Novel pathways for biosynthesis of nucleotide-activated glycerol-manno-heptose precursors of bacterial glycoproteins and cell surface polysaccharides

Miguel A. Valvano,¹ Paul Messner² and Paul Kosma³

Author for correspondence: Miguel A. Valvano. Tel.: +1 519 661 3996. Fax: +1 519 661 3499. e-mail: mvalvano@uwo.ca

¹Department of Microbiology and Immunology and Medicine, University of Western Ontario, London, Ontario, N6A 5C1, Canada
²Zentrum für Ultrastrukturforschung und Ludwig Boltzmann-Institut für Molekulare Nanotechnologie, Universität für Bodenkultur Wien, A-1180 Wien, Austria
³Institut für Chemie, Universität für Bodenkultur Wien, A-1190 Wien, Austria

Keywords: lipopolysaccharide, sugar kinase, sugar phosphate phosphatase, surface layer (S-layer), capsules

Overview

glycerol-manno-Heptose is widely present in the lipopolysaccharide (LPS) of most Gram-negative bacteria. This carbohydrate and its derivatives are sometimes also found in capsules and O antigens as well as in the glycocalyx moieties of bacterial cell surface (S-layer) glycoproteins. Gram-negative bacteria with a heptoseless LPS are highly sensitive to hydrophobic substances and they also display reduced virulence in animal infection models. Thus heptose biosynthetic genes and their products are potentially attractive targets for developing novel antimicrobial compounds. Recent work in our laboratories has elucidated the complete pathways for biosynthesis of the nucleotide-activated heptosyl precursors (Fig. 1), which are utilized in the assembly of the LPS inner-core oligosaccharide and a bacterial S-layer glycoprotein (Kneidinger et al., 2001, 2002). We review the state of the art of these biosynthesis pathways, particularly regarding the enzyme functions involved in the various steps of the biosynthesis of ADP-α-glycerol-β-manno-heptose and GDP-β-glycerol-β-manno-heptose. We also provide a comparative genetic analysis of the biosynthetic genes in sequenced bacterial genomes.

Importance of heptose residues in cell surface polysaccharides and bacterial glycoproteins

LPS is an amphipathic glycolipid located in the outer leaflet of the outer membrane that consists of lipid A and a core oligosaccharide domain. Some micro-organisms also express a hydrophilic surface-exposed O-specific polysaccharide that is found attached to the reducing end of the lipid A-core (Whitfield & Valvano, 1993). Lipid A is made of a 1,6-β-linked glucosamine dimer, which becomes phosphorylated and acylated with a variable number of fatty and hydroxyfatty acid chains (Raetz, 1996). The core oligosaccharide can be subdivided into inner- and outer-core domains. The outer core usually consists of hexoses and hexosamines while the inner core is in most cases composed of two residues of 3-deoxy-α-manno-octulosonic acid, and depending on the particular species, two or three residues of 1-glycerol-β-manno-heptose (Heinrichs et al., 1999). Heptose residues can also be found in the outer core of some LPS molecules, as in the case of Haemophilus ducreyi (Melaugh et al., 1992) and Escherichia coli K-12 (Holst et al., 1991). The structure of the inner core is relatively conserved among many enteric and non-enteric bacteria (Heinrichs et al., 1999), with a few exceptions such as Acinetobacter, Legionella and Rhizobium. The inner core of these micro-organisms lacks heptose and contains unusual sugars (Kadrmas et al., 1996; Kawahara et al., 1987; Knirel et al., 1996), but very little information is currently available on its biosynthesis and genetics. glycerol-manno-Heptose residues have also been identified in the O antigen subunit of the Burkholderia pseudomallei LPS, in the capsule of Burkholderia mallei (DeShazer et al., 2001) and the O antigen from Yokenella regensburgei strains (Jachymek et al., 1999). It should be noted that 1-glycerol-β-manno-heptose units have been found in different anomeric configuration. Thus short heptan chains containing α-configured 1-glycerol-α-manno-heptose units have been detected in the core of Helicobacter pylori and Klebsiella pneumoniae strains (Aspinall et al., 1997; Süskind et al., 1998), whereas...
d-glycero-β-D-manno-heptose residues occur in the O antigen from a *Plesiomonas shigelloides* strain (Czaja et al., 2000).

LPS plays an important role in maintaining the structural integrity of the bacterial outer membrane by interacting with outer-membrane proteins as well as divalent cations (Ferguson et al., 2000; Hancock et al., 1994). Phosphate groups covalently attached to heptose residues in the inner core participate in these ionic interactions, which provide a barrier preventing the passage of hydrophobic substances such as detergents, dyes and antibiotics across the outer membrane (Nikaido, 1994; Nikaido & Vaara, 1985). *E. coli* mutants lacking heptose in the LPS display hypersensitivity to novobiocin, detergents and bile salts (Tamaki et al., 1971). They also have defects in F plasmid conjugation and generalized transduction by the bacteriophage P1 (Curtiss et al., 1968; Havekes et al., 1976; Sherburne & Taylor, 1997). These phenotypes are associated with a reduced amount of outer-membrane proteins, some of which serve as surface receptors for conjugation and bacteriophage attachment (Bayer et al., 1975; Koplow & Goldfine, 1974; Sherburne & Taylor, 1997; Vakharia & Misra, 1996; van Alphen et al., 1976; Verkleij et al., 1976) or as channel components of efflux systems (Fralick & Burns-Kelihier, 1994; Koronakis et al., 1997). The reduced stability of the outer membrane in these mutants is associated at least in part with the absence of phosphate groups, since mutations in genes encoding LPS core kinases also show pleiotropic phenotypes similar to those found in heptose-deficient mutants (Yethon & Whitfield, 2001; Yethon et al., 2000). Usually LPS heptose-deficient mutants in many bacterial species can survive in the laboratory. However, in some microorganisms like *Pseudomonas aeruginosa*, heptoseless mutants have not been isolated and it has been proposed that these residues (or the phosphates covalently attached to them) are essential for bacterial survival in *vitro* (Walsh et al., 2000). Heptose-deficient mutants are generally serum-sensitive and display reduced virulence in experimental infection models (Helander et al., 1988; Zwahlen et al., 1985).

Many bacterial species are characterized by the production of regular protein surface arrays known as S-layers (Messner & Sleytr, 1992; Sára & Sleytr, 2000; Sleytr & Messner, 2000). Recent work has revealed that some S-layer proteins are glycosylated (Messner & Sleytr, 1991; Messner & Schäffer, 2002; Schäffer et al., 2001; Sumer & Wieland, 1995). *glycero-manno*-Heptose is one of the sugar components of the disaccharide repeating unit of the well-characterized S-layer glycan in *Aneurinibacillus thermoautotrophicus* DSM 10155 (Kosma et al., 1995; Wugeditsch et al., 1999). Protein glycosylation in *prokaryotes* is not only limited to S-layer proteins. Pilins, non-piliated adhesins, flagellar filament subunits and secreted exoenzymes in a variety of archaeal and eubacterial micro-organisms are modified by the addition of carbohydrate residues (Messner & Sleytr, 2002; Schäffer et al., 2001). Glycosylation with *glycero-manno*-heptose residues has been demonstrated in the AIDA-I autotransporter adhesin of diarrhoeagenic *E. coli* (Benz & Schmidt, 2001). This modification not only improved the stability of the adhesin but also was essential for the adherent function of the protein (Benz & Schmidt, 2001). Thus glycosylation of bacterial surface proteins may serve a number of functions, including adherence, evasion of host immune responses, and an enhanced resistance to proteolytic attack.
A novel kinase/phosphatase cascade is involved in the biosynthesis of nucleotide-activated glycero-manno-heptose

Eidels & Osborn (1971, 1974) first established that the synthesis of L-glycero-d-manno-heptose utilizes sedoheptulose 7-phosphate, an intermediate in the nonoxidative portion of the pentose phosphate pathway. These investigators proposed a 4-step biosynthesis pathway involving: (i) conversion of sedoheptulose 7-phosphate into d-glycero-d-manno-heptose 7-phosphate by a phosphoheptose isomerase; (ii) formation of d-glycero-d-manno-heptose 1-phosphate by a mutase reaction; (iii) transfer of a nucleotide via a phosphodiester linkage; and (iv) epimerization of the d-glycero-d-manno-heptose residue of the sugar nucleotide to L-glycero-d-manno-heptose. Later on, other investigators identified ADP as the nucleotide sugar residue attached to glycero-manno-heptose in Shigella sonnei and Salmonella enterica (Kocsis & Kontrohr, 1984).

More recently, it was demonstrated that the E. coli hldE gene (formerly rfaE, see below) encodes a bifunctional protein with two distinct functional domains that may be involved in the biosynthesis of d,β-heptose 1-phosphate as well as the activating step (Valvano et al., 2000). One of the HldE domains shares structural features with members of the ribokinase family, while the other domain has conserved features present in nucleotidyltransferases (Valvano et al., 2000). The demonstration of a protein domain corresponding to a putative sugar kinase suggested a different process for nucleotide-activated heptose biosynthesis than the one proposed by Eidels and Osborn and, at the same time, predicted the existence of an additional phosphatase step (Valvano, 1999). A kinase/phosphatase cascade in place of a mutase step was biochemically confirmed following the complete elucidation of the biosynthesis pathway for GDP-d-glycero-α,β-manno-heptose in the Gram-positive bacterium A. thermoaerophilus (Kneidinger et al., 2001) and for ADP-d-glycero-β-d-manno-heptose in E. coli (Kneidinger et al., 2002). These novel steps revealed a major difference from the majority of the classical pathways leading to the formation of nucleotide-activated sugars, which usually involve a mutase step catalysing the intramolecular transfer of a phosphate group from the distal carbon to the C-1 position. This phosphate is subsequently modified to form a phosphodiester linkage with the nucleotide by a nucleotidyltransferase, which results in the synthesis of a nucleotide diphosphate-sugar precursor.

Furthermore, the kinase enzymes determine the anomic specificity of the diphosphate reaction products, and, as expected, the nucleotidyltransferases are also specific for the anomic form of the d,β-heptose 1-phosphate that results from the phosphatase step (Kneidinger et al., 2002). The GDP-activated heptose serves as a precursor for the glycan moiety of A. thermoaerophilus S-layer glycoprotein, which consists of α-L-thamnose and d-glycero-β-d-manno-heptose residues (M. Graninger, B. Kneidinger, K. Bruno & P. Messner, unpublished results; Kneidinger et al., 2001; Kosma et al., 1995). In contrast, the ADP-activated heptose serves as a substrate for the glycosyltransferases involved in the assembly of the LPS inner-core oligosaccharide (Gronow et al., 2001; Zamyatina et al., 2000). Two different biosynthetic pathways, which we have named the d,α-d-heptose and the L,β-d-heptose pathways (Fig. 1), can be distinguished on the basis of the

**Table 1. New gene nomenclature for nucleotide-activated biosynthesis pathway genes**

<table>
<thead>
<tr>
<th>New gene designation</th>
<th>Encoded function</th>
<th>Previous designation(s) (source/reference)</th>
</tr>
</thead>
<tbody>
<tr>
<td>gmbA</td>
<td>Sedoheptulose-7-phosphate isomerase</td>
<td>lpcA (Brooke &amp; Valvano, 1996b)</td>
</tr>
<tr>
<td>gmbB</td>
<td>D-α,β-β-Heptose-1,7-bisphosphate phosphatase</td>
<td>tfrA (Haque et al., 1976)</td>
</tr>
<tr>
<td>hddA</td>
<td>D-α,β-β-Heptose-7-phosphate kinase</td>
<td>yaeD</td>
</tr>
<tr>
<td>hddC</td>
<td>D-α,β-β-Heptose-1-phosphate guanylyltransferase</td>
<td>gmbX (Shih et al., 2001)</td>
</tr>
<tr>
<td>hldE</td>
<td>Bifunctional D-β-β-Heptose-7-phosphate kinase/D-β-β-Heptose-1-phosphate adenylyltransferase</td>
<td>waaE (Burtinck &amp; Woods, 1999; Regue et al., 2001)</td>
</tr>
<tr>
<td>hldA</td>
<td>D-β-β-Heptose-7-phosphate kinase</td>
<td>rfaE (Shih et al., 2001)</td>
</tr>
<tr>
<td>hldC</td>
<td>D-β-β-Heptose-1-phosphate adenylyltransferase</td>
<td>rfaE</td>
</tr>
<tr>
<td>hldD</td>
<td>ADP-d-β-β-Heptose epimerase</td>
<td>rfaD</td>
</tr>
<tr>
<td></td>
<td></td>
<td>waaD</td>
</tr>
<tr>
<td></td>
<td></td>
<td>gmbD (Sozhamannan et al., 1999)</td>
</tr>
</tbody>
</table>

**Biosynthesis of nucleotide-activated heptoses**
specificity of the kinase and the nucleotidyltransferase steps. The isomerase and phosphatase enzymes are common to both pathways and their functions are independent of the anomeric conformation of the sugar phosphate substrates. Thus these enzymes are conserved in a wide range of Gram-positive and Gram-negative bacteria. We have proposed a rational gene nomenclature to account for the differences and similarities between the two pathways (Kneidinger et al., 2002) (Table 1). In the sections that follow, we describe the characteristics of the enzymes involved in the biosynthesis pathways for nucleotide-activated heptose.

**Common steps in the biosynthesis pathways of nucleotide-activated heptose**

**GmhA phosphohexose isomerase**

Using radiolabelled sedoheptulose 7-phosphate and determining the specific labelling of LPS residues in a transketolase mutant of *Salmonella*, Eidels & Osborn (1971) established that sedoheptulose 7-phosphate was a precursor of L-glycero-D-manno-heptose. They later identified and partially purified a phosphohexose isomerase involved in the formation of D-glycero-D-manno-heptose (Eidels & Osborn, 1974; Eidels et al., 1974). Brooke & Valvano (1996b) characterized a locus in *E. coli* that was previously associated with the absence of LPS. The mutations mapped in the region of the *E. coli* chromosome between *ara* and *lac* next to the *proAB* genes, and identified the phosphohexose isomerase gene, currently designated *gmhA* (Table 1; Brooke & Valvano, 1996a; Reeves et al., 1996). Cell-free extracts of the *E. coli* *gmhA* mutant containing the recombinant *gmhA* gene catalysed the isomerization of sedoheptulose 7-phosphate into an aldose phosphate presumed to be D-glycero-D-manno-heptose 7-phosphate (Brooke & Valvano, 1996b). These observations were confirmed with purified GmhA protein (Brooke, 1996). The mass of the denatured GmhA polypeptide is 29 kDa as determined by SDS-PAGE (Brooke & Valvano, 1996b). It is likely that the native protein is a dimer since purified GmhA fractionates in the 60–80 kDa region of a Superose gel column (J. S. Brooke & M. A. Valvano, unpublished results).

Inspection of sequenced bacterial genomes reveals that GmhA is highly conserved in many Gram-negative bacteria as well as in some Gram-positive microorganisms, such as *A. thermoaerophilus* and *Clostridium acetobutylicum*, and also in *Mycobacterium tuberculosis*. The homologue from *Haemophilus influenzae* has been cloned and shown to complement the *E. coli* *gmhA* (formerly *lpca*) mutation (Brooke & Valvano, 1996a). GmhA has amino acid sequence conservation with the C-terminal region of *L*-glutamine:D-fructose-6-phosphate amidotransferases, which are also ketose/aldose isomerases (Golinelli-Pimpaneau et al., 1989). All of these proteins share a domain termed the sugar isomerase (SIS) domain (Bateman, 1999). The C-terminal isomerase domain of the *E. coli* glucosamine-6-phosphate synthase (GlmS) has been crystallized in the presence of glucosamine 6-phosphate and shown to consist of two topologically identical subdomains of equal size. Each of these domains is characterized by an αβα motif that represents the nucleotide-binding motif of the enzyme. The nucleotide-binding motif of GmhA enzymes is likely to be an αβα motif that represents the nucleotide-binding motif of the enzyme.
GmhB from *A. thermoferophilus* and *E. coli* revealed that this enzyme can use either d-glycero-α-d-manno-heptose 1,7-bisphosphate or d-glycero-β-d-manno-heptose 1,7-bisphosphate as substrate (Kneidinger et al., 2001, 2002). Thus it is not surprising to identify homologues of the *A. thermoferophilus* protein in Gram-negative and Gram-positive bacterial genomes from species that synthesize glycerol-heptose-phosphate (Fig. 3). Sequence analysis shows that GmhB belongs to the haloacid dehalogenase (HAD) superfamily of hydrolases. This family comprises phosphatases, epoxide hydrolases and 1,2-haloacid dehalogenases (Koonin & Tatusov, 1994). All of these enzymes contain three highly conserved sequence motifs. Four representative proteins from the HAD family have been characterized structurally (Argradi et al., 2000; Hisano et al., 1996; Ridder et al., 1997; Wang et al., 2001). In general, these proteins reveal a core α/β domain resembling an NAD(P)-binding Rossmann fold (Rossmann et al., 1974) and a four-helix bundle domain. These domains are connected by loops, one of which contains motif I residues and is part of the enzyme’s active site. The first aspartic acid residue in motif I gets phosphorylated, resulting in a phosphoryl-aspartyl enzyme intermediate (Collet et al., 1998). Structural modelling of the *E. coli* GmhB protein also suggests a core α/β domain. The catalytic aspartic acid as well as the motif I are highly conserved among the GmhB homologues (Fig. 3), suggesting a similar structure and mechanism of action as that of the other members of the HAD family.

**D-α-D-Heptose pathway**

**HddA d-glycero-α-d-manno-heptose kinase**

HddA is responsible for the conversion of d-glycero-α-d-manno-heptose 1-phosphate into d-glycero-α-d-manno-heptose 1,7-bisphosphate (Fig. 1; Table 1). Analysis of the reaction products from assays containing sedoheptulose 7-phosphate, ATP and purified GmhA and HddA proteins clearly demonstrated that the

---

**Fig. 3. Features of GmhB phosphatases. Partial alignment showing conserved motifs within the GmhB homologues. Motif I is highly conserved in the HAD family (see text) and contains a catalytic aspartic acid residue (indicated by the asterisk). The proposed structure of the *E. coli* GmhB has been modelled and it is indicated above the alignment. The arrows denote regions with β strands and the cylinders denote regions with α helices. ECOLI, *Escherichia coli*; YEPS, *Yersinia pestis*; HAEIN, *Haemophilus influenzae*; PMULT, *Pasteurella multocida*; VCHOL, *Vibrio cholerae*; ATHER, *Aeruineribacter thermoacidophilus*; BMALL, *Burkholderia mallei*; HSK, *Helicobacter pylori*; CACET, *Clostridium acetobutylicum*; TACID, *Thermoplasma acidophilus*; CJEJU, *Campylobacter jejuni*; TVOLC, * Thermoplasma volcanii*; MLOTI, *Mesorhizobium loti*; SINC, *Streptomyces lincolnensis*.

---

**GmhB heptose-bisphosphate phosphatase**

The removal of a phosphate group at the C-7 position of glycero-manno-heptose 1,7-bisphosphate is the other step common to both the D-α-D-heptose and L-β-D-heptose biosynthesis pathways (Fig. 1; Table 1). The phosphatase activity of GmhB was deduced from the analysis of the reaction products separated by high-performance anionic exchange chromatography (Kneidinger et al., 2001, 2002). Experiments using purified

---

**Fig. 4. Conserved motifs of HddA kinase homologues and other members of the GHMP kinase family using as a reference the structural information from the homoserine kinase (KHSE) from *Methanococcus jannaschii* (Zhou et al., 2000). The KHSE conserved motifs are indicated below the alignment (see text for details). BPSEU, *Burkholderia pseudomallei*; BMALL, *Burkholderia mallei*; ATHER, *Aeruineribacter thermoacidophilus*; CACET, *Campylobacter jejuni*; KIME, *Pyrococcus horikoshii*; KIME.METJ, *Mesorhizobium loti*; KIME.SP, *Streptomyces lincolnensis*.

---

1983
enzyme is specific for the D-α-D-heptose anomer (Kneidinger et al., 2001). A search of the databases reveals that this protein belongs to the GHMP superfamily. GHMP kinases, which include galactokinase, homoserine kinase, mevalonate kinase and phosphomevalonate kinase (Fig. 4), are present in bacteria, archaea and eukaryotes, and they participate in several important metabolic pathways. Galactokinase is involved in the interconversion of galactose into glucose, homoserine kinase catalyses a step in the biosynthesis of threonine, while mevalonate and phosphomevalonate kinases are involved in the biosynthesis of isopentenyl pyrophosphate (Bork et al., 1993; Zhou et al., 2000).

The crystal structure of the homoserine kinase from *Methanococcus jannaschii* and its complex with ADP has recently been solved (Zhou et al., 2000). This analysis revealed that the nucleotide binding engages mostly the N-terminal domain, which contains an unusual left-handed βαβ unit that shares structural similarities with the elongation factor G domain IV (Zhou et al., 2000). This region contains a conserved motif, Pro-X-X-X-Gly-Leu-Gly-Ser-Ser-Ala-Ala, which in all proteins of the GHMP kinase superfamily is located in a predicted loop region and forms an ‘anion hole’ interacting with the α and β phosphates of bound ADP (Bork et al., 1993; Zhou et al., 2000). A comparative alignment of HddA homologues with members of the GHMP kinase superfamily shows that HddA proteins have the same conserved region, except that the last two amino acids are Ser and Ala (Fig. 4). Another conserved region, denoted by the residues (Ser/Thr)-Gly-Ser-Gly-Pro-Ser of the *M. jannaschii* homoserine kinase (Zhou et al., 2000), appears to be important in stabilizing the conformation of the phosphate-binding loop. This region is also partially conserved in HddA homologues (Fig. 4). Altogether, the amino acid se-

![Fig. 5. Examples of the genetic organization of D-α-D- and L-β-D-heptose pathways in different micro-organisms. The genes conserved in both pathways (gmhA and gmhB) are indicated in black. Vertical hatching indicates the genes of the D-α-D pathway, while diagonal hatching indicates the genes of the L-β-D pathway.](image-url)
sequence conservation strongly suggests that HddA has a similar overall structure as that of other members of the GHMP kinase superfamily and probably uses the same nucleotide-binding mechanism.

**HddC d-glycero-\(\alpha\)-d-manno-heptose-1-phosphate guanylyltransferase**

This enzyme belongs to a family of \(\alpha\)-d-mannose-1-phosphate guanylyltransferases. Its activity results in the formation of GDP-activated heptose, which is used as a substrate by specific glycosyltransferases for the assembly of the glycan moiety of bacterial glycoproteins as well as polysaccharide capsules. Homologues of the A. thermoacrophilus HddC are also found in *Cl. acetobutylicum*, *Campylobacter jejuni*, *B. pseudomallei*, *B. mallei* and *M. tuberculosis* (Fig. 5). The presence of HddC in these micro-organisms suggests that the nucleotide precursor for d,\(\alpha\)-heptose is GDP-d-glycero-\(\alpha\)-d-manno-heptose rather than ADP-d-glycero-\(\beta\)-d-manno-heptose. In *Campylobacter* strains, GDP-d-glycero-\(\alpha\)-d-manno-heptose may be a precursor for the 6-deoxy-heptose sugars found in capsular polysaccharides and in the glycan moieties of flagellin glycoproteins (Thibault et al., 2001). *B. pseudomallei* and *B. mallei* express polysaccharide capsules containing d,\(\beta\)-heptose, both of which are important for the virulence of these micro-organisms (DeShazer et al., 2001; Reckseidler et al., 2001).

**L-\(\beta\)-d-Heptose pathway**

**HldE bifunctional kinase/adenylyltransferase**

The *E. coli* bldE gene was recently characterized (Valvano et al., 2000) and shown to complement the heptose-deficient phenotype of the bldE (formerly rfaE) mutation in *Sal. enterica* (Valvano et al., 2000). rfaE is a designation for a locus in *Salmonella* not physically linked to the core oligosaccharide cluster. It was suggested, based on indirect evidence, that this locus encodes the ADP-heptose synthase (Sirisena et al., 1992). The designation of rfaE has now been replaced by *hldE* (Table 1).

The *E. coli* bldE gene encodes a polypeptide of approximately 55 kDa, and comparisons of the predicted amino acid sequence with other proteins in the database showed the presence of two clearly separate domains (Valvano et al., 2000). The N-terminal domain I, which spans amino acids 1–318, shares structural features with members of the ribokinase family. This is a large family of proteins whose function is the phosphorylation of sugars at positions 1 or 6 in the case of hexoses, and 1 or 5 in the case of pentoses (Bork et al., 1993; Sigrell et al., 1998). The C-terminal domain II, which spans amino acids 344–477, has all the conserved features of the cytidylyltransferase superfamily (Bork et al., 1995), including the *aut* gene product of *Ralstonia* (formerly *Alcaligenes*) *eutropha* (Freter & Bowien, 1994; Valvano et al., 2000). Two members of the cytidylyltransferase superfamily, pantoate-\(\beta\)-alanine ligase and acetate:SH-citrate lyase, are ATP transferases (Bork et al., 1995). Also, all the members in this family have structural conservation with the class I tRNA synthetases, which are also ATP transferases (Bork et al., 1995). By subcloning and expressing each domain separately, we showed that domain I was sufficient to complement the rfaE mutation in *Sal. enterica* (Valvano et al., 2000). We also showed that the *aut*::Tn5 mutation in *R. eutropha* determines a heptose-deficient LPS phenotype, which can be complemented only with the expressed domain II. The *aut* mutation has been associated with pleiotropic effects, including autotrophic growth, and changes in cell morphology and colony appearance (Freter & Bowien, 1994). A comparison of the LPS electrophoretic profile of this mutant with the wild-type strain revealed that the mutant has a fast-migrating lipid A–core band that co-migrates with that of the *E. coli* heptose-deficient lipid A–core. Also, the mutant is novobiocin-sensitive while the wild-type strain is resistant to this antibiotic. These experiments demonstrated that both domains of the *E. coli* gene are functionally different, suggesting that the HldE polypeptide is bifunctional. The bifunctional nature of this enzyme was recently demonstrated *in vitro* using purified components of the nucleotide-activated heptose pathway (Kneidinger et al., 2002). Based on genomic sequence comparisons, similar bifunctional proteins are predicted to be present in several Gram-negative microorganisms, including *Ha. influenzae*, *He. pylori*, *Vibrio cholerae* and *Ps. aeruginosa*. In contrast, individual genes encoding domains I and II independently are found in *R. eutropha*, *Neisseria meningitidis* and *Neisseria gonorrhoeae*. In these cases, we propose to use the nomenclature *bldA* and *bldC* to indicate the individual kinase- and adenylyltransferase-encoding genes, respectively (Table 1, see below).

**HldD ADP-L-glycero-\(\alpha\)-manno-heptose 6-epimerase**

The ADP-L-glycero-\(\alpha\)-manno-heptose 6-epimerase, which is encoded by the *bldD* (formerly *waaD* or *rfaD*, see Table 1) gene, has been well characterized. The function of this enzyme in *E. coli* was determined by Coleman (1983) using chromatographic and mass-spectroscopic methods to examine the heptose components of LPS extracted from the *gmbD* mutant. This epimerization reaction is the last enzymic step required for the synthesis of ADP-L-glycero-\(\alpha\)-manno-heptose. The epimerization occurs at the C-6 position of the heptose and involves an oxidation-reduction process that requires NADP⁺ as a cofactor (Ni et al., 2001). HldD has a high structural similarity with UDP-galactose epimerase and it has been classified as a member of the short-chain dehydrogenase/reductase (SDR) superfamily (Deacon et al., 2000). The enzyme is a homopentamer and each monomer is composed of two domains. The large N-terminal domain consists of a modified seven-stranded Rossmann fold which contains the characteristic sequence Gly-Gly-X-Gly-X-Gly (Pegues et al., 1990; Wierenga et al., 1986) and is
associated with NADP binding (Deacon et al., 2000). The C-terminal domain has an α/β structure that is involved in substrate binding.

**Gene organization of heptose biosynthesis pathways**

The heptose biosynthesis genes appear to be present in single copies in all cases examined with the exception of Ca. jejuni NCTC 11168 (Fig. 5). In this micro-organism, a second phosphoheptose isomerase gene, gmhA2, is present within a large cluster that resembles group II capsular polysaccharide gene clusters. It is possible that gmhA2 may be under a different regulation and/or may have been acquired by horizontal transfer.

In the majority of bacterial genomes sequenced to date, the genes encoding the enzymes of the L-β-D-heptose pathway are scattered throughout the chromosome. This is particularly true for the enteric as well as non-enteric bacteria like Ha. influenzae, Ps. aeruginosa, V. cholerae, He. pylori, N. meningitidis, N. gonorrhoeae and others. In most genomes examined, HldE is encoded as a bifunctional enzyme. However, in N. meningitidis, N. gonorrhoeae and R. eutropha, the two domains of the HldE protein are encoded by separate genes, hldA (kinase) and hldC (adenyllytransferase). The hldD gene is the only gene consistently found within core oligosaccharide biosynthesis clusters, usually in association with heptosyltransferase genes, like waaE and/or waaF (Fig. 5), which are known to process the β-anomers of ADP-L-glycero-D-manno-heptose (Gronow et al., 2001; Zamyatina et al., 2000). In some micro-organisms like Ca. jejuni and Mesorhizobium loti, all L-β-D-heptose biosynthesis genes are found clustered (Fig. 5), and the neighbouring genes are also involved in LPS core biosynthesis. In contrast, the four genes for L-β-D-heptose biosynthesis in He. pylori 26695 are part of a large cluster that also includes a flagellin gene, which suggests the intriguing possibility that the flagellin in this strain may be glycosylated.

The genes for the D-α-D-heptose pathway were first described in A. thermoaceticum DSM 10155, where they are located within a cluster containing additional genes involved in the synthesis and transfer of dTDP-rhamnose (Fig. 5). Since dTDP-rhamnose is also a component of the S-layer glycan moiety in A. thermoaceticum DSM 10155 (M. Graninger, B. Kneidinger, K. Bruno & P. Messner, unpublished results; Kneidinger et al., 2001; Kosma et al., 1995), the gene cluster may be involved in protein glycosylation. These genes are also present in My. tuberculosis, albeit in two different locations (Fig. 4), where they may be involved in the synthesis of glycolipids or in protein glycosylation (Romain et al., 1999). Interestingly, in the case of M. tuberculosis strain CDC 1551, the gmhB and gmhA appeared to be fused as a single gene, probably encoding a bifunctional protein with phosphatase and isomerase activities (Table 2). In the Gram-positive anaerobe Cl. acetobutylicum and in the Gram-negatives Ca. jejuni, B. mallei and B. pseudomallei, the D-α-D-heptose pathway genes are present within a large cluster with a gene organization resembling group II capsule genes (Fig. 5). B. pseudomallei produces an exopolysaccharide made of unbranched 2-O-acetyl-6-deoxy-β-D-manno-heptopyranose (Reckseidler et al., 2001). The repeating units of several Ca. jejuni O polysaccharides contain several variants on heptose residues, including D-glycero-D-altro-heptose or modifications of this molecule through methylation, deoxygenation or both (Penner & Aspinall, 1997). Interestingly, genes encoding putative methyltransferases and epimerase/dehydratases are also present within this group II capsule-like cluster in Ca. jejuni (Dorrell et al., 2001). Further research is required to determine whether these rare heptoses arise from precursors of the D-α-D-heptose or the L-β-D-heptose pathways.

**Future areas of research**

The complete elucidation of the nucleotide-activated heptose pathway will permit the undertaking of struc-

<table>
<thead>
<tr>
<th>Pathway</th>
<th>Location</th>
<th>Gene structure</th>
<th>Micro-organism</th>
</tr>
</thead>
<tbody>
<tr>
<td>L-β-D</td>
<td>Individual genes scattered</td>
<td>Bifunctional hldE</td>
<td>Agrobacterium tumefaciens, Buchnera sp., Caulobacter crescentus, Escherichia coli, Haemophilus influenzae, Salmonella typhimurium, Salmonella typhi, Vibrio cholerae, Yersinia pestis</td>
</tr>
<tr>
<td></td>
<td>Individual genes scattered</td>
<td>Individual hldA and hldC</td>
<td>Neisseria meningitidis, Neisseria gonorrhoeae, Ralstonia eutropha</td>
</tr>
<tr>
<td></td>
<td>All within LPS core biosynthesis gene clusters</td>
<td>Bifunctional hldE</td>
<td>Campylobacter jejuni</td>
</tr>
<tr>
<td>D-α-D</td>
<td>All within capsule-like biosynthesis gene clusters</td>
<td>Individual genes</td>
<td>Campylobacter jejuni, Clostridium acetobutylicum, Burkholderia pseudomallei, Burkholderia mallei, Aneurinibacillus thermoaceticum, Mycobacterium tuberculosis H37Rv</td>
</tr>
<tr>
<td></td>
<td>All within putative protein glycosylation gene clusters</td>
<td>Individual genes</td>
<td>Mycobacterium tuberculosis CDC 1551</td>
</tr>
<tr>
<td></td>
<td>All within putative protein glycosylation gene clusters</td>
<td>Bifunctional gmhA/gmhB</td>
<td>Mycobacterium tuberculosis CDC 1551</td>
</tr>
</tbody>
</table>
ture—function studies to understand in detail these enzymes as well as to develop potential inhibitors that may have clinical applications. The availability of these enzymes in a pure form in conjunction with the in vitro reactions established will facilitate the synthesis of nucleotide-activated heptose precursors of more complex sugars and therefore will permit the elucidation of the biosynthesis pathways of methyl and 6-deoxy heptoses. Also, the strong conservation of the heptose biosynthesis genes has a predictive value to delineate and predict nucleotide sugar biosynthesis pathways in bacteria whose genomes are not yet sequenced.

Acknowledgements

We thank Dr Carole Creuzenet for critical reading of the manuscript. The work by the authors has been supported by grants from the Natural Sciences and Engineering Council of Canada (to M.A.V.), and from the Austrian Science Fund, projects P14209-MOB (to P.M.) and P14978-MOB (to P.K.).

References


Bayer, M. E., Koplow, J. & Goldfine, H. (1975). Isolation of polymyxin B-resistant mutants of


NMR investigation of the O-specific polysaccharides in native 2476, 2477, 2478, and 2494: high-resolution magic angle spinning
Re mutants. Salmonella minnesota of heptose transferase(s), from membrane of the cell envelope of heptose-deficient mutants of different from the

Koronakis, V., Koronakis, E., Li, J. & Stauffer, K. (1997). Computer analysis of Isolation of adenosine 5'-glycero-


