Biosynthesis of nucleotide-activated D-glycero-D-manno-heptose*

Bernd Kneidinger‡, Michael Graninger‡, Michael Puchberger§, Paul Kosma§
and Paul Messner‡¶

From the ‡Zentrum für Ultrastrukturforschung und Ludwig Boltzmann-Institut für Molekulare Nanotechnologie, Universität für Bodenkultur Wien, Gregor-Mendel-Str. 33, A-1180 Wien, Austria and §Institut für Chemie, Universität für Bodenkultur Wien, Muthgasse 18, A-1190 Wien, Austria

Running title: GDP-D-glycero-D-manno-heptose Biosynthesis

Keywords: Aneurinibacillus thermoaurophilus/ bacterial glycoprotein/ biosynthesis/ GDP-D-glycero-D-manno-heptose/ S-layer
SUMMARY

The glycan chain repeats of the S-layer glycoprotein of *Aneurinibacillus thermoasterophilus* DSM 10155 contain D-glycero-D-manno-heptose, which has also been described as constituent of lipopolysaccharide cores of Gram-negative bacteria. The four genes required for biosynthesis of the nucleotide-activated form GDP-D-glycero-D-manno-heptose were cloned, sequenced and overexpressed in *Escherichia coli*, and the corresponding enzymes GmhA, GmhB, GmhC and GmhD were purified to homogeneity. The isomerase GmhA catalyzed the conversion of D-sedoheptulose 7-phosphate to D-glycero-D-manno-heptose 7-phosphate, and the phosphokinase GmhB added a phosphate group to form D-glycero-D-manno-heptose 1,7-bisphosphate. The phosphatase GmhC removed the phosphate in the C-7 position, and the intermediate D-glycero-α-D-manno-heptose 1-phosphate was eventually activated with GTP by the pyrophosphorylase GmhD to yield the final product GDP-D-glycero-α-D-manno-heptose. The intermediate and end products were analyzed by high-performance liquid chromatography. Nuclear magnetic resonance spectroscopy was used to confirm the structure of these substances. This is the first report of the biosynthesis of GDP-D-glycero-α-D-manno-heptose in Gram-positive organisms. In addition, we propose a pathway for biosynthesis of the nucleotide-activated form of L-glycero-D-manno-heptose.
INTRODUCTION

The cell surface of many archaea and bacteria is composed of crystalline two-dimensional protein arrays, termed S-layers\(^1\) (1). Frequently, the S-layer proteins are glycosylated (2). The S-layer glycoprotein of the Gram-positive bacterium *Aeurinibacillus thermoaeophilus* DSM 10155, a member of the Bacillus/Clostridium group, is composed of disaccharide repeating units of $\alpha$-L-rhamnose and $D$-$\beta$-$D$-heptose units (3). So far, this is the only Gram-positive bacterium, where $D$-$D$-heptose has been described as constituent of a cellular component. The principal architecture of S-layer glycoproteins resembles that of the LPS of Gram-negative bacteria (4). Both glycoconjugates exhibit a tripartite structural organization, where, in general, conserved core regions connect a glycan chain, composed of identical repeating units, either with the S-layer polypeptide or the lipid A of LPS. It has been proposed that comparable pathways are used for the biosynthesis of these similar glycoconjugates (2). Recently we were able to verify this proposal by investigation of the biosynthesis of nucleotide-activated $D$-rhamnose (5) and $L$-rhamnose (Graninger, M., Kneidinger, B., Bruno, K., Messner, P., unpublished) in *A. thermoaeophilus* strains L420-91\(^T\) and DSM 10155, respectively.

$L$-$D$-heptose is a common constituent of the LPS inner core of enteric and non-enteric bacteria (6). $D$-$D$-heptose, on the other hand, has been described as component of the outer core region of LPS. Numerous bacteria, most importantly the pathogens *Proteus vulgaris* R110/1959 (7), *Haemophilus ducreyi* (8), *Klebsiella pneumoniae* ssp. *pneumoniae* R20 (O1 Lansing:K20+) (9) and *Helicobacter pylori* AF1 and 007 (10) have been shown to contain both types of heptose in LPS: $L$-$D$-heptose in the inner core and $D$-$D$-heptose in the outer core. In *Yersinia*
enterococitica Ye75R both D,D-heptose and L,D-heptose are part of the inner core (11). It
has to be noted, however, that only part of the existing literature on D,D-heptose and
L,D-heptose is listed above, and coverage is confined to those references where complete structural
studies have been performed. L-Rhamnose, the second sugar of the repeating unit of the S-
layer glycan of A. thermoaeophilus DSM 10155 (3), is usually a constituent of the outer core
region as well as of the O-antigen of LPS (6, 12).

The inner core oligosaccharide of most LPS-containing bacteria consists of
3-deoxy-D-manno-oct-2-ulosonic acid and L,D-heptose units. The outer core, however, is composed
predominantly of hexoses and hexosamines (13), but D,D-heptose has also been reported
recently (see above). Impairment of 3-deoxy-D-manno-oct-2-ulosonic acid biosynthesis
usually results in nonviable cells (6). On the other hand, mutants having a defect in the
biosynthesis of L,D-heptose are viable, albeit with certain characteristics, referred to as the
"deep rough" phenotype in Escherichia coli and Salmonella typhimurium. Transduction by
the P1 bacteriophage and F-plasmid conjugation are impaired (14, 15), and the mutants show
increased sensitivity to detergents, bile salts and hydrophobic antibiotics (16). Reduction of
virulence has been reported for Haemophilus influenza (17) and S. typhimurium (18).
Crystallization of the biosynthesis enzymes could lead to the development of enzyme
inhibitors via molecular modeling. These inhibitors can support novel antibiotic therapies to
circumvent drug-resistance.

The biosynthesis of the nucleotide-activated precursor of
3-deoxy-D-manno-oct-2-ulosonic acid, namely CMP-3-deoxy-D-manno-oct-2-ulosonic acid, has been
described previously (13). In contrast, the complete biosynthesis of the nucleotide-activated
form of D/L,D-heptose has not yet been elucidated. Eidels and Osborn (19) have proposed a
four-step pathway for the synthesis of NDP-L,D-heptose: i. D-sedoheptulose 7-phosphate
is converted to $\text{D,D}-\text{heptose 7-phosphate}$ by phosphoheptose isomerase; ii. a mutase catalyzes the second step to form $\text{D,D}-\text{heptose 1-phosphate}$. iii. this intermediate product is activated to NDP-$\text{D,D}-\text{heptose}$ by the action of NDP-heptose synthetase, and iv. in the last step an epimerase catalyzes the formation of the final product NDP-$\text{L,D}-\text{heptose}$. The nucleotide-activated sugar GDP-$\text{D,D}-\text{heptose}$ has been described in baker’s yeast (20), and both ADP-$\text{D,D}-\text{heptose}$ and ADP-$\text{L,D}-\text{heptose}$ have been isolated from Shigella sonnei and Salmonella minnesota mutants (21, 22). Recently, it has been shown that heptosyltransferases I and II from E. coli accept ADP-$\text{L,}\beta-\text{D}-\text{heptose}$ as substrate, and the $\text{D,}\beta-\text{D}$ isomer is also accepted, albeit with tenfold reduced efficiency (23).

The first step in heptose biosynthesis was described in S. typhimurium (24), and the phosphoheptose isomerase gene $\text{gmah}$ from E. coli was cloned and sequenced (25). So far, no gene product, which catalyzes the proposed mutase step has been described. The RfaE protein, also referred to as ADP-heptose synthetase is proposed to be involved in the nucleotide-activating step. This enzyme has been shown to consist of two separate domains in E. coli (26). Domain I displays homology to members of the ribokinase family, whereas domain II is homologous to various cytidylyltransferases. Thus, RfaE is supposed to be a bifunctional enzyme involved in the biosynthesis of $\text{D,D}-\text{heptose 1-phosphate}$ as well as in the nucleotide-activating reaction step. The $\text{rfaD}$ gene product has been proposed to carry out the last step in the biosynthesis of the $\text{L,D}-\text{form}$, namely the epimerization to yield NDP-$\text{L,D}-\text{heptose}$ (27). Recently, RfaD has been crystallized and characterized (28). So far, functional studies have only been performed for the isomerization reaction (24, 25) and for the epimerization step (29). Enzymes that might carry out the mutase reaction, or an alternative phosphatase have not been described.

In this report we show for the first time the overall biosynthesis of a
nucleotide-activated form of D,D-heptose, namely GDP-D-α-D-heptose in Gram-positive bacteria and propose a general reaction pathway for the biosynthesis of activated L,D-heptose in Gram-negative bacteria. In addition, we present a modified nomenclature scheme for genes involved in heptose biosynthesis.
EXPERIMENTAL PROCEDURES

Materials: ATP, GTP, D-sedoheptulose 7-phosphate and dithiothreitol were obtained from Sigma (Sigma Aldrich Fluka GmbH, Wien, Austria).
D-glycero-α-D-manno-heptose 1-phosphate and D-glycero-β-D-manno-heptose 1-phosphate were kindly provided by A. Zamyatina (Universität für Bodenkultur Wien, Wien, Austria). Glutathione (reduced form) was obtained from Merck GmbH (Wien, Austria). GSTrap, HiTrap Chelating, HiPrep Desalting, MonoQ HR5/5 and Sephadex G-10 were purchased from Amersham Pharmacia Biotech (Uppsala, Sweden). The GATEWAY system was obtained from Life Technologies (Wien, Austria), and the pBAD system was from Invitrogen (Groningen, The Netherlands). Ultrafree-MC 10000 ultrafiltration cartridges were purchased from Millipore (Bedford, MA), and pBCKS was obtained from Stratagene (La Jolla, CA).

Bacterial strains, cloning vectors and growth conditions: The E. coli and A. thermoaerophilus strains and cloning vectors used in this study are listed in Table I.
A. thermoaerophilus DSM 10155 was grown in SVIII medium at 55 °C. E. coli DH5α and E. coli TG1 were used for plasmid propagation. For enzyme overexpression E. coli BL21-SI, E. coli BL21-CodonPlus(DE3)-RIL and E. coli LMG194 were used. Protein overexpression was induced by the addition of NaCl, IPTG or L-arabinose, respectively. Media were supplemented with ampicillin, kanamycin or chloramphenicol at concentrations of 100 µg/ml, 50 µg/ml and 34 µg/ml, respectively, when required. Cells were grown at 30 °C or at 37 °C, with or without agitation.
**Analytical techniques:** Nucleotide-activated sugars were analyzed on a CarboPac PA-1 column (Dionex, Sunnyvale, CA) using the method of Palmieri *et al.* (30) with slight modifications. Monosaccharides were analyzed on a CarboPac PA-1 column as described previously (4). Monosaccharide phosphates were analyzed on the CarboPac PA-1 column using a NaOAc-gradient (100 mM to 500 mM) in 100 mM NaOH. Sodium dodecyl polyacrylamide gel electrophoresis was performed according to the original method of Laemmli (31) using slight modifications. Gels were stained with Coomassie Brilliant Blue R250. Protein concentrations were determined by the method of Bradford (32).

**Sequence analysis:** BLAST (33) and MultAlin (34) were used to analyze nucleotide and protein sequences.

**DNA manipulations, PCR and DNA sequencing:** All standard DNA recombinant procedures were performed according to the methods described by Sambrook *et al.* (35) or as recommended by the corresponding manufacturer. PCR was carried out using a PCR Sprint thermocycler (Hybaid, Ashford, UK). DNA sequencing was performed by MWG BIOTECH (Ebersberg, Germany) and AGOWA (Berlin, Germany).

**Cloning of the heptose genes:** In order to clone the genes involved in S-layer glycoprotein glycan biosynthesis, particularly the dTDP-\(\text{L}\)-rhamnose genes, an alignment of 14 RmlA and putative RmlA protein sequences, available in the NCBI database, was carried out using Multalign (34). The highly conserved six amino acid stretch LGDNIF/Y was used to design the degenerate probe 5´-YTI GGI GAY AAY ATH TT-3´ (where I is inosine, H is A/C/T and Y is C/T). This oligonucleotide was 3´-end-labeled with digoxigenin, and
Southern hybridization experiments of completely digested chromosomal DNA of
*A. thermoaerophilus* DSM 10155 yielded specific signals at 35 °C. A 4.5 kb *Bgl*II-fragment
was isolated by preparative agarose gel electrophoresis and cloned into the plasmid pBCKS,
linearized with the endonuclease *Bam*HI. The corresponding construct pRML was sequenced.

**Plasmid construction:** Oligonucleotide primer sequences are given in Table I.

Primers for the amplification of DNA fragments containing either the *gmhA* (primers
GMHAA1 and GMHAA2), the kinase (primers GMHBA1 and GMHBA2) or the
pyrophosphorylase (primers GMHDA1 and GMHDA2) gene were designed with attB1 or
attB2 sites for the insertion into the GATEWAY donor vector pDONR201 by homologous
recombination. The PCR products were cloned into pDONR201 and the resulting plasmids
gGMHA1, gGMHB1 and gGMHD1 were used to transfer the gene sequences into pDEST15
(GST-fusion) or pDEST17 (histidine-fusion) via homologous recombination. The
 corresponding plasmids gGMHA3 (histidine-fusion), gGMHB3 (histidine-fusion) and
gGMHD2 (GST-fusion) were used for overexpression of the fusion proteins in *E. coli*
BL21-SI.

Primers GMHC1F and GMHC2R were used for the amplification of the phosphatase
gene. The PCR product was digested with *Bgl*II and *Eco*RI and ligated into linearized and
dephosphorylated vector pBADB. The resulting plasmid pGMHC1 was used for
histidine-tagged fusion protein overexpression in *E. coli* LMG194.

**Enzyme purification:** *E. coli* BL21-SI cells harboring the histidine-fusion plasmid
gGMHA3 were grown at 37 °C to an optical density at 600 nm of 0.5 in a one-liter culture
volume and expression was induced for 150 min. *E. coli* BL21-SI carrying gGMHB3 was
grown at 30 °C to an optical density at 600 nm of 0.5 in a two-liter culture volume, and expression was induced overnight. *E. coli* LMG194 harboring pgmhC1 was grown at 37 °C to an optical density at 600 nm of 0.5 in a one-liter culture volume, and expression was induced for three hours. Finally, *E. coli* BL21-SI cells carrying the GST-fusion plasmid gGMHD2 were grown at 30 °C to an optical density at 600 nm of 0.6 in a two-liter culture volume, and expression was done overnight. Production of fusion protein was induced by the addition of NaCl to a final concentration of 0.3 M for *E. coli* BL21-SI, or with L-arabinose to a final concentration of 0.2% for *E. coli* LMG194. The cells were disrupted by ultrasonication on ice, cell debris were removed by centrifugation at 31,000 × g, and membrane fractions were collected as a pellet by ultracentrifugation at 331,000 × g. All protein purifications were done on an FPLC system of Amersham Pharmacia Biotech using appropriate amounts of supernatant from the ultracentrifugation step. Purifications using HiTrap Chelating and GSTrap columns were performed as recommended by the manufacturer. For buffer exchange, a HiPrep Desalting column was used. Anion exchange chromatography was carried out in 20 mM Tris-HCl buffer, pH 7.7, using 1 M KCl as eluent. Dithiothreitol was added to a final concentration of 1 mM during all purification steps, except when using HiTrap Chelating columns. Purification of GmhA, the putative kinase and the putative phosphatase was performed using affinity chromatography on a HiPrep Chelating column. GmhB and GmhC were further purified by anion exchange chromatography; the kinase eluted at 300 mM KCl and the phosphatase eluted at 230 mM KCl. The GST-tagged fusion protein of the putative pyrophosphorylase was purified using a GSTrap column. The enzymes were stored at 4 °C or after stabilization with 50% glycerol at −20 °C. The purity of the enzymes was checked by SDS-PAGE analysis (4% stacking gel and 12% separating gel).
Enzyme assays: Ten nmol of D-sedoheptulose 7-phosphate or α-D,D-heptose 1-phosphate were used for enzyme assays. 50 nmol α-D-glucose 1,6-bisphosphate was used in a negative control assay to test the specificity of D,D-heptose 1,7-bisphosphate phosphatase. The assay buffer TH8 contained 20 mM Tris-HCl, pH 8.0 and 10 mM MgCl₂. Appropriate amounts of enzymes were added, and after 45 min incubation at 37 °C the samples were analyzed by HPAEC.

Enzymatic synthesis of D,D-heptose phosphates and GDP-D-α,D-heptose:
One µmol D-sedoheptulose 7-phosphate was incubated with appropriate amounts of GmhA in TH8 buffer at 37 °C until equilibrium was reached. The enzyme was removed by ultrafiltration using Ultrafree-MC 10000 ultrafiltration cartridges. D,D-heptose 7-phosphate was separated from D-sedoheptulose 7-phosphate by preparative HPAEC on a CarboPac PA-1. The sample was desalted using a Carbohydrate Membrane Desalter (CMD; Dionex), lyophilized and investigated by NMR spectroscopy. D,D-heptose 7-phosphate, purified in the first reaction step (approximately 150 nmol), was incubated with 250 nmol ATP and the kinase in 20 mM triethylammonium bicarbonate buffer (pH 8.5; 10 mM MgCl₂) at 37 °C. After removal of the enzyme, the pH was adjusted to 3.0 by addition of Dowex-50 ion exchange resin, the sample was lyophilized and D-α-D-heptose 1,7-bisphosphate was analyzed by NMR spectroscopy. 300 nmol D-sedoheptulose 7-phosphate were converted to D-α-D-heptose 1-phosphate, using appropriate amounts of GmhA, the kinase, the phosphatase and 600 nmol ATP in TH8 buffer at 37 °C. The enzyme mix was removed by ultrafiltration, and D-α-D-heptose 1-phosphate was purified by preparative HPAEC. The desalted (CMD) and lyophilized sample was analyzed...
by NMR spectroscopy. Finally, 2 µmol D-α-D-heptose 1-phosphate were incubated with the pyrophosphorylase and 3 µmol GTP in TH8 buffer at 37 °C. Inorganic pyrophosphatase was used to push the equilibrium reaction towards GDP-D-α-D-heptose. The enzyme was removed, and GDP-D-α-D-heptose was purified on a semipreparative CarboPac PA-1 column. The sample was desalted using Sephadex G-10 and lyophilized. NMR analysis was performed to elucidate the structure of the final reaction product.

NMR spectroscopy: Samples were dissolved in 99.95% D₂O (0.5 ml). Spectra were recorded at 300 K at 300.13 MHz for ¹H and at 75.47 MHz for ¹³C with a Bruker AVANCE 300 spectrometer equipped with a 5 mm QNP-probehead with z-gradients. Data acquisition and processing were performed with the standard XWINNMR software (Bruker).

¹H spectra were referenced to 2,2-dimethyl-2-silapentane-5-sulfonic acid (δ = 0), ¹³C spectra were referenced externally to 1,4-dioxane (δ = 67.40), and ³¹P spectra were referenced externally to H₃PO₄ (δ = 0). HMQC-, ¹H,³¹P-HMQC-, and ¹H,³¹P-HSQC-COSY-spectra were recorded in the phase sensitive mode using TPPI and pulsed field gradients for coherence selection (36, 37). Spectra resulted from a 256 x 4096 data matrix, with 1800 or 800 scans per t1 value, respectively.
RESULTS

Cloning of the heptose operon: The genes coding for enzymes involved in the biosynthesis of a nucleotide-activated sugar precursor generally are clustered within the corresponding gene cluster for a particular bacterial polysaccharide (38). For the genes in the dTDP-L-rhamnose biosynthetic pathway this proposal was verified in Gram-negative as well as in Gram-positive organisms (39, 40, Graninger, M., Kneidinger, B., Bruno, K., Messner, P., unpublished). To clone the genes involved in S-layer glycoprotein glycan biosynthesis of A. thermoaeerophilus DSM 10155, a strategy was applied using a degenerate probe derived from RmlA protein sequences (Fig. 1). Characterization of the rml genes will be published elsewhere. The insert of the cloned construct pRML was sequenced and contained the rmlA gene, incomplete at the 3’-end, and four additional ORFs upstream of rmlA. One of these open reading frames codes for a homologue of GmhA, the first enzyme involved in the biosynthesis of nucleotide-activated heptose. The fourth ORF was incomplete, and its 5’-end was sequenced using a recently described method (41), revealing two additional ORFs. Again, the 5’-end of the most upstream gene was missing.

Sequence analyses: The physical map of the sequenced 6,652 bp DNA fragment is depicted in Fig. 1. According to sequence comparisons with other RmlA proteins, the enzyme lacks only few amino acids on its carboxyl terminus. Upstream of the rmlA gene is a region of 939 bp without any predicted ORF coding for proteins larger than 80 amino acids. The first significant ORF upstream of rmlA codes for a protein of 179 amino acids. A Blast search yielded hypothetical ORFs of Vibrio cholerae, Campylobacter jejuni, E. coli, Helicobacter pylori and Haemophilus influenzae, and the bifunctional HisB protein of Buchnera aphidicola.
and *E. coli*, exhibiting imidazole glycerol phosphate dehydratase and histidinol-phosphate phosphatase activities (42, 43). An alignment of the amino acid sequences encoded by these hypothetical ORFs was performed using Multalign (34, see Fig. 2). The six hypothetical proteins showed remarkable homology, and thus, are expected to carry out the same catalytic reaction (Table II). The next ORF of the cloned fragment encodes a protein with similarities to mannose 1-phosphate guanosyltransferases as well as to other pyrophosphorylases transferring thymidine, adenine, cytidine or uridine nucleotides to sugar phosphates. The third gene upstream of *rmlA* encodes a homologue of GmhA, the sedoheptulose 7-phosphate isomerase, catalyzing the conversion of D-sedoheptulose 7-phosphate to D,D-heptose 7-phosphate. Upstream of *gmhA* lies an ORF coding for a protein similar to hypothetical proteins of *C. jejuni, Mycobacterium tuberculosis* and *Thermoplasma acidophilum* and to galactokinases from several organisms. The remaining two ORFs both show weak similarities to glycosyl transferases. In summary, the 5′-part of the sequenced fragment contains genes coding for GmhA, a putative sugar kinase, a putative phosphatase, a hypothetical pyrophosphorylase and two putative glycosyl transferases. Four of these proteins were hypothesized to function in the biosynthesis of nucleotide-activated heptose (Fig. 3). In the predicted pathway, accommodating the putative activities of the gene products, GmhA converts D-sedoheptulose 7-phosphate to an anomeric mixture of D,D-heptose 7-phosphate. Instead of the originally proposed mutase step to get D,D-heptose 1-phosphate from D,D-heptose 7-phosphate (19) two enzymes may actually catalyze this conversion. The putative kinase adds a phosphate group at the C-1 hydroxyl group of D,D-heptose 7-phosphate to form D,D-heptose 1,7-bisphosphate, and the putative phosphatase removes the phosphate group at the C-7 position to yield D,D-heptose 1-phosphate. Finally, the putative pyrophosphorylase transfers a nucleotide to the heptose unit to give
NDP-\textsubscript{D}D-heptose. To test this hypothesis, we decided to overexpress the corresponding proteins in order to perform functional studies.

*Overexpression of the heptose biosynthesis genes:* Overexpression of these genes from a Gram-positive organism in *E. coli* proved to be a difficult task. Codon usage differs remarkably in these two organisms. Different expression systems had to be used to finally achieve overexpression of all four genes. GmhA expression could be performed using pBAD and GATEWAY systems. For enzyme purification a histidine-tagged protein was expressed with the GATEWAY vector gGMHA3, and purification was carried out using a HiTrap Chelating column. Overnight expression of histidine-tagged putative kinase using the GATEWAY vector gGMHB3 yielded sufficient amount of protein for purification over HiTrap Chelating and MonoQ columns. The putative phosphatase was expressed using the pBAD construct pgmhC1, and the histidine-tagged fusion protein was purified like the kinase. Most difficult was the overexpression of the last enzyme of the putative pathway, namely the pyrophosphorylase. Expression of histine-tagged fusion proteins using the pBAD as well as the GATEWAY systems occurred at a negligible level. However, low amounts of GST-tagged fusion protein could be purified after overnight expression with the GATEWAY vector gGMHD2, using a GSTrap column. The molecular masses of the denatured proteins were in good agreement with the calculated masses, as judged by SDS-PAGE analysis (GmhA, 25 kDa; putative kinase, 41.6 kDa; putative phosphatase, 25.2 kDa; putative pyrophosphorylase, 55.2 kDa; Fig. 4). The purified proteins were used for functional characterization of the biosynthetic pathway of nucleotide-activated \textsubscript{D}D-heptose.

*Synthesis of \textsubscript{D}D-heptose 1-phosphate from \textsubscript{D}-sedoheptulose 7-phosphate:* The
conversion of D-sedoheptulose 7-phosphate to D,D-heptose intermediate products was analyzed by phosphate anion exchange chromatography on a CarboPac PA-1 column combined with NMR experiments. GmhA converted D-sedoheptulose 7-phosphate to an anomic mixture of D,D-heptose 7-phosphate in an equilibrium reaction, where approximately 20% of D-sedoheptulose 7-phosphate were isomerized to D,D-heptose 7-phosphate (Fig. 5A). The retention time of D,D-heptose 7-phosphate (24.5 min) resembled that of D-glucose 6-phosphate (22.8 min). This first intermediate product was hydrolyzed with trifluoroacetic acid, and monosaccharide analysis on a CarboPac PA-1 column identified the sugar as D,D-heptose (not shown). Further characterization was performed using NMR analysis. The $^1$H-NMR spectrum of D,D-heptose 7-phosphate indicated the presence of two anomers and due to the small amount of material, the spectrum could not be fully assigned. However, measurement of a H,P-correlated spectrum indicated the occurrence of a phosphomonoester $^{31}$P signal at $\delta$ 3.30 which was coupled to a proton signal at $\delta$ 3.96. In addition, HMQC correlation indicated the connectivity of this proton to a $^{13}$C-signal at $\delta$ 65.5 and another proton at $\delta$ 4.10. Unambiguous evidence for the phosphate substitution at the terminal hydroxymethyl group was obtained from the downfield shift of the C-7 signal and the upfield-shifted signal of the neighboring C-6 signal at $\delta$ 71.8. The connectivity of H- and C-6 to the phosphate-substituted geminal CH$_2$O-protons was established from an H,P-correlated COSY experiment. These data are in full agreement with similar effects observed in the spectra of $\text{L-glycero-}$D-manno-heptopyranose 7-phosphate (44). To investigate the next reaction step, D-sedoheptulose 7-phosphate was incubated with GmhA, the putative sugar kinase, and ATP as phosphate donor. In HPAEC analysis the D-sedoheptulose 7-phosphate peak decreased, D,D-heptose 7-phosphate disappeared, and a
new peak was detected, displaying a retention time of 31.2 min (Fig. 5B). Again, a peak occurred at a similar retention time compared with the corresponding glucose derivative, namely \( \alpha\-D\)-glucose 1,6-bisphosphate (31.6 min). The signal of \( \alpha\-D\)-glucose 1,6-bisphosphate was very weak and gave rise to a very small peak in the chromatogram. As free hydroxyl groups are electrochemically detected in Dionex-HPAEC, signal height decreases with decreasing number of OH-groups. NMR analysis was used to elucidate the structure of \( D\,D\)-heptose 1,7-bisphosphate. For the NMR experiments \( D\,D\)-heptose 1,7-bisphosphate was synthesized from purified \( D\,D\)-heptose 7-phosphate. The H,P-correlated spectrum of the \( D\-\alpha\-D\)-heptose 1,7-bisphosphate displayed a phosphomonoester \( ^{31}P \) signal at \( \delta \) 1.22 coupled to the H-7 proton signals at \( \delta \) 3.98 and 3.84, respectively. An additional \( ^{31}P \) signal at \( \delta \) -1.44 was coupled to the anomeric proton at \( \delta \) 5.39, thereby proving the presence of the anomeric phosphate substituent. A mixture of \( D\)-sedoheptulose 7-phosphate and ATP was also incubated with only the kinase. The peak pattern did not change, which indicates the specificity of the enzyme. The third step of the hypothetical reaction sequence was supposed to be the removal of the phosphate on the C-7 position. This was investigated in a reaction cascade using the three enzymes GmhA, the putative kinase and the putative phosphatase. \( D\)-sedoheptulose 7-phosphate was incubated with the enzyme mix and ATP. The \( D\)-sedoheptulose 7-phosphate peak decreased substantially, and the peaks corresponding to \( D\,D\)-heptose 7-phosphate and \( D\,D\)-heptose 1,7-bisphosphate disappeared completely. A new peak was detected at 14.2 min, which is close to that of \( \alpha\-D\)-glucose 1-phosphate (11.3 min; Fig. 5C). From the comparison with chemically synthesized standards of \( D\-\alpha\-D\)-heptose 1-phosphate (14.1 min) and \( D\-\beta\-D\)-heptose 1-phosphate (13.2 min) (23) it was possible to identify the
anomeric configuration of the third intermediate product to be the α-anomer (Fig. 5C). Final characterization of D-α-D-heptose 1-phosphate was performed by NMR analysis, which displayed a single $^31$P signal $\delta$ 2.53 coupled to the anomeric proton ($\delta$ 5.32). Furthermore, the spectral characteristics were in good agreement with the synthetic compound as well as with published data (45). To test the specificity of the phosphatase

α-D-glucose 1,6-bisphosphate was used as a substrate for this enzyme. Similar to the kinase negative control the peak pattern did not change. Whereas the first reaction step, catalyzed by GmhA, is an equilibrium reaction, the next two reactions appear to proceed quantitatively.

Activation of D-α-D-heptose using GTP: The final step in the biosynthetic pathway of nucleotide-activated D,D-heptose is the transfer of an NDP residue to D,D-heptose in the C-1 position. Guanosine and adenosine had previously been described as the activating nucleotide (20-22). To test the proposal, which nucleotide is actually used in the Gram-positive organism *A. thermoaerophilus* DSM 10155, D-α-D-heptose 1-phosphate was incubated with the putative pyrophosphorylase and ATP or GTP, respectively. The reaction mixtures were analyzed by HPAEC using a CarboPac PA-1 column. No new peak appeared, when ATP was used in the reaction (Fig. 6A). However, when the guanosine nucleotide was used, the GTP peak decreased, and a new peak with a retention time slower than GMP (13.3 min), but faster than GDP (37 min), was detected (Fig. 6B). The retention time of 25.9 min was similar to that of GDP-D-mannose (28.4 min). The reaction was also carried out with CTP, dTTP and UTP, but no additional peaks could be detected (data not shown). Thus, guanosine diphosphate is the activating nucleotide of D-α-D-heptose in the biosynthesis of the S-layer glycoprotein glycan of *A. thermoaerophilus* DSM 10155. The end product of the pathway was fully characterized using NMR analysis. The structure of the
sugar nucleoside diphosphate, GDP-\(\text{D-glycero-\(\alpha\)-D-manno}\)-heptopyranose was unambiguously confirmed by the NMR data and by comparison with both anomers of synthetic ADP-\(\text{D-glycero-\(\alpha\)-D-manno}\)-heptopyranose (Fig. 7). The proton and carbon NMR signals of position 8 of the guanine unit were observed at \(\delta\) 8.09 and \(\delta\) 138.6, respectively. In addition, the \(1^H\) NMR signals of the \(\beta\)-\(\text{D-ribofuranosyl}\) residues were clearly separated from those of the heptose, which allowed the complete assignment of all proton signals. Furthermore, HMQC experiments allowed a partial detection of the connected carbons (Table III). Both \(^{31}\text{P}\) signals of the pyrophosphate group were detected and correlated to the anomeric proton of the heptose and the H-5 protons of the ribose residue.
In this report we have elucidated for the first time the overall pathway for biosynthesis of a nucleotide-activated D-α-D-heptose from D-sedoheptulose 7-phosphate. The heptose genes of *A. thermoaerophilus* DSM 10155 are clustered and could be identified upstream of the *rmlA* gene. In this organism also all genes belonging to the dTDP-L-rhamnose operon are clustered (Graninger, M., Kneidinger, B., Bruno, K., Messner, P., unpublished). However, the genes for heptose biosynthesis are not clustered in *E. coli*. In *E. coli* K-12, the ADP-L-heptose epimerase gene *rfaD*, the heptosyltransferase II gene *rfaF* and the heptosyltransferase I gene *rfaC* are part of a single operon, whereas *gmhA* and ADP-heptose synthetase genes are located elsewhere. The lack of an organized cluster of heptose biosynthesis genes in prominent laboratory organisms like *E. coli* or *Salmonella enterica* has hampered determination of the pathway, as has the inability to identify a gene encoding an enzyme catalyzing the mutase step, proposed in the biosynthetic pathway of Eidels and Osborn (19).

In addition, until now, genetic complementation of heptose-deficient mutants did not identify an enzyme with phosphatase homology. In the Gram-positive bacterium *A. thermoaerophilus* DSM 10155, however, four enzymes, the genes of which are part of the same operon, are sufficient to synthesize GDP-D-α-D-heptose from D-sedoheptulose 7-phosphate. This operon extends in the upstream direction, but the whole cluster has not been sequenced yet. Upstream of the four heptose genes are hypothetical ORFs, that encode putative glycosyl transferases (Fig. 1), likely involved in S-layer glycan synthesis.

Overexpression of the heptose biosynthesis genes was difficult due to differing codon usage in *E. coli* and the Gram-positive bacterium used in this study. The use of the special expression host *E. coli* BL21-CodonPlus-RIL, which harbors a plasmid carrying extra copies
of rare tRNA genes for arginine (codons AGA and AGG), isoleucine (codon AUA) and leucine (codon CUA), did not enhance expression of the genes. Thus, codons for other amino acids seem to be the reason for the low expression levels. Leucine UUA, valine GUA, threonine ACA, arginine CGA and glycine GGA are used at a threefold higher rate in the *A. thermoaeophilus* DSM 10155 heptose genes in comparison to the *E. coli* K-12 genome. However, the use of different expression systems eventually led to successful overexpression of all four heptose genes. Problems in overexpressing both the kinase and the pyrophosphorylase could only be overcome by overnight expression, although even under these conditions expression of the latter enzyme was still at a very low level.

The four steps of the reaction cascade synthesizing GDP-\(\text{D-}\alpha\text{-D}\)-heptose from \(\text{D}\)-sedoheptulose 7-phosphate in *A. thermoaeophilus* were unambiguously characterized by HPAEC combined with NMR-spectroscopy. \(\text{D}\)-Sedoheptulose 7-phosphate was converted to an anomic mixture of \(\text{D, D}\)-heptose 7-phosphate in an equilibrium reaction. \(\text{D, D}\)-heptose 7-phosphate was further converted to \(\text{D-}\alpha\text{-D}\)-heptose 1,7-bisphosphate in a quantitative reaction step using ATP as phosphate donor. \(\text{H, P}\)-correlation spectroscopy confirmed that there were two phosphate groups on one \(\text{D-}\alpha\text{-D}\)-heptose unit, one phosphorus atom being spin-coupled to the anomic proton of C-1, and the other coupled to the two protons of C-7. Thus, a mutase reaction can be excluded, and synthesis of \(\text{D-}\alpha\text{-D}\)-heptose 1-phosphate from \(\text{D, D}\)-heptose 7-phosphate proceeds via two reaction steps involving a kinase and a phosphatase. The phosphatase reaction was convincingly demonstrated in a three-step reaction cascade, where \(\text{D-}\alpha\text{-D}\)-heptose 1-phosphate was isolated by preparative HPAEC, which was identical to authentic synthetic material. The final product of the heptose biosynthetic pathway is GDP-\(\text{D-glycero-}\alpha\text{-D-manno}\)-heptose.

As described before, neither an enzyme catalyzing the mutase step nor an alternative
monofunctional phosphatase has been reported, so far. A BLAST search using the
*A. thermoaerophilus* phosphatase yielded putative ORFs from *Vibrio cholerae* N16961,
*Campylobacter jejuni* NCTC 11168, *E. coli* K-12, *Helicobacter pylori* 26695 (and also J99)
and *Haemophilus influenzae* Rd. Accession numbers of putative heptose biosynthesis genes
of the completely sequenced bacterial strains are given in Table II. For brevity of this work,
only one of the sequenced strains of *Helicobacter pylori* (26695) and *Neisseria meningitidis*
(MC58) was taken into account, but the genes are found in the genomes of additional strains.
Judging from the high homology throughout the whole protein sequence, all of these enzymes
are expected to carry out the same catalytic reaction (Fig. 2). For the conversion of the
anomeric mixture of D,D-heptose 7-phosphate to D,D-heptose 1-phosphate it seems
reasonable that in Gram-positive and Gram-negative organisms either the α- or the
β-anomer is taken up selectively by a phosphokinase (Fig. 8). The subsequent removal of the
7-phosphate group is expected to be performed by a phosphatase. Such phosphatases can also
be found in *Pseudomonas aeruginosa* PAO1 and *Neisseria meningitidis* MC58 and Z2491,
however, with lower homologies (Table II). To date, all heptose-containing bacteria, that
have been completely sequenced, contain a gene encoding a protein homologous to the
phosphatase GmhC of *A. thermoaerophilus*. *Mycobacterium tuberculosis* H37Rv also has a
gene homologous to this phosphatase despite the fact that there is currently no report of
heptose-containing saccharides in this organism. Homologues of the first three enzymes of
the heptose pathway are part of a single operon, whereas the last enzyme can also be found
elsewhere on the chromosome in this strain.

As far as currently known from Gram-negative bacteria (except *N. meningitidis*, see
below), enzymes for heptose biosynthesis appear to be encoded by four genes, one of which
encodes a bifunctional protein (26). This bifunctional enzyme putatively catalyzes two
non-contiguous steps of the pathway, namely the addition of a phosphate residue to
\textit{D,D}-heptose 7-phosphate to give \textit{D,D}-heptose 1,7-bisphosphate and the transfer of the
nucleotide to \textit{D,D}-heptose 1-phosphate to yield ADP-\textit{D,D}-heptose. The four genes \textit{gmhA},
the bifunctional putative kinase and pyrophosphorylase, the putative phosphatase and
ADP-\textit{L,D}-heptose epimerase can be found on the chromosomes of \textit{C. jejuni} NCTC 11168, \textit{E. coli}
K-12, \textit{Haemophilus influenzae} Rd, \textit{Helicobacter pylori} strains 26695 and J99 and \textit{P. aeruginosa}
PAO1 (Table II). In \textit{E. coli} K-12, \textit{Ha. influenzae} Rd and \textit{P. aeruginosa} PAO1 these genes are
scattered on the chromosome, whereas in \textit{C. jejuni} NCTC 11168 and \textit{He. pylori} 26695 and
J99, the genes are part of a single operon. \textit{Neisseria meningitidis} strains MC58 and Z2491
each possess five heptose biosynthesis genes scattered all over their chromosomes. In this
organism the bifunctional putative kinase and pyrophosphorylase is split into two separate
enzymes. \textit{Vibrio cholerae} N16961 possesses \textit{gmhA} and an unlinked phosphatase gene, but
lacks homologues of the kinase, the guanosyltransferase, the bifunctional putative
kinase/pyrophosphorylase and the epimerase. \textit{Campylobacter jejuni} NCTC 11168 seems to
carry two sets of heptose biosynthesis genes on its chromosome. They are organized in two
operons: genes homologous to \textit{gmhA}, ADP-heptose synthetase, ADP-\textit{L,D}-heptose
epimerase, and a putative phosphatase are part of the same operon, which also contains
glycosyl transferases. A second operon contains a possible sugar-phosphate
nucleotidytransferase, an additional \textit{gmhA} gene, and a putative sugar kinase as well as
possible glycosyl transferases, but lacks a phosphatase homologue. However, it is not known,
whether \textit{C. jejuni} NCTC 11168 contains both \textit{D,D}-heptose and \textit{L,D}-heptose as cellular
constituents. In the LPS core of this strain only the \textit{L,D}-form of heptose has been reported
\cite{46}. Additional glycoconjugates are known to occur in the cell envelope of this organism
\cite{47}. Although there was no indication for the presence of heptose in these glycoconjugates
two separate sets of genes would potentially facilitate for incorporation of \( \text{D, D}-\text{heptose} \) or \( \text{L, D}-\text{heptose} \) into the corresponding polysaccharide structure.

The biosynthetic pathway of heptose seems to diverge during the kinase step (Fig. 8). The kinase GmhB from \textit{A. thermoaeophilus} catalyzes the formation of \( \text{D-} \alpha\text{-D-heptose 1,7-bisphosphate} \). Since homologies from GmhB of \textit{A. thermoaeophilus} to the kinase part of ADP-heptose synthetase from different Gram-negative bacteria are low, in the case of ADP-heptose biosynthesis this intermediate might be \( \text{D-} \beta\text{-D-heptose 1,7-bisphosphate} \).

For the phosphatase step, which does not involve the anomeric phosphate group, the homology to the corresponding genes is high (Fig. 2). The \( \text{D, D}-\text{heptose 1-phosphate} \) guanosyltransferase accepts \( \text{D-} \alpha\text{-D-heptose 1-phosphate} \), whereas the \( \text{D, D}-\text{heptose 1-phosphate} \) adenosyltransferase would accept the putative \( \text{D-} \beta\text{-D-heptose 1-phosphate} \).

Recently, a general nomenclature for bacterial polysaccharide synthesis has been introduced (48). The three-letter gene name \textit{gmh} (\textit{glycerol-manno-heptose}) is used for the heptose biosynthesis genes. As we have shown in \textit{A. thermoaeophilus}, there are four steps to get GDP-activated \( \text{D, D}-\text{heptose} \) from \( \text{D-sedoheptulose 7-phosphate} \). Thus, the suffixes A-D are sufficient to describe this pathway (Fig. 8). In Gram-negative bacteria, however, suffixes A-E would be required to identify the five steps including the epimerization step to make \( \text{L, D}-\text{heptose} \). \textit{gmhA} has been assigned to sedoheptulose 7-phosphate isomerase and \textit{gmhD} has been proposed for the ADP-\( \text{L, D}-\text{heptose} \) epimerase gene. Since the latter enzyme catalyzes the fifth step in the heptose pathway of Gram-negative bacteria, we suggest this gene to be named \textit{gmhE}. \textit{gmhB} with subscript \( \alpha \) or \( \beta \) might be assigned to the gene encoding \( \text{D, D}-\text{heptose 7-phosphate kinase} \), and the proposed name for the gene encoding the \( \text{D, D}-\text{heptose 1,7-bisphosphate phosphatase} \) should be \textit{gmhC}. Finally, \textit{gmhD} with the subscript
NDP (e.g. $gmhD_{\text{ADP}}$ or $gmhD_{\text{GDP}}$) might be assigned to the gene encoding the enzyme carrying out the D,D-heptose 1-phosphate nucleosyltransferase step. As described above, a bifunctional enzyme, currently named ADP-heptose synthetase, putatively catalyzes the addition of a phosphate group to D,D-heptose 7-phosphate and the activation of D,D-heptose 1-phosphate. In our scheme these two reactions are catalyzed by $gmhB$ and $gmhD$, respectively. Therefore, $gmhBD$ would be an appropriate name for the gene encoding this bifunctional enzyme (Fig. 8).

In this study we have shown biosynthesis of GDP-D-\(\alpha\)-D-heptose in a Gram-positive bacterium. To test our proposal for the heptose pathway, the enzymes $gmhBD$ and $gmhC$ from $E.\ coli$ are currently under investigation. As there are no genetic tools presently available for the Gram-positive bacterium $A.\ thermoaerophilus$, knock-out experiments with the D,D-heptose 1,7-bisphosphate phosphatase gene $gmhC$ will be performed in $E.\ coli$ K-12 to assess its involvement in heptose biosynthesis.
REFERENCES

1 Sára, M. and Sleytr, U. B. (2000) J. Bacteriol. 182, 859-868

2 Messner, P. and Schäffer, C. (2000) in Glycomicrobiology (Doyle, R. J., ed) pp. 93-125, Kluwer Academic/Plenum Publishers, New York

3 Kosma, P., Wugeditsch, T., Christian, R., Zayni, S. and Messner, P. (1995) Glycobiology 5, 791-796, Erratum: Glycobiology (1996) 6, 5

4 Schäffer, C., Wugeditsch, T., Neuninger, C. and Messner, P. (1996) Microb. Drug Resist. 2, 17-23

5 Kneidinger, B., Graninger, M., Adam, G., Puchberger, M., Kosma, P., Zayni, S. and Messner, P. (2001) J. Biol. Chem. 276, 5577-5583

6 Raetz, C. R. H. (1990) Annu. Rev. Biochem. 59, 129-170

7 Radziejewska-Lebrecht, J. and Mayer, H. (1989) Eur. J. Biochem. 183, 573-581

8 Melaugh, W., Phillips, N. J., Campagnari, A. A., Karalus, R. and Gibson, B. W. (1992) J. Biol. Chem. 267, 13434-13439

9 Süsskind, M., Brade, L., Brade, H. and Holst, O. (1998) J. Biol. Chem. 273, 7006-7017

10 Knirel, Y. A., Kocharova, N. A., Hynes, S. O., Widmalm, G., Andersen, L. P., Jansson, P. E. and Moran, A. P. (1999) Eur. J. Biochem. 266, 123-131

11 Radziejewska-Lebrecht, J., Shashkov, A. S., Stroobant, V., Wartenberg, K., Warth, C. and Mayer, H. (1994) Eur. J. Biochem. 221, 343-351

12 Wilkinson, S. G. (1983) Rev. Infect. Dis. 5 Suppl. 5, S941-S949

13 Heinrichs, D. E., Valvano, M. A. and Whitfield, C. (1999) in Endotoxin in Health and Disease. (Brade, H., Morrison, D. C., Vogel, S. and Opal, S. eds) pp. 305-330, Marcel Dekker, New York
14 Curtiss, R., Charamella, J., Stallions, D. R. and Mays, J. A. (1968) *Bacteriol. Rev.* **32**, 320-348
15 Havekes, L. M., Lugtenberg, B. J. J. and Hoekstra, W. P. M. (1976) *Mol. Gen. Genet.* **146**, 43-50
16 Tamaki, S., Sato, T. and Matsuhashi, M. (1971) *J. Bacteriol.* **105**, 968-975
17 Zwahlen, A., Rubin, L. G., Connelly, C. J., Inzana, T. J. and Moxon, E. R. (1985) *J. Infect. Dis.* **152**, 485-492
18 Sirisena, D. M., MacLachlan, P. R., Liu, S.-L., Hessel, A. and Sanderson, K. E. (1994) *J. Bacteriol.* **176**, 2379-2385
19 Eidels, L. and Osborn, M. J. (1971) *Proc. Natl. Acad. Sci. USA* **69**, 1673-1677
20 Ginsburg, V., O’Brien, P. J. and Hall, C. L. (1962) *J. Biol. Chem.* **237**, 497-499
21 Kontrohr, T. and Kocsis, B. (1981) *J. Biol. Chem.* **256**, 7715-7718
22 Kocsis, B. and Kontrohr, T. (1984) *J. Biol. Chem.* **259**, 11858-11860
23 Zamyatina, A., Gronow, S., Oertelt, C., Puchberger, M., Brade, H. and Kosma, P. (2000) *Angew. Chem. Int. Ed.* **39**, 4150-4153
24 Eidels, L. and Osborn, M. J. (1974) *J. Biol. Chem.* **249**, 5642-5648
25 Brooke, J. S. and Valvano, M. A. (1996) *J. Biol. Chem.* **271**, 3608-3614
26 Valvano, M. A., Marolda, C. L., Bittner, M., Glaskin-Clay, M., Simon, T. L. and Klena, J. D. (2000) *J. Bacteriol.* **182**, 488-97
27 Coleman, W. G., Jr. (1983) *J. Biol. Chem.* **258**, 1985-1990
28 Deacon, A. M., Ni, Y. S., Coleman, W. G., Jr. and Ealick, S. E. (2000) *Structure Fold. Des.* **8**, 453-462
29 Ding, L., Seto, B. L., Ahmed, S. A. and Coleman, W. G., Jr. (1994) *J. Biol. Chem.* **269**, 24384-24390
30 Palmieri, M. J., Berry, G. T., Player, D. A., Rogers, S. and Segal, S. (1991) *Anal. Biochem.* **184**, 388-393

31 Laemmli, U. K. (1970) *Nature* **227**, 680-685

32 Bradford, M. M. (1976) *Anal. Biochem.* **72**, 248-254

33 Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. J. (1997) *Nucleic Acids Res.* **25**, 3389-3402

34 Corpet, F. (1988) *Nucleic Acids Res.* **16**, 10881-10890

35 Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*. CSH Laboratory Press, Cold Spring Harbor, NY

36 Bax, A., Griffey, R. H. and Hawkins, B. L. (1983) *J. Magn. Res.* **55**, 301-315

37 Summers, M. F., Marzilli, G. L. and Bax, A. (1986) *J. Am. Chem. Soc.* **108**, 4285-4294

38 Reeves, P. R. (1994) *Compr. Biochem.* **27**, 271-314

39 Marumo, K., Lindquist, L., Verma, N., Weintraub, A., Reeves, P. R. and Lindberg, A. A. (1991) *Eur. J. Biochem.* **204**, 539-545

40 Giraud, M.-F. and Naismith, J. H. (2000) *Curr. Opin. Struct. Biol.* **10**, 687-696

41 Kneidinger, B., Graninger, M. and Messner, P. (2001) *BioTechniques* **30**, 248-249

42 Clark, M. A., Baumann, L. and Baumann, P. (1998) *Curr. Microbiol.* **37**, 356-358

43 Carlonagno, M. S., Chiariotti, L., Alifano, P., Nappo, A. G. and Bruni, C. B. (1988) *J. Mol. Biol.* **203**, 585-606

44 Grzeszczyk, B., Holst, O., Müller-Loennies, S. and Zamojski, A. (1998) *Carbohydr. Res.* **307**, 55-67

45 Paulsen, H., Pries, M. and Lorentzen, J. P. (1994) *Liebigs Ann. Chem.* **1994**, 389-397

46 Naess, V. and Hofstad, T. (1984) *J. Gen. Microbiol.* **130**, 2783-2789

47 Szymanski, C. M., Yao, R., Ewing, C. P., Trust, T. J. and Guerry, P. (1999) *Mol.
Reeves, P. R., Hobbs, M., Valvano, M. A., Skurnik, M., Whitfield, C., Coplin, D., Kido, N., Klena, J., Maskell, D., Raetz, C. R. H. and Rick, P. D. (1996) *Trends Microbiol.* **4**, 495-503

Heyndrickx, M., Lebbe, L., Vancanneyt, M., Kersters, K., De Vos, P., Logan, N. A., Forsyth, G., Nazli, S., Ali, N. and Berkeley, R. C. W. (1997) *Int. J. Syst. Bacteriol.* **47**, 808-817
FOOTNOTES

*This work was supported by the Austrian Science Fund (projects P12966-MOB and P14209-MOB to P. M.).

The nucleotide and protein sequences reported in this paper have been submitted to the GenBank/EMBL Data Bank with accession number AF324836.

1 The abbreviations used are: S-layer, surface-layer; D,D-heptose, D-glycero-D-manno-heptose; LPS, lipopolysaccharide; L,D-heptose, L-glycero-D-manno-heptose; ATP, adenosine triphosphate; GTP, guanosine triphosphate; HPAEC, high-performance anion exchange chromatography; ORF, open reading frame; PCR, polymerase chain reaction; NMR, nuclear magnetic resonance spectroscopy.

¶ To whom correspondence should be addressed: Zentrum für Ultrastrukturforschung und Ludwig Boltzmann-Institut für Molekulare Nanotechnologie, Universität für Bodenkultur Wien, Gregor-Mendel-Str. 33, A-1180 Wien, Austria. Phone: +43-1-47654 (ext. 2202); Fax: +43-1-4789112; email: pmessner@edv1.boku.ac.at

ACKNOWLEDGMENTS

We thank Dr. C. Whitfield, University of Guelph, Guelph, Ont., Canada, for stimulating discussions and critical reading of the manuscript, Dr. A. Zamyatina for providing synthetic compounds and NMR spectra, and S. Zayni and A. Scheberl for the excellent technical assistance.
FIGURE LEGENDS

Fig. 1. Physical map of the GDP-D-α-D-heptose operon of *Aeurinibacillus thermoaeophilus* DSM 10155. The arrows indicate the genes and the hypothetical ORFs on the sequence. *gmhA, gmhB, gmhC* and *gmhD* encode sedoheptulose 7-phosphate isomerase, D,D-heptose 7-phosphate kinase, D,D-heptose 1,7-bisphosphate phosphatase, and D,D-heptose 1-phosphate guanosyltransferase, respectively. *rmlA* encodes glucose 1-phosphate thymidyltransferase, part of the pathway for dTDP-L-rhamnose biosynthesis.

Fig. 2. Multiple sequence alignment of putative phosphatase sequences. A. *t.*, *Aeurinibacillus thermoaeophilus* DSM 10155; C. *j.*, *Campylobacter jejuni* NCTC 11168; He. *p.*, *Helicobacter pylori* 26695; Ha. *i.*, *Haemophilus influenzae* Rd; E. *c.*, *Escherichia coli* K-12 and V. *c.*, *Vibrio cholerae* N16961.

Fig. 3. Schematic representation of the biosynthesis pathway of GDP-D-α-D-heptose.

**A**, sedoheptulose 7-phosphate isomerase GmhA; **B**, D,D-heptose 7-phosphate kinase GmhB, **C**, D,D-heptose 1,7-bisphosphate phosphatase GmhC and **D**, D,D-heptose 1-phosphate guanosyltransferase GmhD.

Fig. 4. SDS-polyacrylamide gel electrophoresis analysis of the purified enzymes involved in the biosynthesis of GDP-D-α-D-heptose. **A**, enzymes catalyzing the formation of D-α-D-heptose 1-phosphate from sedoheptulose 7-phosphate; **Lane 1**, molecular mass standards, myosin (200 kDa), β-galactosidase (116.3 kDa), phosphorylase b (97.4 kDa), bovine serum albumin (66.3 kDa), glutamic dehydrogenase (55.4 kDa), lactate dehydrogenase (36.5 kDa),
carbonic anhydrase (31 kDa), trypsin inhibitor (21.5 kDa), and lysozyme (14.4 kDa); Lane 2, isomerase GmhA; Lane 3, kinase GmhB and Lane 4, phosphatase GmhC. B, enzyme activating β-D-heptose 1-phosphate to form GDP-α-D-heptose; Lane 1, see above; Lane 5, pyrophosphorylase GmhD.

Fig. 5. HPAEC analysis of enzymatic synthesis of β-D-heptose 1-phosphate.
A, β-D-sedoheptulose 7-phosphate, converted with the isomerase GmhA.
B, β-D-sedoheptulose 7-phosphate, converted with GmhA, the kinase GmhB and ATP.
C, β-D-sedoheptulose 7-phosphate, converted with GmhA, GmhB, the phosphatase GmhC and ATP.
* marks retention time of synthetic β-D-heptose 1-phosphate. D, Standard: α-D-glucose 1-phosphate, D-glucose 6-phosphate, D-sedoheptulose 7-phosphate, α-D-glucose 1,6-bisphosphate. Peaks with retention times in the range of 30 and 35.5 min originate from adenosine phosphates and dithiothreitol, respectively.
α-D-glucose 1,6-bisphosphate yields a very low signal, and is therefore not visible in the chromatogram.

Fig. 6. HPAEC analysis of the activation of D,D-heptose. A, D-α-D-heptose 1-phosphate, converted with the pyrophosphorylase GmhD and ATP. B, D-α-D-heptose 1-phosphate, converted with GmhD and GTP. C, Standard: NAD+, AMP, GMP, ADP, GDP-D-mannose, GDP, ATP and GTP.

Fig. 7. 1H-NMR-analysis of GDP-D-α-D-heptose. A, expansion plot of ADP-β-D-heptose. B, expansion plot of ADP-α-D-heptose. C, expansion plot of
GDP-\(\mathbf{D}\)-\(\alpha\)-\(\mathbf{D}\)\text{-}heptose. \(\mathbf{D}\), full spectrum of GDP-\(\mathbf{D}\)-\(\alpha\)-\(\mathbf{D}\)\text{-}heptose.

Fig. 8. **Comparison of heptose biosynthetic pathways.** Left column: biosynthesis of GDP-\(\mathbf{D}\)-\(\alpha\)-\(\mathbf{D}\)\text{-}heptose in *A. thermoerophilus* DSM 10155. Right column: putative pathway for biosynthesis of ADP-\(\mathbf{L}\)-\(\beta\)-\(\mathbf{D}\)\text{-}heptose in Gram-negative bacteria. For details see Discussion.
### TABLE I

Bacterial strains, plasmids and oligonucleotides used in this study

| Genotype/description | Source |
|----------------------|--------|
| **Strains**          |        |
| **E. coli**           |        |
| TG1                  | supE·thi-l Δ(lac-proAB) Δ(mcrB-hsdSM)5 (rK·mK·) [F traD36 proAB lacIqZΔM15] |
|                      | Stratagene, La Jolla, CA |
| DH5α                 | K-12 F·Φ80d lacZΔM15 endA1 recA1 hsdR17 (rK·mK·) supE44 thi-l gyrA96 relA1 Δ(lacZYA-argF) U169 |
|                      | LifeTechnologies, Wien, Austria |
| LMG194               | F· ΔlacX74 gal E thi rpsL ΔphoA (Pvu II) Δara714 leu::Tn10 |
|                      | Invitrogen, Groningen, The Netherlands |
| BL21-CodonPlus-RIL   | E. coli B F·ompT hsdS(rB·mB·) dcm+ Tetf gal λ(DE3) endA Hte [argU ileY leuW Camf] |
|                      | Stratagene, La Jolla, CA |
| BL21-SI              | F·ompB hsdSB(rB·mB·) gal dcm endA1 proU::T7 RNAP::malQ-lacZ Tetf |
|                      | LifeTechnologies, Wien, Austria |
**A. thermoaeophilus**

DSM 10155  
**wild type strain**  
Ref. 49  
DSM, Braunschweig, Germany

**Plasmids**

| Plasmid     | Description                                      | Supplier                  | Location                      |
|-------------|--------------------------------------------------|----------------------------|-------------------------------|
| pBCKS       | cloning vector                                   | Stratagene, La Jolla, CA  |
| pDONR201    | cloning vector                                   | LifeTechnologies, Wien, Austria |
| pDEST15     | expression vector for GST-fusion                 | LifeTechnologies, Wien, Austria |
| pDEST17     | expression vector for histidine-fusion           | LifeTechnologies, Wien, Austria |
| pBADB       | expression vector for histidine-fusion           | Invitrogen, Groningen, The Netherlands |
| pRML        | 4.452 bp heptose fragment in pBCKS (BglII in BamHI) | this study               |
| gGMHA1      | gmhA in pDONR201                                 | this study               |
| gGMHB1      | gmhB in pDONR201                                 | this study               |
| gGMHD1      | gmhD in pDONR201                                 | this study               |
| pGMHA3      | gmhA in pDEST17                                 | this study               |
| pGMHB3      | gmhB in pDEST17                                 | this study               |
| pgmhC1      | gmhC in pBADB (BglII and EcoRI)                  | this study               |
| pGMHD2      | gmhD in pDEST15                                 | this study               |

**Oligonucleotide primers**
GMHAA1(forward)a  5′-attB1-GCGAGAATCTCTACTTCCAAGGAATGAACGATTATATTAAAAGTC-3′

GMHAA2(reverse)a  5′-attB2-CCCTACCAGTATAATTGCTTCC-3′

GMHBA1(forward)a  5′-attB1-GCGAGAATCTCTACTTCCAAGGAATGATTTTTCGTTCAAAAGC-3′

GMHBA2(reverse)a  5′-attB2-TCGGCTAGAATTGCTTCC-3′

GMHC1F(forward)b  5′-GGAGATCTACATTACATATGAAAAACAAGG-3′

GMHC2R(reverse)b  5′-CCGAATTCTATCATTTTCATTGCCAGC-3′

GMHDA1(forward)a  5′-attB1-GCGAGAATCTCTACTTCCAAGGAATGGAAGCAATTATACTGG-3′

GMHDA2(reverse)a  5′-attB2-CATCTCTATCCAAAAACAATGC-3′

a attB1 and attB2 are the minimal 25 bp sequences required for homologous recombination in the GATEWAY system, LifeTechnologies, Wien, Austria.

b The underlined sequences represent the BglII and EcoRI restriction sites, respectively.
| Protein | Accession number | Organism | a.a.b |
|---------|------------------|----------|-------|
| GmhA    | this study       | A. thermoaerophilus DSM 10155 | 198   |
|         | CAB73848.1       | C. jejuni NCTC 11168 | 201   |
|         | CAB73403.1       | C. jejuni NCTC 11168 | 186   |
| P5I001  |                   | E. coli K-12 | 192   |
|         | P45093           | Ha. influenzae Rd | 194   |
|         | O25528           | He. pylori 26695 | 192   |
|         | AAF95374         | V. cholerae N16961 | 191   |
|         | AAG07813.1       | P. aeruginosa PAO1 |       |
|         | AAF42407.1       | N. meningitidis MC58 | 197   |
|         | O53635           | M. tuberculosis H37Rv | 196   |
| GmhBα   | this study       | A. thermoaerophilus DSM 10155 | 341   |
|         | CAB73849.1       | C. jejuni NCTC 11168 | 339   |
|         | F70840           | M. tuberculosis H37Rv | 386   |
| GmhBβ?  | AAF41238         | N. meningitidis MC58 | 323   |
| GmhC    | this study       | A. thermoaerophilus DSM 10155 | 179   |
|         | CAB73406.1       | C. jejuni NCTC 11168 | 186   |
|         | P31546           | E. coli K-12 | 191   |
|         | P46452           | Ha. influenzae Rd | 184   |
| Accession Number | Organism/Strain          | Length |
|------------------|-------------------------|--------|
| D64627           | *He. pylori* 26695      | 173    |
| AAF94070.1       | *V. cholerae* N16961    | 186    |
| AAG03396.1       | *P. aeruginosa* PAO1    | 178    |
| AAF42354.1       | *N. meningitidis* MC58  | 187    |
| E70840           | *M. tuberculosis* H37Rv | 190    |
| GmhDGTP          | this study<sup>c,d</sup> | A. thermoaerophilus DSM 10155 | 230 |
| CAB73847.1       | *C. jejuni* NCTC 11168 | 221    |
| A70978           | *M. tuberculosis* H37Rv | 359    |
| GmhDATP          | AAF42395                | *N. meningitidis* MC58  | 168  |
| GmhBDATP         | CAB73404.1              | *C. jejuni* NCTC 11168 | 461  |
| P76658           | *E. coli* K-12          | 477    |
| O05074           | *Ha. influenzae* Rd     | 476    |
| O25529           | *He. pylori* 26695      | 461    |
| AAG08381.1       | *P. aeruginosa* PAO1    | 474    |
| GmhE             | CAB73405.1              | *C. jejuni* NCTC 11168 | 317  |
| P17963<sup>c</sup> | *E. coli* K-12        | 310    |
| P45048           | *Ha. influenzae* Rd     | 308    |
| C64627           | *He. pylori* 26695      | 330    |
| AAG06725.1       | *P. aeruginosa* PAO1    | 332    |
| AAF41240.1       | *N. meningitidis* MC58  | 334    |

<sup>a</sup> Accession number from the National Center for Biotechnology Information database.

<sup>b</sup> Amino acids.
c Functional studies have been performed.

d The nucleotide sequence of the 6.652 bp DNA fragment has been submitted to the GenBank/EMBL Data Bank with accession number AF324836.
TABLE III

\textit{NMR data of GDP-\textit{D}-glycero-\textit{\alpha}-\textit{D}-manno-heptose}

| Atom H/C/P (ppm) | 1  | 2  | 3  | 4  | 5  | 6  | 7a | 7b |
|------------------|----|----|----|----|----|----|----|----|
| Unit             |    |    |    |    |    |    |    |    |
| \textit{\alpha}-Hep-1→\textit{1H} | 5.47 | 4.02 | 3.90 | 3.76 | 3.89 | 3.99 | 3.76 | 3.71 |
| \textit{J (Hz)}  | 1.6 | 3.2 | 9.6 | 9.6 | 3.3 | 3.8 | 12.2 | 6.  |
| \textit{\beta}-Rib-1→\textit{1H} | 5.94 | 4.77 | 4.50 | 4.35 | 4.19 (2x) |    |    |    |
| \textit{J (Hz)}  | 6.3 | 5.2 | 3.7 | n.d.a |    |    |    |    |
| \textit{13C}     | 96.9 | 70.6 | 70.6 | n.d.a | n.d.a | 73.1 | 62.5 |    |
| \textit{\alpha}-Hep-1-P-\textit{31P} | -13.2 |    |    |    |    |    |    |    |
| \textit{J (Hz)}  | 7.8, 20.5 |    |    |    |    |    |    |    |
| \textit{-P-3-Rib} \textit{31P} | -10.9 |    |    |    |    |    |    |    |
| \textit{Atom H/C/P (ppm)} | 8  |    |    |    |    |    |    |    |
| \textit{Guanine} | \textit{1H} | 8.09 |    |    |    |    |    |    |
| \textit{13C} | 138.6 |    |    |    |    |    |    |    |
a Not determined.
Fig. 1 of Kneidinger et al.
Fig. 2 of Kneidinger et al.
Fig. 3 of Kneidinger et al.
Fig. 5 of Kneidinger et al.
Fig. 6 of Kneidinger et al.
Fig. 8 of Kneidinger et al.
