6). However,

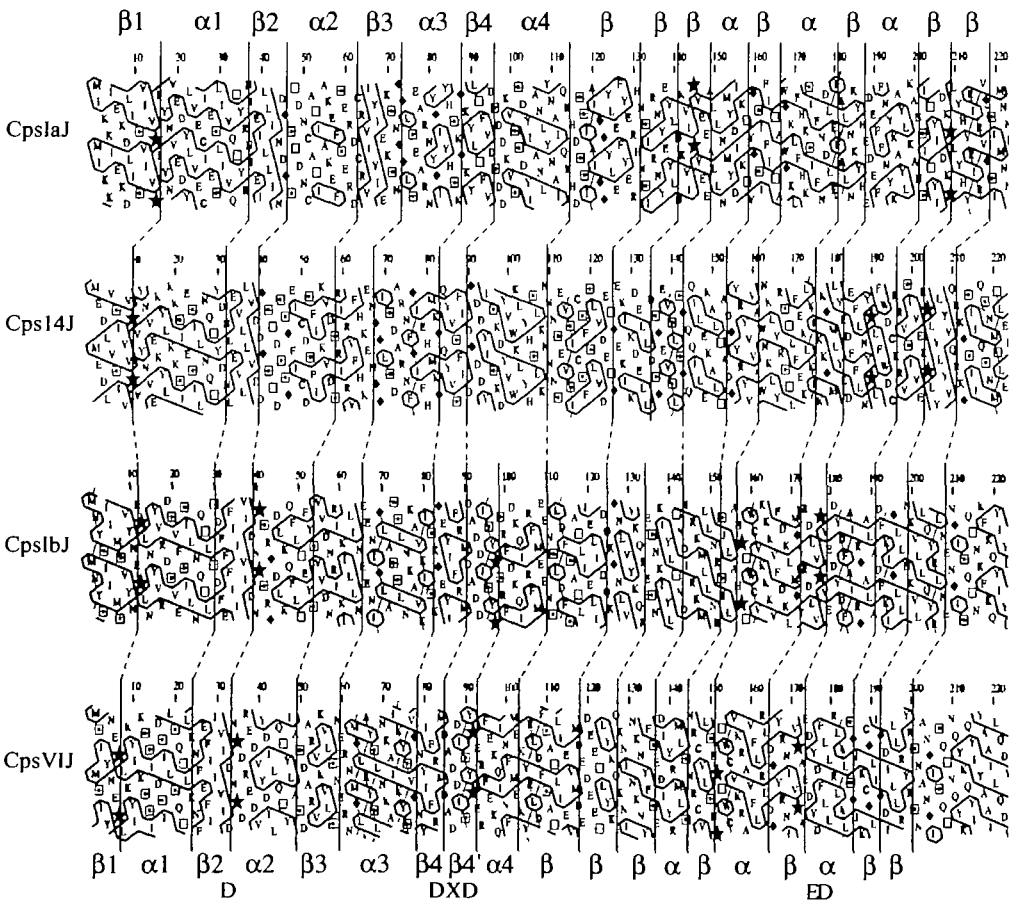

Fig. 6. **HCA plots of β1,3- and β1,4-galactosyltransferases derived from various *S. agalactiae* and *S. pneumoniae* type 14.** The HCA plots were obtained from the drawhca server on the Internet. This program writes protein sequences on a duplicated α-helical net and circles clusters of hydrophobic amino acids (Ile, Leu, Met, Phe, Tyr, Trp, Val). The one-letter code is used for amino acids except for Gly, Pro, Ser, and Thr, which are represented by diamonds, stars, squares with dots, and open squares, respectively. Vertical lines are drawn to indicate structurally conserved features. The predictions of β-strands and α-helices are based on the observed association of specific hydrophobic cluster shapes with secondary structures and results obtained with the Jpred2 secondary structure prediction program on the Internet. The DXD motif, ED motif, and the aspartic acid residue following the β2 region are indicated below the plots.

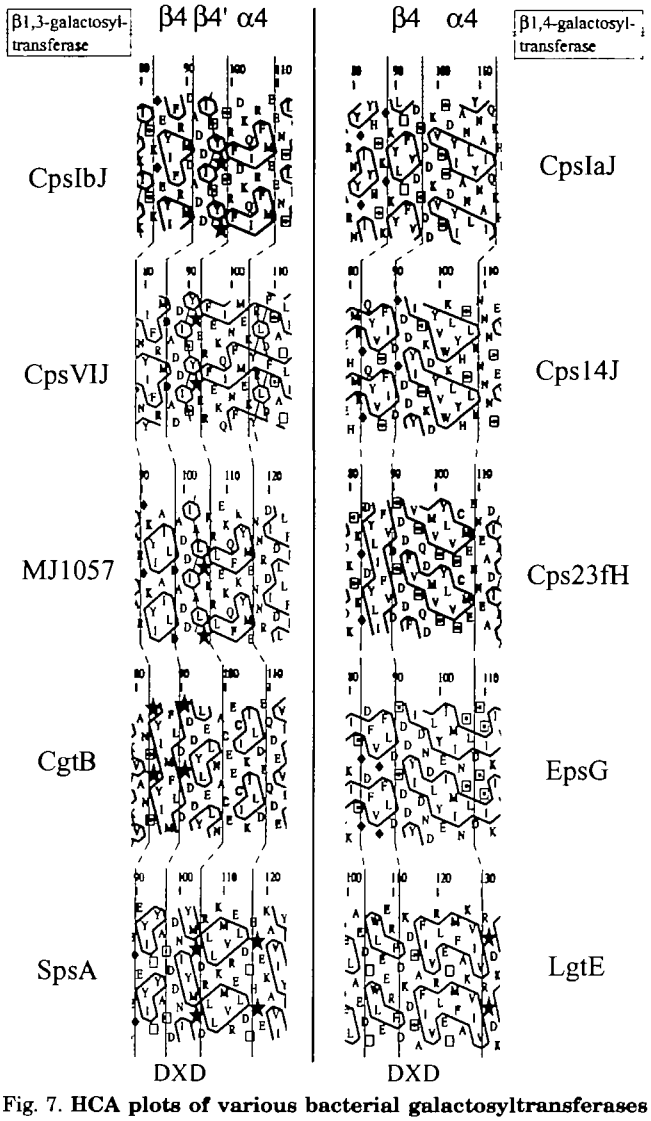

Fig. 7. **HCA plots of various bacterial galactosyltransferases near the DXD motif.** The DXD motif is indicated below the plots.

the C-terminal one-third of the proteins do not show similarity in structure (data not shown). Furthermore, as shown in Fig. 7, there seems to be a difference in the region containing the DXD motif between β1,3- and β1,4-galactosyltransferases. While the motifs of β1,4-enzymes exist between a typical β-strand and an α-helix, a short sequence is inserted immediately downstream of the motifs of β1,3-enzymes. This short sequence is flanked by a proline residue [CpsIbJ, CpsVIJ, MJ1057, and SpsA of *Bacillus subtilis* (*30*)] or by a loop structure of hydrophilic amino acid residues (CgtB). However, no such short sequence could be observed and the motif was directly followed by the α-helix structure in β1,4-galactosyltransferases, including EpsG of *Lactococcus lactis* (*31*) and LgtE of *Neisseria gonorrhoeae* (*32*).

DISCUSSION

Since *S. agalactiae* strains have similar but distinct CPs, they are suitable for studying the structure–function relationship of sugar transferases. In fact, of nine *S. agalactiae* strains, *cps* DNA sequences of type Ia (*13*), Ib (this study), III (*11*), IV (GenBank accession number AF355776), V (AF349539), and VI (AF337958) are available. In the present study, we have cloned and analyzed a novel β1,3-galactosyltransferase gene from *S. agalactiae* type Ib.

Breton *et al.* divided glycosyltransferases into five families, based on sequence alignment and HCA (*33*). The alignment of proteins shown in Fig. 5 reveals that bacterial galactosyltransferases including CpsIbJ retain amino acids characteristic of family E glycosyltransferases. Furthermore, interesting structural features of these bacterial β1,3- and β1,4-galactosyltransferases were identified by HCA. As shown in Fig. 6, the N-terminal halves of *S. agalactiae* galactosyltransferases share a similar type of fold (β1α1β2α2β3α3β4) originally reported as domain A (*29*). In addition, an aspartic acid residue exists in the loop structure following the β2 region in CpsIbJ, as is the case with other galactosyltransferases (Fig. 6) (*29*).

The DXD motif is found at the C-terminal end of a β-structure (β4) (Fig. 6). In β1,4-galactosyltransferases, this motif is followed by an α-helix structure (α4). However, α4 of β1,3-galactosyltransferases seems to be disrupted by a proline residue, unlike the same region in β1,4-galactosyltransferases. In fact, amino acids from Val^{100} to Ile^{112} (VSSDYIANLYNAI) form an α-helix structure in CpsIaJ as shown in Fig. 7. However, the corresponding sequence of CpsIbJ [from Ile^{96} to Met^{108} (ISYPSRFDKQIRFM)] contains a proline residue. Thus, there are four amino acid residues (ISYP) between DXDD and α4 (Fig. 7). Recently, the crystal structure of SpsA from *Bacillus subtilis* was reported (*30*). The enzyme is a glycosyltransferase implicated in the synthesis of the spore coat. As shown in Fig. 7, the structure around the DXD motif is similar between CpsIbJ and SpsA. According to the crystal structure of SpsA, this region of the protein folds ββα as predicted by HCA. Therefore, three dimensional structures of CpsIbJ and CpsIaJ galactosyltransferases were estimated using an automated homology modeling method (*19*) with SpsA as the reference protein. The predicted three dimensional model showed that CpsIbJ has a short β-strand composed of three amino acid residues just downstream of the DXD motif as observed in SpsA protein. On the other hand, the putative β-strand was not found in the CpsIaJ protein (data not shown). Together, the results, shown in Fig. 7, suggest that three amino acid residues between the DXDD sequence and the proline residue may form a β-strand, and the fold around DXDD may be ββα in CpsIbJ. This putative β-strand structure followed by the proline residue or a hydrophilic region was found in all bacterial β1,3-galactosyltransferases reported to date (Fig. 7), and several bacterial glycosyltransferases including ExoO and ExoU (*26*). But no bacterial β1,4-galactosyltransferase contains this β-strand and formed βα fold around the DXD motif as far as we have studied. Since the DXD motif is known to associate with UDP-galactose through a Mn^{2+} or Mg^{2+} ion, the structural difference in this region may affect the selectivity of the sugar linkage. However, we cannot explain the specificity of β1,3- and β1,4-galactosyltransferases just by this fact, since several other bacterial glycosyltransferases that form β1,6- and β1,3-linkages contain this putative short β-strand just downstream of the motif. Furthermore, the C-terminal halves of β1,3- and β1,4-galactosyltransferases

have divergent structures. Thus, it is possible that these regions play a crucial role in generating specific linkages as well as in acceptor recognition.

Several β1,3-galactosyltransferases have been identified in humans (*34*). DXD and ED motifs are also observed in human enzymes, although the location of these motifs in the enzyme molecules differs and the homology of the entire amino acid sequence is low between mammalian and bacterial enzymes. However, HCA and Jpred analysis showed that the ββα structures observed around the DXD motifs of bacterial β1,3-enzymes are also present in all human β1,3-galactosyltransferases (data not shown). On the other hand, human β1,4-galactosyltransferases do not have βα structures observed around the DXD motifs of bacterial enzymes. Together, secondary structures around the DXD motifs are conserved among bacterial and human β1,3-galactosyltransferases.

CpsIbJ is the first example of a bacterial β1,3-galactosyltransferase that can produce LNT from LNT2. LNT is a useful starting material for the production of Galβ1-3GlcNAc-containing type I oligosaccharides such as sialyl Le^a, Le^b, and Le^d. Therefore, we think that CpsIbJ may be a good tool for oligosaccharide production in glycoengineering.

We are grateful to Dr. Michio Ohta, Department of Bacteriology, Graduate School of Medicine, Nagoya University for providing us with *S. agalactiae* OI2 strain. We also thank Dr. Tsuyoshi Shirai, Department of Biotechnology, Graduate School of Engineering, Nagoya University for helpful discussion on the structural prediction of proteins.

REFERENCES

1. Boyer, K.M., Gadzala, C.A., Burd, L.I., Fisher, D.E., Paton, J.B., and Gotoff, S.P. (1983) Selective intrapartum chemoprophylaxis of neonatal group B streptococcal early-onset disease. I. Epidemiologic rationale. *J. Infect. Dis.* **148**, 795–801
2. Wessels, M.R., DiFabio, J.L., Benedi, V.-J., Kasper, D.L., Michon, F., Brisson, J.-R., Jelikova, J., and Jennings, H.J. (1991) Structural determination and immunochemical characterization of the type V group B *Streptococcus* capsular polysaccharide. *J. Biol. Chem.* **266**, 6714–6719
3. Kogan, G., Uhrín, D., Brisson, J.-R., Paoletti, L.C., Kasper, D.L., von Hunostein, C., Orefici, G., and Jennings, H.J. (1994) Structure of the type VI group B *Streptococcus* capsular polysaccharide determined by high resolution NMR spectroscopy. *J. Carbohydr. Chem.* **271**, 8786–8790
4. Kogan, G., Brisson, J.-R., Kasper, D.L., von Hunolstein, C., Orefici, G., and Jennings, H.J. (1995) Structural elucidation of the novel type VII group B *Streptococcus* capsular polysaccharide by high resolution NMR spectroscopy. *Carbohydr. Res.* **277**, 1–9
5. Kogan, G., Uhrín, D., Brisson, J.-R., Paoletti, L.C., Blodgett, A.E., Kasper, D.L., and Jennings, H.J. (1996) Structural and immunochemical characterization of the type VIII group B *Streptococcus* capsular polysaccharide. *J. Biol. Chem.* **271**, 8786–8790
6. Boulnois, G.J. and Jann, K. (1989) Bacterial polysaccharide capsule synthesis, export and evolution of structural diversity. *Mol. Microbiol.* **3**, 1819–1823
7. Boulnois, G.J. and Roberts, I.S. (1990) Genetics of capsular polysaccharide production in bacteria. *Curr. Top. Microbiol. Immunol.* **150**, 1–18
8. Kolkman, M.A.B., Wakarchuk, W., Nuijten, P.J.M., and van der Zeijst, B.A.M. (1997) Capsular polysaccharide synthesis in *Streptococcus pneumoniae* serotype 14: molecular analysis of the complete *cps* locus and identification of genes encoding glycosyltransferases required for the biosynthesis of the tetrasaccharide subunit. *Mol. Microbiol.* **26**, 197–208
9. Ramirez, M. and Tomasz, A. (1998) Molecular characterization of the complete 23F capsular polysaccharide locus of *Streptococcus pneumoniae*. *J. Bacteriol.* **180**, 5273–5278
10. Rubens, C.E., Heggen, L.M., Haft, R.F., and Wessels, M.R. (1993) Identification of *cpsD*, a gene essential for type III capsule expression in group B *streptococci*. *Mol. Microbiol.* **8**, 843–855
11. Chaffin, D.O., Beres, S.B., Yim, H.H., and Rubens, C.E. (2000) The serotype of type Ia and III group B streptococci is determined by the polymerase gene within the polycistronic capsule operon. *J. Bacteriol.* **182**, 4466–4477
12. Cieslewicz, M.J., Kasper, D.L., Wang, Y., and Wessels, M.R. (2001) Functional analysis in type Ia group B *Streptococcus* of a cluster of genes involved in extracellular polysaccharide production by diverse species of Streptococci. *J. Biol. Chem.* **276**, 139–146
13. Yamamoto, S., Miyake, K., Koike, Y., Watanabe, M., Machida, Y., Ohta, M., and Iijima, S. (1999) Molecular characterization of type-specific capsular polysaccharide biosynthesis genes of *Streptococcus agalactiae* type Ia. *J. Bacteriol.* **181**, 5176–5184
14. von Hunolstein, C., Nicolini, L., D'Ascenzi, S., Volpe, C., Alfarone, G., and Orefici, G. (1993) Sialic acid and biomass production by *Streptococcus agalactiae* under different growth conditions. *Appl. Microbiol. Biotechnol.* **38**, 458–462
15. Wakarchuk, W.W., Martin, A., Jennings, M.P., Moxon, E.R., and Richards, J.C. (1996) Functional relationships of the genetic locus encoding the glycosyltransferase enzymes involved in expression of the lacto-*N*-neotetraose terminal lipopolysaccharide structure in *Neisseria meningitidis*. *J. Biol. Chem.* **271**, 19166–19173
16. Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
17. Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Smith, J.A., Seidman, J.G., and Struhl, K. (eds.) (1987) *Current Protocols in Molecular Biology*, John Wiley & Sons, New York, NY
18. Cuff, J.A., Clamp, M.E., Siddiqui, A.S., Finlay, M., and Barton, G.J. (1998) Jpred: a consensus secondary structure prediction server. *Bioinformatics* **14**, 892–893
19. Ogata, K. and Umeyama, H. (2000) An automated homology modeling method consisting of database searches and simulated annealing. *J. Mol. Graph. Model.* **18**, 258–272
20. Lee, Y.C. (1990) High-performance anion-exchange chromatography for carbohydrate analysis. *Anal. Biochem.* **189**, 151–162
21. Smith, H.E., Damman, M., Van Der Velde, J., Wagenaar, F., Wisselink, H.J., Stockhofe-Zurwieden, N., and Smits, M.A. (1999) Identification and characterization of the *cps* locus of *Streptococcus suis* serotype 2: the capsule protects against phagocytosis and is an important virulence factor. *Infect. Immun.* **67**, 1750–1756
22. Stingele, F., Neeser, J.-R., and Mollet, B. (1996) Identification and characterization of the *eps* (exopolysaccharide) gene cluster from *Streptococcus thermophilus* sfi6. *J. Bacteriol.* **178**, 1680–1690
23. Gilbert, M., Brisson, J.-R., Karwaski, M.-F., Michnievicz, J., Cunningham, A.-M., Wu, Y., Young, N.M., and Wakarchuk, W.W. (2000) Biosynthesis of ganglioside mimics in *Campylobacter jejuni* OH4384. Identification of the glycosyltransferase genes, enzymatic synthesis of model compounds, and characterization of nanomole amounts by 600-MHz ^1H and ^{13}C NMR analysis. *J. Biol. Chem.* **275**, 3896–3906
24. Bult, C.J., White, O., Olsen, G.J., Zhou, L., Fleischmann, R.D., Sutton, G.G., Blake, J.A., Fitzgerald, L.M., Clayton, R.A., Gocayne, J.D., Kerlavage, A.R., Dougherty, B.A., Tomb, J.-F., Adams, M.D., Reich, C.I., Overbeek, R., Kirkness, E.F., Weinstock, K.G., Merrick, J.M., Glodek, A., Scott, J.L., Geoghagen, N.S.M., Weidman, J.F., Fuhrmann, J.L., Nguyen, D., Utterback, T.R., Kelley, J.M., Peterson, J.D., Sadow, P.W., Hanna, M.C., Cotton, M.D., Roberts, K.M., Hurst, M.A., Kaine, B.P., Borodovsky, M., Klenk, H.-P., Fraser, C.M., Smith, H.O., Woese,

C.R., and Venter, J.C. (1996) Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. *Science* **273**, 1058–1073
25. Ünligil, U.M. and Rini, J.M. (2000) Glycosyltransferase structure and mechanism. *Curr. Opin. Struct. Biol.* **10**, 510–517
26. Keenleyside, W.J. and Whitfield, C. (1996) A novel pathway for O-polysaccharide biosynthesis in *Salmonella enterica* serovar borreze. *J. Biol. Chem.* **271**, 28581–28592
27. Lemesle-Varloot, L., Henrissat, B., Gaboriaud, C., Bissery, V., Morgat, A., and Mornon, J.P. (1990) Hydrophobic cluster analysis: procedures to derive structural and functional information from 2-D-representation of protein sequences. *Biochimie* **72**, 555–574
28. Callebaut, I., Labesse, G., Durand, P., Poupon, A., Canard, L., Chomilier, J., Henrissat, B., and Mornon, J.P. (1997) Deciphering protein sequence information through hydrophobic cluster analysis (HCA): current status and perspectives. *Cell. Mol. Life Sci.* **53**, 621–645
29. Saxena, I.M., Brown, R.M. Jr., Fevre, M., Geremia, R.A., and Henrissat, B. (1995) Multidomain architecture of β-glycosyl transferases: implications for mechanism of action. *J. Bacteriol.* **177**, 1419–1424
30. Charnock, S.J. and Davis, G.J. (1999) Structure of nucleotide-diphospho-sugar transferase, SpsA from *Bacillus subtilis*, in native and nucleotide-complexed forms. *Biochemistry* **38**, 6380–6385
31. van Kranenburg, R., van Swam, I.I., Marugg, J.D., Kleerebezem, M., and de Vos, W.M. (1999) Exopolysaccharide biosynthesis in *Lactococcus lactis* NIZO B40: functional analysis of the glycosyltransferase genes involved in synthesis of the polysaccharide backbone. *J. Bacteriol.* **181**, 338–340
32. Gotschlich, E.C. (1994) Genetic locus for the biosynthesis of the variable portion of *Neisseria gonorrhoaea* lipooligosaccharide. *J. Exp. Med.* **180**, 2181–2190
33. Breton, C., Bettler, E., Joziasse, D.H., Geremia, R.A., and Imberty, A. (1998) Sequence-function relationships of prokaryotic and eukaryotic galactosyltransferases. *J. Biochem.* **123**, 1000–1009
34. Amado, M., Almeida, R., Schwientek, T., and Clausen, H. (1999) Identification and characterization of large galactosyltransferase gene families: galactosyltransferases for all functions. *Biochim. Biophys. Acta* **1473**, 35–53