5,7-DIAMINO-3,5,7,9-TETRADEOXYNON-2-ULOSONIC ACIDS IN BACTERIAL GLYCOPOLYMERS: CHEMISTRY AND BIOCHEMISTRY

BY YURIY A. KNIREL,[a] ALEXANDER S. SHASHKOV,[a] YURY E. TSVETKOV,[a] PER-ERIK JANSSON,[b] AND ULRICH ZÄHRINGER[c]

[a]N. D. Zelinsky Institute of Organic Chemistry, Russian Academy of Sciences, Moscow, Russia; [b]Karolinska Institute, Clinical Research Center, Huddinge University Hospital, Huddinge, Sweden; and [c]Research Center Borstel, Center for Medicine and Biosciences, Borstel, Germany

I. INTRODUCTION

Ald-2-ulosonic acids are important components of natural glycoconjugates. Sialic acids, namely *N*- and *O*-acyl derivatives of 5-amino-3,5-dideoxy-D-*glycero*-D-*galacto*-non-2-ulosonic acid (neuraminic acid, Neu), generally occur in glycoconjugates of vertebrates and play a significant role in recognition, regulation, and protection.[1,2] A deamino analogue of neuraminic acid, 3-deoxy-D-*glycero*-D-*galacto*-non-2-ulosonic acid (Kdn),

has also been found in a variety of animal tissues.[3] 3-Deoxy-D-*manno*-oct-2-ulosonic acid (Kdo) is an essential component of lipopolysaccharides (LPSs) of Gram-negative bacteria that functions to link the carbohydrate portion to the lipid moiety.[4,5] In rare cases, Kdo in LPS is replaced with a 3-hydroxylated analogue, D-*glycero*-D-*talo*-oct-2-ulosonic acid.[5,6] Neu, Kdn, Kdo, hex-2-ulosonic, 3- and 4-deoxyhex-2-ulosonic, and 3-deoxyhept-2-ulosaric acids have been identified in several bacterial polysaccharides.[6–8]

A new class of ald-2-ulosonic acids, 5,7-diamino-3,5,7,9-tetradeoxynon-2-ulosonic acids, reported as "sialic acids of a new type," was discovered in 1984.[9] Members of this class contain an additional amino group at C-7, additional deoxygenation at C-9 (methyl group) and may exhibit configurational differences as compared with neuraminic acid. Unlike sialic acids, the new class of ald-2-ulosonic acids appears to be unique to microorganisms.

A short review article[10] in 1987 highlighted the occurrence and chemical properties of two isomers of 5,7-diamino-3,5,7,9-tetradeoxynon-2-ulosonic acids. Since that time, the configuration of one of these sugars was revised and new isomers have been discovered.[11,12] Sugars of this class have been reported as components of a variety of bacterial LPSs as well as several capsular polysaccharides (CPSs)[13,14] and glycoproteins.[15,16] Further, various isomers have been synthesized and NMR spectroscopic data accumulated for identification of such sugars.[11,17,18]

The present article focuses on the occurrence and characterization of derivatives of 5,7-diamino-3,5,7,9-tetradeoxynon-2-ulosonic acids and presents experimental approaches used to identify them and to elucidate the structures of the bacterial polysaccharides that contain the nonulosonic acids. Recent data on the biosynthesis of these sugars and their role in immune recognition and epitope specificity of bacterial glycopolymers are discussed.

II. NATURAL OCCURRENCE, BIOSYNTHESIS, AND BIOLOGICAL SIGNIFICANCE

1. Occurrence and Structural Features of the Natural Sugars

The first 5,7-diamino-3,5,7,9-tetradeoxynon-2-ulosonic acid was discovered simultaneously in the O-chain polysaccharides (OPSs) of LPSs of *Pseudomonas aeruginosa* O7 and O9[19] and *Shigella boydii* type 7 (for a review see Ref. 10). The nonulosonic acid was identified as the L-*glycero*-L-*manno* isomer (1) and called pseudaminic acid.[20–24] Later, pseudaminic acid was found in LPSs of several other Gram-negative bacteria[25–29]

TABLE I
Occurrence of Derivatives of 5,7-Diamino-3,5,7,9-tetradeoxynon-2-ulosonic Acids (1–4) in Bacterial Polysaccharides

N-Acyl Substituents[a]			
at N-5	at N-7	Bacterial Source[19]	References
L-*glycero*-L-*manno* isomer (pseudaminic acid, Pse, **1**)			
Ac	Ac	*Escherichia coli* O136	26
Ac	Ac	*Proteus vulgaris* O39	28
Ac	Ac	*Pseudoalteromonas atlantica* T9	29
Ac	Fo	*Pseudomonas aeruginosa* O7a,7b,7d	21,23
		and O7a,7d (immunotype 6)	
		Pseudoalteromonas distincta KMM 638	27
*R*3Hb	Fo	*Pseudomonas aeruginosa* O7a,7b,7c	23
Ac	*R*3Hb	*Pseudomonas aeruginosa* and O9a,9b	20,22
		Shigella boydii type 7	20,24
		Sinorhizobium fredii HH103	14
Ac	*S*3Hb	*Sinorhizobium fredii* HH103	14
Am	Ac	*Vibrio cholerae* O:2	25
D-*glycero*-D-*galacto* isomer (legionaminic acid, Leg, **2**)			
Ac	Ac	*Vibrio alginolyticus* 945-80	11,33
		Acinetobacter baumannii O24	11,35
*S*3Hb	Ac	*Acinetobacter baumannii* O24	11,35
Am	Ac	*Legionella pneumophila* serogroup 1	11,31
		Pseudomonas fluorescens ATCC 49271	11,32
		Vibrio salmonicida NCMB 2262	11,34
L-*glycero*-D-*galacto* isomer (8-epilegionaminic acid, 8eLeg, **3**)			
Ac	Ac	*Pseudomonas aeruginosa* O12	22
4Hb	Ac	*Yersinia ruckeri* O1	36
*R*3Hb	Ac	*Salmonella arizonae* O61	37
Am	Ac	*Morganella morganii* KF 1676 (RK 4222)	38
Ac	Am	*Shewanella putrefaciens* A6	39
D-*glycero*-D-*talo* isomer (4-epilegionaminic acid, 4eLeg, **4**)			
Ac	Ac	*Legionella pneumophila* serogroup 1	11,12
Am	Ac	*Legionella pneumophila* serogroup 2	40

[a]Am, acetimidoyl; Fo, formyl; *R*3Hb and *S*3Hb, (*R*)- and (*S*)-3-hydroxybutanoyl; 4Hb, 4-hydroxybutanoyl.

(Table I) and CPSs of *Sinorhizobium*.[13,14] Recently, pseudaminic acid or its enantiomer has been reported as a component of surface bacterial glycoproteins: pilin of *P. aeruginosa*[15] and flagellins of *Campylobacter jejuni* and *Campylobacter coli*.[16]

In 1987, a second isomer of this class was found in OPS of *P. aeruginosa* O12 and thought to have the D-*glycero*-L-*galacto* configuration.[30] The same, or mistakenly inferred as the same, sugar was reported to be present in a

number of other bacterial polysaccharides,[31–37] including LPS of *Legionella pneumophila* serogroup 1,[31] and was named legionaminic acid. In 1996, the configuration of legionaminic acid was revised to the opposite configuration.[32,34] Using synthetic models (Section III.1), it was finally determined that two isomers differing in the configuration at C-8 exist in different bacterial polysaccharides.[11] The D-*glycero*-D-*galacto* isomer (**2**) that has the same configuration as neuraminic acid kept the name legionaminic acid. It was present in *L. pneumophila*[31] and several other bacteria[32–35] (Table I). The other, L-*glycero*-D-*galacto* isomer (**3**) found in *P. aeruginosa* O12[30] and four more bacteria[36–39] (Table I) was called 8-epilegionaminic acid. The data that prompted the aforementioned revisions have been summarized.[11]

A fourth sugar was found only in LPS of *L. pneumophila*.[12,40] It was initially identified as the L-*glycero*-D-*talo* isomer[12] but later revised to the D-*glycero*-D-*talo* isomer and named 4-epilegionaminic acid (**4**).[11,40] Two more stereoisomers of the nonulosonic acids were found in several other *L. pneumophila* isolates[40] but their configurations have not yet been determined.

The naturally occurring isomers **1–4** have the same configuration at C-6, whereas the configurations of the other chiral centers may be different (Fig. 1). In aqueous solution, the free sugars exist in the $^2C_5$ pyranose form as mixtures of α and β anomers with a predominance of the thermodynamically more stable anomer having an equatorial carboxyl group. According to the Nomenclature of Carbohydrates,[41] the reference atom for the definition of the anomeric configuration is the highest-numbered atom of the group of chiral centers next to the anomeric center that is specified by a single configurational prefix. For 5,7-diamino-3,5,7,9-tetradeoxynon-2-ulosonic acids the reference atom is C-7, which has the L configuration in pseudaminic acid **1** and the D configuration in legionaminic acid **2** and the epimers **3** and **4**. Correspondingly, Fig. 1 shows the α anomer of **1** and β anomers of **2–4**. In natural glycopolymers, pseudaminic and legionaminic acids may be α-linked or β-linked, and the anomers with an axial carboxyl group occur more frequently; 4- and 8-epilegionaminic acids are known only as α-glycosides.

Only *N*-acylated 5,7-diamino-3,5,7,9-tetradeoxynon-2-ulosonic acids are found in Nature. Although *N*-acetyl groups predominate, other substituents, such as formyl, (*R*)- and (*S*)-3-hydroxybutanoyl, 4-hydroxybutanoyl, and acetimidoyl groups, are not uncommon. *N*-Acetyl and *N*-acetimidoyl groups are found in all known natural isomers, whereas other *N*-acyl groups are restricted to one or two isomers (Table I). Sometimes the nonulosonic acids are *O*-acetylated at position 4[22,24] or 8,[12,31,32,40] 8-*O*-acetylation being nonstoichiometric in *Pseudomonas fluorescens*[32] and some *L. pneumophila* serogroups.[40]
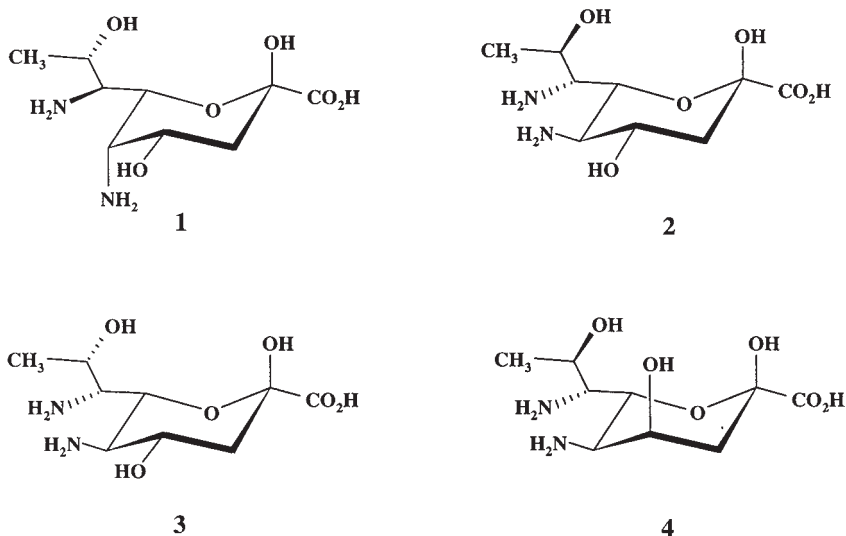
FIG. 1. Naturally occurring isomers of 5,7-diamino-3,5,7,9-tetradeoxynon-2-ulosonic acids **1**–**4**. [Shown are thermodynamically more stable anomers with an equatorial carboxyl group, which correspond to the α anomer of **1** (L-*glycero*-L-*manno* isomer, pseudaminic acid, Pse) and the β anomer of **2** (D-*glycero*-D-*galacto* isomer, legionaminic acid, Leg), **3** (L-*glycero*-D-*galacto* isomer, 8-epilegionaminic acid, 8eLeg), and **4** (D-*glycero*-D-*talo* isomer, 4-epilegionaminic acid, 4eLeg).]

2. Structures of Bacterial Glycopolymers

The known polysaccharides containing 5,7-diamino-3,5,7,9-tetradeoxynon-2-ulosonic acids are diverse. They vary from homopolymers[14,31,32,40] to heteropolysaccharides that contain pentasaccharide repeating units.[12,29] Homopolysaccharides are known for all naturally occurring isomers except 8-epilegionaminic acid. Pseudaminic, legionaminic, and 8-epilegonaminic acids, but not 4-epilegonaminic acid, have been reported as constituents of heteropolysaccharides.

a. Pseudaminic Acid.—The CPS of *Sinorhizobium fredii* HH103 (**5**) is a homopolymer of 5-*N*-acetyl-7-*N*-(3-hydroxybutanoyl)-α-pseudaminic acid, the 3-hydroxybutanoyl group being present as a 3:1 mixture of the *R* and *S* isomers[14] (Fig. 2). This polysaccharide is unusual because the sugar residues are linked through a glycosidic linkage to the hydroxyl group of the *N*-(3-hydroxybutanoyl) group, and thus represents a copolymer of pseudaminic acid and 3-hydroxybutanoic acid.

OPSs of *P. aeruginosa* O9 (**6**, **7**) include the same pseudaminic acid derivative but with the (*R*)-3-hydroxybutyryl group at N-5[22] (Fig. 2).
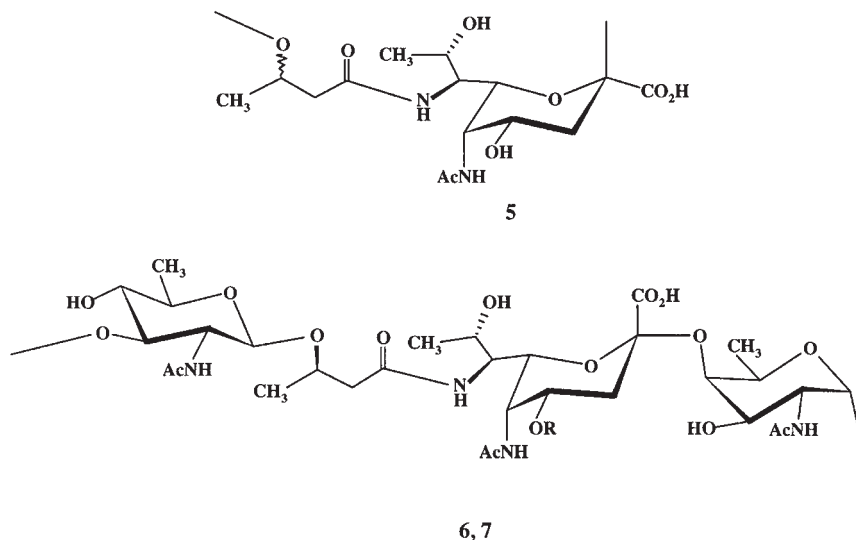
5

6, 7

FIG. 2. Structures of CPS of *Sinorhizobium fredii* HH103 (**5**)[14] and OPSs of *Pseudomonas aeruginosa* O9 (**6,7**)[22] [**6** R = H (subgroup O9a), **7** R = Ac (subgroup O9a,9b)].

The repeating units are connected by both glycosidic and amidic bonds through the *N*-(3-hydroxybutanoyl) group, namely by the same linkage as occurs in the polysaccharide of *S. fredii*.[14] Two other components of the trisaccharide repeating units are 2-acetamido-2,6-dideoxy-D-glucose (D-QuiNAc) and 2-acetamido-2,6-dideoxy-D-galactose (D-FucNAc). Position 4 of pseudaminic acid is *O*-acetylated in serogroup O9a,9b but not in serogroup O9a (Fig. 2).

OPSs of *P. aeruginosa* serogroup O7 (**8–10**) include a derivative with a formyl group at N-7 and either an acetyl or (*R*)-3-hydroxybutanoyl group at N-5.[23] This variability, along with the presence or absence of the *O*-acetyl group at O-4 of the neighboring D-FucNAc residue, is the basis for subdivision of the O7 serogroup into three subgroups (Fig. 3).

Pilin of *P. aeruginosa* strain 1244, a ∼ 16 kDa glycoprotein of the somatic pili, contains an O-linked pseudaminic acid-containing trisaccharide (**11**) attached to a serine residue in each pilin monomer (Fig. 3).[15] The trisaccharide has the same structure as the repeating unit of OPS of *P. aeruginosa* O7a,7b,7c except for the absence of the *O*-acetyl group in the D-FucNAc residue. Further studies should show if the OPS repeating unit of *P. aeruginosa* 1244 has the same structure as that of *P. aeruginosa* O7a,7b,7c and, thus, as the pilin trisaccharide.

Single O-linked residues of pseudaminic acid derivatives (or its enantiomer) modify flagellin, the major structural protein of the

→4)-α-Pse$p$7Fo-(2→4)-β-D-Xyl$p$-(1→3)-β-D-Fuc$p$NAc-(1→
    5   4
    | |
   R^1 R^2
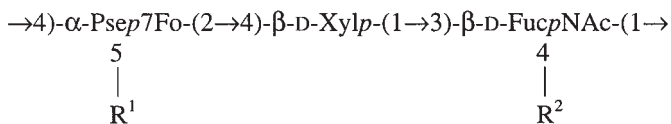
8 R^1 = (R)-3-hydroxybutanoyl, R^2 = Ac (subgroup O7a,7b,7c)

9 $R^1 = R^2$ = Ac (subgroup O7a,7b,7d)

10 R^1 = Ac, R^2 = H (subgroup O7a,7d or immunotype 6)

α-Pse$p$5(3Hb)7Fo-(2→4)-β-D-Xyl$p$-(1→3)-β-D-Fuc$p$NAc-(1→O)-Ser   **11**

FIG. 3. Structures of OPSs of *Pseudomonas aeruginosa* O7 (**8–10**)[23] and the serine-linked trisaccharide chain of pilin of *P. aeruginosa* 1244 (**11**).[15] (Pse, pseudaminic acid; FucNAc, 2-acetamido-2,6-dideoxygalactose.)

flagellar filament of *Campylobacter jejuni* and *Campylobacter coli*.[16] In the ∼ 65 kDa glycoprotein of *C. jejuni* 81-176, 19 of the total 107 Ser/Thr residues are glycosylated, and, hence, this is one of the most extensively modified prokaryotic proteins identified to date. Together with 5,7-di-*N*-acetylpseudaminic acid, 5,7-di-*N*-acetyl-8-*O*-acetyl, 5-*N*-acetimidoyl-7-*N*-acetyl, and 5,7-di-*N*-glyceroyl derivatives were tentatively identified. The anomeric configuration of the nonulosonic acid has not been assigned.

Structures of pseudaminic acid-containing OPSs from other bacteria (**12–16**) are shown in Fig. 4. OPSs of *Vibrio cholerae* O:2[25] (**13**) and *Escherichia coli* O136[26] (**14**) show marked structural similarities.

5,7-Di-*N*-acetylpseudaminic acid was also found in OPS of *Pseudoalteromonas atlantica* T9[29] and CPS of *Sinorhizobium* sp. NGR 234,[13] and 7-*N*-acetyl-5-*N*-(3-hydroxybutanoyl)pseudaminic acid in CPS of *Sinorhizobium meliloti* AK631.[13] The data reported on the monosaccharides in the *Sinorhizobium* CPS are limited, and the full structure of none of the three polysaccharides has been determined.

b. Legionaminic Acid.—OPSs of all *L. pneumophila* serogroups are α (2 → 4)-linked homopolymers of 5-acetimidoylamino-7-acetamido-3,5,7,9-tetradeoxynon-2-ulosonic acids[31,40] (structures **17** and **18** in Fig. 5). In serogroup 1, OPS is polylegionaminic acid, which is 8-*O*-acetylated in some strains[31] (structure **17**) and mostly non-*O*-acetylated in the others.[42,43] Accordingly, using mAbs that recognize 8-*O*-acetyllegionaminic acid (see Section II.4), serogroup 1 strains were divided into the Pontiac group and

Shigella boydii type 7[24]

→2)-β-D-Gal*f*-(1→3)-α-D-Glc*p*NAc-(1→8)-β-Pse*p*4Ac5Ac7(*R*3Hb)-(2→

6)-α-D-Gal*p*-(1→6)-α-D-Glc*p*-(1→   **12**

Vibrio cholerae O:2[25]

→4)-β-D-Qui*p*NAc-(1→4)-β-Pse*p*5Am7Ac-(2→4)-β-D-Gal*p*-(1→   **13**

Escherichia coli O136[26]

→4)-β-D-Glc*p*NAc-(1→4)-β-Pse*p*5Ac7Ac-(2→4)-β-D-Gal*p*-(1→   **14**

Proteus vulgaris O39[28]

→3)-α-D-Glc*p*NAc-(1→8)-β-Pse*p*5Ac7Ac-(2→3)-α-L-Fuc*p*NAc-(1→   **15**

Pseudoalteromonas distincta KMM 638[27]

→4)-α-Pse*p*5Ac7Fo-(2→4)-β-D-Qui*p*NAc-(1→   **16**
                                        3
                                        ↑
                                        1
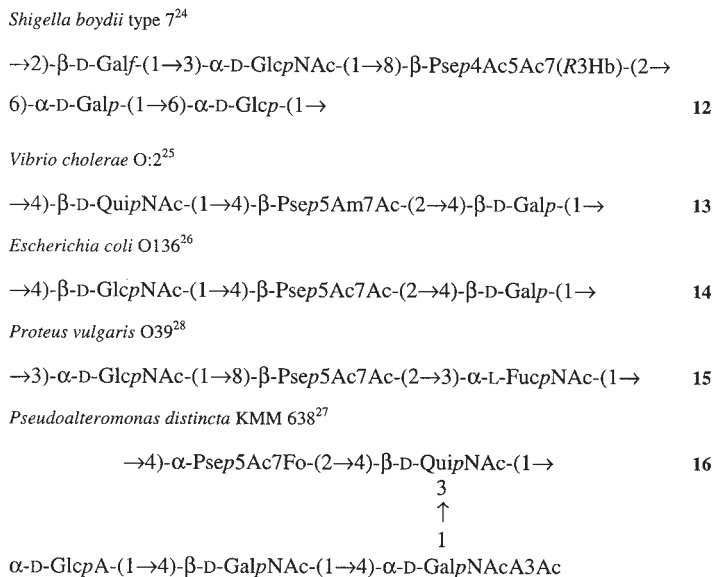α-D-Glc*p*A-(1→4)-β-D-Gal*p*NAc-(1→4)-α-D-Gal*p*NAcA3Ac

FIG. 4. Structures of OPSs containing pseudaminic acid (Pse). (QuiNAc, 2-acetamido-2,6-dideoxyglucose.)

the non-Pontiac group. The first three legionaminic acid residues next to the core in short-chain OPS (< 10 legionaminic acid residues) are 8-*O*-acetylated in both groups.[42]

Derivatives of legionaminic acid that are mono- or di-*N*-methylated at the 5-*N*-acetimidoyl group (**19** and **20**) are minor components of LPS of *L. pneumophila* serogroup 1[44] (Fig. 5). The monomethyl derivative **19** occurs as two stereoisomers, *E* and *Z*, whereas only one isomer was observed for the dimethyl derivative **20**. Molecular modeling data suggested that this is the sterically less-hindered *E* isomer.[44] The *N*-methylated derivatives are present in LPS from both Pontiac and non-Pontiac groups and can be 8-*O*-acetylated in the Pontiac group. A single residue of **19** and **20** is located exclusively in long-chain OPS (> 10 legionaminic acid residues), most likely, close to the LPS core. *N*-Methylation is rare in bacterial polysaccharides,[6] and no *N*-methylated acetimidoylamino (acetamidino) group has been found in other naturally occurring monosaccharides.

OPS of *P. fluorescens* has the same structure **17** as that of *L. pneumophila* serogroup 1 strain Philadelphia 1[11,31] but the degree of *O*-acetylation is lower (~ 75%).[32] *P. fluorescens* also produces an LPS with one monomer of the *O*-chain attached to the core oligosaccharide[32,45] (structure **21** in Fig. 6). 5-*N*-Acetimidoyl-7-*N*-acetyllegionaminic acid was also found in the
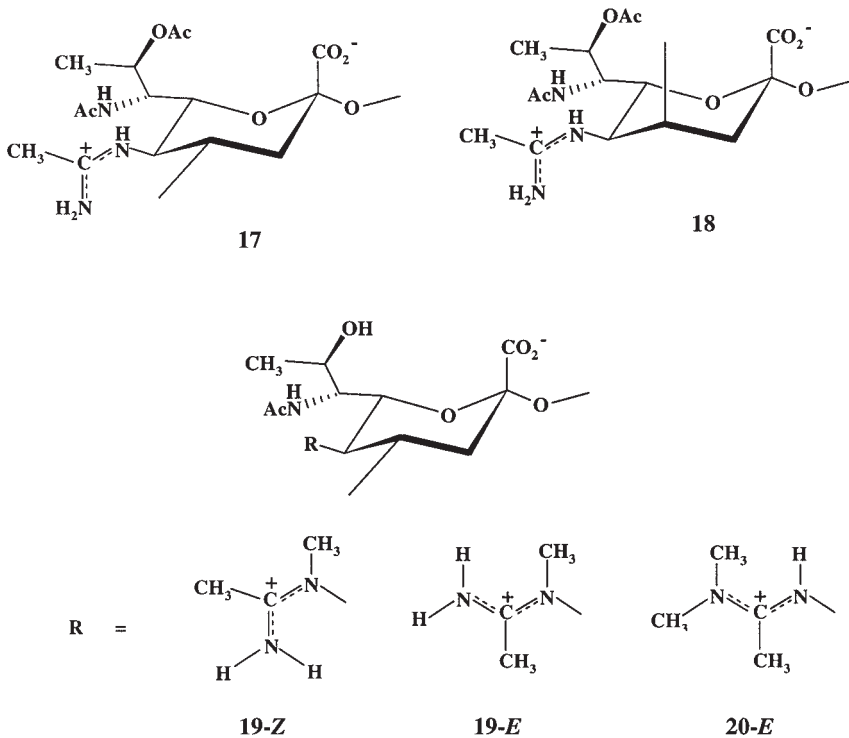
FIG. 5. Structures of OPSs of *Legionella pneumophila* serogroup 1 strain Philadelphia 1 (**17**)[11,31] and serogroup 2 (**18**)[40] and of the minor components of OPS of serogroup 1 (**19** and **20**).[44]
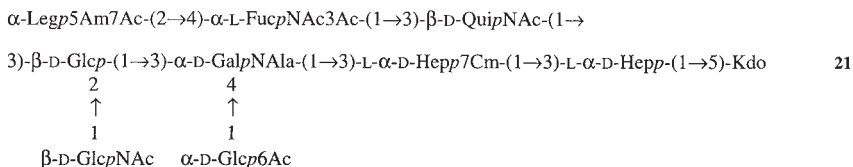
oligosaccharide portion of short-chain LPS produced by *Vibrio salmonicida* (structure **22** in Fig. 6).

OPS of *Acinetobacter baumannii* O24 (**23**) is a heteropolysaccharide that contains a 5-*N*-acetyl derivative of legionaminic acid in about half of its tetrasaccharide repeating units and a 5-*N*-[(*S*)-3-hydroxybutanoyl] derivative in the remaining units[35] (Fig. 6).

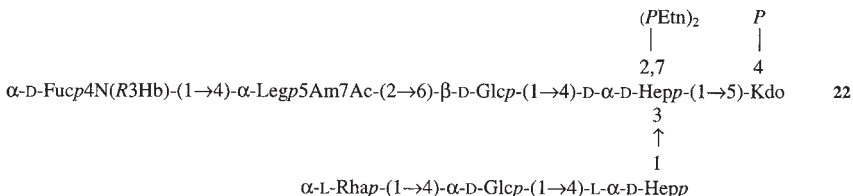5,7-Di-*N*-acetyllegionaminic acid was found in OPS of *Vibrio alginolyticus*,[11,33] whose structure remains unknown. Short-chain LPSs of *Vibrio parahaemolyticus* O2 and KX-V212 (O-untypeable strain) contain 5,7-di-*N*-acetyl and 5-*N*-acetyl-7-*N*-(*N*-acetyl-L-alanyl) derivatives of legionamnic acid, respectively.[46] The former was suggested to be involved in defining the serological specificity of the bacterium.[46]

c. 8-Epilegionaminic Acid.—Polysaccharides containing 8-epilegionaminic acid can be divided into two structural groups. One group that contains

Pseudomonas fluorescens ATCC 49271[11,32,45]

α-Leg*p*5Am7Ac-(2→4)-α-L-Fuc*p*NAc3Ac-(1→3)-β-D-Qui*p*NAc-(1→

3)-β-D-Glc*p*-(1→3)-α-D-Gal*p*NAla-(1→3)-L-α-D-Hep*p*7Cm-(1→3)-L-α-D-Hep*p*-(1→5)-Kdo   **21**
     2   4
     ↑   ↑
     1   1
  β-D-Glc*p*NAc   α-D-Glc*p*6Ac

Vibrio salmonicida NCMB 2262[11,34]

                                               (*P*Etn)$_2$ *P*
                                             | |
                                           2,7   4
α-D-Fuc*p*4N(*R*3Hb)-(1→4)-α-Leg*p*5Am7Ac-(2→6)-β-D-Glc*p*-(1→4)-D-α-D-Hep*p*-(1→5)-Kdo   **22**
                                                 3
                                             ↑
                                             1
               α-L-Rha*p*-(1→4)-α-D-Glc*p*-(1→4)-L-α-D-Hep*p*

Acinetobacter baumannii O24[11,35]

→6)-α-D-Glc*p*NAc-(1→3)-α-L-Fuc*p*NAc-(1→3)-α-D-Glc*p*NAc-(1→4)-β-Leg*p*7Ac-(2→   **23**
                                                    5
                                                  |
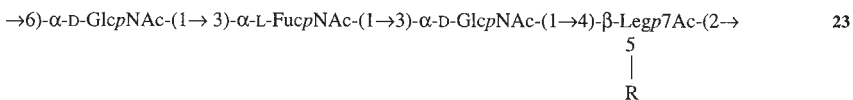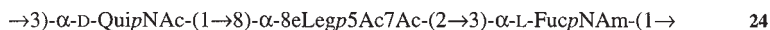   R

FIG. 6. Structures of the LPS-derived oligosaccharides and OPS containing legionaminic acid (Leg). [D-α-D-Hep and L-α-D-Hep, D-*glycero*- and L-*glycero*-α-D-*manno*-heptose; Kdo, 3-deoxy-D-*manno*-oct-2-ulosonic acid; Fuc4N*R*3Hb, 4-[(*R*)-3-hydroxybutanamido]-4,6-dideoxygalactose; Ala, L-alanyl; Cm, carbamoyl; *P*Etn, ethanolamine phosphate; R is acetyl in some repeating units and (*S*)-3-hydroxybutanoyl in the others.]
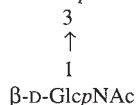
OPSs of *P. aeruginosa* O12[30] (**24**), *Yersinia ruckeri* O1[36] (**25**), and *Salmonella arizonae* O61[37] (**26**) is characterized by the presence of similar repeating trisaccharides in the main chain that contain an 8-substituted derivative of 8-epilegionaminic acid with different substituents at N-5, along with 3-substituted 2-acetimidoylamino-2,6-dideoxy-L-galactose (L-FucNAm) and 3-substituted D-GlcNAc or D-QuiNAc, all α-linked (Fig. 7). In the branched OPS of *Yersinia ruckeri* O1 (**25**), substitution with a lateral GlcNAc residue is nonstoichiometric.[36] These OPSs, especially that of *S. arizonae* O61[37] (**26**), show some structural similarity to OPS of *Proteus vulgaris* O39[30] (**15**) (Fig. 4).

    The second structural group is represented by OPSs from *Morganella morganii*[38] (**27, 28**) and *Shewanella putrefaciens*[39] (**29**), each having a disaccharide repeating unit that contains derivatives of 8-epilegionaminic acid and a unique branched sugar called shewanellose (Fig. 8). In OPSs of *M. morganii*, the branched sugar occurs either in the pyranose or furanose form.[38] The polysaccharides of the two bacteria differ in the linkage

Pseudomonas aeruginosa O12[30]

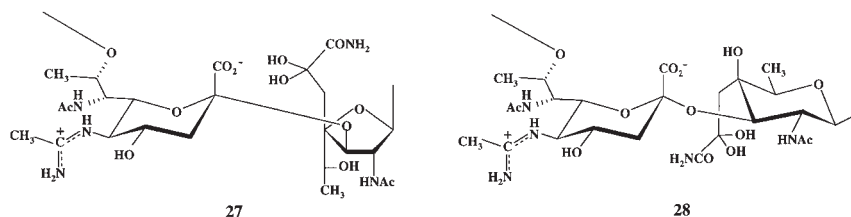→3)-α-D-QuipNAc-(1→8)-α-8eLegp5Ac7Ac-(2→3)-α-L-FucpNAm-(1→   **24**

Yersinia ruckeri O1[36]

→3)-α-D-GlcpNAc-(1→8)-α-8eLegp5(4Hb)7Ac-(2→3)-α-L-FucpNAm-(1→   **25**
       3
       ↑
       1
β-D-GlcpNAc

Salmonella arizonae O61[37]

→3)-α-D-GlcpNAc-(1→8)-α-8eLegp5(*R*3Hb)7Ac-(2→3)-α-L-FucpNAm-(1→   **26**

FIG. 7. Structures of OPSs containing 8-epilegionaminic acid (8eLeg).

Morganella morganii KF 1676[38]

27   **28**

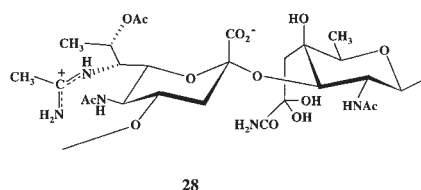*Shewanella putrefaciens* A6[39]

28

FIG. 8. Structures of OPSs containing 8-epilegionaminic acid and shewanellose.

position, the location of N-linked substituents (acetyl and acetimidoyl groups) and the presence or absence of an *O*-acetyl group in 8-epilegionaminic acid.

d. 4-Epilegionaminic Acid and Unidentified Isomers.—OPS of *L. pneumophila* serogroup 2 is an 8-*O*-acetylated homopolymer of 5-*N*-acetimidoyl-7-*N*-acetyl-4-epilegionaminic acid[40] (**18**) (Fig. 5). In most other serogroups except serogroups 7 and 13 and some strains of serogroup 5,

OPS is the same homopolysaccharide that is 8-*O*-acetylated to varying degrees (from < 10 to > 90%).[40] In OPSs of serogroups 5 and 13, 4-epilegionaminic acid is a minor constituent, whereas the major is an unidentified isomer. OPS of serogroup 7 is composed of yet another isomer, whose configuration also remains to be determined.

5,7-Di-*N*-acetyl-4-epilegionaminic acid is present in LPS of *L. pneumophila* serogroup 1[11,12] but the site of attachment of this sugar to LPS has not been determined.

3. Biosynthesis

In contradistinction to sialic acids,[1,2] little is known about biosynthesis of 5,7-diamino-3,5,7,9-tetradeoxynon-2-ulosonic acids. Some data are available on biosynthesis of derivatives of pseudaminic acid[16] and legionaminic acid[47] in human pathogens *C. jejuni* and *L. pneumophila*.

a. Pseudaminic Acid.—Genomic sequencing of *C. jejuni* NCTC 11168 revealed the presence of multiple alleles of genes encoding proteins predicted to be involved in Neu5Ac biosynthesis.[48] Whilst one set comprises genes of biosynthesis of Neu5Ac found in the LPS core of *C. jejuni*, the other two sets are involved in modifications of flagellin, a protein glyco-sylated with single residues of di-*N*-acetylpseudaminic acid and several other derivatives, including 5-*N*-acetimidoyl-7-*N*-acetylpseudaminic acid[16] (Section II.2). Mutations in genes termed *ptmB* and *neuB3* encoding a putative acylneuraminate cytidylyl transferase (CMP-Neu5Ac synthase) and a putative Neu5Ac synthase, respectively, affected flagellin (Ref. 16 and references cited therein). The *neuB2* and *neuB3* gene products showed homology to NeuB, *E. coli* K1 Neu5Ac synthase that catalyzes condensation of mannosamine and enolpyruvate phosphate to form sialic acid. Therefore, the enzyme encoded by either of the *neuB* genes might be involved in the condensation of a C_6 precursor (e.g., a 2,4-diacetamido-2,4,6-trideoxyhexose) and a C_3 precursor (an activated pyruvate form) to give Pse5Ac7Ac in a similar fashion.

In the genome region of *C. jejuni* 81-176, open reading frame (ORF) Cj1316c encoding a predicted protein of 43.7 kDa is located adjacent to ORF Cj1317 corresponding to the *neuB3* gene. The ORF Cj1316c gene termed *pseA* was suggested to be involved in biosynthesis of the acetamidino group on pseudaminic acid.[16] While insertional inactivation of *neuB3* caused loss of motility, mutation of *pseA* resulted in a motile phenotype. Flagellin from the *pseA* mutant no longer contained *N*-acetimidoyl groups, all Pse5Am7Ac residues present in the wild-type flagellin being replaced by Pse5Ac7Ac residues. The *pseA* gene product shares significant similarity
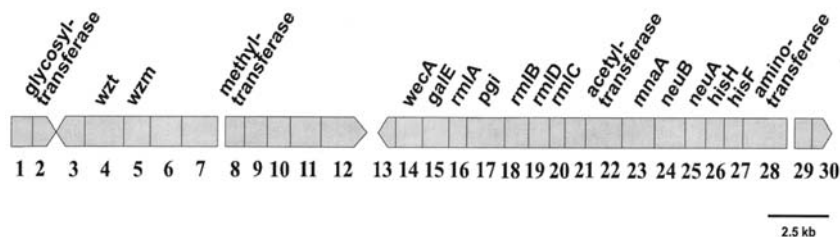
FIG. 9. Lipopolysaccharide biosynthesis locus of *Legionella pneumophila* serogroup 1 strain RC1 (OLDA) (taken from Ref. 46 and modified[16,44]).

to two proteins involved in LPS biosynthesis in *P. aeruginosa* IATS serotype O5[49] and *L. pneumophila*.[47] The *wbpG* gene of *P. aeruginosa* O5 is thought to encode an aminotransferase responsible for synthesis of the acetimidoyl group at N-3 of a 2,3-diamino-2,3-dideoxy-D-mannuronic acid residue in OPS.[49] The gene termed ORF 28 in *L. pneumophila* is part of an LPS locus involved in the synthesis of polylegionaminic acid[47] (see below), whose monomer carries an acetamidino group at C-5.[31] It remains unknown at which step in the biosynthesis of pseudaminic acid derivatives the substitution of the *N*-acetyl group with *N*-acetimidoyl group might occur.

b. Legionaminic Acid.—For studies of biosynthesis of legionaminic acid, wild-type strains of *L. pneumophila* serogroup 1 and isogenic LPS mutants were used,[47] all containing a homopolymer OPS composed of 5-*N*-acetimidoyl-7-*N*-acetyllegionaminic acid[31] (structure **17** in Fig. 5).

Complementation of a spontaneous LPS mutant (mutant 137) with a genomic library prepared from *L. pneumophila* serogroup 1 strain 5097 (OLDA) restored the wild-type LPS phenotype. A 32.6 kb DNA fragment that contained 30 ORFs (Fig. 9) was found to be responsible for the complementation. The locus includes ORFs with significant homology to genes encoding enzymes required for the biosynthesis of LPS or polysaccharide capsule of Gram-negative bacteria. The entire locus is present in *L. pneumophila* serogroup 1 strains whereas only its parts are present in strains of other serogroups, which contain as OPS homopolymers of 4-epilegionaminic acid or isomers not yet identified (Section II.2). In addition to genes likely involved in LPS core biosynthesis and LPS assembly, there were genes that encode putative enzymes of legionaminic acid biosynthesis (*mnaA*, *neuB*, and *neuA*).

The amino acid sequence deduced from ORF 23 (*mnaA*) shared homology with bacterial *N*-acetylglucosamine 2-epimerases that convert *N*-acetyl-glucosamine into *N*-acetylmannosamine, a precursor of sialic acids. ORF 9 and ORF 24 (*neuB*) both produce polypeptides that exhibited 25 and 31%

identity, respectively, to the SiaC protein of *Neisseria meningitidis*, an enzyme that mediates condensation of *N*-acetylmannosamine and enolpyruvate phosphate. The reason for the existence of two genes that encode SiaC homologues is unknown. The predicted polypeptide encoded by ORF 25 (*neuA*) revealed striking homology to other bacterial CMP-Neu5Ac synthases. This class of enzymes catalyzes the synthesis of CMP-*N*-acetylneuraminic acid, which is subsequently polymerized to form polysialic acid.

The *neuA* (ORF 25) and *neuB* (ORF 24) gene products were functionally characterized by complementation of *E. coli* K1 mutants EV5 and EV24 that are defective in K1 polysialic acid capsule biosynthesis. The EV5 mutant was complemented to wild-type capsule production by introduction of *neuA*, and the EV24 mutation complemented by the *neuB* gene.

Based on these findings, it is likely that ORFs 23, 24, and 25 encode enzymes that play a role in biosynthesis and polymerization of legionaminic acid and that the biosynthetic pathway is similar to that of bacterial sialic acid-containing polysaccharides.[1] ORFs that encode putative legionaminic acid transferase or polymerase were not found in the 32.6-kb locus.

The ORF 28 gene product showed similarity to PseA, a putative aminotransferase responsible for the synthesis of the *N*-acetimidoyl group on pseudaminic acid in *C. jejuni* flagellin.[16] ORF 28 might be thus involved in conversion of the 5-*N*-acetyl group on legionaminic acid into the 5-*N*-acetimidoyl group.

ORFs 8–12 comprise an operon that is present only in serogroup 1 strains of *L. pneumophila*.[47] The deduced amino acid sequence of ORF 8 revealed homology (45–52%) to bacterial methyltransferases. Expression of the ORF 8 gene in LPS mutant 137 restored binding of mAb 2625 (Section II.4), which is specific for *N*-methylated legionaminic acid residues **19** and **20** in long-chain OPS[44] (Section II.1, Fig. 5). Therefore, it was suggested that the ORF 8 gene encodes methyltransferase that is responsible for *N*-methylation of the acetamidino group at position 5 of legionaminic acid.

Interestingly, the LPS epitope recognized by mAb 2625[43] and *N*-methylation of legionaminic acid[44] is lost upon phase variation in *L. pneumophila* serogroup 1. Phase variation was found to be dependent upon chromosomal excision and replication of a high-copy plasmid.[50] DNA sequencing of the 30 kb episome did not reveal any genes with homology to LPS or another surface polysaccharide biosynthesis genes. Therefore, phase variation of serogroup 1 LPS must be mediated by an indirect mechanism, possibly by involvement of a regulatory factor whose normal function might be inhibited by a gene product encoded on the episome.

Using Tn5 mutagenesis, complementation analysis, and DNA-sequencing experiments, a gene designated *lag-1* (lipopolysaccharide-associated gene)

that is involved in 8-*O*-acetylation of legionaminic acid was identified in Pontiac-group strains of *L. pneumophila* serogroup 1.[51,52] *lag-1* mutants, CS332 and TF 3/1 from strains Philadelphia 1[51] and Corby,[52] respectively, lost the ability to produce 8-*O*-acetylated polylegionaminic acid and to bind mAb 2 and mAb 3/1 (see Section II.4). Complementation of both mutants with a wild-type *lag-1* gene restored 8-*O*-acetylation and mAb binding. Introduction and expression of this gene in the non-Pontiac-group OLDA strain that normally does not carry *lag-1* resulted in production of 8-*O*-acetylated polylegionaminic acid.[42]

The Philadelphia 1 *lag-1* gene encodes a 357-amino-acid protein that exhibited strong homology (54% identity) with Oac, a membrane-anchored *O*-acetyltransferase from bacteriophage SF6 of *Shigella flexneri*.[51] The *lag-1* locus from strain Corby exhibited 89.6% similarity at the nucleotide sequence level and 91.3% similarity in amino acid sequence to the Philadelphia 1 *lag-1* gene. The mutant *lag-1* gene present in the Corby TF 3/1 LPS mutant contained a single nucleotide change at position 169 that caused an amino acid change from serine to a leucine in a highly conserved motif present in many bacterial *O*-acetyltransferases. This change is apparently responsible for the loss of enzymatic activity of Lag-1 encoded by the mutant gene.[52]

Upstream and downstream genes adjacent to the Corby *lag-1* gene correspond to ORF 2 and ORF 3 on the 32.6 kb LPS biosynthesis locus described in strain OLDA[52] (Fig. 9). The *lag-1* gene is found at various chromosomal locations in different Pontiac-group isolates,[53] which suggested that *lag-1* may be contained in an unstable genetic element.

Interestingly, in short-chain OPS, the first three legionaminic acid residues located next to the LPS core are 8-*O*-acetylated in all *L. pneumophila* serogroup 1 strains, whether they contain *lag-1* or not.[42] This suggested that there must be another gene different from *lag-1* that encodes a second *O*-acetyltransferase with specificity for both the position of the legionaminic acid residue and OPS chain length. *lag-1*-Independent *O*-acetylation seems to prevent *N*-methylation of legionaminic acid in short-chain OPS (see above), whereas *lag-1*-dependent *O*-acetylation of long-chain OPS does not interfere with *N*-methylation. This is likely due to the notion that *lag-1*-dependent *O*-acetylation occurs after *N*-methylation. A likely candidate for the second *O*-acetyltransferase-encoding gene is ORF 22 found in the 32.6 kb locus (Fig. 9). A portion of ORF 22 that corresponded to amino acid positions 100–210 of the deduced ORF 22 gene product exhibited homology to cytoplasmic acetyltransferases but no homology to Lag-1.[47] However, it is not excluded that ORF 22 rather play a role in *O*-acetylation of LPS core monosaccharides.

4. Role in the Immunospecificity of Bacterial Antigens

LPSs carry the major immunogenic determinants on the cell surface of Gram-negative bacteria and, in many instances, serve as the basis for their serological classification. In this section, contribution of LPSs that contain 5,7-diamino-3,5,7,9-tetradeoxynon-2-ulosonic acids to the immuno-specificity and serological classification of two human pathogens, viz. L. pneumophila and P. aeruginosa, are reviewed.

a. Legionella pneumophila.—Derivatives of 5,7-diamino-3,5,7,9-tetra-deoxynon-2-ulosonic acid isomers that are present in OPSs of all L. pneumophila serogroups,[11,31,40] constitute multiple antigenic determinants and serve as the basis for serotyping. Strains of L. pneumophila are divided into at least 15 serogroups with rabbit polyclonal antisera[54] and 64 subgroups with mAbs.[55] Serogroups 1 and 7 exhibited no serological cross-reactivity with serogroup-specific mAbs.[55] This finding is consistent with the presence of legionaminic acid[11,31] and an unidentified isomer of legion-aminic acid[40] in OPSs of serogroups 1 and 7, and only in these OPSs, respectively. In contrast, all other serogroups exhibited cross-reactivity with various serogroup-specific mAbs.[55] This is likely due to the existence of common or cross-reactive epitopes formed by 4-epilegionaminic acid, which is present in OPSs of all serogroups except 1 and 7.[40] The chemical basis for classification of L. pneumophila strains to different cross-reactive (non-1, non-7) serogroups remains unknown.

Clinical isolates of L. pneumophila serogroup 1 can be subdivided into 15 mAb subgroups[55] but the epitope specificity of only a few of these mAbs has been determined. One of them is mAb 3/1 (or mAb 2), which recognizes a major epitope on polylegionaminic acid associated with the 8-O-acetyl group.[56] The reactivity is lost after chemical removal of the O-acetyl group[56] or as a result of a mutation in the lag-1 gene[51,52] that makes this mAb a useful tool for differentiation between functional lag-1-positive (Pontiac group) and -negative (non-Pontiac group) strains (see also Section II.3). The loss of the mAb 3/1 epitope in serogroup 1 OPS was accompanied by binding of some other LPS-specific mAbs that were unable to bind to OPS that contained the mAb 3/1 epitope.[52,56] This suggested that the 8-O-acetyl group that binds mAb 3/1 blocks the access to some other mAb epitopes present on polylegionaminic acid.

Binding of mAb 2625 to serogroup 1 isolates was found to correlate with N-methylation of the 5-acetimidoylamino group on a single legionaminic acid residue in OPS[44] (structures **19** and **20** in Fig. 5). Saturation-transfer-difference NMR spectroscopy showed that the binding of mAb 2625 is mainly mediated via the N-methylated 5-acetimidoylamino group in both

19 and **20** and via the closely located 7-acetamido group of the same legionaminic acid residue, thus confirming the *N*-methylated derivatives of legionaminic acid to represent the major epitope of mAb 2625.[57] In Pontiac group strains, this epitope is masked by the presence of the 8-*O*-acetyl group of legionaminic acid but becomes accessible to the antibody binding after chemical de-*O*-acylation of LPS.[44] The mAb 2625 epitope is lost upon inactivation of ORF 8 that encodes a putative methyltransferase[47] (see Section II.3). Binding of mAb 2625 interferes with binding of mAb LPS-1, an antibody that recognizes a highly *O*-acetylated outer-core region of serogroup 1 LPS. This suggested that the mAb 2625 epitope is located close to the core region.[44]

The 8-*O*-acetyl-associated epitope of legionaminic acid was also found in *P. fluorescens* ATCC 49271, which has the same OPS structure **17** as *L. pneumophila* serogroup 1[32,56] (Fig. 5). This epitope, possibly along with other, yet uncharacterized cross-reactive epitopes, is likely responsible for the serological cross-reactivity exhibited by *P. fluorescens* and distantly related *L. pneumophila*.

b. *Pseudomonas aeruginosa*.—The presence or absence of the *O*-acetyl group at position 4 of pseudaminic acid is responsible for the subdivision of *P. aeruginosa* serogroup O9 into two subgroups, O9a and O9a,9b[22] (Section II.2, Fig. 2). Epitope O9b in the latter is evidently associated with the 4-*O*-acetyl group.

A role of the 5-*N*-acyl group of pseudaminic acid in serospecificity was demonstrated in studies of *P. aeruginosa* serogroup O7 strains, which could be divided into three subgroups: O7a,7b,7c, O7a,7b,7d, and O7a,7d (Section II.2, Fig. 3). It was concluded that epitopes O7c and O7d are linked to the 5-*N*-[(*R*)-3-hydroxybutanoyl] and 5-*N*-acetyl group, respectively.[23]

A remarkable antigenic similarity between LPS and pilin of *P. aeruginosa* 1244 was recognized using a polyclonal typing serum to LPS of *P. aeruginosa* O7 and mAb 11.14.[16] This was substantiated by modification of pilin with an O-linked pseudaminic acid-containing trisaccharide (**11**), which is identical or closely related to the OPS repeating unit (Section II.2, Fig. 3). It was demonstrated that the mAb 11.14-reactive epitope is present on the pilus surface under physiological conditions,[16] and, therefore, like the O-chain polysaccharide of LPS, the oligosaccharide chain of pilin may contribute to the O7 immunospecificity. It is likely that both antigens have a common biosynthetic origin, and thus the pseudaminic acid derivative in OPS of *P. aeruginosa* O7 occupies the terminal nonreducing position most accessible to specific antibodies.

A cross-reactivity was reported between *P. aeruginosa* serogroups having LPSs that contain 5,7-diamino-3,5,7,9-tetradeoxynon-2-ulosonic acids and some enterobacteria.[58,59] A close structural similarity of OPSs of *P. aeruginosa* O12 and *Salmonella arizonae* O61,[58] which both contain 8-epilegionaminic acid (**24** and **26**), is evidently responsible for the serological relatedness of these bacteria.[37] In contrast, OPSs of serologically related[59] bacteria *P. aeruginosa* O9a and *Shigella boydii* type 7 are different (compare structures **6**[22] and **12**,[24] Section II.2) and have a β-linked pseudaminic acid derivative as the only monosaccharide in common. Most likely, in this case the cross-reactivity is due to the presence of the common sugar at the terminal nonreducing end of the OPS chain in both bacteria.

5. Possible Biological Significance

a. Polylegionaminic Acid.—*Legionella pneumophila* is an intracellular pathogen that invades and multiplies in mononuclear phagocytes of mammalian hosts. Legionaminic acid in OPS of *L. pneumophila* serogroup 1 has the same configuration and the same $^2C_5$ ring conformation as neuraminic acid, a common constituent of mammalian host-cell-surface glycoconjugates. Therefore, it is possible that the structural similarities between polylegionaminic acid present on the *L. pneumophila* cell surface and sialic acid-containing glycoconjugates on host cells enable legionellae to escape the immune response of an infected host. On the other hand, the similarity between legionaminic and neuraminic acids might make *L. pneumophila* the target of eukaryotic sialidases. It is possible that the *N*-acetyl, *N*-acetimidoyl, and *O*-acetyl groups enhance the resistance of polylegionaminic acid to eukaryotic sialidases and aid the intracellular persistence of the microorganism in infected host cells.[60]

Due to the negative charge of the carboxyl groups and positive charge of the *N*-acetimidoyl groups, polylegionaminic acid may be involved in binding of ions and regulation of the outer-membrane permeability. Attraction and repulsion between polylegionaminic acid and other cell-surface components may stabilize the conformation of the macromolecules and the outer membrane as a whole. On the other hand, the presence of the deoxy groups and *N*- and *O*-acyl substituents in polylegionaminic acid makes LPS of *L. pneumophila* highly hydrophobic.[60] The hydrophobicity of the bacterial surface may promote the adherence to alveolar macrophages, an early step of pulmonary infection by the bacterium.

The majority of clinical isolates, in particular strains associated with *L. pneumophila* outbreaks, were found to bind mAb 2 and mAb 3/1. This has led to the notion that the 8-*O*-acetyl-associated epitope of legionaminic

acid is associated with virulence.[56] However, infection experiments showed no discernible difference in uptake or intracellular multiplication of *lag-1*-positive strains and the corresponding isogenic *lag-1* mutants in monocyte-like U937 cells, guinea pigs alveolar macrophages, and free-living amoebae.[52,61] Nevertheless, 8-*O*-acetylation of polylegionaminic acid may contribute to virulence by promotion of the spread of the bacterium during *L. pneumophila* outbreaks. Aerosolation of contaminated water, mainly the condensing water of cooling towers, has been shown to be the major source of transmission during outbreaks of community-acquired Legionnaires' disease. 8-*O*-Acetylation increases the hydrophobicity of LPS and thus enhances the ability of the microorganism to form stable aerosols. In support of this idea, it was found that mAb 3/1 bound to 26 of 30 clinical isolates from community-acquired cases but only to 10 of 23 nosocomial isolates.[52]

b. Glycoproteins.—Pilin of *P. aeruginosa*[15] and *Campylobacter* flagellins[16] are involved in carrying out motility and other surface protein-dependent functions of bacteria involved in pathogenesis. Mutations of *neuB3*, a putative gene involved in biosynthesis of pseudaminic acid in *C. jejuni* flagellin, resulted in loss of motility.[62]

Glycosylation of the proteins has the potential to influence the interaction of the cell with its environment. The presence of negatively charged pseudaminic acid derivatives on the surface would lower the isoelectric point, influence solubility, and likely increase ionic interactions.[15] In *C. jejuni* flagellin, these may be further controlled by introduction of a basic acetamidino group to the sugar.[16]

Pseudaminic acid derivatives may function as a biological mask protecting sensitive protein structures from proteolytic cleavage. A structural similarity between pseudaminic and sialic acids may play a role in immune avoidance by protecting the proteins from complement binding and phagocytosis, or from the host B-cell response. This may also account for binding of a sialic acid-specific lectin to *Campylobacter* flagellins.[16]

III. CHEMICAL SYNTHESIS AND STRUCTURE DETERMINATION

The main approaches used for identification of natural isomers of 5,7-diamino-3,5,7,9-tetradeoxynon-2-ulosonic acids and determination of structures of bacterial poly- and oligosaccharides-containing derivatives of these sugars were ^{1}H and ^{13}C NMR spectroscopy. This could be applied to both intact natural glycopolymers and carbohydrate portions obtained after delipidation of LPS as well as to oligosaccharide fragments and

monosaccharides prepared by chemical degradations. Complementary and sometimes essential information could be obtained using mass spectrometry of oligosaccharides and open-chain derivatives of monosaccharides. The most complex problem was the full configuration assignment. Whereas the relative configuration of the chiral centers within the pyranose ring (C-4–C-6) could be easily established using ^{1}H NMR spectroscopy (see later), determination of the configuration in the side chain (C-7–C-8) was accomplished after chemical synthesis of various stereoisomers with fully defined configurations.

1. Chemical Synthesis

By analogy with the synthesis of *N*-acetylneuraminic acid,[63] di-*N*-acetyl derivatives of 5,7-diamino-3,5,7,9-tetradeoxynon-2-ulosonic acids could be obtained by condensation of 2,4-diacetamido-2,4,6-trideoxyhexoses with oxaloacetic acid under basic conditions. Four chiral centers in the C$_6$ precursors, C-2–C-5, correspond to the centers C-5–C-8 in the target C$_9$ products, and the fifth asymmetric center, C-4, is formed upon condensation. At present, derivatives of twelve 2,4-diamino-2,4,6-trideoxy-hexoses with the D-*gluco*, D-*manno*, L-*allo*, D-*galacto*, D- and L-*altro*, D- and L-*talo*, D- and L-*gulo*, D- and L-*ido* configurations have been prepared by multistep chemical syntheses.[11,17,18,64,65]

2,4-Diacetamido-2,4,6-trideoxy-D-mannose (**30**), -L-gulose (**31**), -D-talose (**32**), and -L-allose (**33**) were used in the synthesis of 5,7-diacetamido-3,5,7,9-tetradeoxynon-2-ulosonic acids.[11,17,18] The initial C$_6$ compounds possess the same L,L configuration at C-2 and C-3 (corresponding to C-5 and C-6 in the C$_9$ products), whereas the configurations at C-4 and C-5 vary, thus adopting all possible stereochemical combinations at C-7 and C-8 of the nonulosonic acids (D,D; D,L; L,D; and L,L; respectively). The reaction products were isolated by anion-exchange chromatography and isomers separated by reversed-phase HPLC. The results of condensation of **30–33** with oxaloacetic acid are summarized in Fig. 10.

Compounds **30** and **31** having the *threo* configuration of the C-3–C-4 fragment afforded pairs of the C-4 epimers in nearly equal amounts (**34–37**). The products from **30** are derivatives of legionaminic acid (**34**) and 4-epilegionaminic acid (**35**), and one of the products from **31** is a derivative of 8-epilegionaminic acid (**36**). Compounds **32** and **33** having the *erythro* configuration of the C-3–C-4 fragment yielded the expected sugars with an equatorial HO-4 as the major products (**38** and **40**) but no corresponding compounds with an axial HO-4. Instead, the isomers **39**, **41**, and **42** with an axial AcNH-5 were isolated as minor products, which obviously resulted from a base-induced epimerization at C-2 in the starting monosaccharides
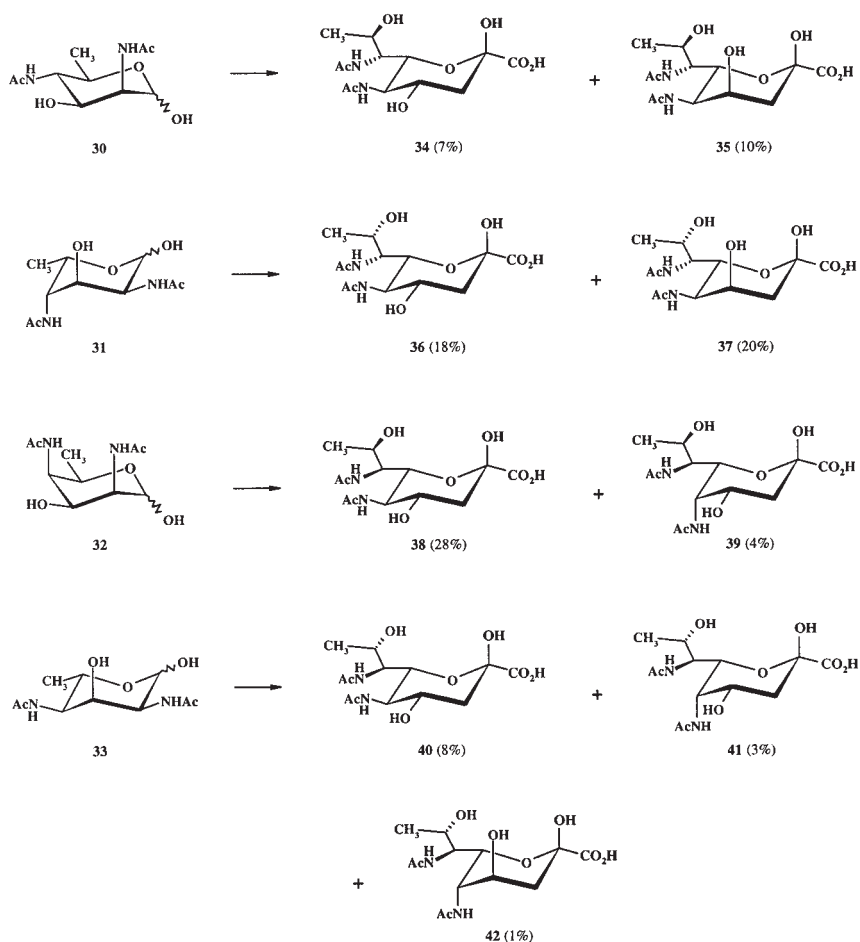
Fig. 10. Chemical syntheses of 5,7-diacetamido-3,5,7,9-tetradeoxynon-2-ulosonic acids (**34–42**) from 2,4-diacetamido-2,4,6-trideoxyhexoses (**30–33**).[11,17,18] (The reaction conditions: 1.75 mol oxaloacetic acid, 0.4 mol $Na_2B_4O_7$, pH 10.5, room temperature. The yields of the products are given in parentheses.)

before condensation. A minor product **41** is a derivative of pseudaminic acid. The nonulosonic acids were not obtained in crystalline form and were characterized by optical rotation (Table II) and NMR spectroscopy (Section III.4, Tables III and IV).

   To sum up, derivatives of nine isomers of 5,7-diamino-3,5,7,9-tetradeoxynon-2-ulosonic acids having the D-*glycero*-D-*galacto* (**34**),

TABLE II
Ratio of α and β Anomers and Specific Optical Rotation of Synthetic
5,7-Diacetamido-3,5,7,9-tetradeoxynon-2-ulosonic Acids

Compound	Configuration	α/β Ratio[a]	[α]$_D$ (°, Water)	References
34	D-*glycero*-D-*galacto*	1:18	+ 27.2	11,18
35	D-*glycero*-D-*talo*	1:5.4	− 12.5	11,18
36	L-*glycero*-D-*galacto*	1:19	+ 15.4	17,18
37	L-*glycero*-D-*talo*	1:8	− 19.2	17,18
38	D-*glycero*-L-*altro*	13.3:1	− 14.3	18
40	L-*glycero*-L-*altro*	8.3:1	− 48.2	18
39	D-*glycero*-L-*manno*	12.5:1	− 39.0	18
41	L-*glycero*-L-*manno*	7.5:1	− 56.9	18
42	L-*glycero*-L-*gluco*	4.0:1	− 76.0	18

[a]The ratios of α and β anomers are given for solutions in D$_2$O at 30 °C.

D-*glycero*-D-*talo* (**35**), L-*glycero*-D-*galacto* (**36**), L-*glycero*-D-*talo* (**37**), D-*glycero*-L-*altro* (**38**), D-*glycero*-L-*manno* (**39**), L-*glycero*-L-*altro* (**40**), L-*glycero*-L-*manno* (**41**), and L-*glycero*-L-*gluco* (**42**) configuration have been synthesized, including all four isomers that have been identified as natural compounds. Comparison of the data of the synthetic and natural sugars enabled confirmation of the identity of pseudaminic acid, whereas the configurations of the other isomers were confirmed in some polysaccharides and revised in others.[11,18]

2. Preparation of Monosaccharides and Oligosaccharides from Bacterial Polysaccharides

In most structural studies of the polysaccharides, oligosaccharides containing the nonulosonic acids were prepared in order to have better-resolved NMR spectra and for investigation by MS. In some examples, a monosaccharide or a monosaccharide glycoside were also prepared by chemical degradations.

a. Solvolysis.—Independent of the anomeric configuration, the glycosidic linkage of the nonulosonic acids was found to be completely stable toward solvolysis, which makes this method most useful for preparation of oligosaccharides with linked residues of the nonulosonic acids.[22–24,30,36,37,39] Thus, a disaccharide with pseudaminic acid at the nonreducing end (**43**) was obtained by solvolysis with anhydrous hydrogen fluoride of OPS of *P. aeruginosa* O9a[22] (**6**) (Fig. 11). When solvolysis was performed in the presence of methanol, the products were methyl glycosides of the same disaccharide[22] (**44**). These and other methyl glycosides were found to be

more convenient for isolation by reversed-phase HPLC and subsequent NMR spectroscopic studies compared to free oligosaccharides.[22,23,30,37] Various oligosaccharides, from di- to tetra-saccharide, or oligosaccharide methyl glycosides were prepared by solvolytic cleavage of some other heteropolysaccharides containing pseudaminic acid [*P. aeruginosa* O7a,7d[23] (**10**) and *S. boydii* type 7[24] (**12**)] or 8-epilegionaminic acid [*P. aeruginosa* O12[30] (**24**), *Y. ruckerii* O1[36] (**25**), *S. arizonae* O61[37] (**26**), and *S. putrefaciens* A6[39] (**27**)].

Varying the reaction conditions gave rise to larger or smaller oligosaccharides. For instance, the α-D-Qui*p*NAc-(1 → 8)-α-8eLeg*p*5Ac7Ac-(2 → 3)-α-L-Fuc*p*NAm-(1 → OMe) trisaccharide or the α-8eLeg*p*5Ac7Ac-(2 → 3)-α-L-Fuc*p*NAm-(1 → OMe) disaccharide were obtained as the predominant products from OPS of *P. aeruginosa* O12 (**24**) by solvolysis with anhydrous hydrogen fluoride in methanol at 20 and 40 °C, respectively.[30]

In a recent study of the pseudaminic acid-containing OPS of *P. vulgaris* O39 (**15**), hydrogen fluoride was successfully replaced with trifluoro-methanesulfonic (triflic) acid.[28] Again, solvolysis with triflic acid could be performed to give the α-D-Glc*p*NAc-(1 → 8)-β-Pse*p*5Ac7Ac-(2 → 3)-L-FucNAc trisaccharide under mild conditions (−4 °C for 2 h) or the β-Pse*p*5Ac7Ac-(2 → 3)-L-FucNAc disaccharide under more drastic conditions (20 °C for 16 h).

b. Acid Hydrolysis.—Compared to solvolysis, acid hydrolysis of the polysaccharides proceeded differently and enabled preparation of oligo-saccharides with the nonulosonic acids at the reducing end. The rate of hydrolysis of the ketosidic linkage was found to depend significantly on the anomeric configuration. When the carboxyl group is equatorial, as in α-pseudaminic acid[23,27] or β-legionaminic acid,[35] hydrolysis proceeded so easily that the polysaccharide chain depolymerized during delipidation of LPS with dilute aqueous acetic acid[23,27,35] or sodium acetate buffer at pH 4.2.[21] Thus, no OPS could be isolated from the α-pseudaminic acid-containing LPS of *P. aeruginosa* O7a,7d (immunotype 6) (**10**) since aqueous 1% acetic acid at 100 °C cleaved the polysaccharide to the repeating unit trisaccharide[23] (**46**) (Fig. 12). The CPS from *S. fredii* HH103 (**5**) was also cleaved under mild acidic conditions, and the resulting mixture of 5-*N*-acetyl-7-*N*-(3-hydroxybutanoyl)pseudaminic acid and its oligomers was fractionated by gel-permeation chromatography on Sephadex G-50 and Sephadex G-15.[14]

The hydrolytic cleavage of the glycosidic linkage of the sugars with an axial carboxyl group required more drastic conditions, but even these allowed selective degradation. For instance, selective cleavage of the

TABLE III
^{1}H NMR Data of Synthetic 5,7-Diacetamido-3,5,7,9-tetradeoxynon-2-ulosonic Acids[a]

Configuration (Compound)		H-3eq ($J_{3eq,3ax}$)	H-3ax ($J_{3ax,4}$)	H-4 ($J_{3eq,4}$)	H-5 ($J_{4,5}$)	H-6 ($J_{5,6}$)	H-7 ($J_{6,7}$)	H-8 ($J_{7,8}$)	H-9 ($J_{8,9}$)
D-glycero-D-galacto	α	2.73	1.71	3.82	3.68	3.93		3.94	1.16
(34)		(12.9)	(11.9)	(4.7)		(10.3)			
	β	2.31	1.87	3.98	3.72	4.31	3.91	3.85	1.16
		(13.1)	(11.7)	(4.8)	(10.3)	(10.5)	(1.9)	(8.9)	(6.2)
D-glycero-D-talo	α	2.69	1.94	4.10	3.86	4.55	3.88	4.00	1.20
(35)		(14.4)	(3.5)	(3.0)	(2.9)	(10.8)	(2.3)	(8.6)	(6.4)
	β	2.19	2.14	4.13	3.90	4.63	3.92	3.92	1.18
		(14.9)	(3.4)	(3.0)	(2.9)	(10.8)	(<2)[b]	(8.7)[b]	(5.5)
L-glycero-D-galacto	α	2.69	1.71	3.82	3.67	3.85	3.93	4.00	1.20
(36)		(13.0)	(12.0)	(4.9)	(10.4)	(10.4)	(1.7)[b]	(6.4)	(6.5)[b]
	β	2.32	1.86	3.95	3.73	4.16	3.95	3.91	1.18
		(13.1)	(11.5)	(4.8)	(10.2)	(10.3)	(2.0)	(6.4)	(6.2)
L-glycero-D-talo	α	2.65	1.94	4.08	3.84	4.47	3.89	4.08	1.28
(37)		(14.4)	(2.7)	(3.6)	(2.7)	(10.5)	(2.0)[b]		(6.3)
	β	2.18[c]	2.13[c]	4.11	3.89	4.48	3.95	3.96	1.21
		(14.9)	(3.3)	(2.9)	(2.8)	(10.6)	(1.3)	(6.5)[b]	(5.7)
D-glycero-L-altro	α	2.34	1.93	3.99	3.86	3.90	3.92	4.42	1.08
(38)		(13.2)	(11.6)	(4.8)	(9.4)	(10.1)	(2.8)	(<1)	(6.4)
	β	2.71	1.75	3.82	3.81	3.56	3.91	4.40	1.08
		(12.5)[b]	(11.1)[b]	(4.5)[b]	(9.8)[b]	(9.8)[b]	(5.1)[b]	(1.8)[b]	(6.3)[b]

L-*glycero*-L-*altro* (**40**)	α	2.32 (13.3)	1.93 (12.5)	3.94 (4.4)	3.91 (10.2)	3.94 (10.2)	4.16 (2.8)	4.06 (5.8)	1.18 (6.4)
	β	2.70 (12.8)	1.72 (12.2)	3.79 (4.5)	3.85 (10.7)	3.57 (10.7)	4.13 (3.3)	4.07 (6.0)	1.21 (6.3)
D-*glycero*-L-*manno* (**39**)	α	2.03 (13.4)	1.82 (12.3)	4.25 (5.0)	4.29 (4.3)	4.27 (1.8)	3.82 (10.1)	4.12 (1.2)	1.09 (6.6)
	β	2.50 (13.2)	1.64 (12.8)	4.10 (4.9)	4.22 (4.5)	4.14 (2.1)	3.86 (10.4)	4.25 (1.6)	1.10 (6.4)
L-*glycero*-L-*manno* (**41**)	α	2.01 (12.8)	1.80 (12.0)	4.20 (4.6)	4.27 (3.7)	4.08 (1.0)[b]	4.17 (10.7)	4.10 (3.3)	1.10 (6.5)
	β	2.48 (13.0)	1.62 (12.9)	4.08 (4.7)	4.29 (3.6)	3.96 (2.4)	4.15 (10.5)	4.18 (3.4)	1.12
L-*glycero*-L-*gluco* (**42**)	α	1.95 (15.2)	2.13 (3.6)[b]	4.00 (3.7)	3.91 (3.3)	4.42 (2.1)	4.22 (10.5)	4.16 (3.5)	1.15 (6.6)
	β	2.48 (14.8)	1.92 (2.9)	4.00 (3.1)	3.85 (2.6)	4.36 (2.2)	4.15 (10.3)	4.24 (4.0)	1.21 (6.5)

[a]Chemical shifts in ppm at 500 MHz, coupling constants in Hz for solutions in D_2O at 30 °C. Signals for the *N*-acetyl groups are at δ 1.95–2.08. Refs. 11, 17, 18.
[b]Data of sodium salt.
[c]Assignment could be interchanged.

TABLE IV
^{13}C NMR Data of Synthetic 5,7-Diacetamido-3,5,7,9-tetradeoxynon-2-ulosonic Acids[a]

Configuration (Compound)		C-1	C-2	C-3	C-4	C-5	C-6	C-7	C-8	C-9
D-glycero-D-galacto	α		97.2	41.4	69.2	53.4	73.2	54.4	68.0	20.4
(34)	β	174.4[b]	96.6	40.3	68.4	53.9	70.9	54.4	67.5	20.4
D-glycero-D-talo	α			40.2	66.9	49.7	70.3	55.0	68.2	19.8
(35)	β	174.7[b]	96.3	37.6	67.1	49.7	66.5	54.7	67.2	20.4
L-glycero-D-galacto	α	173.1	97.3	41.2	68.9	53.7	75.2	54.4	69.4	19.8
(36)	β	174.3	96.5	40.4	68.3	54.1	72.9	54.4	69.3	19.8
L-glycero-D-talo	α			40.0	66.6	50.1	72.6	54.9	69.7	19.9
(37)	β	174.5[b]	96.1	37.7	66.9	50.0	68.5	54.9	69.2	19.9
D-glycero-L-altro	α	173.7	96.1	39.9	67.6	54.7	75.7	53.9	66.6	20.2
(38)	β	172.9	97.2	41.1	68.7	54.7	77.3	54.3	66.5	20.0
L-glycero-L-altro	α	173.7	96.1	39.9	68.2	55.3	73.7	55.5	67.6	19.7
(40)	β	173.0	97.0	41.2	69.1	54.9	76.0	55.8	67.8	19.7
D-glycero-L-manno	α	174.7	96.8	35.5	66.1	49.9	70.3	54.4	66.0	20.0
(39)	β			36.8	68.0	49.1	73.0	54.4	67.4	20.0
L-glycero-L-manno	α	174.9	97.0	35.6	66.1	49.9	71.4	54.0	68.1	16.7
(41)	β			36.7	67.3	49.9	74.3	53.8	68.2	16.7
L-glycero-L-gluco	α	174.6	96.4	33.3	67.1	48.9	67.1	53.8	67.9	16.6
(42)	β			35.7	67.4	48.5	71.8	54.4	68.6	17.0

[a]Chemical shifts in ppm at 125 MHz for solutions in D$_2$O at 30 °C. Signals for the N-acetyl groups are at δ 22.9–23.5 (CH$_3$) and 174.2–175.9 (CO). Refs. 11, 17, 18.
[b]Assignment of the signals for CO$_2$H of the sugar and CO of the N-acetyl groups could be interchanged.

glycosidic linkage of α-8-epilegionaminic acid in OPSs of *Y. ruckeri* O1[36] (**25**) and *S. arizonae* O61[37] (**26**) was performed with 0.1 M HCl at 95 or 100 °C for 5 h and resulted in tri- and tetrasaccharides. Similar hydrolysis of the α-D-Glc*p*NAc-(1 → 8)-α-8eLeg*p*5(*R*3Hb)7Ac-(2 → 3)-α-L-Fuc*p*NAm-(1 → OMe) trisaccharide obtained by solvolysis of the polysaccharide **26** afforded the α-D-Glc*p*NAc-(1 → 8)-8eLeg*p*5(*R*3Hb)7Ac disaccharide and α-L-Fuc*p*NAm-(1 → OMe) glycoside.[37]

In contrast, the homopolymers of 5-N-acetimidoyl-7-N-acetyl-α-legionaminic (**17**) and -4-epilegionaminic (**18**) acids present in LPS of *L. pneumophila* were found to be stable toward acid hydrolysis.[31,40] Attempts to isolate monomers by methanolysis or solvolysis with hydrogen fluoride also failed, and only minor amounts of chiral alcohol derivatives for determination of the absolute configurations by GLC–MS could be prepared by acid alcoholysis with (*S*)-2-butanol.[11,40] Conversion of the N-acetimidoyl group to the N-acetyl group was necessary for the cleavage of polylegionaminic acid **17** but not poly(4-epilegionaminic) acid **18**. The conversion in the polymer **17** required rather drastic alkaline conditions (0.1 M NaOH, 100 °C, 5 h).[31] Other oligo- and polysaccharides containing
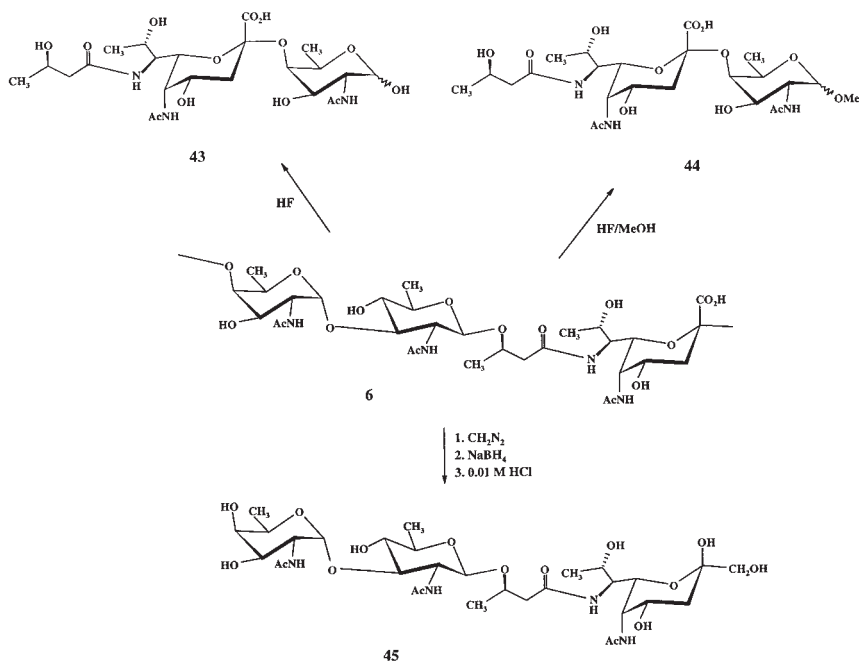
FIG. 11. Selective cleavages of OPS of *Pseudomonas aeruginosa* O9a (**6**).[22]

N-acetimidoyl derivatives of the nonulosonic acids, including the polymer **18**, could be hydrolyzed to the corresponding *N*-acetyl derivatives by treatment with a weak aqueous base, e.g., 12% ammonia (80 °C, 16 h or 45 °C, 4 h),[32,40] 5% triethylamine (60 °C, 16 h),[25] or Na_2CO_3 at pH 12 (20 °C, 14 days).[39]

Carboxyl reduction of a nonulosonic acid with an axial carboxy group facilitated the cleavage of the glycosidic linkage. While the β-pseudaminic acid-containing OPS of *P. aeruginosa* O9a (**6**) was stable toward hydrolysis with 0.01 *M* hydrochloric acid (100 °C, 2 h), the carboxyl-reduced polysaccharide smoothly afforded under these conditions a trisaccharide with the carboxyl-reduced sugar at the reducing end[22] (**45**) (Fig. 11). The preferred carboxyl-reduction procedure was found to be borohydride reduction of a methyl ester of the nonulosonic acid, whereas the use of the carbodiimide method resulted in a complex mixture of products.[22]

In most cases, the nonulosonic acids present in heteropolysaccharides could not be released as monosaccharides by direct acid hydrolysis because of their instability under the drastic conditions that were necessary to cleave the aldosidic linkage of the neighboring sugar residue. A combination of solvolysis with hydrogen fluoride and acid hydrolysis of the resulting
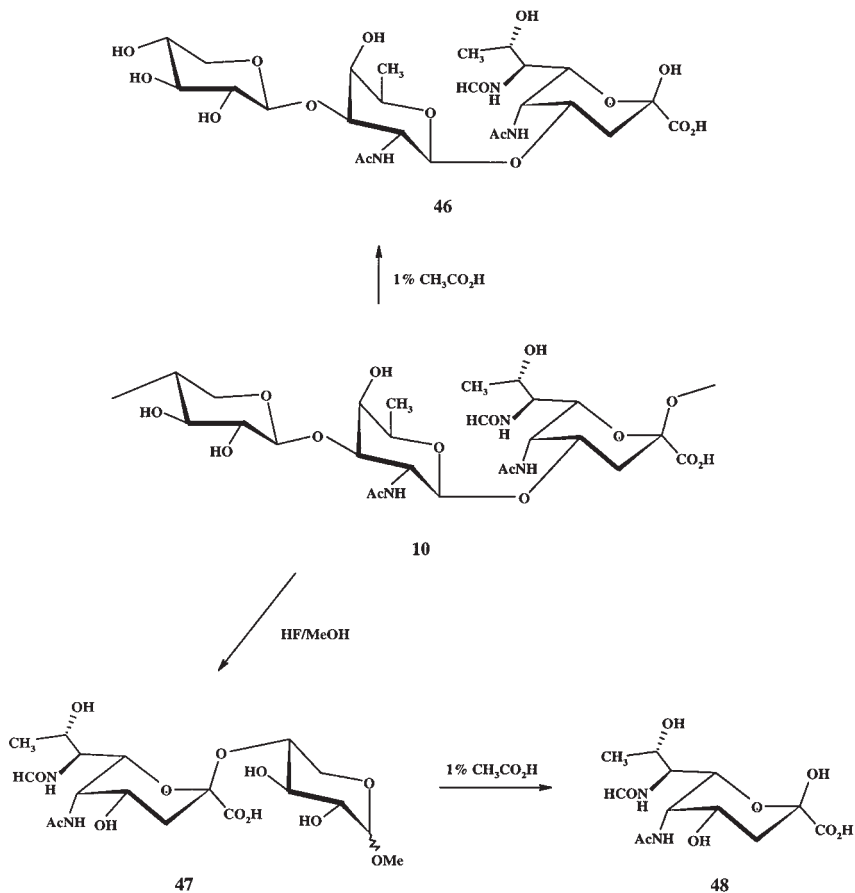
Fig. 12. Selective cleavages of OPS of *Pseudomonas aeruginosa* O7a,7d (**10**) and preparation of 5-*N*-acetyl-7-*N*-formylpseudaminic acid (**48**).[23]

oligosaccharide(s) containing the nonulosonic acid at the nonreducing end provided a way to overcome this difficulty. Thus, 5-*N*-acetyl-7-*N*-formylpseudaminic acid (**48**) was obtained as monosaccharide by hydrolysis with 1% acetic acid (100 °C, 1 h) of disaccharide glycosides (**47**) prepared by solvolysis of LPS of *P. aeruginosa* O7a,7d (**10**) with hydrogen fluoride in methanol[23] (Fig. 12).

Solvolysis with the same mixture of OPS of *S. putrefaciens* A6 (**29**) followed by mild alkaline treatment of the product (**49**) and hydrolysis of the resultant disaccharide glycoside (**50**) with 0.5 *M* CD₃CO₂D in D₂O (100 °C, 6.5 h, with NMR control) released a 2,8-anhydro derivative of 8-epilegionaminic acid with a deuterated methylene group (**51**), together

with a methyl furanoside of shewanellose (**52**)[39] (Fig. 13). The same derivative was obtained from synthetic 8-epilegionaminic acid by hydrolysis under similar conditions.[39]

c. Smith Degradation.—Three repeated Smith degradations of OPS of *S. boydii* type 7 (**12**) gave, via two oligosaccharides, an ethylene glycol glycoside of β-pseudaminic acid.[24] Following conversion of the 5-*N*-acetimidoyl group into the 5-*N*-acetyl group, Smith degradation of OPS of *V. cholerae* O2 (**13**) gave β-D-Qui*p*NAc-(1 → 4)-β-Pse*p*5Ac7Ac-(2 → 2)-threitol.[25] Methanolysis of the same modified polysaccharide gave a β-Pse*p*5Ac7Ac-(2 → 4)-D-Gal disaccharide derivative, which upon saponi-fication and Smith degradation yielded a tetritol glycoside of β-pseudaminic acid.[25]

d. Cyclization in Oligosaccharides.—Under certain conditions, the nonulosonic acids can form interresidue spiro-lactones or-lactams. Thus, lactonization between the residues of di-*N*-acetyllegionaminic acid and *N*-acetylfucosamine to a 1,4-dioxane structure (**53**) was observed on dephosphorylation of the oligosaccharide **21** from *P. fluorescens* ATCC 49271 with aqueous 48% hydrofluoric acid (30 °C, 4 h) following treatment with aqueous ammonia[32] (Fig. 14).

Acetylation of a disaccharide glycoside (**54**), prepared by solvolysis of OPS of *P. aeruginosa* O12 (**24**) with hydrogen fluoride in methanol (see above), was accompanied by lactam formation between residues of di-*N*-acetyl-8-epilegionaminic acid and *N*-acetimidoylfucosamine.[30] Upon de-*O*-acetylation of the resulting tricyclic disaccharide **55** with 1 *M* NaOMe in methanol (20 °C, 48 h), the *N*-acetimidoyl group was cleaved with no effect on the six-membered lactam ring to yield disaccharide **56** (Fig. 15).

3. Mass Spectrometry

a. Soft Ionization MS.—Soft ionization FAB, ESI, or MALDI mass spectra were obtained for a number of the nonulosonic acid-containing oligosaccharides derived from the glycopolymers (Section III.2). Thus, the FAB mass spectrum of the disaccharide **49** from OPS of *S. putrefaciens* strain A6 (Fig. 13) demonstrated not only the nature of the higher sugars but also that the 8-epilegionaminic acid residue had one *N*-acetyl and one *N*-acetimidoyl group.[39] The negative-mode ESI mass spectrum of the oligosaccharide fraction obtained on acid treatment of OPS of *Pseudoalteromonas distincta* KMM 638 showed the presence of a pentasaccharide containing all components of the OPS repeating unit and confirmed the presence of one *N*-acetyl and one *N*-formyl group on the
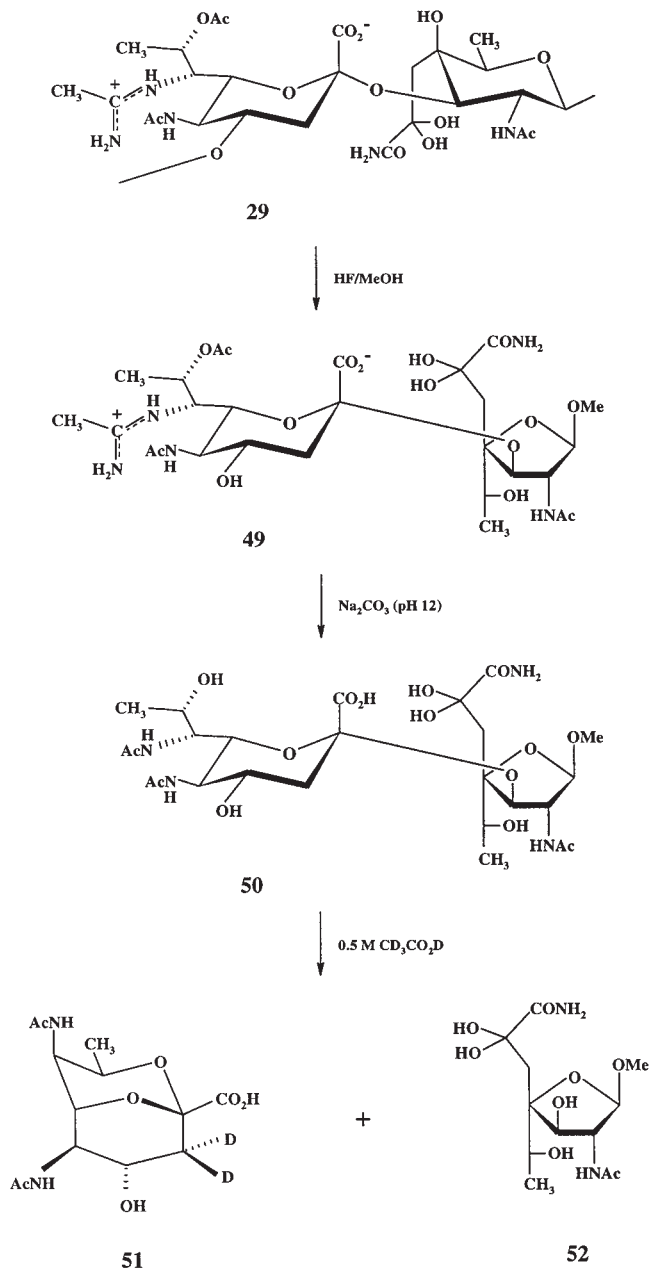
Fig. 13. Selective cleavage of OPS of *Shewanella putrefaciens* A6 (**29**) and modification and cleavage of the resultant disaccharide (**49**).[39]
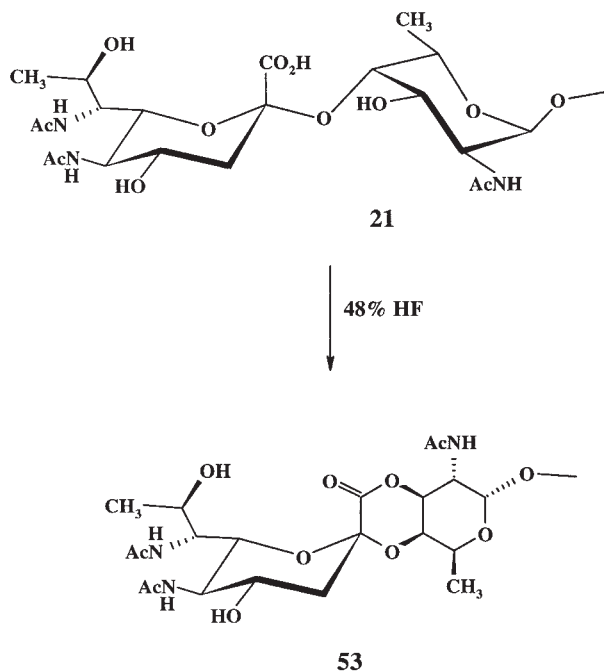
FIG. 14. Lactonization in the oligosaccharide from LPS of *Pseudomonas fluorescens* ATCC 49271 (**21**) upon treatment with aqueous 48% hydrofluoric acid.[32] (For the full structure of the oligosaccharide **21**, see Fig. 6.)

pseudaminic acid residue.[27] MALDI-TOF MS on an oligomeric fraction derived from the homopolysaccharide **5** of *S. fredii* HH103 (Fig. 2) gave the molecular mass of the repeating unit established as the average difference between clusters formed from several sodium adduct ions.[14]

Positive-mode ESI MS–MS analysis was applied to tryptic glyco-peptides from *C. jejuni* flagellin for identification of the glycosyl groups.[16] Second-generation fragment ions with m/z 317, 316, and 409 formed by collision-induced dissociation (CID) were consistent with di-*N*-acetyl, *N*-acetimidoyl-*N*-acetyl, and di-*N*-glyceroyl derivatives of pseudaminic acid. The product ion with m/z 316 showed prominent losses of NH_3 and $C_2N_2H_6$ confirming the presence of an acetamidino group, and the occurrence of side-chain fragment ions common to both Pse5Ac7Ac and Pse5Am7Ac suggested that the acetamidino group was located at position 5 rather than 7.

FAB with CID was applied for sequence determination of the oligosaccharide **22** derived from LPS of *V. salmonicida*[34] (Fig. 6). Using the pseudomolecular ion, $[M + H]^+$, as precursor ion, several fragments due to cleavage of the glycosidic linkages, including the ketosidic linkage of
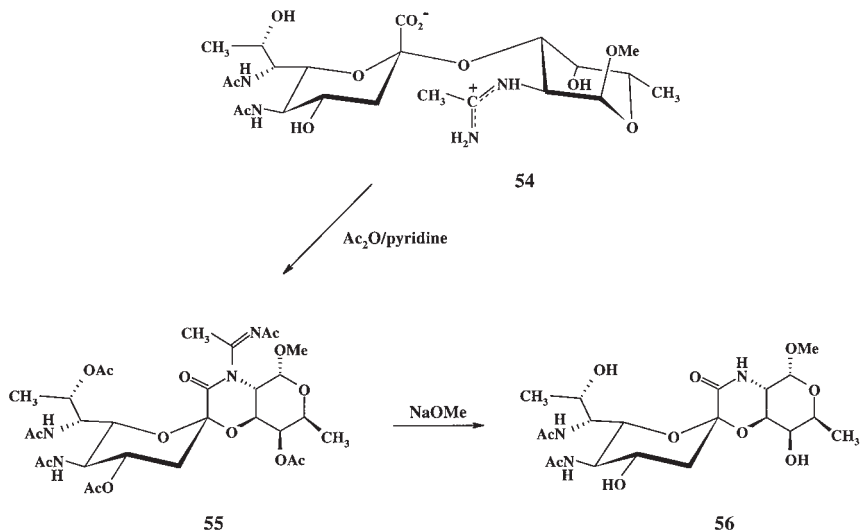
FIG. 15. Formation of lactam upon acetylation of the disaccharide (**54**) from OPS of *Pseudomonas aeruginosa* O12 (structure **24**, Fig. 7) and cleavage of the *N*-acetylacetimidoyl group upon de-*O*-acetylation of the resultant disaccharide (**55**).[30]

legionaminic acid, (B- and Y-type fragments) were observed in the B/E-linked scan FAB mass spectrum (Fig. 16). Positive-mode FAB MS of a tetrasaccharide with 8-epilegionaminic acid at the reducing end (**57**) from OPS of *Y. ruckerii* O1 (Fig. 7, structure **24**) showed, in addition to $[M+H]^+$, B- and Y-type fragment ions (Fig. 16), and thus in this case no CID was necessary.[36] The monosaccharide sequence in the O-linked trisaccharide chain of *P. aeruginosa* pilin with pseudaminic acid at the nonreducing end was confirmed by positive-mode ESI MS–MS of the oligosaccharide–serine **11** isolated from proteolytically digested pure pili.[15] Upon CID, the pseudomolecular ion $[M+H]^+$ afforded B-type fragment ions due to the successive losses of Ser, FucNAc-Ser, and Xyl-FucNAc-Ser (Fig. 16).

b. Electron Impact MS.—In early studies, molecular masses of oligosaccharides were determined by electron impact MS of the acetylated derivatives. For instance, the chemical modification of the disaccharide **54** from OPS of *P. aeruginosa* O12, namely the formation of the spiro-lactam **55** and cleavage of the *N*-acetimidoyl group on alkaline treatment (Fig. 15), were monitored in this way.[30] Fragmentation was observed due to losses of the C-8–C-9 fragment, of the methoxycarbonyl group of 8-epilegionaminic acid, and cleavage of the ketosidic linkage.
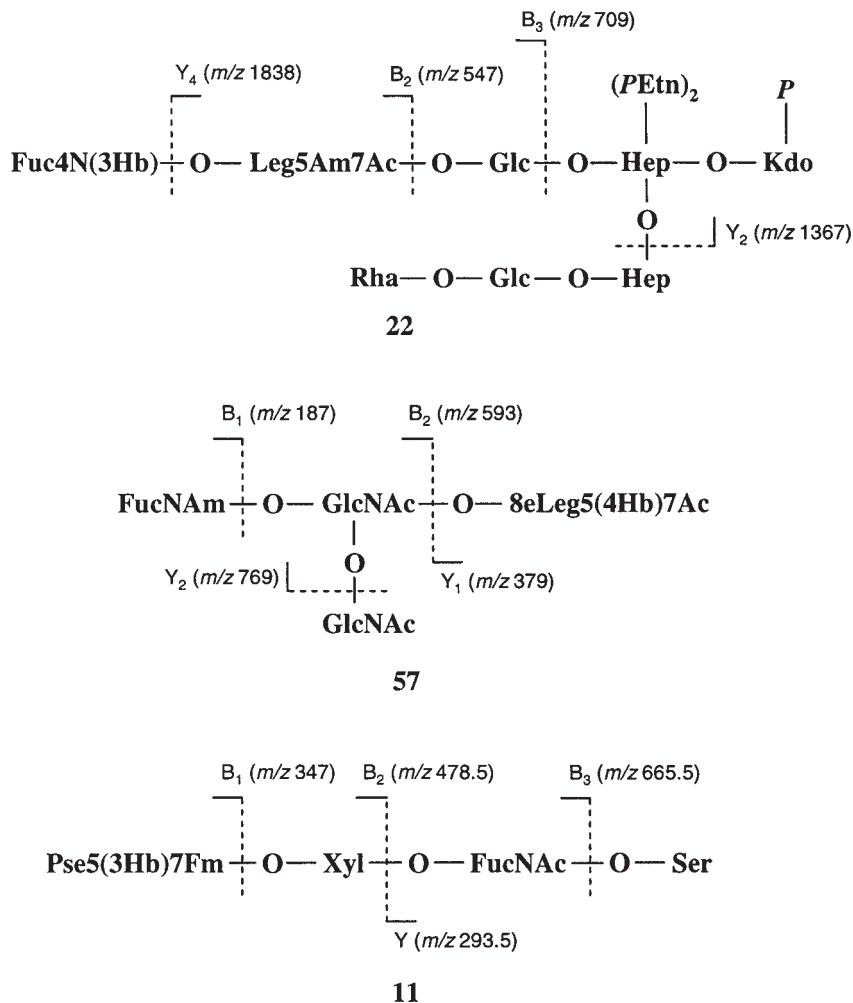
B_3 (*m/z* 709)

Y_4 (*m/z* 1838) B_2 (*m/z* 547) $(PEtn)_2$ *P*

Fuc4N(3Hb)—O— Leg5Am7Ac—O— Glc—O—Hep—O— Kdo

O Y_2 (*m/z* 1367)

Rha— O— Glc— O—Hep

22

B_1 (*m/z* 187) B_2 (*m/z* 593)

FucNAm—O— GlcNAc—O— 8eLeg5(4Hb)7Ac

Y_2 (*m/z* 769) O Y_1 (*m/z* 379)

GlcNAc

57

B_1 (*m/z* 347) B_2 (*m/z* 478.5) B_3 (*m/z* 665.5)

Pse5(3Hb)7Fm —O—Xyl—O— FucNAc —O— Ser

Y (*m/z* 293.5)

11

FIG. 16. Fragmentation in FAB MS–MS of the oligosaccharide (**22**) from LPS of *Vibrio salmonicida*,[34] FAB MS of the oligosaccharide (**57**) from OPS of *Yersinia ruckerii* O1,[36] and ESI MS–MS of the trisaccharide-serine (**11**) from pilin of *Pseudomonas aeruginosa* 1244.[15] (For abbreviations, see Figs. 3, 5, and 7. Pseudomolecular ions $[M + H]^+$ at *m/z* 1867.6 and 771.5 were used as precursor ions in analysis of **22** and **11**, respectively.)

Electron impact MS of open-chain monomeric derivatives was useful for both determination of the position of acylamino groups and linkage analysis. The mass spectrum of the carbonyl-reduced and methylated derivative of the monosaccharide **48** from *P. aeruginosa* O7a,7d (Fig. 12) showed the expected fragmentation pattern, which demonstrated the
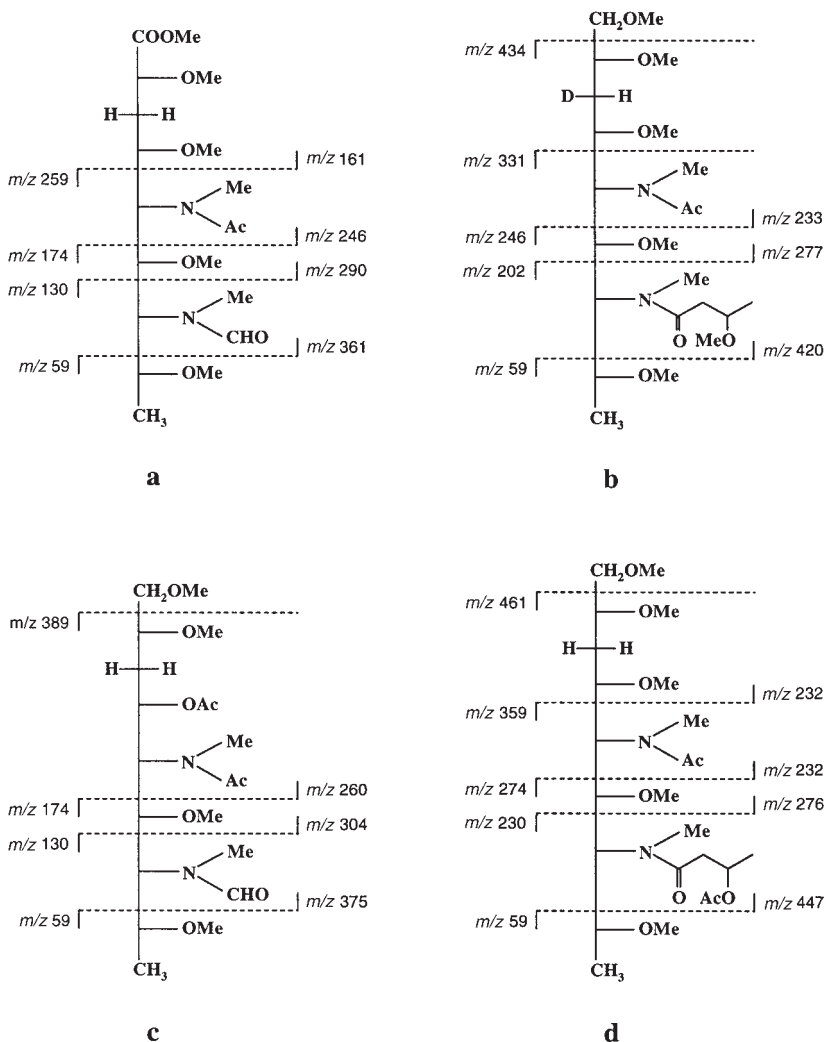
FIG. 17. Electron impact MS fragmentation of the fully (a, b) and partially (c, d) methylated reduced open-chain derivatives of the nonulosonic acids. [Taken from Refs. 9 (d), 14 (b), and 23 (a, c); configuration is not shown.]

position of the amino and deoxy groups and confirmed the identity of the N-acyl substituents[23] (Fig. 17a). The pseudaminic acid derivative from CPS of S. fredii HH103[14] (Fig. 2) and di-N-acetyl-4-epilegionaminic acid from LPS of L. pneumonphila serogroup 1[12] were carboxyl-reduced before methylation. The mass spectra obtained showed the same general

fragmentation pattern (e.g., see Fig. 17b for the pseudaminic acid derivative).

When the monosaccharide was not available directly after hydrolysis or solvolysis of the polysaccharide, a monomeric derivative for MS analysis could be obtained by a combination of chemical modifications and cleavages. Thus, in order to obtain an open-chain derivative of pseudaminic acid, OPS of *S. boydii* type 7 (Fig. 4, structure **12**) was carboxyl-reduced, hydrolyzed under mild acid conditions, the products were carbonyl-reduced, solvolyzed with anhydrous hydrogen fluoride, and methylated.[9,24]

Linkage analysis was performed on oligosaccharides obtained by a selective cleavage. Thus, carbonyl reduction, carboxyl methylation, and carboxyl reduction of the trisaccharide **46** from OPS of *P. aeruginosa* O7a,7d (Fig. 12) followed by methylation, solvolysis with anhydrous hydrogen fluoride, and acetylation gave a pseudaminic acid derivative that was *O*-acetylated at position 4 and consequently showed the linkage site[23] (Fig. 17c).

Similar methylation analysis demonstrated the glycosidic linkage between a sugar and an *N*-(3-hydroxybutanoyl) substituent of pseudaminic acid in two polysaccharides (Fig. 2). The carboxyl-reduced trisaccharide **45** from OPS of *P. aeruginosa* O9a (Fig. 11) was carbonyl-reduced and methylated.[9] The open-chain derivative obtained after solvolysis with hydrogen fluoride and acetylation gave the expected electron impact MS fragmentation, which demonstrated *O*-acetylation of 3-hydroxybutanoic acid (Fig. 17d). An alternative sequence was applied to a disaccharide from the CPS **5** of *S. fredii* HH103 (Fig. 2), namely, carbonyl-reduction, methylation, carboxyl-reduction, hydrolysis with dilute trifluoroacetic acid, and a standard conversion to the alditol acetate with *O*-acetyl groups at position 1 of the alditol and position 3 of 3-hydroxybutanoic acid.[14]

4. NMR Spectroscopy and Conformational Analysis

a. ^{1}H NMR Spectroscopy.—The ^{1}H NMR spectra of reducing 5,7-diamino-3,5,7,9-tetradeoxynon-2-ulosonic acid derivatives contained two series of signals belonging to the major and minor anomers with an equatorial and an axial carboxy group, respectively. In each series of the synthetic di-*N*-acetyl derivatives[11,17,18] (Table III), the signal for the CH_3 group (H-9) was at δ 1.05–1.28 (3H, d, $J_{8,9}$ 6.1–6.6 Hz), two signals for the CH_2 group (H-3) were in the region δ 1.5–2.7, and the resonance region for H-4–H-8 was δ 3.4–4.6.

Both ^{1}H NMR chemical shifts and the *J*-splitting showed a significant dependence on the relative configuration of the nonulosonic acids. The $^3J_{H,H}$ coupling constants for the protons within the pyranose ring (H-3–H-6)

were similar to those observed in hexopyranoses.[66] The $J_{6,7}$ coupling constant depended on the configuration at C-5: it was small (1.3–3.3 Hz) when NH-5 was equatorial, and large (10.1–10.7 Hz) when axial (Table III). These values were indicative of the *syn-* and *trans-*like relationship for H-6 and H-7, respectively.

The chemical shifts were influenced also by the anomeric configuration. In the nonulosonic acids with an equatorial OH-4, the difference between the ^{1}H NMR resonances for H-3ax and H-3eq was 0.86–1.02 ppm for the anomer with an axial carboxy group but only 0.21–0.46 ppm for the other anomer, independently of whether NH-5 was axial or equatorial. When OH-4 was axial and NH-5 equatorial, a typical difference of 0.71–0.75 or 0.05 ppm was observed for the anomers with an axial or an equatorial carboxyl group, respectively. In the glycosidically linked nonulosonic acids, the differences were not the same but in most examples were useful for determination of the anomeric configuration of naturally occurring isomers in OPSs and LPSs.[14,22,23,25,30,31,34–36,40]

The nature of *N*-acyl substituents at N-5 and N-7 and the presence of an *O*-acetyl group at O-4 or O-8 influenced ^{1}H NMR chemical shifts of the nearby protons. These features could be determined by characteristic chemical shifts and, for *N*-(hydroxybutanoyl) groups, by the corresponding spin systems. The most significant displacement (downfield by 0.9–1.3 ppm) was observed for the signal of the proton at the acetoxylated carbon compared to that at the corresponding hydroxylated carbon. Therefore, comparison of the spectra of the initial and chemically de-*O*-acetylated compounds was used for determination of the *O*-acetylation sites in the nonulosonic acids.[12,22,39,40]

b. ^{13}C NMR Spectroscopy.—The ^{13}C NMR spectra of each anomer of synthetic 5,7-diacetamido-3,5,7,9-tetradeoxynon-2-ulosonic acids (Table IV) showed nine signals from the sugar moiety, including those of the anomeric carbon C-2 (δ 96.1–97.3), carboxyl group C-1 (δ 174.2–175.9), CH_3 (C-9, δ 16.6–20.4), and CH_2 (C-3, δ 33.3–41.4) groups, three CHOH groups (C-4, C-6, and C-8, δ 66.1–77.3), and two CHNH groups (C-5 and C-7, δ 48.5–55.8), as well as signals for two *N*-acetyl groups (CH_3 at δ 22.9.0–23.5 and CO at δ 174.2–175.9).[11,17,18]

In sugars with different chirality at C-4 or C-5, predictable chemical shift differences were observed for carbons involved in the pyranose ring (compare data for hexopyranoses[67]). Although epimeric differences at C-7 and C-8 cause less predictable changes in the ^{13}C NMR chemical shifts, some particular regularities could be tracked for the D-*galacto* and D-*talo* isomers, which enabled revision of the configurations of legionaminic and 4-epilegionaminic acids.[11] For instance, in sugars with the D configuration

at C-8, the C-6 and C-8 signals appeared upfield by 2.0–2.3 and 1.4–2.0 ppm, respectively, compared to the corresponding L epimer. In the *glycero*-L-*manno* isomers, a significant difference was observed for the C-9 signal, which appeared at δ 20.0 in the D epimer at C-8 but at δ 16.7 in the L epimer (Table IV). This finding was useful for determination of the configuration of pseudaminic acid.

The anomeric configuration had a significant influence on the C-3 and C-6 resonances. Compared to the other anomer, in the anomer with an equatorial carboxy group both signals appeared upfield by 0.8–1.3 and 1.6–2.9 ppm when OH-4 was equatorial or by 2.3–2.6 and 3.8–4.7 ppm, respectively, when OH-4 was axial. The regularity for the C-6 chemical shift was applied for analysis of the anomeric configurations of the naturally occurring nonulosonic acids.[22,23,35,37,40] Another approach could be measuring $J_{C-1,H-3ax}$ and $J_{C-2,H-3ax}$ long-range coupling constants[26] using regularities found for N-acetylneuraminic acid.

If an N-acetimidoyl group was present instead of an N-acetyl group, the ^{13}C NMR signal of the corresponding nitrogen-bearing carbon appeared downfield by 2.9–5.9 ppm (α effect).[25,31,40] In contrast, compared to an N-acetyl group, an N-formyl group caused a small upfield α effect of 1–1.5 ppm.[23] Since the N-formyl group occurs as two stereoisomers (major Z and minor E), the ^{13}C NMR spectrum, as well as the ^{1}H NMR spectrum, showed two sets of signals for the group itself and for the neighboring sugar atoms.[27] An influence similar to that of an acetyl group on the ^{13}C NMR chemical shifts was found for 3- and 4-hydroxybutanoyl groups,[14,20,22,36,37] except for a 3-hydroxybutanoyl group at N-5 of pseudaminic acid, which caused a downfield β effect on C-4 of ~1 ppm[23] (Table V).

In accordance with published data on O-acetylated hexopyranoses,[69] an α effect of O-acetylation at O-4 or O-8 on the ^{13}C NMR chemical shift was positive (2.3–3.2 ppm), whereas β effects of a similar magnitude on the neighboring carbons were negative[22,24,31,40] (Table V).

In addition to N-acylation and O-acetylation, the mode of glycosylation was yet another feature to be taken in consideration for naturally occurring derivatives: whether the sugar was glycosidically linked or had a free anomeric center and whether it was glycosylated itself or occupied a terminal nonreducing position. Comparison of the ^{13}C NMR chemical shifts of the polysaccharides and isolated oligo- and monosaccharides (Table V) as well as the synthetic di-N-acetyl derivatives (Table IV) showed that in glycosides C-2 resonated downfield by 2–7 ppm compared to the corresponding free nonulosonic acids. In pseudaminic acid, the effect was larger when the carboxyl group was axial (4–6 vs. 2–3 ppm) (Table V).

TABLE V
^{13}C NMR Data of Derivatives of 5,7-Diamino-3,5,7,9-tetradeoxynon-2-ulosonic Acids from Natural Glycopolymers[a]

Derivative	C-2	C-3	C-4	C-5	C-6	C-7	C-8	C-9	Bacterial Source[19]	Reference
Pseudaminic acid (Pse)										
α-Pse5Ac7(3Hb)	97.7	36.0	66.6	50.1	71.3	54.2	68.4	16.8	*Sinorhizobum fredii*	14
α-Pse5Ac7Fo	97.6	35.8	66.4	49.9	71.1	52.7	67.6	16.3	*Pseudomonas aeruginosa* O7a,7d	23
α-Pse5Ac7Fo-(2 →	99.8	36.8	66.3	49.9	72.2	53.9	68.3	17.6	*Pseudomonas aeruginosa* O7a,7d	23
α-Pse5Ac7(3Hb)-(2 →	99.7	37.2	66.4	49.8	73.1	55.1	69.2	18.3	*Sinorhizobum fredii*	14
β-Pse5Ac7Ac-(2 →		36.8	67.9	49.6	74.7	54.9	69.5	17.4	*Proteus vulgaris* O39	28
β-Pse5Ac7(R3Hb)-(2 →	102.1	36.8	67.6	49.0	74.8	54.7	69.2	17.7	*Pseudomonas aeruginosa* O9a	22
β-Pse4Ac5Ac7(R3Hb)-(2 →	103.0	34.2	70.3	46.4	74.3	54.5	69.4	17.7	*Pseudomonas aeruginosa* O9a,9b	22
→ 4)-α-Pse5Ac7Fo[b]	97.1	34.8	72.3	46.6	71.2	52.3	67.5	16.3	*Pseudomonas aeruginosa* O7a,7b,7d	23
→ 4)-α-Pse5(R3Hb)7Fo[b]	96.3	34.8	73.2	46.7	71.2	52.3	67.5	16.3	*Pseudomonas aeruginosa* O7a,7b,7c	23
→ 4)-α-Pse5Ac7(ZFo)-(2 →	97.6	35.3	72.9	47.1	71.7	52.8	68.2	16.9	*Pseudoalteromonas distincta*	27
→ 4)-α-Pse5Ac7(EFo)-(2 →[b]	97.6	35.3	72.9	47.0	70.8	54.2	66.7	16.9	*Pseudoalteromonas distincta*	27
→ 4)-α-Pse5Ac7Fo-(2 →[b]	99.7	35.3	72.2	46.5	71.8	53.5	68.2	17.5	*Pseudomonas aeruginosa* O7a,7b,7d	23
→ 4)-α-Pse5(R3Hb)7Fo-(2 →[b]	99.6	35.6	73.4	46.6	71.8	53.5	68.3	17.7	*Pseudomonas aeruginosa* O7a,7b,7c	23
→ 4)-β-Pse5Ac7Ac-(2 →[b]	102.7	36.0	73.5	45.9	75.0	54.5	69.5	17.6	*Escherichia coli* O136	26
→ 8)-β-Pse5Ac7Ac-(2 →[c]		36.2	67.9	49.3	73.4	53.8	75.0	14.0	*Proteus vulgaris* O39	28
→ 8)-β-Pse5Ac7(R3Hb)-(2 →[c]	101.5	36.8	67.5	49.7	73.4	54.0	74.4	14.4	*Shigella boydii* type 7	24
→ 8)-β-Pse4Ac5Ac7(R3Hb)-(2 →[c]	102.5	34.2	70.7	47.1	74.5	54.0	74.5	14.3	*Shigella boydii* type 7	24
Legionaminic acid (Leg)										
β-Leg5Ac7Ac	97.7	40.9	67.9	54.1	70.8	54.5	68.3	20.4	*Vibrio alginolyticus*	33
→ 4)-β-Leg5R7Ac[c,d]	96.6	36.9	72.7	51.3	70.5	54.0	67.3	20.1	*Acinetobacter baumannii* O24	35
→ 4)-β-Leg5R7Ac-(2 →[c,d]	100.2	37.5	72.7	51.1	70.6	54.0		20.3	*Acinetobacter baumannii* O24	35
→ 4)-α-Leg5Am7Ac-(2 →[e]	∼101	37.6	74.4	54.5	71.1	55.0	67.4	19.1	*Vibrio salmonicida*	34

→ 4)-α-Leg-(2 → [e]	101.4	38.7	72.3	52.3	71.7	55.8	67.5	19.5	*Legionella pneumophila serogroup 1*	68
→ 4)-α-Leg5Ac7Ac-(2 → [e]	101.5	40.0	71.6	51.3	73.5	55.1	68.2	20.0	*Legionella pneumophila serogroup 1*	31
→ 4)-α-Leg5Am7Ac-(2 → [e]	101.8	39.2	71.6	54.2	72.4	55.2	67.7	19.4	*Legionella pneumophila serogroup 1*	31
→ 4)-α-Leg5Am7Ac8Ac-(2 → [e]	101.6	40.6	71.4	55.1	72.4	52.6	70.8	17.3	*Legionella pneumophila serogroup 1*	31
8-Epilegionaminic acid (8eLeg)										
α-8eLeg5Ac7Ac-(2 →	104.0	41.8	69.3	53.8	74.8	54.4	70.1	19.7	*Pseudomonas aeruginosa O12*	30
α-8eLeg5Ac7Am8Ac-(2 →	99.2	42.6	67.3	53.5	73.6	57.7	73.2	17.3	*Shewanella putrefaciens*	39
→ 8)-β-8eLeg5(*R*3Hb)7Ac[c]	96.6	40.7	68.1	53.8	71.6	54.0	73.0	15.6	*Salmonella arizonae O61*	37
→ 8)-β-8eLeg5(4Hb)7Ac[c]	97.2	40.6	68.2	53.5	70.8	53.8	72.0	14.6	*Yersinia ruckeri O1*	36
→ 4)-α-8eLeg5Ac7Ac-(2 → [b]	99.7	42.6	78.3	52.2	75.8	55.6	70.0	20.1	*Shewanella putrefaciens*	39
→ 4)-α-8eLeg5Ac7Am8Ac-(2 → [b]	99.6	42.5	77.0	51.8	74.1	57.8	73.3	17.3	*Shewanella putrefaciens*	39
→ 8)-α-8eLeg5Am7Ac-(2 → [b]	99.2	42.5	68.5	57.4	74.0	55.2	79.4	19.4	*Morganella morganii*	38
→ 8)-α-8eLeg5Ac7Ac-(2 → [c]	104.6	42.3	69.5	53.8	73.5	54.5	73.5	15.0	*Pseudomonas aeruginosa O12*	30
→ 8)-α-8eLeg5(*R*3Hb)7Ac-(2 → [c]	104.7	42.5	69.2	53.6	73.5	54.4	73.5	15.3	*Salmonella arizonae O61*	37
→ 8)-α-8eLeg5(4Hb)7Ac-(2 → [c]		42.8	69.8	54.0	74.0	55.4	74.0	16.0	*Yersinia ruckeri O1*	36
4-Epilegionaminic acid (4eLeg)										
α-4eLeg5Ac7Ac8Ac		39.7	67.5	49.8	69.5	52.2	70.4	17.1	*Legionella pneumophila serogroup 1*	12
β-4eLeg5Ac7Ac8Ac	96.5	37.9	67.4	49.6	66.4	52.3	71.5	17.1	*Legionella pneumophila serogroup 1*	12
→ 4)-α-4eLeg5Am7Ac-(2 → [e]	102.0	40.4	70.3	52.6	69.9	55.8	68.4	19.5	*Legionella pneumophila serogroup 2*	40
→ 4)-α-4eLeg5Am7Ac8Ac-(2 → [e]	101.9	40.5	69.8	52.7	69.8	52.8	70.7	17.4	*Legionella pneumophila serogroup 2*	40

[a]All sugars are in the pyranose form. For abbreviations for the *N*-acyl substituents, see note to Table I. Data of the sugar carboxyl group (C-1) and acyl substituents are not shown.

[b]Substituted with a β-D-aldopyranose.

[c]Substituted with an α-D-aldopyranose.

[d]R is acetyl or (*S*)-3-hydroxybutanoyl.

[e]Substituted with an ald-2-ulosonic acid.

In agreement with the general observation of small α-glycosylation effects in ketosides, an α-$(2 \to 4)$-ketosidic linkage between two legionaminic acid residues caused a relatively small positive α effect of 2.4 ppm on C-4 and small negative β effects of ~ 1 ppm on C-3 and C-5.[31] In 4-epilegionaminic acid, the effects were markedly larger ($+3.4$ ppm on C-4 and -2.9 ppm on C-5[27]) showing a dependence on the orientation of OH-4.

Glycosylation effects of aldopyranoses demonstrated a clear dependence on the anomeric configuration of the glycosylating sugar. For instance, glycosylation of legionaminic acid at position 4 by an α-D-aldopyranose residue caused a relatively small α effect on C-4 ($+4.8$ ppm), a relatively large negative β effect on C-3 (-3.8 ppm), and a small positive β effect on C-5 ($+1.2$ ppm).[34] Glycosylation of 8-epilegionaminic acid at the same position by a β-D-aldopyranose residue resulted in significantly different shifts: $+9.7$ ppm for C-4, $+1.4$ ppm for C-3, and -1.7 ppm for C-5.[39] In pseudaminic acid substituted at position 4 by a β-D-aldopyranose residue, a large negative β effect on C-5 and a smaller β effect on C-3 could be predicted,[70] and the experimental values were approximately -3 and -1 ppm, respectively.[23,26]

An α effect on C-8 resulting from glycosylation of pseudaminic acid at position 8 by an α-D-aldopyranose residue was in the range 3.6–5.5 ppm. A negative β effect on C-9 was relatively large in magnitude (3.4–4.2 ppm) and that on C-7 was also negative and small (~ 1 ppm).[24,28] Glycosylation of 8-epilegionaminic acid at the same position caused similar glycosylation effects when the glycosylating sugar was an α-D-aldopyranose (α effect of 3.4–5.0 ppm on C-8, β effects of -4.7 to -4.8 ppm on C-9, and 0 to -0.4 ppm on C-7[22,37]) but different effects when it was substituted by a β-D-aldopyranose ($+9.3$, -0.3, and $+1.3$ ppm for C-8, C-9, and C-7, respectively[38]). An appreciable γ effect (-1.3 to -1.6 ppm) was observed on C-6 of both pseudaminic and 8-epilegionaminic acids independent of the configuration of the glycosyl group.

The dependence of the ^{13}C NMR glycosylation effects, especially β effects, on the anomeric configuration of a glycosylating aldopyranose implies their dependence also on the relative absolute configuration of the glycone and aglycone.[70,71] As a result, the regularities in the effects of glycosylation of the nonulosonic acids at both pyranose ring (O-4) and the side chain (O-8), together with the effects on C-1 of the glycosylating aldose,[70,71] enabled determination of the absolute configurations at C-6 and C-8, respectively, provided that the general configuration was established. This approach was used for the determination of the absolute configurations of pseudaminic,[23] legionaminic,[34] and 8-epilegionaminic[22] acids.

c. NOE Spectroscopy and Molecular Modeling.—Yet another useful tool for the analysis, including determination of the position of the *N*-acyl groups and the configuration of the nonulosonic acids, is nuclear Overhauser effect (NOE) spectroscopy. The location of an acetimidoyl group at N-5 of legionaminic acid in the oligosaccharide **22** from LPS of *V. salmonicida* (Fig. 6) was proved by correlation in a NOESY spectrum between signals for NH-5 and H-5 of the sugar and between NH-5 and the acetimidoyl methyl group.[34] A NOESY experiment in combination with a ^{1}H,^{13}C heteronuclear multiple bond correlation (HMBC) experiment enabled determination of the structure and stereochemistry of *N*-methylated acetamidino groups at position 5 of legionaminic acid (**19** and **20**) in OPS of *L. pneumophila*[44] (Fig. 5).

An example of a stereochemical assignment using a NOESY experiment was the determination of the axial orientation of H-6 by an NOE correlation between H-4 and H-6, which could hardly be done from the $J_{5,6}$ coupling constant in the nonulosonic acids with an equatorial H-5, as in pseudaminic acid.[26] NOEs observed on preirradiation of protons linked to nitrogen (NH-5 and NH-7) were helpful for the determination of the side-chain configuration and conformation. Thus, the *erythro* configuration of the C-6–C-7 fragment of pseudaminic acid was established by NOEs between NH-5 and H-7 and between NH-7 and H-5 in combination with a relatively large $J_{6,7}$ coupling constant (~ 10 Hz) showing the *trans* relationship between H-6 and H-7 (see earlier). NOE spectroscopy in combination with molecular mechanics calculations was used to confirm the *threo* configuration of the C-6–C-7 fragment of legionaminic acid in LPS of *V. salmonicida*, that is to distinguish between the *glycero*-D-*galacto* and *glycero*-L-*altro* configuration.[34] An observation of an NOE between NH-5 and H-7 and no NOE between NH-5 and NH-7 was in accord with the calculated distances in the former but not in the latter. The experiments could be performed either in an H_2O–D_2O mixture[26] or applied to a fully *O*-acetylated methyl ester in an organic solvent[20] since *O*-acetylation does not significantly change the conformation of the molecules.[18]

Conformational studies of methyl (5,7-diacetamido-2,4,8-tri-*O*-acetyl-3,5,7,9-tetradeoxynon-2-ulopyranos)onates having equatorial OH-4 and NH-5 and different configurations in the side chain showed a *syn* relationship for H-6–H-7 in all four isomers[18] (Fig. 18). This followed from small $^3J_{\text{H-6,H-7}}$ coupling constant values (1–2.5 Hz) and a strong H-6,H-7 correlation in the NOESY spectra. An HMBC experiment revealed a strong H-6,C-8 correlation in the D-*glycero*-L-*altro* and L-*glycero*-L-*altro* isomers (Fig. 18a and b), thus showing a relatively large $^3J_{\text{H-6,C-8}}$ coupling constant value and, correspondingly, the *trans* orientation of H-6 and C-8 around the C-6–C-7 bond in the predominant rotamer.
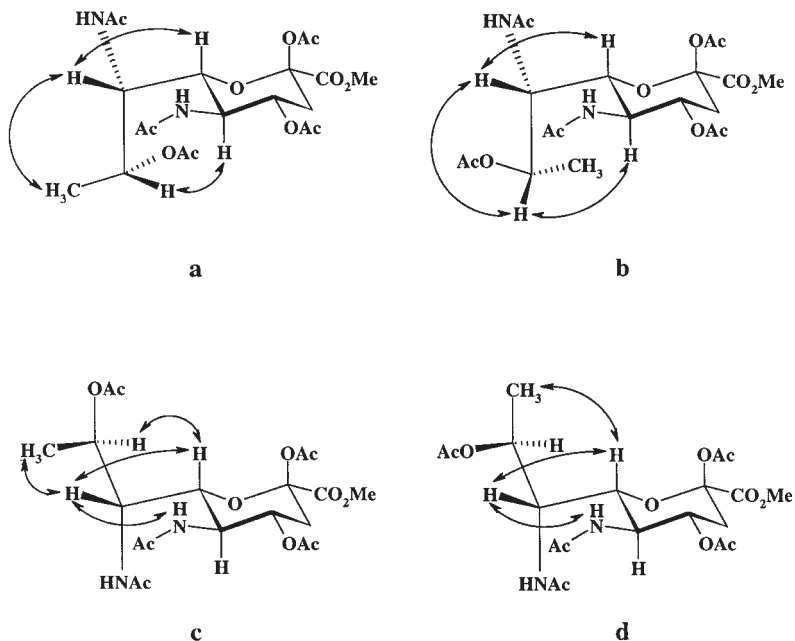
FIG. 18. Selected NOE correlations and conformation of the side chain in methyl (5,7-diacetamido-2,4,8-tri-O-acetyl-3,5,7,9-tetradeoxynon-2-ulopyranos)onates having the D-*glycero*-L-*altro* (a), L-*glycero*-L-*altro* (b), D-*glycero*-D-*galacto* (c), and L-*glycero*-D-*galacto* (d) configuration[18]

The predominant side-chain conformers were different for the four isomers studied (Fig. 18). The most populated rotamer around the C-7–C-8 bond inferred from NOESY data for the D-*glycero*-L-*altro* isomer (Fig. 18a) was in agreement with a small $^3J_{H-7,H-8}$ value of ~ 1 Hz. In contrast, that for the L-*glycero*-L-*altro* isomer (Fig. 18b) was inconsistent with a relatively large $^3J_{H-7,H-8}$ value (5.0 and 6.8 Hz for the α and β anomers, respectively). This contradiction could be accounted for by a significant contribution of a rotamer with a small H-7–C-7–C-8–H-8 dihedral angle. A relatively large $^3J_{H-7,H-8}$ value (~ 7 Hz) and only a weak H-7,H-8 NOE correlation indicated the *trans* orientation of H-7 and H-8 in the D-*glycero*-D-*galacto* and L-*glycero*-D-*galacto* isomers (Fig. 18c and d). NOESY data showed a spatial proximity of H-9 and H-7 in the former (Fig. 18c), whereas in the latter H-9 was more close to H-6 than to H-7 (Fig. 18d).

The conformation of CPS from *S. fredii* HH103 (Fig. 2, structure 5) was investigated using NOE spectroscopy and molecular modeling.[72] The alternating glycosidic and amidic linkages between 3-hydroxybutanoic acid (*R* and *S* isomers) and pseudaminic acid, together with flexibility in the side

chain of pseudaminic acid, should result in a considerable degree of freedom for rotation between the sugar pyranose rings. Coupling constant values observed for the free monosaccharide indicated *trans*- and *syn*-like relationships for H-6−H-7 and H-7−H-8, respectively, resulting in a conformer with an extended sugar side chain and an extended 3-hydroxybutanoic acid. This was true for both the oligosaccharide fragments and the polysaccharide. "Non-trivial" NOEs observed in the polysaccharide between H-4 of 3-hydroxybutanoic acid and H-6 and H-8 of pseudaminic acid and between H-2a,2b of 3-hydroxybutanoic acid and H-6, H-7, and H-9 of pseudaminic acid were in agreement with the *J* data.

Starting from 27 geometries, molecular mechanics calculations and molecular dynamics simulations were made first on the monosaccharide, then on a disaccharide, and finally on a tetrasaccharide and an octasaccharide in order to simulate the conformations of the polysaccharide of *S. fredii* HH103. The calculations included five of the six torsion angles between the pyranose rings and yielded five low-energy minima, which varied only within 3 kJ, that is, they had almost the same energy. In the glycosidic linkage the φ angle was primarily in a staggered conformation (about $60°$), whereas the ψ angle values were around $-120°$ showing a nearly eclipsed disposition between the bonds C-2–O-2 and C-3′–H-3′. This minimized the steric interactions between the sugar ring and both the methyl group and the rest of the side chain. The conformation of the octasaccharide was rather stable and the chain adopted a pseudohelical structure. The average distances between protons found from the molecular dynamics simulations explained satisfactorily most, but not all, of the observed NOEs. A model with alternating (*R*)- and (*S*)-3-hydroxybutanoic acid residues, *SRSR*, could explain the small discrepancies.

IV. Concluding Remarks

Following the discovery of pseudaminic acid in 1984, 5,7-diamino-3,5,7,9-tetradeoxynon-2-ulosonic acid isomers have been found in over 20 bacterial species belonging to different families of Gram-negative bacteria, including enteric bacteria, pseudomonas, vibrios, rhizobia, legionellae, and marine bacteria. The nonulosonic acids have been found primarily in the O-chain polysaccharides of the lipopolysaccharides and several capsular polysaccharides. Recent identification of derivatives of pseudaminic acid in glycoproteins indicate that 5,7-diamino-3,5,7,9-tetradeoxynon-2-ulosonic acids may be more common in bacteria than previously believed.

The novel sugars appear to be constituents of important bacterial cell-surface glycopolymers that contribute to pathogenesis. They are structurally related to sialic acids, which are essential components of animal

glycoproteins and glycolipids. Moreover, initial data on the biosynthesis of 5,7-diamino-3,5,7,9-tetradeoxynon-2-ulosonic acids suggest that it is similar to the sialic acid pathway. The similarity of the bacterial nonulosonic acids to sialic acids may contribute to bacterial virulence by dampening the immune response to invading bacteria.

Recent advances in NMR spectroscopy and mass spectrometry has contributed much to the progress in the chemistry of the nonulosonic acids. Classical chemical approaches, such as various degradation processes and the synthesis of model compounds, were useful in isolation and characterization of 5,7-diamino-3,5,7,9-tetradeoxynon-2-ulosonic acid isomers, including solving the questions of chirality. It is likely we will witness the discovery of new isomers and new derivatives in the near-future. Establishing the configuration of new isomers and the nature and location of acyl groups will become easier as analytical techniques improve and data grow. In addition, more sophisticated conformational analysis will enable studies of the interactions between glycopolymers that contain the nonulosonic acids and other biomolecules.

REFERENCES

1. R. Schauer, *Adv. Carbohydr. Chem. Biochem.*, 40 (1982) 131–234.
2. R. Schauer, J. P. Kamerling J. Montreuil, J. F. G. Vliegenthart, and H. Schachter (Eds.), *Glycoproteins II*, Elsevier, Amsterdam, 1997, pp. 241–400.
3. S. Inoue, Y. Inoue Y. Inoue, Y. C. Lee, and F. A. Troy II (Eds.), *Sialobiology and Other Novel Forms of Glycosylation*, Gakushin Publishing Co., Osaka, 1999, pp. 57–67.
4. F. M. Unger, *Adv. Carbohydr. Chem. Biochem.*, 38 (1981) 323–388.
5. O. Holst, H. Brade, S. M. Opal, S. N. Vogel, and D. C. Morrison (Eds.), *Endotoxin in Health and Disease*, Marcel Dekker, New York, 1999, pp. 115–154.
6. B. Lindberg and S. Dumitriu (Ed.), *Polysaccharides: Structural Diversity and Functional Versatility*, Marcel Dekker, New York, 1998, pp. 237–273.
7. A. S. Shashkov, G. M. Streshinskaya, L. N. Kosmachevskaya, L. I. Evtushenko, and I. B. Naumova, *Mendeleev Commun.*, (2000) 167–168.
8. A. M. Gil-Serrano, M. A. Rodríguez-Carvajal, P. Tejero-Mateo, J. L. Espartero, J. Thomas-Oates, J. E. Ruiz-Sainz, and A. M. Buendía-Clavería, *Biochem. J.*, 334 (1998) 585–594.
9. Y. A. Knirel, E. V. Vinogradov, V. L. L'vov, N. A. Kocharova, A. S. Shashkov, B. A. Dmitriev, and N. K. Kochetkov, *Carbohydr. Res.*, 133 (1984) C5–C8.

10. Y. A. Knirel and N. K. Kochetkov, *FEMS Microbiol. Rev.*, 46 (1987) 381–384.
11. Y. E. Tsvetkov, A. S. Shashkov, Y. A. Knirel, and U. Zähringer, *Carbohydr. Res.*, 331 (2001) 233–237.
12. Y. A. Knirel, H. Moll, J. H. Helbig, and U. Zähringer, *Carbohydr. Res.*, 304 (1997) 77–79.
13. B. L. Reuhs, D. P. Geller, J. S. Kim, J. E. Fox, V. S. K. Kolli, and S. G. Pueppke, *Appl. Environ. Microbiol.*, 64 (1998) 4930–4938.
14. A. M. Gil-Serrano, M. A. Rodríguez-Carvajal, P. Tejero-Mateo, J. L. Espartero, M. Menendez, J. Corzo, J. E. Ruiz-Sainz, and A. M. Buendía-Claveria, *Biochem. J.*, 342 (1999) 527–535.
15. P. Castric, F. J. Cassels, and R. W. Carlson, *J. Biol. Chem.*, 276 (2001) 26479–26485.
16. P. Thibault, S. M. Logan, J. F. Kelly, J.-R. Brisson, C. P. Ewing, T. J. Trust, and P. Guerry, *J. Biol. Chem.*, 276 (2001) 34862–34870.
17. Y. E. Tsvetkov, A. S. Shashkov, Y. A. Knirel, L. V. Backinowsky, and U. Zähringer, *Mendeleev Commun.*, (2000) 90–92.
18. Y. E. Tsvetkov, A. S. Shashkov, Y. A. Knirel, and U. Zähringer, *Carbohydr. Res.*, 335 (2001) 221–243.
19. Here and everywhere *Pseudomonas aeruginosa* O-serogroups are referred to according to the classification of Lányi and Bergan, whereas in the original papers the classification of Lányi was used. See Y. A. Knirel, *CRC Crit. Rev. Microbiol.*, 17 (1990) 273–304.
20. Y. A. Knirel, E. V. Vinogradov, A. S. Shashkov, N. K. Kochetkov, V. L. L'vov, and B. A. Dmitriev, *Carbohydr. Res.*, 141 (1985) C1–C3.
21. Y. A. Knirel, N. A. Kocharova, A. S. Shashkov, and N. K. Kochetkov, *Carbohydr. Res.*, 145 (1986) C1–C4.
22. Y. A. Knirel, E. V. Vinogradov, A. S. Shashkov, B. A. Dmitriev, N. K. Kochetkov, E. S. Stanislavsky, and G. M. Mashilova, *Eur. J. Biochem.*, 157 (1986) 129–138.
23. Y. A. Knirel, N. A. Kocharova, A. S. Shashkov, B. A. Dmitriev, N. K. Kochetkov, E. S. Stanislavsky, and G. M. Mashilova, *Eur. J. Biochem.*, 163 (1987) 639–652.
24. V. L. L'vov, A. S. Shashkov, and B. A. Dmitriev, *Bioorg. Khim.*, 13 (1987) 223–233.
25. L. Kenne, B. Lindberg, E. Schweda, B. Gustafsson, and T. Holme, *Carbohydr. Res.*, 180 (1988) 285–294.
26. M. Staaf, A. Weintraub, and G. Widmalm, *Eur. J. Biochem.*, 263 (1999) 656–661.
27. J. Muldoon, A. S. Shashkov, S. N. Senchenkova, S. V. Tomshich, N. A. Komandrova, L. A. Romanenko, Y. A. Knirel, and A. V. Savage, *Carbohydr. Res.*, 330 (2001) 231–239.
28. A. N. Kondakova, A. V. Perepelov, B. Bartodziejska, A. S. Shashkov, S. N. Senchenkova, M. Wykrota, Y. A. Knirel, and A. Rozalski, *Carbohydr. Res.*, 333 (2001) 241–249.
29. A. V. Pereplov, A. S. Shashkov, V. I. Torgov, E. L. Nazarenko, R. P. Gorshkova, E. P. Ivanova, N. M. Gorshkova, G. Widmalm, unpublished data.
30. Y. A. Knirel, E. V. Vinogradov, A. S. Shashkov, B. A. Dmitriev, N. K. Kochetkov, E. S. Stanislavsky, and G. M. Mashilova, *Eur. J. Biochem.*, 163 (1987) 627–637.
31. Y. A. Knirel, E. T. Rietschel, R. Marre, and U. Zähringer, *Eur. J. Biochem.*, 221 (1994) 239–245.
32. Y. A. Knirel, J. H. Helbig, and U. Zähringer, *Carbohydr. Res.*, 283 (1996) 129–139.
33. E. L. Nazarenko, A. S. Shashkov, Y. A. Knirel, E. P. Ivanova, and Y. S. Ovodov, *Bioorg. Khim.*, 16 (1990) 1426–1429.
34. P. Edebrink, P.-E. Jansson, J. Bøgwald, and J. Hoffman, *Carbohydr. Res.*, 287 (1996) 225–245.
35. S. R. Haseley and S. G. Wilkinson, *Eur. J. Biochem.*, 250 (1997) 617–623.
36. L. M. Beynon, J. C. Richards, and M. B. Perry, *Carbohydr. Res.*, 256 (1994) 303–317.
37. E. V. Vinogradov, A. S. Shashkov, Y. A. Knirel, N. K. Kochetkov, J. Dabrowski, H. Grosskurth, E. S. Stanislavsky, and E. V. Kholodkova, *Carbohydr. Res.*, 231 (1992) 1–11.

38. M. Kilcoyne, A. S. Shashkov, S. N. Senchenkova, Y. A. Knirel, E. V. Vinogradov, J. Radziejewska-Lebrecht, R. Galimska-Stypa, and A. V. Savage, *Carbohydr. Res.*, 337 (2002) 1697–1702.

39. A. S. Shashkov, V. I. Torgov, E. L. Nazarenko, V. A. Zubkov, N. M. Gorshkova, R. P. Gorshkova, G. Widmalm, *Carbohydr. Res.*, 337 (2002), 1119–1127.

40. Y. A. Knirel, S. N. Senchenkova, N. A. Kocharova, A. S. Shashkov, J. H. Helbig, and U. Zähringer, *Biochemistry (Moscow)*, 66 (2001) 1035–1041.

41. Nomenclature of Carbohydrates, *Carbohydr. Res.*, 297 (1997) 18.

42. O. Kooistra, E. Lüneberg, B. Lindner, Y. A. Knirel, M. Frosch, and U. Zähringer, *Biochemistry*, 40 (2001) 7630–7640.

43. E. Lüneberg, U. Zähringer, Y. A. Knirel, D. Steinmann, M. Hartmann, I. Steinmetz, M. Rohde, J. Köhl, and M. Frosch, *J. Exp. Med.*, 188 (1998) 49–60.

44. O. Kooistra, E. Lüneberg, Y.A. Knirel, M. Frosch, and U. Zähringer, *Eur. J. Biochem.*, 269 (2002) 560–572.

45. Y. A. Knirel, H. Grosskurth, J. H. Helbig, and U. Zähringer, *Carbohydr. Res.*, 279 (1995) 215–226.

46. N. Hashii, Y. Isshiki, T. Iguchi, K. Hisatsum, and S. Kondo, *Cardohydr. Res.*, 338 (2003) 1055–1062.

47. E. Lüneberg, N. Zetzmann, D. Alber, Y. A. Knirel, O. Kooistra, U. Zähringer, and M. Frosch, *Int. J. Microbiol. Med.*, 290 (2000) 37–49.

48. J. Parkhill, B. W. Wren, and K. Mungall, *Nature*, 403 (2000) 665–668.

49. H. L. Rocchetta, L. L. Burrows, and J. S. Lam, *Microbiol. Mol. Biol. Rev.*, 63 (1999) 523–553.

50. E. Lüneberg, B. Mayer, N. Daryab, O. Kooistra, U. Zähringer, M. Rohde, J. Swanson, and M. Frosch, *Mol. Microbiol.*, 39 (2001) 1259–1271.

51. C. H. Zou, Y. A. Knirel, J. H. Helbig, U. Zähringer, and C. S. Mintz, *J. Bacteriol.*, 181 (1999) 4137–4141.

52. P. C. Lück, T. Freier, C. Steudel, Y. A. Knirel, E. Lüneberg, U. Zähringer, and J. H. Helbig, *Int. J. Med. Microbiol.*, 291 (2001) 345–352.

53. C. S. Mintz, unpublished data.

54. D. J. Brenner, A. G. Steigerwalt, P. Epple, W. F. Bibb, R. M. McKinney, R. W. Starnes, J. M. Colville, R. K. Selander, P. H. Edelstein, C. W. Moss, *J. Clin. Microbiol.*, 26 (1988) 1695–1703.

55. J. H. Helbig, J. B. Kurtz, M. C. Pastoris, C. Pelaz, and P. C. Lück, *J. Clin. Microbiol.*, 35 (1997) 2841–2845.

56. J. H. Helbig, P. C. Lück, Y. A. Knirel, and U. Zähringer, *Epidemiol. Infect.*, 115 (1995) 71–78.

57. O. Kooistra, L. Herfurth, E. Lüneberg, M. Frosch, T. Peters, and U. Zähringer, *Eur. J. Biochem.*, 269 (2002) 573–582.

58. B. Lányi, M. M. Adam, and S. Vörös, *Acta Microbiol. Acad. Sci. Hung.*, 19 (1972) 259–265.

59. B. Lányi, S. Vörös, and M. M. Adam, *Acta Microbiol. Acad. Sci. Hung.*, 20 (1973) 249–254.

60. U. Zähringer, Y. A. Knirel, B. Lindner, J. H. Helbig, A. Sonesson, R. Marre, and E. T. Rietschel, *Prog. Clin. Biol. Res.*, 392 (1995) 113–139.

61. C. S. Mintz and C. H. Zou, *FEMS Microbiol. Lett.*, 93 (1992) 249–254.

62. D. Linton, A. V. Karlyshev, P. G. Hitchen, H. R. Morris, A. Dell, N. A. Gregson, and B. W. Wren, *Mol. Microbiol.*, 35 (2000) 1120–1134.

63. M. J. How, M. D. A. Halford, M. Stacey, and E. Vickers, *Carbohydr. Res.*, 11 (1969) 313–320.

64. K. Čapek, J. Čapkova, J. Jary, Y. A. Knirel, and A. S. Shashkov, *Coll. Czech. Chem. Commun.*, 52 (1987) 2248–2259. See also references cited therein.

65. A. Banaszek, Z. Pakulski, and A. Zamojski, *Carbohydr. Res.*, 279 (1995) 173–182.
66. C. Altona and C. A. G. Haasnoot, *Org. Magn. Reson.*, 13 (1980) 417–433.
67. K. Bock and C. Pedersen, *Adv. Carbohydr. Chem. Biochem.*, 41 (1983) 27–66.
68. O. Kooistra, Ph. D. Thesis. Medical University of Lübeck, Germany, 2001.
69. P.-E. Jansson, L. Kenne, and E. Schweda, *J. Chem. Soc., Perkin Trans. 1*, (1987) 377–383.
70. A. S. Shashkov, G. M. Lipkind, Y. A. Knirel, and N. K. Kochetkov, *Magn. Reson. Chem.*, 26 (1988) 735–747.
71. P.-E. Jansson, L. Kenne, and G. Widmalm, *Carbohydr. Res.*, 188 (1989) 169–191.
72. M. A. Rodriguez-Carvajal, M. Bernabe, J. L. Espartero, P. Tejero-Mateo, A. Gil-Serrano, and J. Jiménez-Barbero, *J. Mol. Graphics Mod.*, 18 (2000) 135–142.