Published in final edited form as:
Biochemistry. 2009 February 17; 48(6): 1162–1172. doi:10.1021/bi802211k.

Identification of Undecaprenyl Phosphate-β-D-Galactosamine in *Francisella novicida* and Its Function in Lipid A Modification

Xiaoyuan Wang[1,2], **Anthony A. Ribeiro**[3], **Ziqiang Guan**[1], and **Christian R. H. Raetz**[*,1]

1*Department of Biochemistry, Duke University Medical Center, Durham, NC 27710*

2*State Key Laboratory of Food Science and Technology, Jiangnan University, Wuxi, Jiangsu 214122, China*

3*Duke NMR Center and Department of Radiology, Duke University Medical Center, Durham, NC 27710*

Abstract

Francisella tularensis is a highly infectious pathogen that causes tularemia. *Francisella* lipid A contains an unusual galactosamine (GalN) unit, attached to its 1-phosphate moiety. Two genes, *flmF2* and *flmK*, are required for the addition of GalN to *Francisella* lipid A, but the relevant enzymes and the GalN donor substrate have not been characterized. We now report the purification and identification of a novel minor lipid from *Francisella novicida* that functions as the GalN donor. Based on electrospray ionization mass spectrometry (ESI/MS) and NMR spectroscopy, we propose that this compound is undecaprenyl phosphate-β-D-GalN. Approximately 0.5 mg of pure lipid was obtained from 10 g of *F. novicida* by chloroform/methanol extraction, followed by DEAE-cellulose chromatography, mild alkaline hydrolysis, and thin layer chromatography. ESI/MS in the negative mode revealed a molecular ion $[M-H]^-$ at *m/z* 1006.699, consistent with undecaprenyl phosphate-GalN. ^{31}P-NMR spectroscopy showed a single phosphorus atom in phosphodiester linkage. Selective inverse decoupling difference spectroscopy demonstrated that the undecaprenyl phosphate group is attached to the anomeric carbon of the sugar. ^{1}H-NMR studies showed the presence of a polyisoprene chain and a sugar consistent with a β-D-GalN unit. Heteronuclear multiple quantum coherence (HMQC) analysis confirmed that nitrogen is attached to C-2 of the sugar. Purified undecaprenyl phosphate-β-D-GalN supports the *in vitro* modification of lipid IV_A by membranes of *E. coli* cells expressing FlmK, an orthologue of *E. coli* ArnT, the enzyme that transfers 4- amino-4-deoxy-L-arabinose to lipid A in polymyxin-resistant strains. The discovery of undecaprenyl phosphate-β-D-GalN suggests *Francisella* modifies lipid A with GalN on the periplasmic surface of the inner membrane.

Lipopolysaccharide (LPS) makes up the outer leaflet of the outer membranes of most Gram-negative bacteria (1). It typically consists of three covalently linked domains: the hydrophobic lipid A moiety, the core oligosaccharide, and the O-antigen polymer (2-4). Lipid A from enteric organisms (Fig. 1A), like *Escherichia coli* or *Salmonella typhimurium*, activates the Toll-like receptor 4 (TLR4) of the mammalian innate immune system, which triggers an inflammatory response and helps clear localized infections (5-8). However, a more generalized response to lipid A in the context of an overwhelming systemic infection, accompanied by over-production of cytokines, can contribute to Gram-negative septic shock and death (9,10). Low concentrations of lipid A (pM to nM) are usually sufficient to activate the TLR4/MD2 receptor complex (11,12).

*Author to whom correspondence should be addressed: Christian R. H. Raetz, Department of Biochemistry, Duke University Medical Center, Box 3711, Durham, NC 27710, USA. Telephone: 919-684-5178; Fax: 919-684-8885; E-mail:raetz@biochem.duke.edu.

Some bacteria, such as the highly infectious human pathogen *Francisella tularensis*, synthesize lipid A molecules with relatively long acyl chains (13-16) (Fig. 1B) that do not activate TLR4/MD2 (17-19). In addition, wild-type *Francisella* lipid A is covalently modified (Fig. 1B) to prevent killing by cationic antimicrobial peptides of the innate immune system or polymyxin (19). *Francisella* contains an unusual galactosamine (GalN) unit connected to the phosphate group at the 1-position of its lipid A (Fig. 1B), which neutralizes its negative charge (14,15). Furthermore, the 4′-phosphate group and the 3′-hydroxyacyl chain, which are present in *E. coli* lipid A, are missing entirely in wild-type *Francisella* lipid A (Figs. 1A and 1B), eliminating the net negative charge (14,15). The 4′-phosphate group of *Francisella* lipid A is removed on the periplasmic surface of the inner membrane by a selective phosphatase, termed LpxF (20). Deletion of *lpxF* in *Francisella novicida*, a mouse specific model organism for tularemia, results in the retention of both the lipid A 4′-phosphate moiety and the 3′-hydroxyacyl chain (Fig. 1C) (19), suggesting an obligatory order of processing. *F. novicida* mutants lacking *lpxF* are hypersensitive to polymyxin and are highly attenuated in the mouse infection model (19). Another unusual feature of *F. novicida* is the fact that >90% of its lipid A is in a “free” state, *i.e.* not linked to a core oligosaccharide and O-antigen (15).

The biosynthetic origin of the GalN unit attached to lipid A in *Francisella* has not been characterized at the enzyme level. However, two genes, *flmK* and *flmF2*, are necessary for the addition of the GalN moiety to lipid A (15,21), as judged by analysis of the lipid A isolated from the appropriate mutants. The *flmK* gene is an orthologue of *E. coli arnT* (previously known as *pmrK*) (15,22); the latter encodes the enzyme that transfers 4-amino-4-deoxy-L-arabinose (L-Ara4N) to lipid A on the outer surface of the inner membrane in polymyxin-resistant strains (23,24). ArnT uses undecaprenyl phosphate-α-L-Ara4N as its donor substrate (Fig. 2) (23,24). By analogy, we reasoned that FlmK might utilize undecaprenyl phosphate-β-D-GalN as its sugar donor (Fig. 2). The *flmF2* gene is an orthologue of *E. coli arnC*(*pmrF*) (4,22,25), suggesting that it might encode an enzyme required for the biosynthesis of undecaprenyl phosphate-β-D-GalN.

Here we report the purification and characterization of the novel prenol lipid undecaprenyl phosphate-β-D-GalN from *F. novicida* and show that it can donate its GalN unit to the 1-phosphate group of lipid A-like precursor molecules. Analysis of this lipid by electrospray-ionization mass spectrometry (ESI/MS) and high resolution NMR spectroscopy strongly supports its structure as undecaprenyl phosphate-β-D-GalN. Small amounts of dodecaprenyl phosphate-β-D-GalN were also detected. The purified lipid functions *in vitro* as an effective GalN donor substrate in the modification of the precursor lipid IV_A (4), catalyzed by recombinant FlmK expressed in *E. coli*. In the accompanying manuscript, we report an *in vitro* system for the enzymatic synthesis of undecaprenyl phosphate-β-D-GalN from UDP-GalNAc and undecaprenyl phosphate.

Experimental procedures

Materials

Glass backed 0.25 mm Silica Gel 60 thin layer chromatography (TLC) plates, chloroform, ammonium acetate, and sodium acetate were obtained from EMD Chemicals Inc. (Gibbstown, NJ), while pyridine, methanol, and formic acid were from Mallinckrodt Baker, Inc. (Phillipsburg, NJ). Trypticase soy broth, yeast extract and tryptone were purchased from Difco/Becton Dickinson and Co. (Franklin Lakes, NJ). DEAE-cellulose (DE52) was purchased from Whatman (Florham Park, NJ). $^{32}P_i$ was purchased from PerkinElmer Life and Analytical Sciences Inc. (Waltham, MA). D_2O, CD_3OD and $CDCl_3$ were from Sigma-Aldrich (St. Louis, MO).

Purification of a Putative GalN Donor Lipid from F. novicida

Two liters of *F. novicida* strain XWK2, previously constructed in our laboratory by deletion of the *arnT* orthologue of *F. novicida* U112 (15), were grown at 37 °C in 3% Trypticase soy broth and 0.1% cysteine (TSB-C broth), harvested by centrifugation, and washed with phosphate-buffered saline (26). About 10 g wet cells were extracted for 1 hr at room temperature with 1.8 liters of a single-phase Bligh-Dyer mixture (27) and centrifuged to remove insoluble debris. The supernatant (containing mostly phospholipids and free lipid A) was converted to a two-phase Bligh-Dyer system (27). The lower phase was dried by rotary evaporation. About 360 mg crude lipids were obtained.

The crude lipids were dissolved in 60 ml chloroform/methanol/water (2:3:1, v/v/v) and applied to a 30 ml DEAE-cellulose column (Whatman DE52 in the acetate form), equilibrated with the same solvent (28). The column was washed with 10 bed volumes of chloroform/methanol/water (2:3:1, v/v/v). The lipid donor emerged in the run-through fractions, together with phosphatidylethanolamine, phosphatidylcholine and free lipid A, as judged by ESI/MS. The run-through fractions were converted to a two-phase Bligh-Dyer system. The lower phase was recovered and dried by rotary evaporation. About 250 mg of lipid was recovered.

To remove the phosphatidylethanolamine and phosphatidylcholine, the entire sample was subjected to mild alkaline hydrolysis (29). The lipids were dissolved in 114 ml of a single-phase Bligh-Dyer mixture, containing NaOH at a final concentration of 0.2 M, and incubated at room temperature for 1 h. Next, the mixture was neutralized with concentrated HCl and converted to a two-phase Bligh-Dyer system. About 50 mg of solids were recovered from the lower phase, consisting mainly of the base-stable GalN donor lipid and free fatty acids, as judged by ESI/MS analysis (data not shown).

The hydrolyzed material was dissolved in 114 ml chloroform/methanol (95:5, v/v) and applied onto a 20 ml Bio-SilA silicic acid column (Bio-Rad, Richmond, CA), equilibrated in chloroform/methanol (95:5, v/v). The column was washed with 114 ml of chloroform/methanol (95:5, v/v). The fatty acids mostly emerged in the run-through, but the GalN donor lipid remained bound to the column, as judged by ESI/MS. The GalN donor lipid was eluted from the column with 114 ml of a single-phase Bligh-Dyer mixture. Following conversion to a two-phase Bligh-Dyer system, about 3.6 mg of lipid was recovered from the lower phase.

To purify the GalN donor lipid further, preparative TLC was employed. The sample was dissolved in 1 ml chloroform/methanol (4:1, v/v) and applied to a 20 × 20 cm TLC plate. Chloroform/pyridine/88% formic acid/water (30:70:16:10, v/v/v/v) was used for chromatography. While the plates were drying at room temperature, two lipid bands were seen transiently as white zones. The bands were marked with a pencil and scraped off after the plates were dry. The silica chips were extracted with a single-phase Bligh-Dyer mixture (7.6 ml for 1 hour at room temperature). The chips were removed by centrifugation. The supernatant was passed through a Pasteur pipette column fitted with a glass wool plug, and the run-through was converted into a two-phase Bligh-Dyer system. The phases were separated by centrifugation, and the lower phase was dried under a stream of N_2. The material recovered from the lower band observed while drying the TLC plate was confirmed to be the putative GalN donor lipid by ESI/MS. About 0.5 mg of pure material was recovered from 10 g of *F. novicida* wet cells, which represents about 0.1 % of the total lipid.

ESI/MS and MS/MS Analysis

All ESI/MS spectra were acquired on a QSTAR XL quadrupole time-of-flight tandem mass spectrometer (ABI/MDS-Sciex, Toronto, Canada) equipped with an ESI source. Lipid samples were dissolved in chloroform/methanol (2:1, v/v), containing 1 % piperidine, at about 25 μg/

ml and subjected to ESI/MS in the negative ion mode (12,15). Nitrogen was used as collision gas for the MS/MS experiments (12). Data acquisition and analysis were performed using the instrument's Analyst QS software.

NMR Analysis

The entire 0.5 mg sample of the putative GalN donor lipid was dissolved in 0.2 ml $CDCl_3$/CD_3OD/D_2O (1:2:0.8, v/v/v) and transferred into a 3 mm NMR tube. Proton and carbon chemical shifts are reported relative to the tetramethylsilane (TMS) scale at 0.00 ppm. The ^2H signal of CD_3OD was used as a field frequency lock with the residual signal of CD_3OD serving as the secondary reference at 49.5 ppm for carbon spectra. ^1H NMR spectra were recorded at the Duke NMR Center on Varian Inova 800 or 600 NMR spectrometers, equipped with Varian cryogenic probes and Dell 370 computers. ^1H NMR spectra at 800 MHz were obtained with a 7.2 kHz spectral window, a 67 ° pulse field angle (4.5 μsec), a 4.5 sec acquisition time, and a 1 sec relaxation delay. The spectra were digitized using 64k points to obtain a digital resolution of 0.225 Hz/pt. The two-dimensional NMR experiments COSY, heteronuclear multiple bond coherence (HMBC) and heteronuclear multiple quantum coherence (HMQC) were performed at 800 MHz as previously described (30,31). ^1H-decoupled ^{31}P NMR spectra were recorded at 202.3 MHz on a Varian Inova 500 spectrometer with a spectral window of 12143.3 Hz digitized into 25,280 data points (digital resolution of 1Hz/point or ~0.005 ppm/point), a 600 pulse flip angle (8 ms), and a 1.6-s repeat time. ^{31}P chemical shifts were referenced to 85% H_3PO_4 at 0.00 ppm. Inverse decoupled difference spectra were recorded as ^1H-detected ^{31}P-decoupled heteronuclear NMR experiments, as previously described (30,31).

Expression Cloning of the F. novicida flmK Gene in E. coli

Genomic DNA was isolated from *F. novicida* U112 using the Easy DNA kit (Invitrogen). The *flmK* gene was amplified by PCR and cloned behind the *lac* promoter in the vector pWSK29 (32). The forward PCR primer (5′-GCGTCTAGAAGGAGATATACCatgaataaatctaaaaactc-3′) was designed with a clamp region, a *Xba*I restriction site (*underlined*), a ribosome binding site, and a region matching the coding strand, starting at the translation initiation site (*lower case*). The reverse primer (5′-GCGACTCGAGctattttacaacactgacaacaac-3′) was designed with a clamp region, an *Xho*I restriction site (*underlined*), and a region matching the anti-coding strand that included the stop codon. The PCR was performed using *Pfu* polymerase and genomic DNA as the template. Amplification was carried out in a 100 μl reaction mixture containing 100 ng of template, 250 ng primers, and 2 units of *Pfu* polymerase. The reaction was started at 94 °C for 1 min, followed by 25 cycles of denaturation (30 seconds at 94 °C), annealing (30 seconds at 55 °C), and extension (45 seconds at 72 °C). After the 25th cycle, a 10 min extension time was used. The reaction product was analyzed on a 1% agarose gel. The desired band was excised and gel purified. The PCR product was then digested using *Xba*I and *Xho*I, and ligated into the expression vector pWSK29, which had been similarly digested and treated with shrimp alkaline phosphatase. The ligation mixture was transformed into *E. coli* XL1-blue cells and screened for positive inserts by selection of colonies on LB plates containing 30 μg/ml ampicillin. The desired plasmid, designated pFlmK, was isolated from a positive transformant, and the *flmK* insert was confirmed by DNA sequencing. The plasmid pFlmK was then re-transformed into *E. coli* XL1-Blue to form XL1-Blue/pFlmK.

In vitro Assay of FlmK with the Purified GalN Donor Lipid

The substrate [4′-^{32}P]-lipid IV_A was generated from [γ-^{32}P]-ATP and the tetraacyl-disaccharide 1-phosphate lipid A precursor, using the overexpressed lipid A 4′-kinase present in membranes of *E. coli* BLR(DE3)/pLysS/pJK2 (33,34).

Lipid IV_A carrier was isolated from the triple mutant *E. coli* MKV15b, which synthesizes lipid A lacking secondary acyl chains (35). Briefly, the cells were grown, harvested by centrifugation, and washed with phosphate-buffered saline. Glycerophospholipids were extracted with a single phase Bligh-Dyer mixture. The lipid IV_A was recovered from the LPS present the cell residue by hydrolysis at 100 °C in sodium acetate buffer, pH 4.5, in the presence of 1% SDS (31), followed by Bligh-Dyer extraction. The lipid IV_A was then purified by chromatography on a DEAE-cellulose column, followed by a C18 reverse phase column, as described previously (36,37).

Membranes of *E. coli* XL1-Blue expressing the FlmK protein were prepared from IPTG-induced cells, as previously described (20), and were assayed at 30 °C. The reaction mixture contained 50 mM potassium phosphate, pH 6.4, 0.05% Triton X-100, and 10 μM [4′-^{32}P]-lipid IV_A (3000-6000 cpm/nmol) and 0.06 mg/ml membranes. The reaction was terminated by spotting 5 μl samples onto a silica TLC plate. The plate was developed in the solvent of chloroform/pyridine/88% formic acid/water (50:50:16:5,v/v/v/v). After drying and overnight exposure of the plate to a PhosphorImager screen, product formation was detected and quantified with a Storm 840 PhosphorImager, equipped with ImageQuant software (GE Healthcare, Chalfont St. Giles, Buckinghamshire, U.K.).

Results

Purification and ESI/MS Analysis of a Putative GalN Donor Lipid from F. novicida

A GalN residue is attached to the 1-phosphate group of lipid A in several *Francisella* subspecies, including the mouse model organism *F. novicida* (Fig. 1) (14,15). The *flmK* gene is required for the addition of the GalN moiety to lipid A in *F. novicida* (15,21). Deletion and insertion mutants of *flmK* in *F. novicida* synthesize lipid A molecules lacking the GalN unit, as judged by the loss of a 161 amu substituent (15,21). D-GalN is similar in structure and charge to L-Ara4N (Fig. 2). The L-Ara4N residue is transferred to lipid A on the outer surface of the inner membrane in polymyxin-resistant mutants of *E. coli* and *Salmonella* by ArnT (4,23,24). The L-Ara4N donor substrate is undecaprenyl phosphate-L-Ara4N (Fig. 2) (4,23,24). Given the sequence similarity of *E. coli* ArnT and *F. novicida* U112 FlmK (29% identity and 49% similarity over 335 residues), we explored the possibility that the GalN unit found on the lipid A of *F. novicida* might arise from the analogous donor substrate, undecaprenyl phosphate-β-D-GalN (Fig. 2). We therefore extracted total lipids from *F. novicida* with a single-phase Bligh-Dyer mixture and analyzed them by ESI/MS, as shown in Fig. 3A. Small overlapping peaks were observed at *m/z* 1006.698 and *m/z* 1007.603 (Fig. 3A, inset), suggesting a mixture of an undecaprenyl phosphate-hexosamine (predicted $[M-H]^-$ at *m/z* 1006.726) and an undecaprenyl phosphate-hexose (predicted $[M-H]^-$ at *m/z* 1007.710). A smaller peak near *m/z* 1074.748 (Fig. 3A inset) indicated the presence of a dodecaprenyl phosphate-hexosamine. ESI/MS analysis of crude *E. coli* lipids revealed a peak near *m/z* 1007.7, suggestive of an undecaprenyl phosphate-hexose, but not at *m/z* 1006.7 (data not shown).

To characterize the putative GalN donor lipid in more detail, the total lipids were fractionated on a DEAE cellulose column. The putative undecaprenyl phosphate-hexosamine emerged in the run-through, as judged by ESI/MS (not shown), together with phosphatidylethanolamine, phosphatidylcholine and free lipid A. The pooled run-through fractions were subjected to a mild alkaline hydrolysis to deacylate the glycerophospholipids (29). The free fatty acids were separated from the undecaprenyl phosphate-hexosamine by silicic acid column chromatography. To obtain homogeneous material, the resulting sample (3.6 mg) was subjected to preparative TLC. Approximately 0.5 mg of pure undecaprenyl phosphate-hexosamine was obtained from 10 g of *F. novicida* wet cells (~0.1% of the total lipid). The purified lipid was analyzed by ESI/MS (Fig. 3B). All the phospholipids and free lipid A had been removed, and the two predominant remaining molecular ions at *m/z* 1006.699 and *m/z*

1074.764 corresponded within error to the predicted $[M-H]^-$ values expected for undecaprenyl phosphate-hexosamine (1006.726) and dodecaprenyl phosphate hexosamine (1074.789), respectively. The contaminating lipid that was seen near *m/z* 1007.603 in the crude material (Fig. 3A) was removed during chromatography on DEAE cellulose (not shown). The new small peak at *m/z* 1022.698 can be attributed to the incorporation of a single oxygen atom into the undecaprenyl phosphate-hexosamine, most likely because of chemical oxidation during preparative TLC.

ESI/MS/MS Analysis of the Putative GalN Donor Lipid

The prominent molecular ion $[M-H]^-$ at *m/z* 1006.699 (Fig. 3B) of the purified lipid was subjected to ESI/MS/MS analysis (Fig. 4). The peaks at *m/z* 78.957 $[PO_3]^-$ and *m/z* 96.966 $[H_2PO_4]^-$ confirmed the presence of a phosphate group. The peak at *m/z* 258.033 had the value expected for a phosphorylated hexosamine derivative, suggesting that the putative GalN moiety is connected to the phosphate group present in the purified lipid (Figs. 2 and 4). The ion near *m/z* 324 (not labeled) had the mass predicted for a phosphorylated hexosamine plus an isoprene unit, consistent with the phosphate group connecting the GalN moiety and the proximal isoprene unit of the undecaprenyl chain (Fig. 2). A prominent ion derived by neutral loss of the GalN unit was seen at *m/z* 845.640, which also corresponds to the $[M-H]^-$ of undecaprenyl phosphate. The ESI/MS/MS analysis of the molecular ion at *m/z* 1074.764 in Fig. 3B was very similar (not shown), except that a large peak at *m/z* 913.7 replaced the peak at *m/z* 845.640, confirming that the parent ion at *m/z* 1074.764 arises from a dodecaprenyl phosphate-hexosamine. In summary, the ESI/MS/MS study supports the hypothesis that the purified lipid consists mostly of undecaprenyl phosphate-GalN (Fig. 2) and some dodecaprenyl phosphate-GalN.

Characterization of the Purified Lipid by NMR Spectroscopy

^{31}P NMR spectroscopy of the purified lipid revealed a single phosphorus resonance near 0.752 ppm (Fig. 5A), consistent with the presence of a phosphodiester linkage. Subtraction of two ^1H NMR spectra obtained with on and off resonance selective decoupling of the 0.752 ppm phosphate signal (24,30,31) revealed simultaneous changes at the anomeric H-1 signal of the D-GalN moiety and at the αCH_2OP signal of the proximal isoprene residue (Fig. 5B), consistent with the proposed structure shown in Fig. 2.

The 800-MHz ^1H NMR spectrum of the purified lipid showed six resolved sugar protons (Fig. 5C). The assignments of the individual protons of the putative GalN residue were derived from a two-dimensional COSY analysis (Fig. 6 and Table 1). The peak at 5.06 ppm ($J_{1,2}$ = 8.2 Hz) is assigned to the anomeric H-1 proton, based on its chemical shift and the result of subtracting the two ^1H NMR spectra obtained with on and off resonance selective decoupling of the 0.752 ppm phosphate signal (Fig 5B). The COSY cross peak from H-1 locates the H-2 signal at 3.22 ppm (Fig. 6). H-2 connects to H-3 at 3.77 ppm, which, in turn, is coupled to H-4 at 3.91 ppm. Further tracing of the COSY cross peaks locate the remaining GalN protons, i.e. H-5 at 3.67 ppm, H-6a at 3.79 ppm and H-6b at 3.73 ppm (Table 1).

The 800-MHz ^1H NMR spectrum of the purified lipid also revealed a ten-proton prenyl CH peak near 5.1 ppm, a resolved αCH near 5.4 ppm, a resolved αCH_2OP near 4.4 ppm, a CH_2 envelope near 2.0 ppm and twelve CH_3 signals near 1.6 to 1.7 ppm, verifying the presence of the undecaprenyl chain (Fig. 5C, 6, 7B and Table 2) (24). The two-dimensional COSY analysis revealed a strong cross peak between the αCH and αCH_2OP groups of the proximal isoprene unit (Figs. 2 and 6). Furthermore, the αCH and αCH_2OP of the proximal isoprene unit each showed a weak cross peak to the resolved methyl signal at 1.73 ppm (Fig. 6). The major unresolved polyisoprene CH signals at 5.10 ppm similarly showed strong cross-peaks to the upfield CH_2 resonances and weaker (long range) cross-peaks to the upfield CH_3 signals (Fig.

6). The low field shift of H-1 (5.06 ppm) and the large 8.2 Hz coupling constant ($J_{1,2}$) revealed in the ^{31}P decoupled ^{1}H difference spectrum (Fig. 5B) indicates that H-1 of the presumptive D-GalN residue is in the axial (β) position, so that the undecaprenyl phosphate chain must be situated equatorially (Fig. 2).

To confirm the assignments derived from the ^{1}H NMR spectroscopy, the purified lipid was further analyzed by HMQC and HMBC two-dimensional NMR experiments. The partial two-dimensional HMQC ^{1}H-^{13}C correlation map reveals six relevant ^{1}H-^{13}C single-bond cross peaks and one impurity correlation (X) in the sugar region (Fig. 7A). The GalN H-1 signal reveals the anomeric carbon resonance at 96.1 ppm (C-1), consistent with a β-linked D-GalN sugar residue (38). The H-6a and H-6b multiplets correlate to a single carbon signal at 62.0 ppm (C-6), while the H-3, H-4 and H-5 multiplets connect to carbon resonances at 70.5 (C-3), 68.1 (C-4) and 77.1 (C-5) ppm, respectively, which correspond to oxygen-substituted carbons of sugars. However, nitrogen-substituted carbons of amino sugars resonate near 50-55 ppm (38). The H-2 multiplet shows a prominent cross peak near 55.3 ppm, confirming C-2 as the site of the amino group substitution (Fig. 2). Fig. 7A also shows the direct bond correlations from the major unresolved methine proton signals of the undecaprenyl chain to unresolved olefinic carbon signals near 126 ppm, and from the methine and the proximal methylene protons of the α-isoprene unit to carbon resonances at 122.5 and 64.0 ppm, respectively.

The ^{1}H-^{13}C HMQC of the purified lipid revealed structural information for the polyisoprene unit similar to that of undecaprenyl phosphate-L-Ara4N, previously characterized in our laboratory (24). The CH_2 protons of the undecaprenyl chain yield 4 distinct ^{1}H-^{13}C HMQC cross peaks (Fig. 7B and Table 2). Based on the multibond correlations discussed below, the major carbon peak at 27.5 ppm is assigned to the proximal CH_2 groups of the bulk isoprene units (*i.e.* those that are adjacent to a CH proton as shown in Table 2), while the 33.2 ppm peak of about equal intensity is assigned to the CH_2 groups of the bulk isoprene units that are *trans* relative to a methine proton (Table 2) (24). The small carbon cross peak at 40.5 ppm arises from the *cis*-CH_2 groups of the ω-1 and ω-2 isoprene units. The small, partially resolved cross peak at 32.8 ppm arises from the *trans* CH_2 group of the α isoprene unit (Fig. 7B and Table 2).

The CH_3 protons yield five HMQC peaks (Fig. 7B). The *cis* CH_3 protons from the α isoprene unit (1.73 ppm), the seven *cis* CH_3 of the interior β to ω-3 units (1.67 ppm), and the *cis* CH_3 of the ω unit (1.66 ppm) yield distinct carbon cross peaks at 24.2, 24.4 and 26.5 ppm respectively (Table 2). The three *trans* methyl groups of the ω-2, ω-1 and ω isoprene units yield three overlapped carbon cross peaks near 17.0 ppm (Fig. 7B and Table 2).

Additional HMBC multibond correlations (not shown) in general verify the GalN and undecaprenyl assignments derived from COSY and HMQC. For example, the resolved H-4 at 3.91 ppm shows distinct multibond correlations to C-2 (55.2 ppm) and C-3 (71.0 ppm). The HMBC correlations yield a complete analysis of the α isoprene unit. The *cis* CH_2OP proton signal at 4.44 ppm shows distinct multibond correlations to carbon resonances at 122.5 ppm (CH from the α isoprene unit) and 142.3 ppm (quaternary carbon resonance from the α isoprene unit). Scrutiny of the cross peaks from the 1.73 ppm methyl proton signal (*cis* CH_3 of α isoprene unit) locates the corresponding multibond correlations to 33.2 (*trans* CH_2 of α isoprene unit), 122.5 and 142.3 ppm, respectively, thus verifying the assignment of the 1.73 ppm signal as the *cis* methyl group of the α isoprene unit.

The HMBC correlations seen for the bulk undecaprenyl protons of undecaprenyl-phosphate-D-GalN are identical to those previously observed for undecaprenyl-phosphate-L-Ara4N (24).

In summary, the NMR data supports the hypothesis, also derived from the mass spectrometry, that the purified lipid has the structure undecaprenyl phosphate-β-D-GalN, as shown in Fig. 2.

In vitro Assay for FlmK Activity Dependent on the Purified GalN Donor Lipid

We next tested the purified undecaprenyl phosphate-β-D-GalN as a GalN donor in an *in vitro* FlmK assay with lipid IV_A (4) as the acceptor substrate (Scheme 1). The assay conditions were similar to those previously reported for detecting L-Ara4N transferase activity catalyzed by ArnT (23,24). The *flmK* gene of *F. novicida* (an orthologue of *arnT*) was cloned into the expression vector pWSK29 to generate pFlmK. The latter was transformed in *E. coli* XL1-Blue. Membranes were prepared from XL1-Blue/pWSK29 and XL1-Blue/pFlmK, and were assayed with [4′-^{32}P]-lipid IV_A as the acceptor substrate in the presence or absence of undecaprenyl phosphate-β-D-GalN (Fig. 8 and Scheme 1). Efficient formation of a more slowly migrating [4′-^{32}P]-lipid IV_A derivative, consistent with the addition of an amino sugar (23, 24), was dependent upon both FlmK and undecaprenyl phosphate-β-D-GalN (Fig. 8). When the purified GalN donor lipid was omitted, no activity was observed, consistent with the fact that *E. coli* does not synthesize its own undecaprenyl phosphate-β-D-GalN. The involvement of an undecaprenyl phosphate-linked sugar donor strongly suggests that the addition of the D-GalN moiety to lipid A occurs on the periplasmic surface of the inner membrane in *F. novicida*.

Discussion

The lipid A moiety of LPS is a relatively conserved structure that makes up the outer monolayer of the outer membrane of most Gram-negative bacteria (3,4,39). Hexa-acylated lipid A of *E. coli* and related enteric bacteria potently activates the TLR4/MD2 receptor of the innate immune system (11,12), whereas lipid A from photosynthetic bacteria like *Rhodobacter sphaeroides*, or tetra-acylated precursors, such as lipid IV_A of *E. coli* (Scheme 1), are TLR4/MD2 antagonists (40,41). Alternatively, lipid A species derived from *Francisella tularensis* are relatively inactive, both as TLR4 agonists and as antagonists (18,19). Changes in the lengths and numbers of the acyl chains (Fig. 1), and the presence or absence of the phosphate groups, are critical in determining these biological activities (11,42,43). The recent elucidation of the structure of the TLR4/MD2 receptor with a bound antagonist should shed additional light on this interesting structure-activity relationship (8,44).

The attachment of cationic sugars to the lipid A phosphate groups, which include 4-amino-4-deoxy-L-arabinose (L-Ara4N) in enteric bacteria like *E. coli* and D-galactosamine in *Francisella* (Figs. 1 and 2), reduces the net negative charge of lipid A (4). These modifications confer bacterial resistance to cationic antimicrobial peptides and polymyxin, which cannot bind to the modified lipid (1). Mutants that are unable to modify their lipid A with cationic sugars therefore tend to be less virulent and are hypersensitive to polymyxin (19,21,45).

The origin of the L-Ara4N moiety and the mechanism of its attachment to lipid A have been fully established in *E. coli* and *Salmonella typhimurium* (4). The critical last step is the transfer of the L-Ara4N residue to lipid A on the outer surface of the inner membrane by ArnT, a complex polytopic membrane protein that is distantly related to eukaryotic protein mannosyl transferases (4,23). The L-Ara4N donor substrate for ArnT is undecaprenyl phosphate-L-Ara4N (Fig. 2) (24). Sugar nucleotides are not available in the periplasm.

F. novicida and related pathogens attain their resistance to cationic antimicrobial peptides in two ways. First, LpxF efficiently removes the 4′-phosphate group on the periplasmic side of the inner membrane (20); mutants lacking the *lpxF* gene retain their 4′-phosphate group (Fig. 1C), are hypersensitive to polymyxin, and are avirulent in the mouse infection model (19). Second, *F. novicida* efficiently modifies its lipid A 1-phosphate group with a GalN residue (Fig. 1), which resembles L-Ara4N (Fig. 2) in structure and charge (14,15). Deletion of *flmK*, which encodes an inner membrane protein with significant sequence similarity to *E. coli* ArnT, results in loss of the GalN substituent of *F. novicida* lipid A and attenuation of virulence (15,

21), suggesting by analogy that undecaprenyl phosphate-GalN should be the donor substrate for FlmK.

To study the enzymatic synthesis of GalN-modified lipid A in *F. novicda*, it was necessary to first identify the GalN donor substrate unequivocally and to develop an *in vitro* assay for the proposed FlmK transferase (Scheme 1). We therefore searched for the novel minor lipid undecaprenyl phosphate-GalN in chloroform extracts of *F. novicida*. ESI/MS revealed the presence of a compound with the expected $[M-H]^-$ near *m/z* 1006.5 (Fig. 3), which is not observed in total *E. coli* lipids (data not shown). Following purification on DEAE cellulose, mild alkaline hydrolysis to remove glycerophospholipids and TLC, we obtained 0.5 mg of pure compound from 10 g of *F. novicida* with the predicted mass spectrum and MS/MS fragmentation pattern (Figs. 3 and 4). The NMR analysis was consistent with the presence of an undecaprenyl phosphate chain attached via a phosphodiester linkage to the anomeric carbon of the β-GalN unit (presumed to be the D-isomer) (Figs. 5-7 and Tables 1 and 2). We estimate that undecaprenyl phosphate-β-D-GalN represents about 0.1% of the total *F. novicida* lipid, an amount that is comparable to the abundance of undecaprenyl phosphate-L-Ara4N in polymyxin-resistant strains of *E. coli* and *Salmonella typhimurium* (24). Importantly, our undecaprenyl phosphate-β-D-GalN preparation is an efficient GalN donor for recombinant *F. novicida* FlmK expressed in *E. coli*, as judged by *in vitro* assays of membranes with lipid IV_A as the acceptor substrate (Fig. 8).

The *flmK* (Ftn_0546) gene of *F. novicida* maps next to *flmF2* (Ftn_0545) (46), the latter gene encoding a protein with significant similarity to *E. coli* ArnC (4), which along with FlmK is required for the modification of *F. novicida* lipid A with GalN (21). *E. coli* ArnC is an enzyme that transfers *N*-formyl-L-Ara4N from UDP-*N*-formyl-L-Ara4N to undecaprenyl phosphate (25). By analogy, it is reasonable to propose that FlmF2 catalyzes the condensation of UDP-GalNAc (or perhaps UDP-GalN) and undecaprenyl phosphate to generate the corresponding undecaprenyl phosphate-sugar. In the accompanying manuscript we describe an *in vitro* system for assaying FlmF2, and we demonstrate this enzyme catalyzes the efficient formation of undecaprenyl phosphate-GalNAc from UDP-GalNAc and undecaprenyl phosphate. We also describe an additional gene, encoding a novel deacetylase, which converts undecaprenyl phosphate-GalNAc to undecaprenyl phosphate-GalN, the novel prenol lipid identified in the present study.

Acknowledgments

We thank Dr. Francis Nano (University of Victoria, Canada) for providing *F. novicida* U112. We thank Dr. David Six for helpful discussions and reading the manuscript.

This research was supported by NIH Grant R-37-GM-51796 to C. R. H. Raetz, the LIPID MAPS Large Scale Collaborative Grant GM-069338, and the NSFC Project 30770114 and 111 Project 111-2-06 to X. Wang. AAR is partially supported by NIH R01 A1057588 and NCI P30-CA-14236. NMR instrumentation in the Duke NMR Center was funded by the NSF, the NIH, the NC Biotechnology Center and Duke University.

List of Abbreviations

L-Ara4N, 4-amino-4-deoxy-L-arabinose; ESI/MS, electrospray ionization/mass spectrometry; GalN, galactosamine; HMBC, heteronuclear multiple bond coherence; HMQC, heteronuclear multiple quantum coherence; LPS, lipopolysaccharide; TLC, thin layer chromatography; TMS, tetramethylsilane.

References

(1). Nikaido H. Molecular basis of bacterial outer membrane permeability revisited. Microbiol. Mol. Biol. Rev 2003;67:593–656. [PubMed: 14665678]

(2). Brade, H.; Opal, SM.; Vogel, SN.; Morrison, DC. Marcel Dekker, Inc.; New York: 1999. p. 950

(3). Raetz CRH, Whitfield C. Lipopolysaccharide endotoxins. Annu. Rev. Biochem 2002;71:635–700. [PubMed: 12045108]

(4). Raetz CRH, Reynolds CM, Trent MS, Bishop RE. Lipid A modification systems in gram-negative bacteria. Annu. Rev. Biochem 2007;76:295–329. [PubMed: 17362200]

(5). Poltorak A, He X, Smirnova I, Liu MY, Huffel CV, Du X, Birdwell D, Alejos E, Silva M, Galanos C, Freudenberg M, Ricciardi-Castagnoli P, Layton B, Beutler B. Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: mutations in Tlr4 gene. Science 1998;282:2085–2088. [PubMed: 9851930]

(6). Miller SI, Ernst RK, Bader MW. LPS, TLR4 and infectious disease diversity. Nat. Rev. Microbiol 2005;3:36–46. [PubMed: 15608698]

(7). Akira S, Uematsu S, Takeuchi O. Pathogen recognition and innate immunity. Cell 2006;124:783–801. [PubMed: 16497588]

(8). Kim HM, Park BS, Kim JI, Kim SE, Lee J, Oh SC, Enkhbayar P, Matsushima N, Lee H, Yoo OJ, Lee JO. Crystal structure of the TLR4-MD-2 complex with bound endotoxin antagonist Eritoran. Cell 2007;130:906–917. [PubMed: 17803912]

(9). Hawkins LD, Christ WJ, Rossignol DP. Inhibition of endotoxin response by synthetic TLR4 antagonists. Curr. Top. Med. Chem 2004;4:1147–1171. [PubMed: 15279606]

(10). Russell JA. Management of sepsis. N. Engl. J. Med 2006;355:1699–1713. [PubMed: 17050894]

(11). Loppnow H, Brade H, Dürrbaum I, Dinarello CA, Kusumoto S, Rietschel ET, Flad HD. IL-1 induction capacity of defined lipopolysaccharide partial structures. J. Immunol 1989;142:3229–3238. [PubMed: 2651523]

(12). Raetz CRH, Garrett TA, Reynolds CM, Shaw WA, Moore JD, Smith DC Jr. Ribeiro AA, Murphy RC, Ulevitch RJ, Fearns C, Reichart D, Glass CK, Benner C, Subramaniam S, Harkewicz R, Bowers-Gentry RC, Buczynski MW, Cooper JA, Deems RA, Dennis EA. $(Kdo)_2$-lipid A of *Escherichia coli*, a defined endotoxin that activates macrophages via TLR-4. J. Lipid Res 2006;47:1097–1111. [PubMed: 16479018]

(13). Vinogradov E, Perry MB, Conlan JW. Structural analysis of *Francisella tularensis* lipopolysaccharide. Eur. J. Biochem 2002;269:6112–6118. [PubMed: 12473106]

(14). Phillips NJ, Schilling B, McLendon MK, Apicella MA, Gibson BW. Novel modification of lipid A of *Francisella tularensis*. Infect. Immun 2004;72:5340–5348. [PubMed: 15322031]

(15). Wang X, Ribeiro AA, Guan Z, McGrath S, Cotter R, Raetz CRH. Structure and biosynthesis of free lipid A molecules that replace lipopolysaccharide in *Francisella tularensis* subsp. *novicida*. Biochemistry 2006;45:14427–14440. [PubMed: 17128982]

(16). Shaffer SA, Harvey MD, Goodlett DR, Ernst RK. Structural heterogeneity and environmentally regulated remodeling of Francisella tularensis subspecies novicida lipid A characterized by tandem mass spectrometry. J. Am. Soc. Mass Spectrom 2007;18:1080–1092. [PubMed: 17446084]

(17). Ancuta P, Pedron T, Girard R, Sandstrom G, Chaby R. Inability of the *Francisella tularensis* lipopolysaccharide to mimic or to antagonize the induction of cell activation by endotoxins. Infect. Immun 1996;64:2041–2046. [PubMed: 8675305]

(18). Hajjar AM, Harvey MD, Shaffer SA, Goodlett DR, Sjostedt A, Edebro H, Forsman M, Bystrom M, Pelletier M, Wilson CB, Miller SI, Skerrett SJ, Ernst RK. Lack of in vitro and in vivo recognition of *Francisella tularensis* subspecies lipopolysaccharide by Toll-like receptors. Infect. Immun 2006;74:6730–6738. [PubMed: 16982824]

(19). Wang X, Ribeiro AA, Guan Z, Abraham SN, Raetz CRH. Attenuated virulence of a *Francisella* mutant lacking the lipid A 4′-phosphatase. Proc. Natl. Acad. Sci. U S A 2007;104:4136–4141. [PubMed: 17360489]

(20). Wang X, McGrath SC, Cotter RJ, Raetz CRH. Expression cloning and periplasmic orientation of the *Francisella novicida* lipid A 4′-phosphatase LpxF. J. Biol. Chem 2006;281:9321–9330. [PubMed: 16467300]

(21). Kanistanon D, Hajjar AM, Pelletier MR, Gallagher LA, Kalhorn T, Shaffer SA, Goodlett DR, Rohmer L, Brittnacher MJ, Skerrett SJ, Ernst RK. A *Francisella* mutant in lipid A carbohydrate modification elicits protective immunity. PLoS Pathog 2008;4:e24. [PubMed: 18266468]

(22). Gunn JS, Lim KB, Krueger J, Kim K, Guo L, Hackett M, Miller SI. PmrA-PmrB-regulated genes necessary for 4-aminoarabinose lipid A modification and polymyxin resistance. Mol. Microbiol 1998;27:1171–1182. [PubMed: 9570402]

(23). Trent MS, Ribeiro AA, Lin S, Cotter RJ, Raetz CRH. An inner membrane enzyme in *Salmonella typhimurium* and *Escherichia coli* that transfers 4-amino-4-deoxy-L-arabinose to lipid A. Induction in polymyxin resistant mutants and role of a novel lipid-linked donor. J. Biol. Chem 2001;276:43122–43131. [PubMed: 11535604]

(24). Trent MS, Ribeiro AA, Doerrler WT, Lin S, Cotter RJ, Raetz CRH. Accumulation of a polyisoprene-linked amino sugar in polymyxin resistant mutants in *Salmonella typhimurium* and *Escherichia coli*. Structural characterization and possible transfer to lipid A in the periplasm. J. Biol. Chem 2001;276:43132–43144. [PubMed: 11535605]

(25). Breazeale SD, Ribeiro AA, McClerren AL, Raetz CRH. A formyltransferase required for polymyxin resistance in *Escherichia coli* and the modification of lipid A with 4-amino-4-deoxy-L-arabinose. Identification and function of UDP-4-deoxy-4-formamido-L-arabinose. J. Biol. Chem 2005;280:14154–14167. [PubMed: 15695810]

(26). Dulbecco R, Vogt M. Plaque formation and isolation of pure lines with poliomyelitis viruses. J. Exp. Med 1954;99:167–182. [PubMed: 13130792]

(27). Bligh EG, Dyer JJ. A rapid method of total lipid extraction and purification. Can. J. Biochem. Physiol 1959;37:911–917. [PubMed: 13671378]

(28). Zhou Z, Lin S, Cotter RJ, Raetz CRH. Lipid A modifications characteristic of *Salmonella typhimurium* are induced by NH_4VO_3 in *Escherichia coli* K12. Detection of 4-amino-4-deoxy-L-arabinose, phosphoethanolamine and palmitate. J. Biol. Chem 1999;274:18503–18514. [PubMed: 10373459]

(29). Kanjilal-Kolar S, Raetz CRH. Dodecaprenyl phosphate-galacturonic acid as a donor substrate for lipopolysaccharide core glycosylation in *Rhizobium leguminosarum*. J. Biol. Chem 2006;281:12879–12887. [PubMed: 16497671]

(30). Ribeiro AA, Zhou Z, Raetz CRH. Multi-dimensional NMR structural analyses of purified lipid X and lipid A (endotoxin). Magn. Res. Chem 1999;37:620–630.

(31). Zhou Z, Ribeiro AA, Raetz CRH. High-resolution NMR spectroscopy of lipid A molecules containing 4-amino-4-deoxy-L-arabinose and phosphoethanolamine substituents. Different attachment sites on lipid A molecules from NH4VO3-treated *Escherichia coli* versus kdsA mutants of *Salmonella typhimurium*. J. Biol. Chem 2000;275:13542–13551. [PubMed: 10788469]

(32). Wang RF, Kushner SR. Construction of versatile low-copy-number vectors for cloning, sequencing and gene expression in *Escherichia coli*. Gene 1991;100:195–199. [PubMed: 2055470]

(33). Garrett TA, Kadrmas JL, Raetz CRH. Identification of the gene encoding the *Escherichia coli* lipid A 4′-kinase. Facile synthesis of endotoxin analogs with recombinant LpxK. J. Biol. Chem 1997;272:21855–21864. [PubMed: 9268317]

(34). Basu SS, York JD, Raetz CRH. A phosphotransferase that generates phosphatidylinositol 4-phosphate (PtdIns-4-P) from phosphatidylinositol and lipid A in *Rhizobium leguminosarum*. A membrane-bound enzyme linking lipid A and PtdIns-4-P biosynthesis. J. Biol. Chem 1999;274:11139–11149. [PubMed: 10196199]

(35). Vorachek-Warren MK, Ramirez S, Cotter RJ, Raetz CRH. A triple mutant of *Escherichia coli* lacking secondary acyl chains on lipid A. J. Biol. Chem 2002;277:14194–14205. [PubMed: 11830595]

(36). Raetz CRH, Purcell S, Meyer MV, Qureshi N, Takayama K. Isolation and characterization of eight lipid A precursors from a 3-deoxy-D-*manno*-octulosonic acid-deficient mutant of *Salmonella typhimurium*. J. Biol. Chem 1985;260:16080–16088. [PubMed: 3905804]

(37). Hampton RY, Golenbock DT, Raetz CRH. Lipid A binding sites in membranes of macrophage tumor cells. J. Biol. Chem 1988;263:14802–14807. [PubMed: 3170565]

(38). Agrawal PK. NMR spectroscopy in the structural elucidation of oligosaccharides and glycosides. Phytochemistry 1992;31:3307–3330. [PubMed: 1368855]

(39). Zähringer, U.; Lindner, B.; Rietschel, ET. Chemical structure of lipid A: recent advances in structural analysis of biologically active molecules. In: Brade, H.; Opal, SM.; Vogel, SN.; Morrison, DC., editors. Endotoxin in Health and Disease. Marcel Dekker, Inc.; New York: 1999. p. 93-114.

(40). Takayama K, Qureshi N, Beutler B, Kirkland TN. Diphosphoryl lipid A from *Rhodopseudomonas sphaeroides* ATCC 17023 blocks induction of cachectin in macrophages by lipopolysaccharide. Infect. Immun 1989;57:1336–1338. [PubMed: 2784418]

(41). Golenbock DT, Hampton RY, Qureshi N, Takayama K, Raetz CRH. Lipid A-like molecules that antagonize the effects of endotoxins on human monocytes. J. Biol. Chem 1991;266:19490–19498. [PubMed: 1918061]

(42). Rietschel ET, Kirikae T, Schade FU, Mamat U, Schmidt G, Loppnow H, Ulmer AJ, Zähringer U, Seydel U, Di Padova F, Schreier M, Brade H. Bacterial endotoxin: molecular relationships of structure to activity and function. FASEB Journal 1994;8:217–225. [PubMed: 8119492]

(43). Persing DH, Coler RN, Lacy MJ, Johnson DA, Baldridge JR, Hershberg RM, Reed SG. Taking toll: lipid A mimetics as adjuvants and immunomodulators. Trends Microbiol 2002;10:S32–37. [PubMed: 12377566]

(44). Ohto U, Fukase K, Miyake K, Satow Y. Crystal structures of human MD-2 and its complex with antiendotoxic lipid IVa. Science 2007;316:1632–1634. [PubMed: 17569869]

(45). Gunn JS, Ryan SS, Van Velkinburgh JC, Ernst RK, Miller SI. Genetic and functional analysis of a PmrA-PmrB-regulated locus necessary for lipopolysaccharide modification, antimicrobial peptide resistance, and oral virulence of *Salmonella enterica* serovar *typhimurium*. Infect. Immun 2000;68:6139–6146. [PubMed: 11035717]

(46). Rohmer L, Fong C, Abmayr S, Wasnick M, Larson Freeman TJ, Radey M, Guina T, Svensson K, Hayden HS, Jacobs M, Gallagher LA, Manoil C, Ernst RK, Drees B, Buckley D, Haugen E, Bovee D, Zhou Y, Chang J, Levy R, Lim R, Gillett W, Guenthener D, Kang A, Shaffer SA, Taylor G, Chen J, Gallis B, D'Argenio DA, Forsman M, Olson MV, Goodlett DR, Kaul R, Miller SI, Brittnacher MJ. Comparison of *Francisella tularensis* genomes reveals evolutionary events associated with the emergence of human pathogenic strains. Genome Biol 2007;8:R102. [PubMed: 17550600]

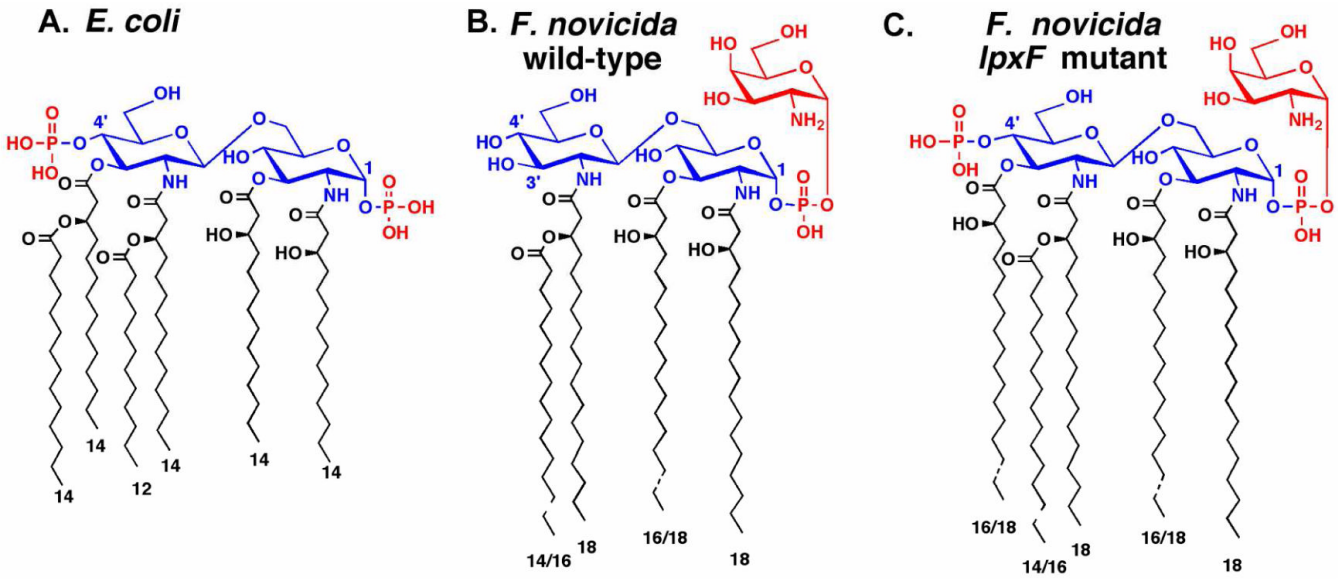


Figure 1.

Lipid A from *E. coli*, *F. novicida*, and the *F. novicida lpxF* mutant. **A.** The predominant lipid A moiety of wild-type *E. coli* LPS is a hexa-acylated disaccharide of glucosamine, substituted with phosphate groups at the 1- and 4′-positions (4). In polymyxin-resistant mutants of *E. coli*, an L-Ara4N residue is attached to the 4′-phosphate group (not shown) by ArnT, a glycosyltransferase that uses undecaprenyl phosphate-α-L-Ara4N (Fig. 2) as its donor substrate (4). **B**. The major lipid A species of *F. novicida* is a tetra-acylated disaccharide of glucosamine lacking the 3′-hydroxyacyl chain and 4′-phosphate group (14,15). An alpha-linked D-GalN residue, which appears to be unique to strains of *Francisella*, is attached to the 1-phosphate group (14,15). Over 90% of *Francisella* lipid A is termed “free” in the sense that it is not linked to a core oligosaccharide and O-antigen (15). However, some LPS is also synthesized, and its lipid A is modified with GalN in the same manner (14). **C**. Free lipid A of a *Francisella* mutant lacking LpxF, the 4′-phosphatase, retains both the 4′-phosphate group and the 3′-hydroxyacyl chain, consistent with an obligatory order of processing (19). The active site of the inner membrane protein LpxF faces the periplasm (20). The gene encoding the 3′-deacylase has not been identified, but it may be an outer membrane protein.

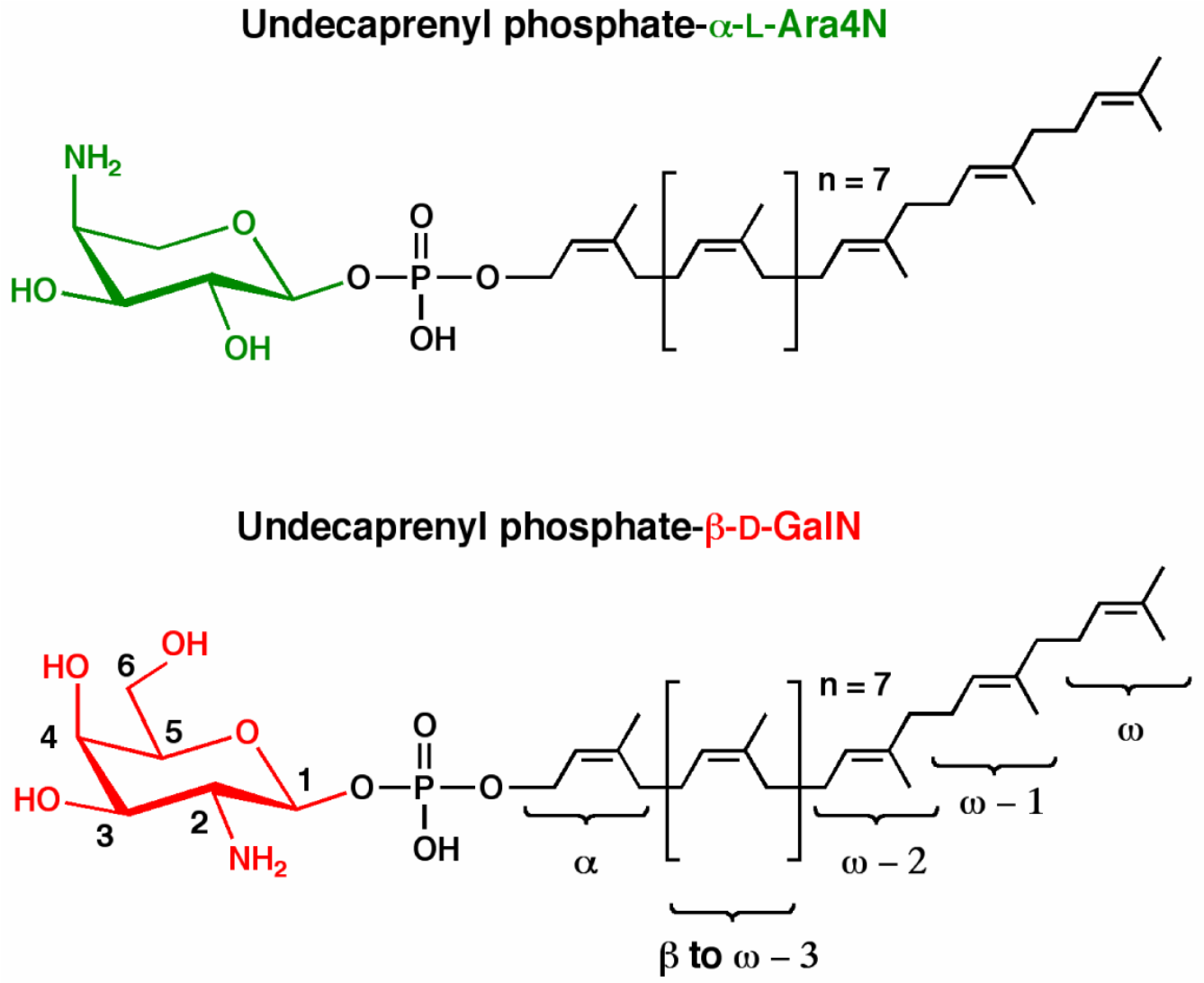


Figure 2.

Proposed structure of undecaprenyl phosphate-β-D-GalN. The numbering scheme is used for the NMR analysis presented in Figs. 5 to 7. The structure of undecaprenyl phosphate-α-L-Ara4N present in polymyxin-resistant strains of *E. coli* is shown for comparison (24).

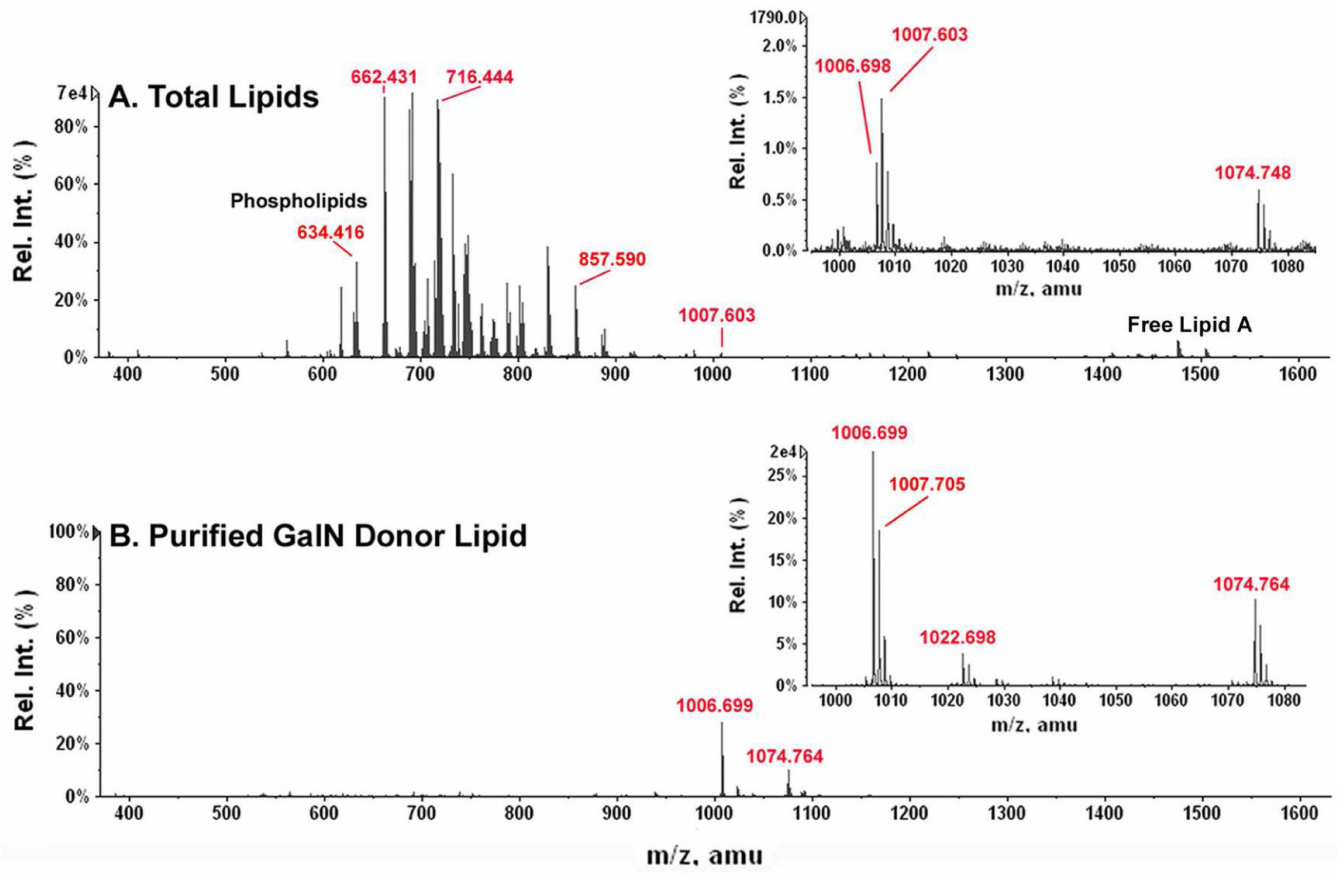


Figure 3.

Negative ion ESI/MS analysis of total *F. novicida* lipids and of the purified GalN donor. **A.** The spectrum of the total lipids from *F. novicida* strain XWK2 (15) reveals major peaks between *m/z* 500-900, arising mainly from glycerophospholipids. The peaks near *m/z* 1476 and 1504 are attributed to free lipid A; this strain lacks the usual GalN substituent because of an insertion mutation in the GalN transferase *flmK*(*arnT*) (15). However, the spectrum of wild-type *F. novicida* U112 lipids is virtually identical, except that the free lipid A is seen near *m/z* 1637 and 1655 (not shown) (15). ***Inset***. The $[M-H]^-$ ion of the proposed undecaprenyl phosphate-α-D-GalN (Fig. 2) donor substrate is predicted at *m/z* 1006.726, and a peak is observed at *m/z* 1006.698. However, an additional overlapping peak, probably arising from an undecaprenyl phosphate-glucose, is seen at *m/z* 1007.603, partially obscuring the undecaprenyl phosphate-α-D-GalN isotopic peak containing one ^{13}C atom. **B.** The spectrum of the purified GalN donor lipid shows a major peak, interpreted as $[M-H]^-$, at *m/z* 1006.699. ***Inset***. The overlapping putative undecaprenyl phosphate-hexose $[M-H]^-$ peak at *m/z* 1007.603 seen in panel A was removed by the purification; what remains at *m/z* 1007.705 can now be attributed to an undecaprenyl phosphate-α-D-GalN isotopic species containing one ^{13}C atom. The smaller peak at *m/z* 1074.764 is consistent with the $[M-H]^-$ of a dodecaprenyl phosphate-GalN derivative. The peak at *m/z* 1022.698 might arise by chemical oxidation of undecaprenyl phosphate-α-D-GalN generated during TLC.

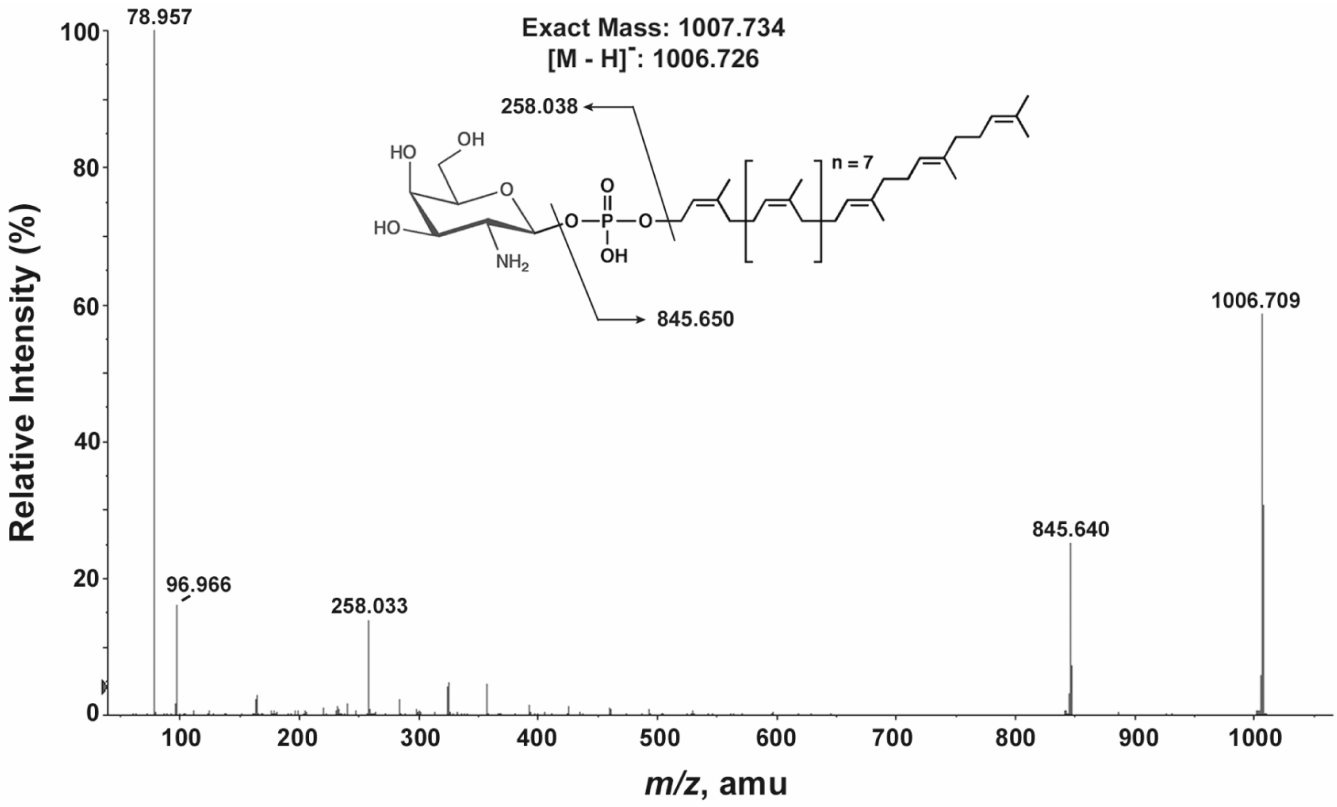


Figure 4.

ESI/MS/MS analysis of purified undecaprenyl phosphate-β-D-GalN. The MS/MS analysis of the dominant $[M-H]^-$ ion at *m/z* 1006.699 in the panel B of Fig. 3 is shown together with the proposed fragment ion assignments and the covalent structure of the intact lipid.

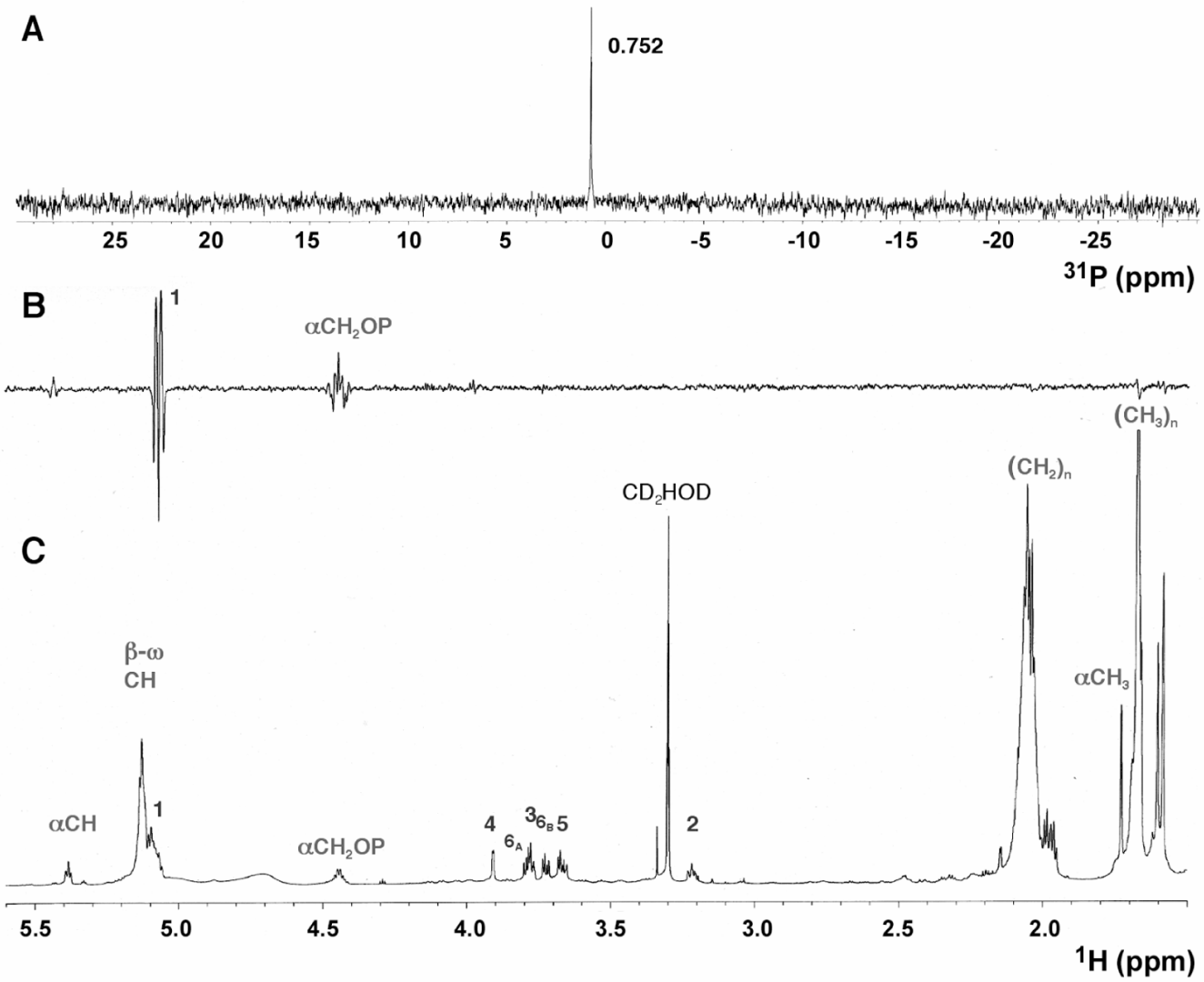


Figure 5.

^{31}P NMR spectrum and selective inverse ^{31}P-decoupled-^{1}H-detected difference spectroscopy of the purified lipid. **A.** The ^{31}P NMR spectrum of the purified lipid is consistent with the presence of one monophosphodiester (0.752 ppm). **B.** The selective inverse decoupled difference ^{1}H NMR spectrum obtained with on and off resonance ^{31}P decoupling of the 0.752 ppm phosphorus signal (30,31) showed that the anomeric carbon of GalN is linked via a phosphodiester bridge to the proximal isoprene unit of the undecaprenyl group. **C.** The one-dimensional 800-MHz ^{1}H NMR spectrum of the donor lipid shows six GalN and various isoprene methine, methylene and methyl proton signals (see Tables 1 and 2 for details). Note that an impurity peak overlaps H-5 near 3.66 ppm. The proposed structure of undecaprenyl phosphate-β-D-GalN is shown in Fig. 2 along with the numbering scheme.

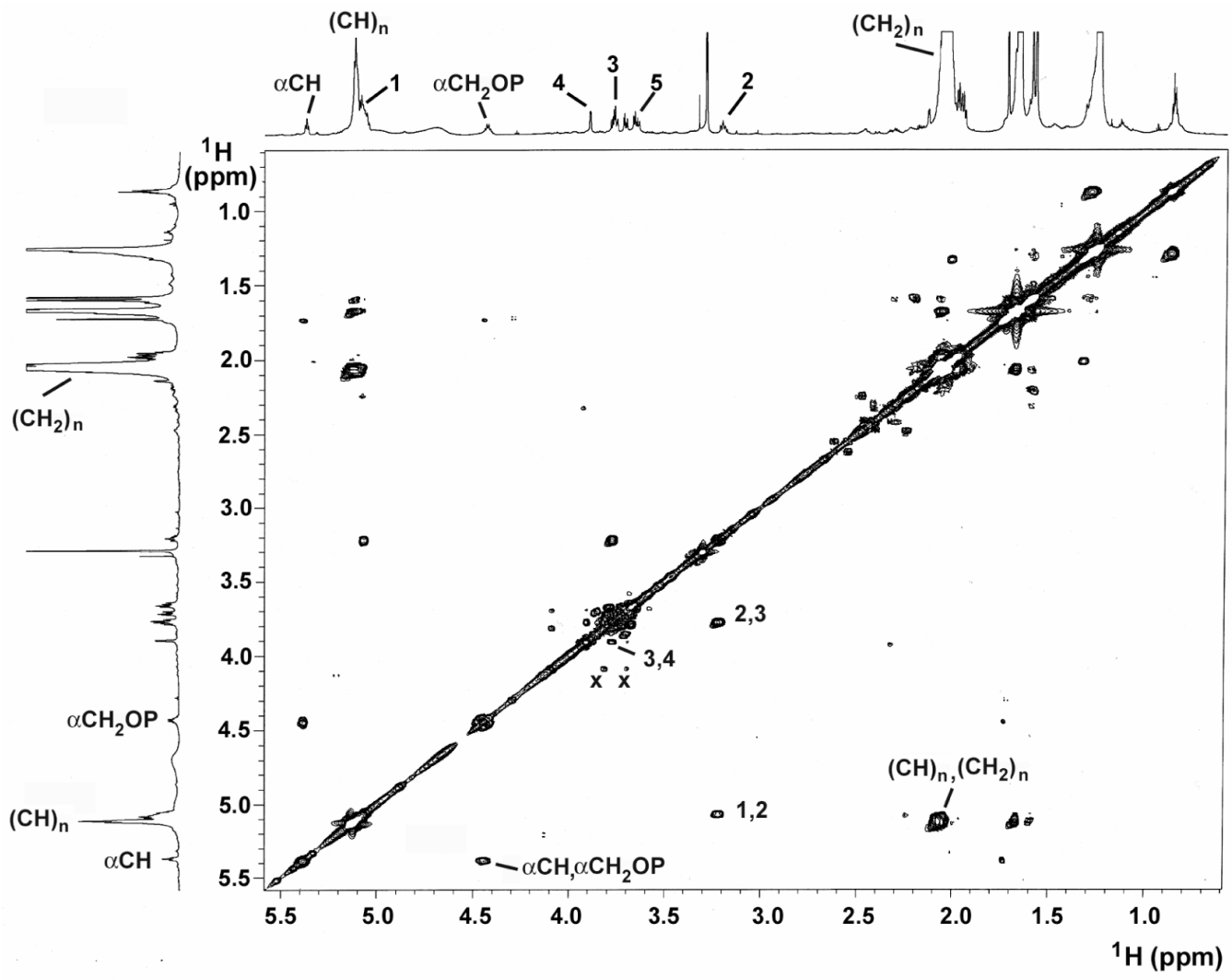


Figure 6.

Two-dimensional ^{1}H-^{1}H COSY analysis of the purified lipid at 800 MHz. The COSY experiment establishes the connectivities between the key GalN sugar and undecaprenyl ^{1}H resonances shown in Fig. 5C. Resonances designated X are due to impurities.

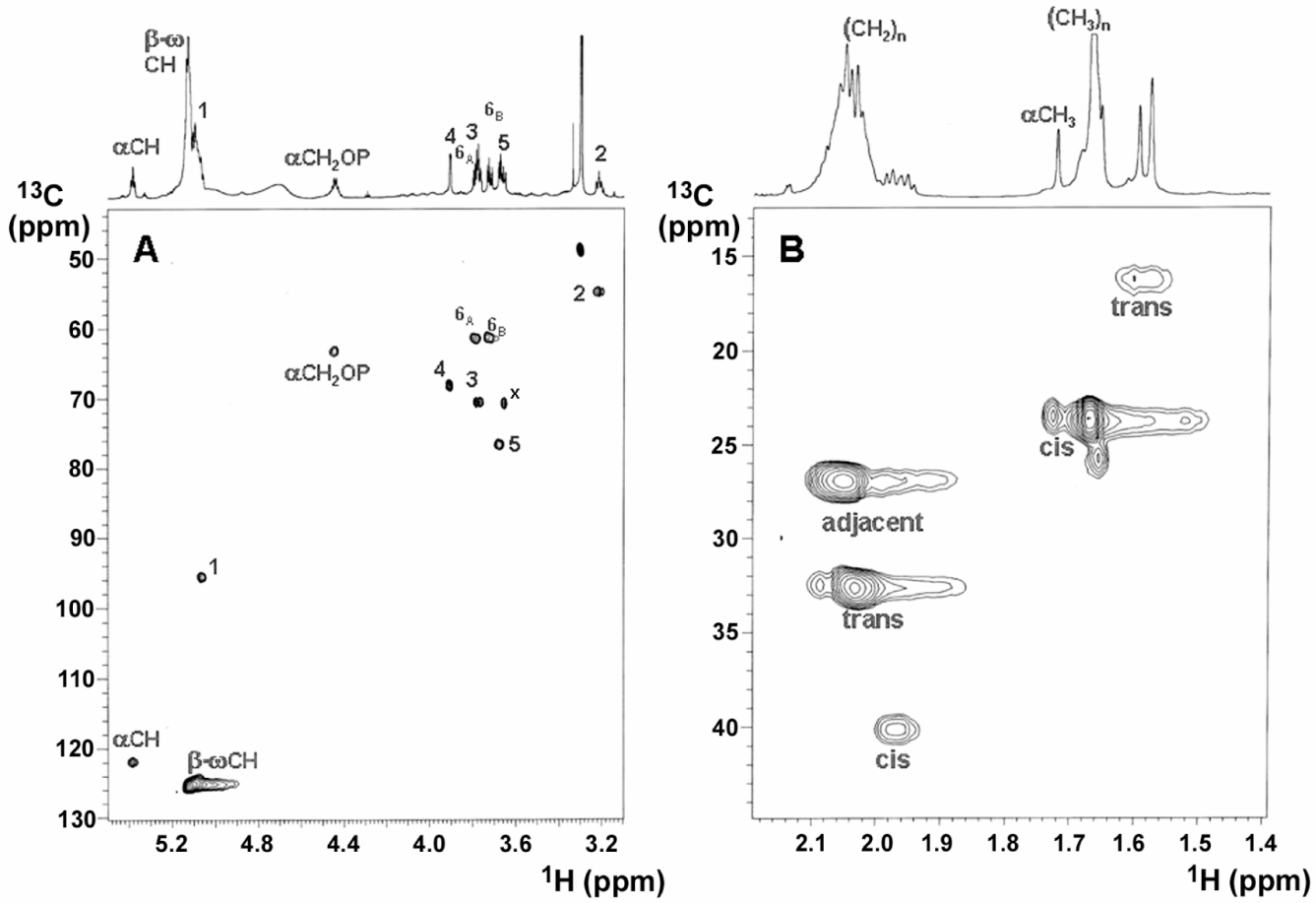


Figure 7.

Partial HMQC spectra of the purified lipid donor. **Panel A.** Single bond ^{1}H-^{13}C correlations of the GalN moiety, part of the proximal isoprene unit, and the unresolved olefinic protons of the undecaprenyl chain are shown. Resonances designated X are due to residual solvent or impurities. **B.** This expansion shows the partially resolved isoprenyl CH_2 and CH_3 groups and their directly bonded carbon atoms. The *cis*, *trans*, and *adjacent* configurations are defined in Table 2 and are the same as those used previously (24).

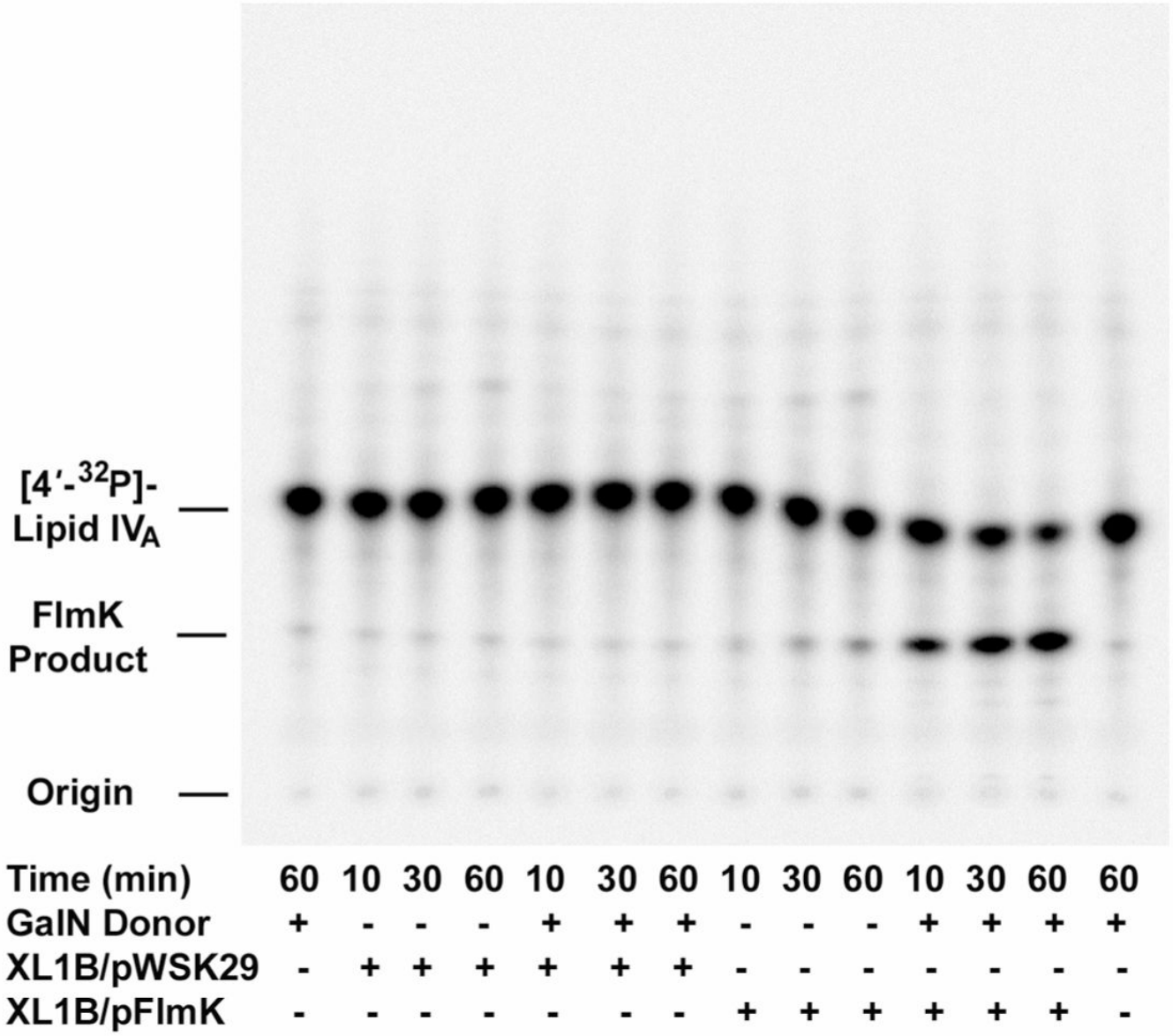
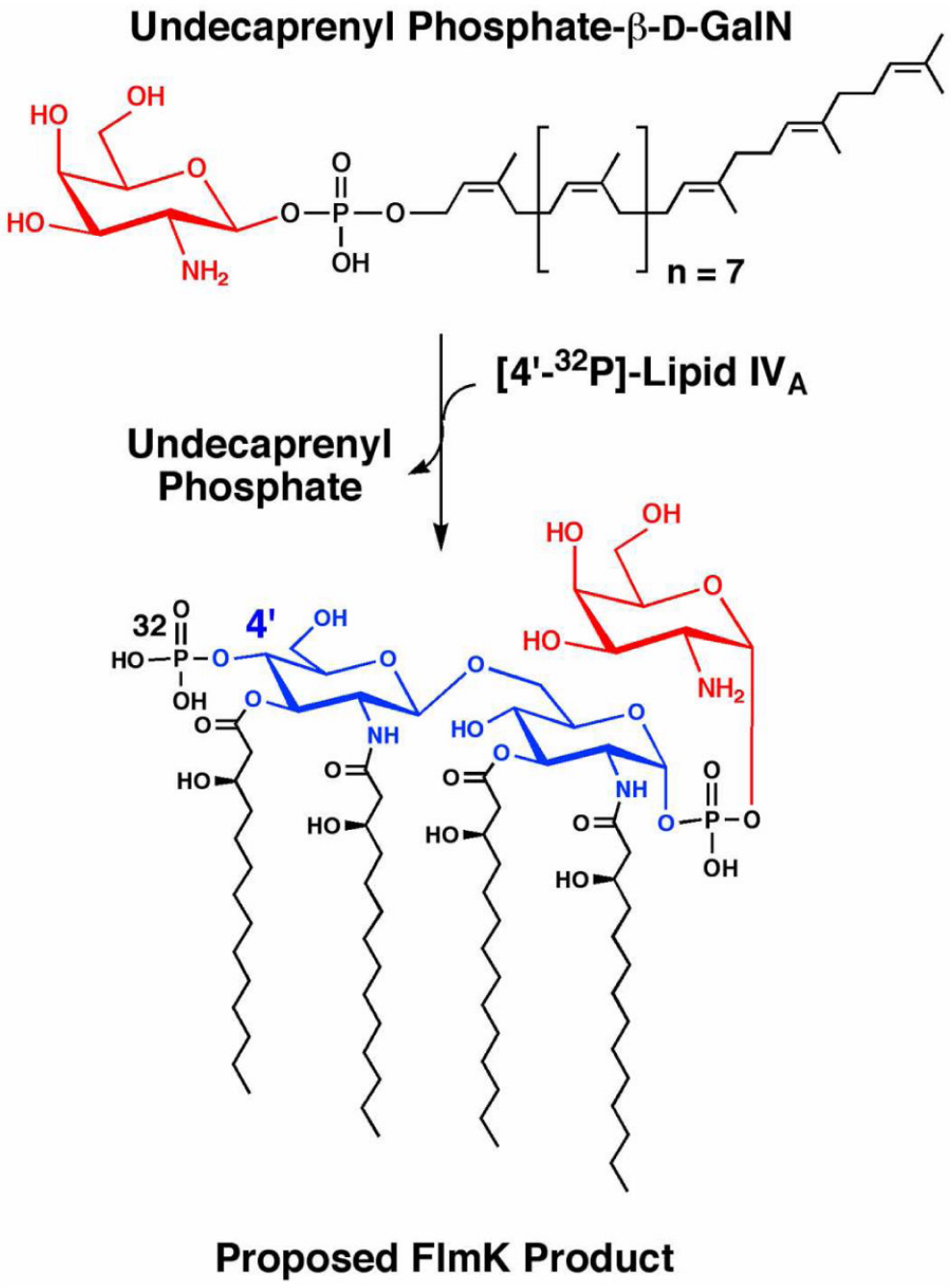


Figure 8.

In vitro assay of FlmK activity with purified undecaprenyl phosphate-β-D-GalN. FlmK was assayed using *E. coli* derived [4′-^{32}P]-lipid IV_A as the acceptor substrate (Scheme 1) under standard assay conditions for the indicated times. The purified undecaprenyl phosphate-β-D-GalN was added at 100 μM, as indicated. The migration of the FlmK product in this TLC system is consistent with the addition of an amino sugar to [4′-^{32}P]-lipid IV_A, as shown previously for L-Ara4N addition to lipid IV_A (23,24).



Scheme 1.
Reaction catalyzed by *F. novicida* FlmK with the model substrate lipid IV_A.

Table 1

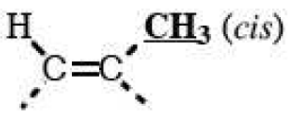
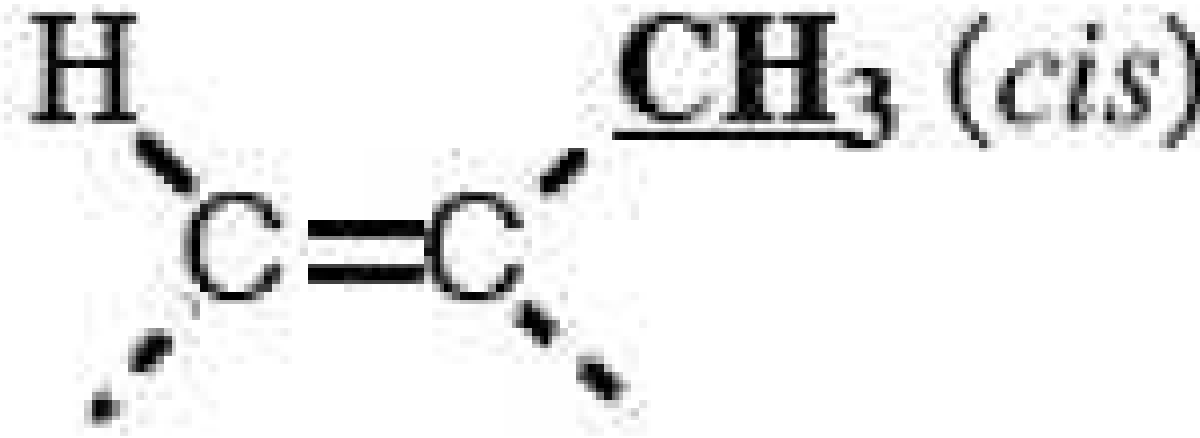
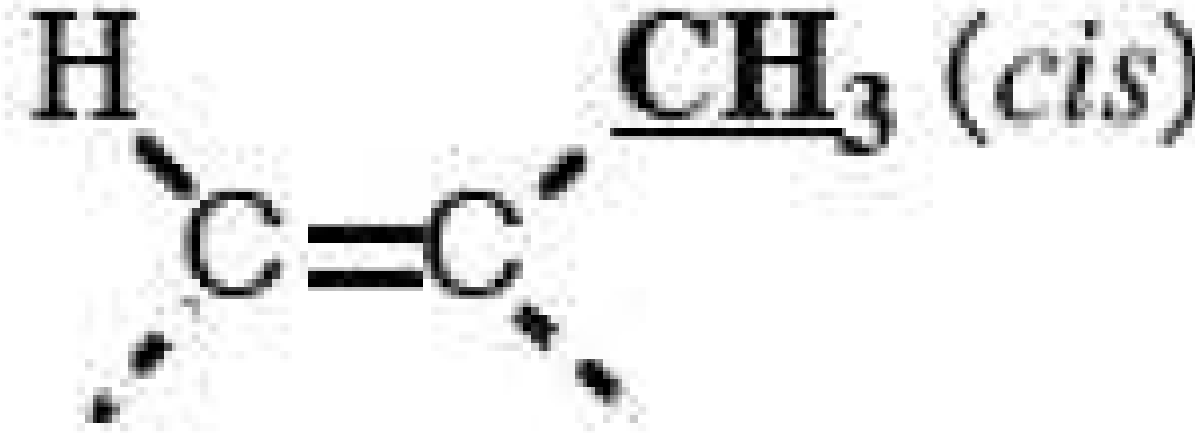
^{1}H and ^{13}C NMR assignments (ppm from TMS) and coupling constants (J, Hz) of the GalN residue in undecaprenyl phosphate-GalN

GalN	**H-1** ($J_{1,2}$)	**H-2** ($J_{2,3}$)	**H-3** ($J_{3,4}$)	**H-4** ($J_{4,5}$)	**H-5** ($J_{4,5}$)	**H-6a** ($J_{5,6a}$) ($J_{6a,6b}$)	**H-6b** ($J_{5,6a}$) ($J_{6a,6b}$)
δH	5.065 (8.2)	3.217	3.77 (3.1)	3.907	3.671	3.789 (6.7)	3.725 (5.3)
	C-1	C-2	C-3	C-4	C-5	C-6	
δC	96.12	55.26	70.46	68.66	77.08	61.98	

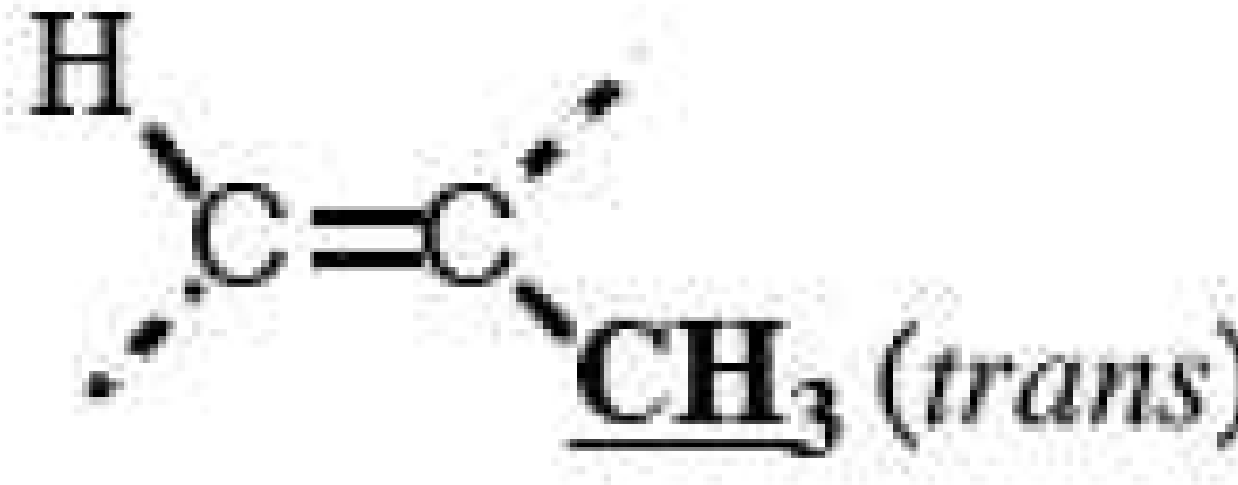
^{1}H chemical shifts (ppm from TMS) are from 1D ^{1}H spectra with a digital resolution of 0.2 Hz per point. Coupling constants ($J_{H,H}$, Hz) were measured from the ^{31}P-decoupled ^{1}H NMR spectra. ^{13}C chemical shifts (ppm from TMS) were estimated from the 2D HMQC spectra.

Table 2

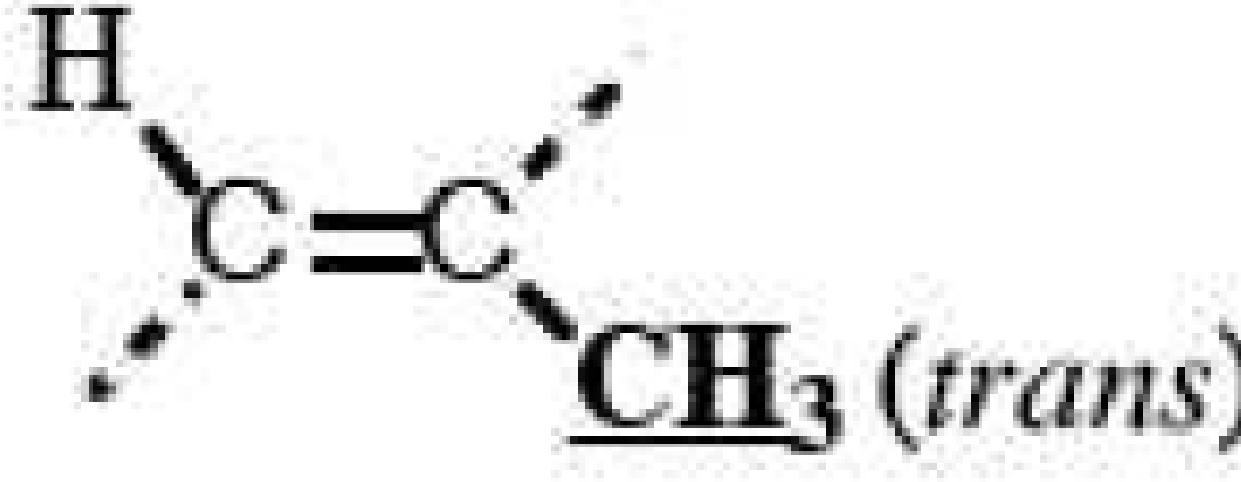
^{1}H and ^{13}C NMR assignments (ppm from TMS) the undecaprenyl chain in undecaprenyl-phosphate-GalN

Isoprene Unit	Methyl Groups
α	
β to ω-3	
ω	

ω-1, ω-2



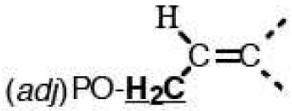
ω



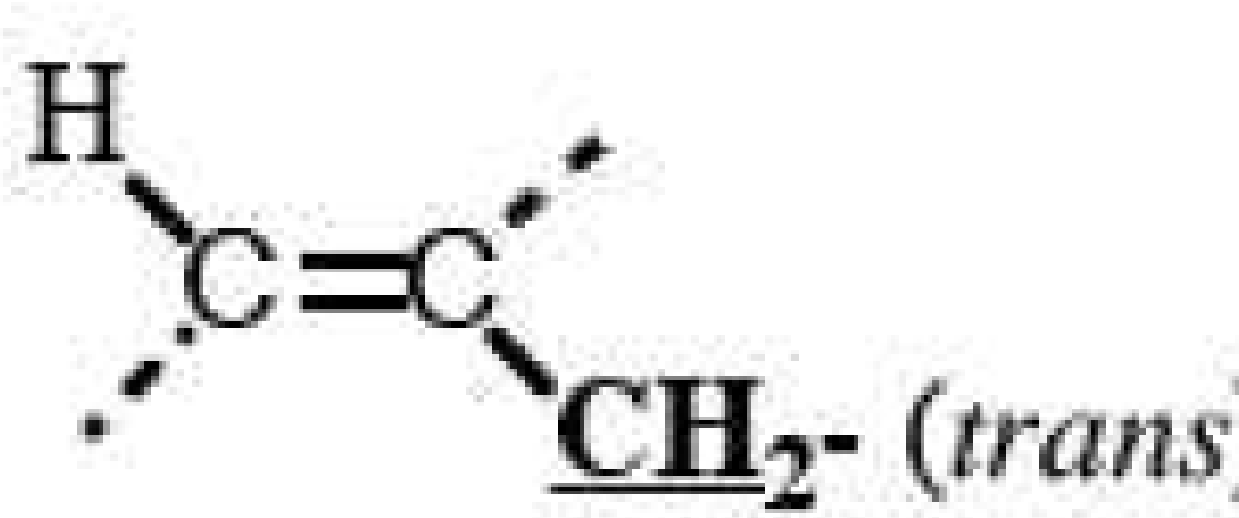
Isoprene Unit

Methylene Groups

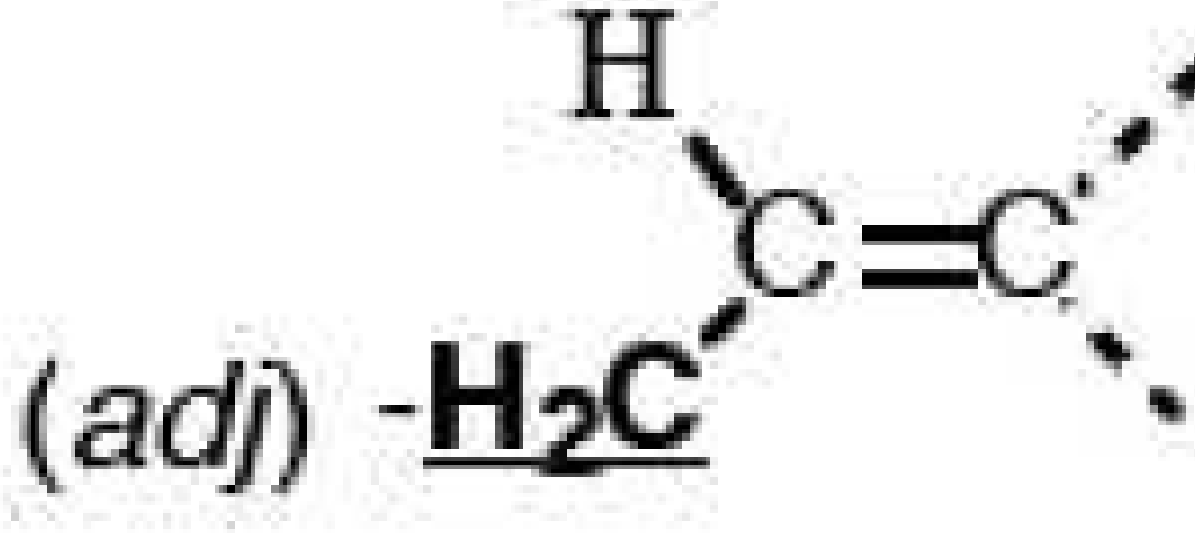
α



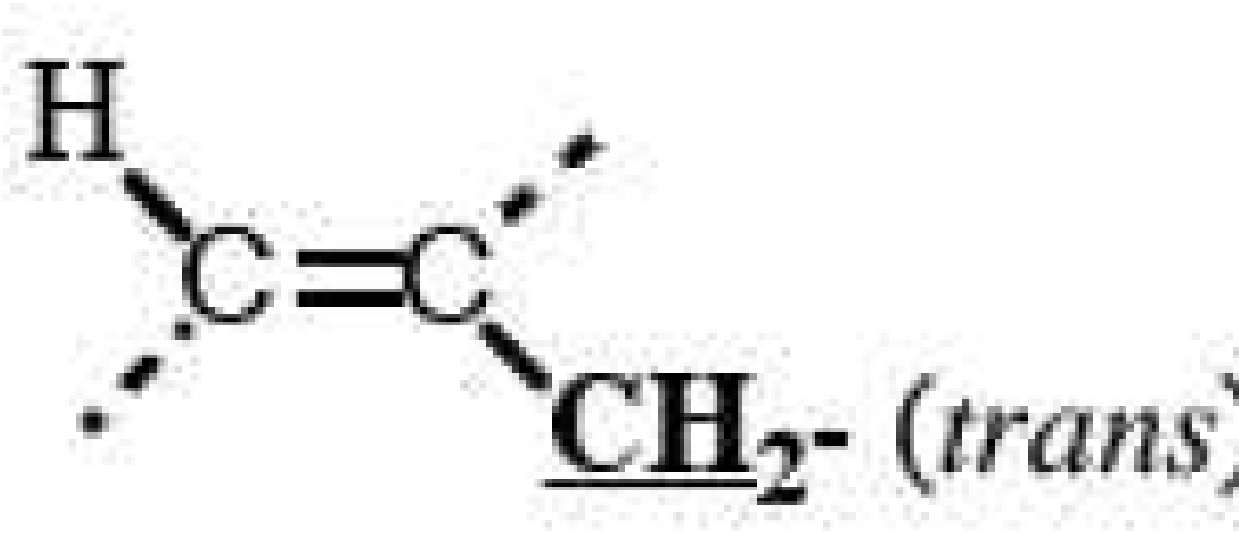
α



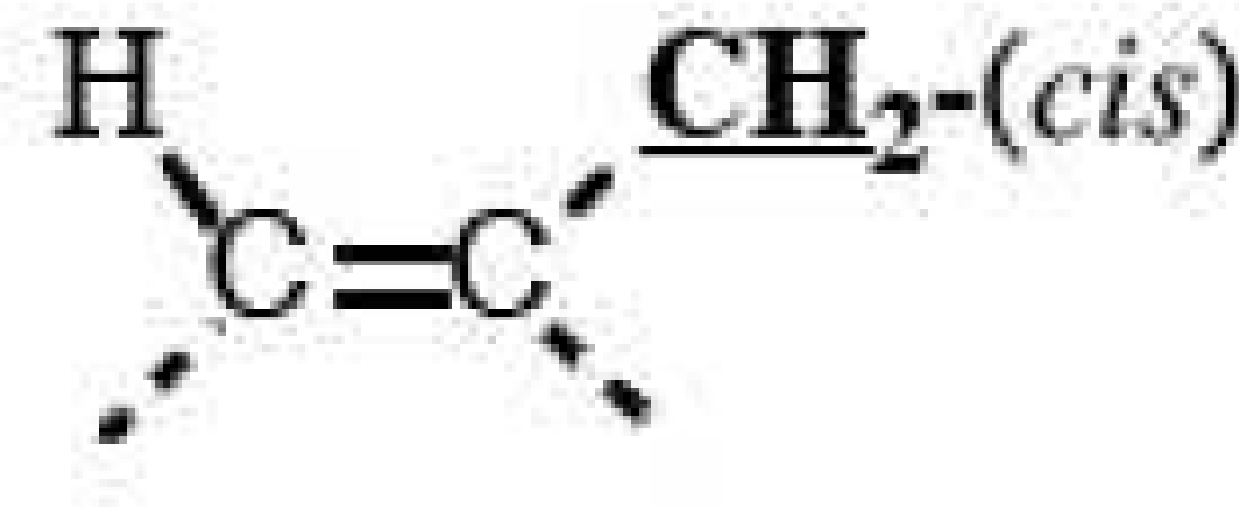
β to ω



β to ω-3



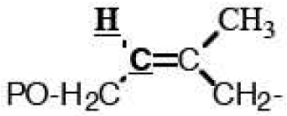
ω-1 to ω-2



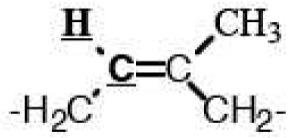
Isoprene Unit

Methine Groups

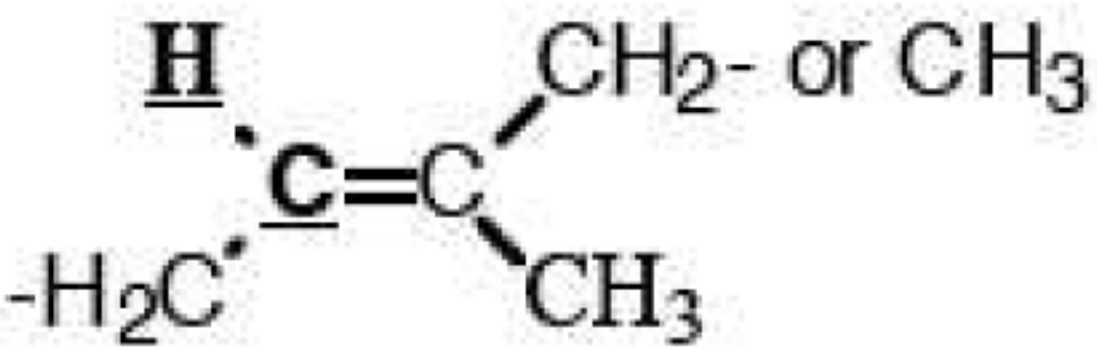
α



β to ω-3



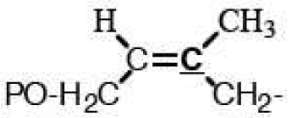
ω to ω-2



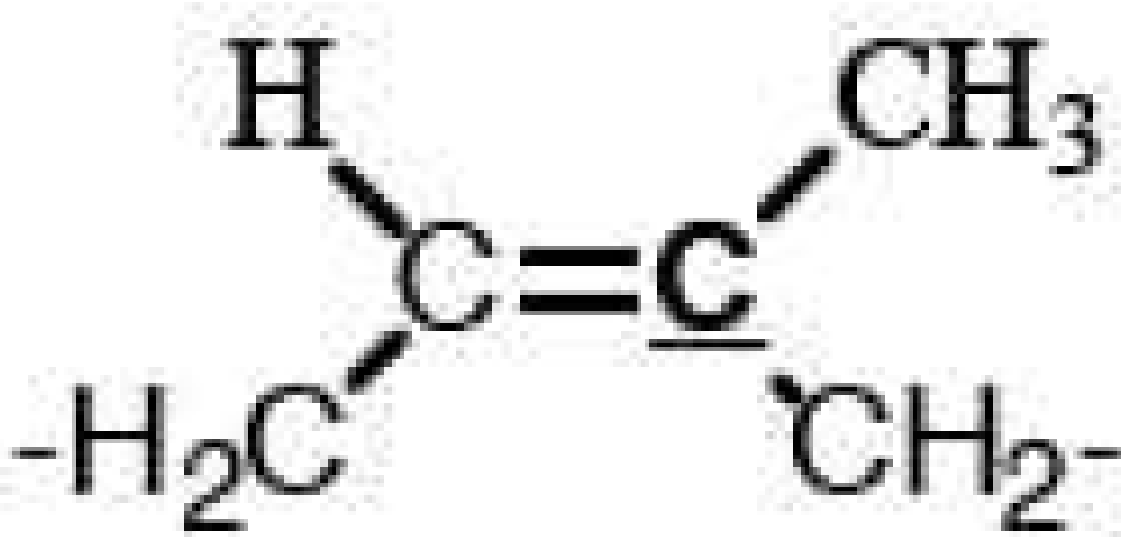
Isoprene Unit

Quaternary Carbons

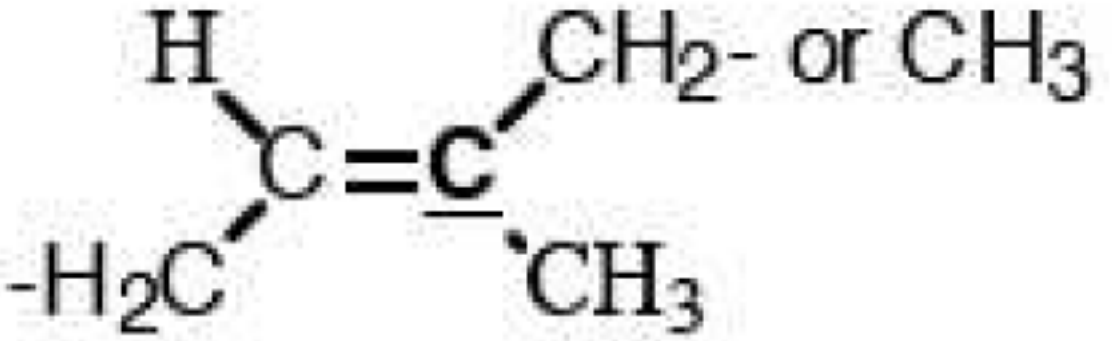
α



β to ω-3



ω to ω-2



Abbreviations: *adj*, adjacent in relation to a methine proton of an isoprene unit; likewise, *cis* and *trans* are used throughout to designate the configurations of various groups in relation to a methine proton of an isoprene unit, as shown above.