



Article

# The Complete Structure of the Core Oligosaccharide from *Edwardsiella tarda* EIB 202 Lipopolysaccharide

Marta Kaszowska <sup>1,\*</sup>, Elena de Mendoza-Barberá <sup>2,†</sup>, Anna Maciejewska <sup>1</sup>, Susana Merino <sup>2</sup>, Czeslaw Lugowski <sup>1,3</sup> and Juan M. Tomás <sup>2</sup>

<sup>1</sup> Department of Immunochemistry, Hirszfeld Institute of Immunology and Experimental Therapy, Polish Academy of Sciences, R. Weigla 12, PL-53-114 Wrocław, Poland;

aniaaugustyniuk@iitd.pan.wroc.pl (A.M.); lugowski@iitd.pan.wroc.pl (C.L.)

<sup>2</sup> Department of Microbiology, University of Barcelona, Diagonal 643, 08071 Barcelona, Spain; edemendoza@ub.edu (E.d.M.-B.); smerino@ub.edu (S.M.); jtomas@ub.edu (J.M.T.)

<sup>3</sup> Department of Biotechnology and Molecular Biology, University of Opole, PL-45-035 Opole, Poland

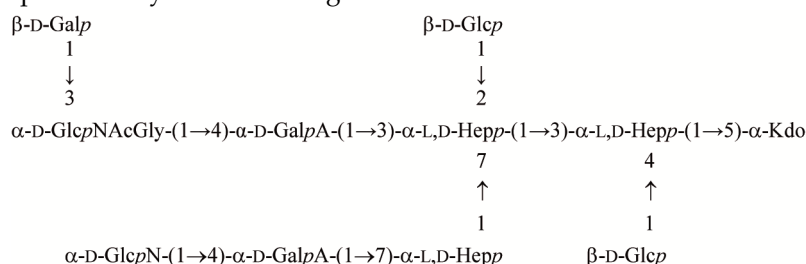
\* Correspondence: marta.kaszowska@iitd.pan.wroc.pl; Tel.: +48-71-370-99-27

† These authors contributed equally to this work.

Academic Editor: William Chi-shing Cho

Received: 18 February 2017; Accepted: 24 May 2017; Published: 31 May 2017

**Abstract:** The chemical structure and genomics of the lipopolysaccharide (LPS) core oligosaccharide of pathogenic *Edwardsiella tarda* strain EIB 202 were studied for the first time. The complete gene assignment for all LPS core biosynthesis gene functions was acquired. The complete structure of core oligosaccharide was investigated by <sup>1</sup>H and <sup>13</sup>C nuclear magnetic resonance (NMR) spectroscopy, electrospray ionization mass spectrometry MS<sup>n</sup>, and matrix-assisted laser-desorption/ionization time-of-flight mass spectrometry. The following structure of the undecasaccharide was established:



The heterogeneous appearance of the core oligosaccharide structure was due to the partial lack of  $\beta$ -D-Galp and the replacement of  $\alpha$ -D-GlcpNAcGly by  $\alpha$ -D-GlcpNGly. The glycine location was identified by mass spectrometry.

**Keywords:** *Edwardsiella tarda*; core oligosaccharide; MALDI-TOF MS; ESI MS<sup>n</sup>; NMR; genomic

## 1. Introduction

*Edwardsiella tarda* is a Gram-negative bacterium and a pathogen of farmed fish. It is the etiological agent of a systematic disease called edwardsiellosis, which has been reported to affect a wide range of freshwater and marine fish [1,2]. In addition to fish, *E. tarda* is also an occasional human pathogen and known to cause both gastroenteritis and extraintestinal infections in humans [3,4]. A number of virulence-associated systems and factors, such as the type III and type VI secretion systems, the LuxS/AI-2 quorum sensing system, and hemolysin systems, have been identified in *E. tarda* [5]. Additionally, a sialidase shows a potential pathogenicity and immunogenicity [6].

In Gram-negative bacteria, the lipopolysaccharide (LPS) is one of the major structural and immunodominant molecules of the outer membrane. It consists of three moieties: lipid A, core

oligosaccharide, and O-specific polysaccharide (O-antigen). The O-antigen is the external component of LPS, and its structure consists of different number of repeating units. The O-specific polysaccharide chains are transferred to lipid A-core to form LPS, in a step involving WaaL, the putative bifunctional enzyme named O-antigen ligase. Another interesting feature is the high chemical variability shown by the O-antigen LPS, leading to a similar genetic variation in the genes involved in their biosynthesis, the so-called *wb* cluster (for a review, see [7]). Despite the emerging importance of this pathogenic microorganism, until now only four LPS structures of *E. tarda* strains were investigated [8–11].

In studies of several *Enterobacteriaceae* such as *Escherichia coli*, *Salmonella enterica*, and *Klebsiella pneumoniae*, genes involved in LPS core biosynthesis are usually found clustered in a region of the chromosome, the *waa* gene cluster [12,13]. On the other hand, a careful analysis of several full sequenced genomes suggested that genes for the LPS core biosynthesis may not be clustered and may be distributed between several regions, e.g., as in *Yersinia pestis* [14] or *Proteus mirabilis* [15]. In other cases, only a single gene involved in LPS core biosynthesis is out of the *waa* gene cluster, for instance, *Plesiomonas shigelloides* [16]. Nothing is known about the genomics or the LPS core structure from any *E. tarda* strain, besides the role played by the *waaL* (O-antigen ligase) characterized from strain EIB 202 [17]. *E. tarda* strain EIB 202 was isolated from moribund fish *Scophthalmus maximum* in a marine culture farm in China [18], and its full genome sequenced [19].

Here, the chemical structure of the core oligosaccharide in a pathogenic strain of *E. tarda* EIB 202 to proceed with the genomics of the core biosynthesis is reported for the first time.

## 2. Results

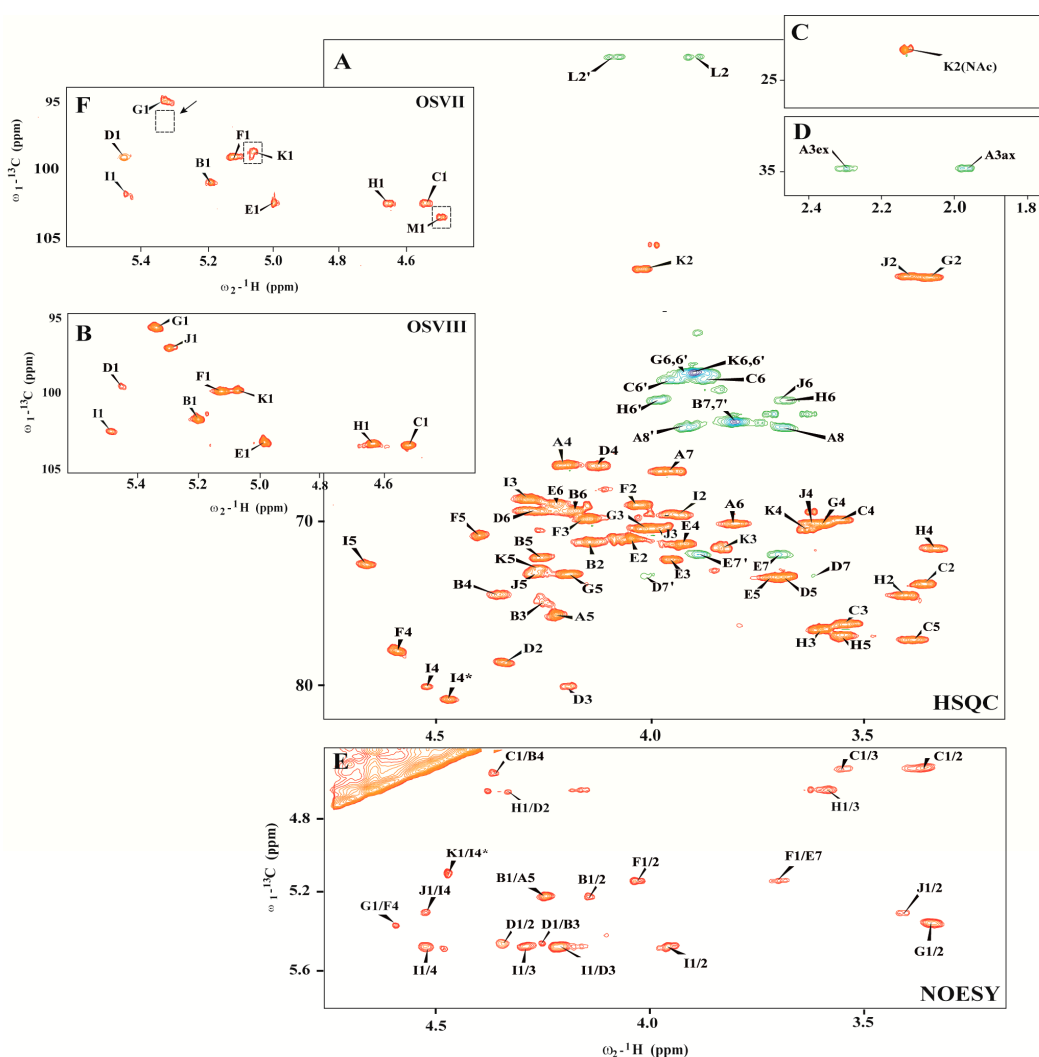
### 2.1. Isolation of the Core Oligosaccharide

LPS of *E. tarda* EIB 202 was isolated from bacterial mass with a yield of 0.5%. The mild acid hydrolysis of the LPS yielded eight polysaccharide (PS) and oligosaccharide (OS) fractions: PSI–VI consisting of a core oligosaccharide substituted by several repeating units, and OSVII and OSVIII—the unsubstituted core oligosaccharide fractions. The high yield of PSI–VI suggested the smooth (S-LPS) type of *E. tarda* EIB 202 LPS. The data presented herein concern the OSVIII fraction. The differences between OSVII and OSVIII fractions are presented herein based on MALDI-TOF MS (matrix assisted laser desorption/ionization-time of flight mass spectrometry) and ESI MS<sup>n</sup> (electrospray ionization mass spectrometry) analysis.

### 2.2. Structure Analysis of the Core Oligosaccharide Fractions

The chemical analyses of OSVIII showed the presence of 2,3,7-trisubstituted L,D-Hepp, 3,4-disubstituted L,D-Hepp, 7-substituted L,D-Hepp, two terminal D-GlcpN, two terminal D-Glcp, two 4-substituted D-GalpA, and 5-substituted Kdop. The analyses of OSVII showed the presence of monosaccharides identified for OSVIII and additionally two sugar residues: the terminal D-Galp, and 3-substituted D-GlcpNAc was identified instead of the terminal D-GlcpN in OSVIII.

The <sup>1</sup>H NMR (nuclear magnetic resonance) spectrum of the OSVIII contained main signals for nine anomeric protons, and signals characteristic for the deoxy protons of Kdop residue belongs to part of the core oligosaccharide (residues A–J). The HSQC-DEPT (heteronuclear single-quantum correlation-distortionless enhancement by polarization transfer) spectra obtained for the OSVIII fraction contained signals for nine major anomeric protons and carbons, and Kdo spin systems, respectively (Figure 1 and Table 1).



**Figure 1.** (A–D) Selected regions of the  $^1\text{H}$ - $^{13}\text{C}$  HSQC-DEPT (heteronuclear single-quantum correlation–distortionless enhancement by polarization transfer) and (E)  $^1\text{H}$ - $^1\text{H}$  NOESY (nuclear overhauser spectroscopy) spectra of the fraction OSVIII of *Edwardsiella tarda* EIB 202 lipopolysaccharide (LPS); (F) anomeric region of the  $^1\text{H}$ - $^{13}\text{C}$  HSQC-DEPT spectrum of the fraction OSVII with marked difference in comparison with OSVIII. The cross-peaks are labeled as shown in the text.

Residue A was identified as the 5-substituted Kdo on the basis of characteristic deoxy proton signals at  $\delta_{\text{H}}$  1.96 ppm (H-3 $_{ax}$ ) and  $\delta_{\text{H}}$  2.29 ppm (H-3 $_{eq}$ ), as well as a high chemical shift of the C-5 signal ( $\delta_{\text{C}}$  75.7 ppm). Residue B ( $\delta_{\text{H}}/\delta_{\text{C}}$  5.19/101.7 ppm,  $^1J_{\text{C-1,H-1}} \sim 175$  Hz) was recognized as the 3,4-disubstituted L-glycero- $\alpha$ -D-manno-Hepp residue on the basis of the small vicinal couplings between H-1, H-2, and H-3 and relatively high chemical shifts of the C-3 ( $\delta_{\text{C}}$  74.7 ppm) and the C-4 ( $\delta_{\text{C}}$  74.4 ppm) signals. Residue C ( $\delta_{\text{H}}/\delta_{\text{C}}$  4.51/103.3 ppm,  $^1J_{\text{C-1,H-1}} \sim 163$  Hz) and residue H ( $\delta_{\text{H}}/\delta_{\text{C}}$  4.63/103.2 ppm,  $^1J_{\text{C-1,H-1}} \sim 166$  Hz) were recognized as the  $\beta$ -D-Glcp on the basis of the large vicinal couplings between all protons in the sugar ring. Residue D ( $\delta_{\text{H}}/\delta_{\text{C}}$  5.44/99.6 ppm,  $^1J_{\text{C-1,H-1}} \sim 175$  Hz) was recognized as the 2,3,7-trisubstituted L-glycero- $\alpha$ -D-manno-Hepp residue from the  $^1\text{H}$  and  $^{13}\text{C}$  chemical shift values, small vicinal couplings between H-1, H-2, and H-3 and relatively high chemical shift values of the C-2 ( $\delta_{\text{C}}$  78.6 ppm), C-3 ( $\delta_{\text{C}}$  80.0 ppm), and C-7 ( $\delta_{\text{C}}$  73.3 ppm) signals. Residue E ( $\delta_{\text{H}}/\delta_{\text{C}}$  4.98/103.2 ppm,  $^1J_{\text{C-1,H-1}} \sim 173$  Hz) was recognized as the 7-substituted L-glycero- $\alpha$ -D-manno-Hepp from the  $^1\text{H}$  and  $^{13}\text{C}$  chemical shifts, the small vicinal couplings between H-1, H-2, and H-3 and the relatively high chemical shift value of the C-7 ( $\delta_{\text{C}}$  72.0 ppm) signal. Residues F ( $\delta_{\text{H}}/\delta_{\text{C}}$  5.12/99.8 ppm,  $^1J_{\text{C-1,H-1}} \sim 175$  Hz) was

recognized as the 4-substituted  $\alpha$ -D-GalpA based on the characteristic five proton spin system, the high chemical shifts of the H-3 ( $\delta_{\text{H}}$  4.16 ppm), H-4 ( $\delta_{\text{H}}$  4.62), H-5 ( $\delta_{\text{H}}$  4.41), and C-4 ( $\delta_{\text{C}}$  77.9 ppm) signals, the large vicinal couplings between H-2 and H-3 and small vicinal coupling between H-3, H-4, and H-5. Residue **I** ( $\delta_{\text{H}}/\delta_{\text{C}}$  5.47/102.5 ppm,  $^1J_{\text{C-1,H-1}} \sim 174$  Hz) was also recognized as the 4-substituted  $\alpha$ -D-GalpA residue based on the similar characteristic five proton spin system. Residues **G** ( $\delta_{\text{H}}/\delta_{\text{C}}$  5.33/95.6 ppm,  $^1J_{\text{C-1,H-1}} \sim 176$  Hz) and **J** ( $\delta_{\text{H}}/\delta_{\text{C}}$  5.29/97.0 ppm,  $^1J_{\text{C-1,H-1}} \sim 176$  Hz) were recognized as the terminal  $\alpha$ -D-GlcpN due to the large coupling between H-1, H-2, and H-3 and the small vicinal coupling between H-3, H-4, and H-5, as well as the chemical shift value of the C-2 ( $\delta_{\text{C}}$  55.1 and  $\delta_{\text{C}}$  55.1 for **G** and **J**, respectively). The 1D  $^{31}\text{P}$  NMR spectrum showed no indication of phosphate groups in the OSVIII.

Additionally, the residue **K** ( $\delta_{\text{H}}/\delta_{\text{C}}$  5.06/99.8 ppm,  $^1J_{\text{C-1,H-1}} \sim 165$  Hz) was recognized as the terminal  $\alpha$ -D-GlcpNAc from a low  $^{13}\text{C}$  chemical shift of the C-2 signal ( $\delta_{\text{C}}$  54.6 ppm), and the large vicinal couplings between all ring protons. The *N*-acetyl group at  $\delta_{\text{H}}/\delta_{\text{C}}$  2.13/23.2 ppm ( $\delta_{\text{C}}$  175.9 ppm) was identified. The presence of heterogeneity in OSVIII was due to partial replacement of  $\alpha$ -D-GlcpN (residue **J**) by  $\alpha$ -D-GlcpNAc (residue **K**). The last sugar residue **M** ( $\delta_{\text{H}}/\delta_{\text{C}}$  4.45/103.6 ppm,  $^1J_{\text{C-1,H-1}} \sim 163$  Hz), identified only in OSVII, was recognized as the terminal  $\beta$ -D-Galp due to the large vicinal couplings between H-1, H-2, and H-3 and the small vicinal couplings between H-3, H-4, and H-5. The terminal residue **M** in OSVII is linked to C-3 of  $\rightarrow 3$ - $\alpha$ -D-GlcpNAc (residue **K**) corresponding to the terminal form of this residue in the OSVIII.

In the HSQC-DEPT spectra of OSVIII (at  $\delta_{\text{H}}/\delta_{\text{C}}$  3.90/41.8 ppm), additional negative  $\text{CH}_2$  signals were detected. These resonances showed correlation with a carbonyl carbon signals at  $\delta_{\text{C}}$  168.0 ppm in the HMBC (heteronuclear multiple bond correlation) spectra, suggesting the presence of glycine (residue **L**). This residue was also confirmed by mass spectrometry.

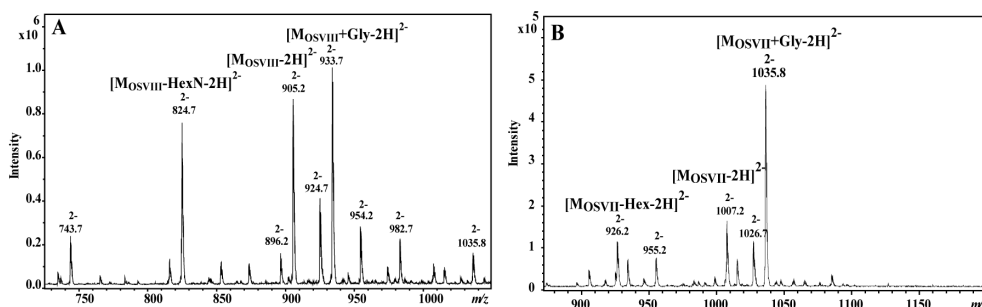
The monosaccharide sequence in OSVIII was established using a NOESY (nuclear overhauser spectroscopy) and HMBC experiments. NOESY spectra showed strong inter-residue cross-peaks between the following transglycosidic protons: H-1 of **B**/H-5 of **A**, H-1 of **D**/H-3 of **B**, H-1 of **C**/H-4 of **B**, H-1 of **E**/H-7,7' of **D**, H-1 of **F**/H-7,7' of **E**, H-1 of **G**/H-4 of **F**, H-1 of **H**/H-2 of **D**, H-1 of **I**/H-3 of **D**, and H-1 of **J**/H-4 of **I**. (Figure 1E). The HMBC spectrum of OSVIII confirmed substitution positions of all of the monosaccharide residues (data not shown). Additionally, NOESY showed the cross-peak between H-1 of **K** and H-4 of **I**\*, and it also provided evidence for the heterogeneity of OSVIII with the presence of  $\alpha$ -D-GlcpNAc and as a substitution of 4-substituted  $\alpha$ -D-GalpA (residue **I**) in OSVIII.

The OSVIII and OSVII fractions were analyzed by ES MS<sup>n</sup> and MALDI-TOF MS/MS. Ten sugar residues: two Glc, three Hep, two GalA, two GlcN, and Kdo, give together a monoisotopic mass of 1812.567 Da ( $M_{\text{OSVIII}}$ ). Eleven sugar residues from OSVII give together a monoisotopic mass of 2016.630 Da. The negative ESI MS<sup>n</sup> (Figure 2A) mass spectrum of OSVIII showed the main ion at  $m/z$  905.2 [ $M_{\text{OSVIII}}-2\text{H}$ ]<sup>2-</sup> correspond to core oligosaccharide (OSVIII), and ions correspond to core substituted with the glycine (Gly) residue at  $m/z$  933.7 [ $M_{\text{OSVIII}}+\text{Gly}-2\text{H}$ ]<sup>2-</sup> and  $m/z$  924.7 [ $M_{\text{OSVIII}}+\text{Gly}-\text{H}_2\text{O}-2\text{H}$ ]<sup>2-</sup>. The negative ESI MS<sup>n</sup> mass spectrum of OSVII showed the main ion at  $m/z$  1035.8 [ $M_{\text{OSVII}}+\text{Gly}-2\text{H}$ ]<sup>2-</sup>, which represented the complete core oligosaccharide substituted by Gly (Figure 2B).

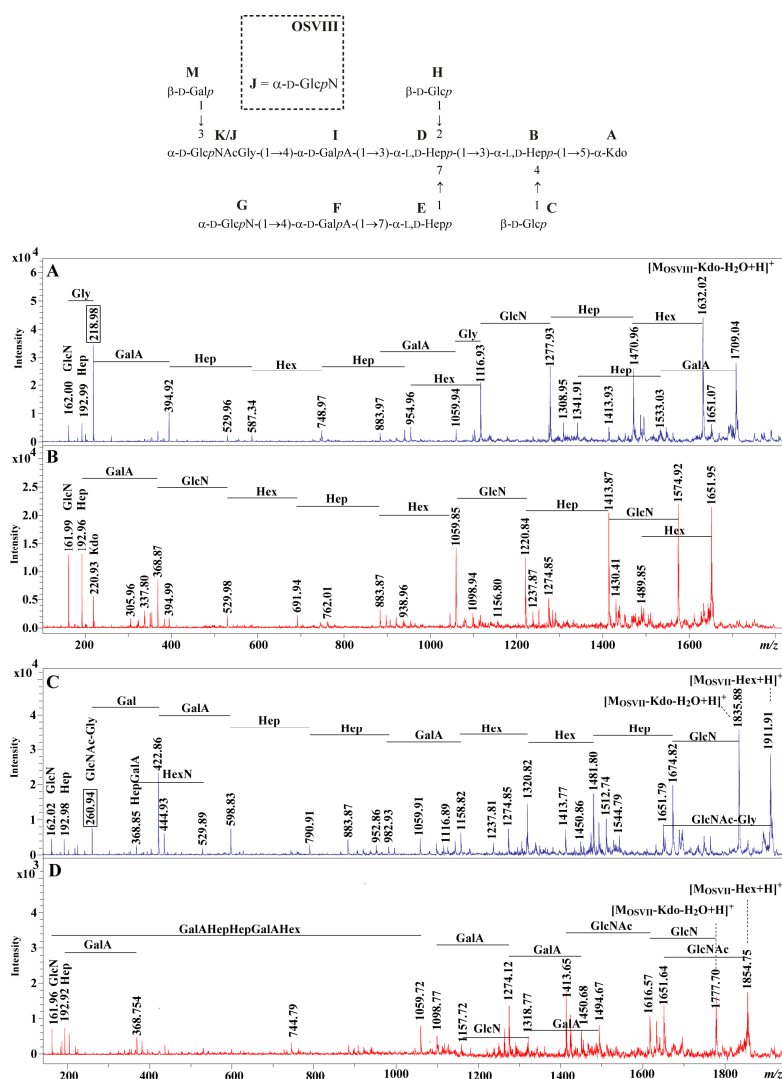
**Table 1.** <sup>1</sup>H and <sup>13</sup>C NMR (nuclear magnetic resonance) chemical shifts of the core oligosaccharide of *E. tarda* EIB 212 LPS.

Residues	Oligosaccharide		Chemical Shifts (ppm)							
	OSVIII	OSVII	H1/C1	H2/C2	H3(ax,eq)/C3 H4/C4		H5/C5	H6,6'/C6	H7,7'/C7	H8,8'/C8 (NAc)
A →5)-Kdo	*	*	- nd	- 97.7	1.96, 2.29 34.7	4.21 66.5	4.25 75.7	3.81 70.1	3.97 66.9	3.69, 3.92 64.3
B →3,4)-L-glycero-α-D-manno-Hepp-(1→	*	*	5.19 101.7	4.15 71.3	4.25 74.7	4.35 74.4	4.26 72.2	4.18 69.4	3.80 63.9	
C →β-D-Glcp-(1→	*	*	4.51 103.3	3.36 73.9	3.55 76.2	3.56 69.9	3.55 76.2	3.87, 3.95 61.5		
D →2,3,7)-L-glycero-α-D-manno-Hepp-(1→	*	*	5.44 99.6	4.34 78.6	4.19 80.0	4.12 66.6	3.67 73.4	4.28 69.3	3.62, 4.01 73.3	
E →7)-L-glycero-α-D-manno-Hepp-(1→	*	*	4.98 103.2	4.05 71.1	3.95 72.3	3.93 71.4	3.72 73.4	4.23 69.1	3.69, 4.88 72.0	
F →4)-α-D-GalpA-(1→	*	*	5.12 99.8	4.05 69.0	4.16 69.9	4.62 77.9	4.41 70.9	176.7		
G α-D-GlcpN-(1→	*	*	5.33 95.6	3.34 55.1	4.02 70.4	3.60 76.6	4.19 73.2	3.91 <sup>a</sup> 60.0		
H α-D-Glcp-(1→	*	*	4.63 103.2	3.40 74.5	3.56 76.3	3.33 71.6	3.60 76.5	3.69, 3.98 62.7		
I →4)-α-D-GalpA-(1→	*	*	5.47 102.5	3.93 69.6	4.26 72.2	4.48 80.9	4.67 72.5	- 175.6		
J α-D-GlcpN-(1→	*		5.29 97.0	3.39 55.1	3.99 70.5	3.63 70.5	4.27 72.8	3.69, 3.97 62.5		
K α-D-GlcpNAc-(1→		*	5.06 99.8	4.02 54.6	3.83 71.6	3.63 70.2	4.28 73.2	3.90 <sup>a</sup> 61.0		2.13 23.2
L Gly	*	*	168.0	3.90, 4.09 41.8						
M →β-D-Galp-(1→		*	4.45 103.6	3.61 71.0	3.73 72.7	3.93 71.4	3.59 75.1	3.74, 3.80 63.2		

ax: Axial position; eq: Equatorial position; nd: Not resolved.



**Figure 2.** ESI (electrospray ionization) mass spectra of the core oligosaccharide fractions (A) OSVIII and (B) OSVII of *E. tarda* EIB 202.



**Figure 3.** MALDI-TOF (matrix assisted laser desorption/ionization-time of flight) MS/MS fragmentation mass spectra of the ions (A) at  $m/z$  1870.16  $[M_{OSVIII}+Gly+H]^+$ , (B) at  $m/z$  1813.15  $[M_{OSVIII}+H]^+$ , (C) at  $m/z$  2074.19  $[M_{OSVII}+Gly+H]^+$ , and (D) at  $m/z$  2017.19  $[M_{OSVII}+H]^+$  of *E. tarda* EIB 202, differing in the presence of glycine.

The location of glycine was determined by the positive ion mode MALDI-TOF MS/MS. The ions at  $m/z$  1813.15  $[M_{OSVIII}+H]^+$ , at  $m/z$  1870.16  $[M_{OSVIII}+Gly+H]^+$ , at  $m/z$  2017.19  $[M_{OSVII}+H]^+$ , and at

$m/z$  2074.19 [ $M_{OSVIII}+Gly+H$ ]<sup>+</sup> were selected for further fragmentations by the use of positive ion mode MALDI-TOF MS/MS. The main daughter ions detected in the MALDI-TOF MS/MS spectra were explained. In Figure 3A, the ion at  $m/z$  162.00 corresponds to GlcN, while the ion at  $m/z$  218.98 was explained by the GlcN–Gly. The similar pair of fragment ions at  $m/z$  1116.93 and at  $m/z$  1059.94 with the mass difference corresponding to the glycine residue was also identified. These ions were not identified on the spectrum of the ion  $m/z$  1813.15 fragmentation (Figure 3B). In Figure 3C, the daughter ion at  $m/z$  260.94 was subsequently attributed to the GlcNAc–Gly fragment. The similar pair of fragment ions at  $m/z$  1911.91 and at  $m/z$  1651.79 with the mass difference corresponding to the glycine residue was also identified. These ions were not identified on the spectrum of the ion at  $m/z$  2017.19 fragmentation (Figure 3D). These observations indicate that glycine substitutes GlcN (residue J) in OSVIII and GlcNAc (residue K) in OSVII. The positions of glycine in OSVIII and OSVII were not determined.

### 2.3. Organization of the *E. tarda* Strain EIB 202 *waa* Gene Cluster

In most *Enterobacteriaceae* studied so far, the genes involved in core LPS biosynthesis were found clustered (*waa* gene cluster). When we inspected the currently available *E. tarda* strain EIB 202 genome, we found a clear region with the *waa* gene cluster (proteins encoded ETAE\_0083 to ETAE\_0072). This *waa* region, like in the majority of *Enterobacteriaceae*, is started by the *hldE* (encoded protein ETAE\_0083), which codifies for the ADP-L-glycero-D-mannoheptose-6-epimerase, and the end flanked by the *coaD* (encoded protein ETAE\_0071) codifying for phosphopantetheine adenylyltransferase [20].

Despite the genome annotation, it seems that more of the genes are shared by different *Enterobacteriaceae* mainly *K. pneumoniae* or *P. shigelloides*, which were previously characterized by us [7,14]. Table 2 shows proteins encoded from *E. tarda* EIB 202 *waa*.

**Table 2.** Characteristics of the proteins encoded from *E. tarda* EIB 202 *waa*.

Protein	Homologus Protein	Predicted Function	% Identity/Similarity
ETAE_0083	HldE <i>Enterobacteriaceae</i>	ADP-L-glycero-D-manno Heptose-6-epimerase	85/91
ETAE_0082	WaaF <i>Enterobacteriaceae</i>	ADP-heptose:LPS heptosyl transferase II	76/86
ETAE_0081	WaaC <i>Enterobacteriaceae</i>	ADP-heptose:LPS heptosyl transferase I	78/84
ETAE_0080	WaaL <i>Klebsiella pneumoniae</i>	O-antigen ligase	29/46
ETAE_0079	Similar only to WabK <i>Klebsiella pneumoniae</i>	unknown	34/53
ETAE_0078	WapC <i>Plesiomonas shigelloides</i>	UDP-galacturonic a transferase $\alpha(1\rightarrow7)$ to HepIII acid	76/89
ETAE_0077	WapB <i>Plesiomonas shigelloides</i>	UDP-N-acetyl glucosamine $\alpha(1\rightarrow4)$ to GalAII	61/82
ETAE_0076	WabN <i>Klebsiella pneumoniae</i>	Protein deacetylase	77/89
ETAE_0075	WaaQ <i>Klebsiella pneumoniae</i>	ADP-heptose: LPS heptosyl	72/83
ETAE_0074	WabG <i>Klebsiella pneumoniae</i>	UDP-galacturonic acid transferase $\alpha(1\rightarrow3)$ to HepIII	78/87
ETAE_0073	WabH <i>Klebsiella pneumoniae</i>	UDP-N-acetyl glucosamine transferase $\alpha(1\rightarrow4)$ to GalAI	70/83
ETAE_0072	WaaA <i>Enterobacteriaceae</i>	3-deoxy-D-manno-octulosonic acid transferase	88/96
ETAE_0071	WapE <i>Plesiomonas shigelloides</i>	UDP-galactose transferase $\beta(1\rightarrow4)$ to l-HepI	82/97
ETAE_0070	CoaD <i>Enterobacteriaceae</i>	Phosphopantetheine adenylyltransferase	81/89

### 3. Discussion

Here, the chemical structure and genomics of the complete undecasaccharide core structure of *E. tarda* EIB 202 LPS are presented for the first time. This core oligosaccharide is heterogeneous. The heterogeneity corresponded to the partial lack of  $\beta$ -D-Galp and the replacement of  $\alpha$ -D-GlcpNAcGly by  $\alpha$ -D-GlcpNGly. The functions of the genes found in the *waa* gene cluster from the *E. tarda* strain EIB 202 seems to be in agreement with the chemical structure of the LPS-core. The *E. tarda* core LPS structure is highly similar to that of *K. pneumoniae* at least up to the outer-core residue GlcNI and *P. shigelloides* 302-73 in practically up to the last monosaccharide residue that links the O-antigen LPS [7,14]. WabH and WapB are enzymes that transfer GlcNAc to a GalA in different acceptor substrates of LPS-core in an  $\alpha(1\rightarrow4)$  linkage. WabG and WapC are enzymes that transfer GalA to a Hep also in different LPS-core substrates and with different linkage,  $\alpha(1\rightarrow3)$  and  $\alpha(1\rightarrow7)$ , respectively. It is important to note that besides performing the same enzymatic functions the acceptor substrate differences determine that the enzymes showed very little homology; furthermore, WabG and WapC are more similar among them (26 identity and 47% similarity) than WabH and WapB are (24% identity and 46% similarity), except that

the latter ones showed identical linkage. This point indicates the importance of the substrates in the enzymatic reactions to build-up the LPS-core molecules. *K. pneumoniae* WabK is a glycosyltransferase that incorporates a Glc residue in a  $\beta(1\rightarrow4)$  to GlcN. *E. tarda* waa ORF4, according to the *E. tarda* strain EIB 202 LPS-core established, as well as their low homology but unique to WabK, could be the galatosyltransferase that incorporates a Gal residue in a  $\beta(1\rightarrow4)$  to GlcNAc. *E. tarda* waa Orf5 encoding for ETAE\_0079 (Table 2). All genes from the *E. tarda* waa cluster were found in the *E. ictaluri* genomes, except for *wapB* and *wapC*, which seems to be unique for the species *E. tarda*. All of the other genes from the *E. tarda* waa cluster (*hldE*, *waaF*, *waaC*, ETAE\_0079, *waaN*, *waaQ*, *wabG*, *wabH*, *waaA*, and *wapE*) show 98% or more identity to the related *E. ictaluri* LPS-core biosynthetic genes according to their genomes. Of course, *E. ictaluri* waaL, which is the O-antigen ligase, is a bifunctional enzyme recognizing the O-antigen LPS and the LPS-core, as it usual shows a reduced identity (56%) compared to *E. tarda* waaL. Nevertheless, the *E. ictaluri* waaL-encoded protein shows the typical transmembrane domains (data not shown).

The *E. tarda* LPS motif  $\beta$ -Glc-(1 $\rightarrow$ 2)- $\alpha$ -L-HepII seems not to be encoded by any of the glycosyltransferases found in the waa cluster. This LPS motif is identical to a previously studied by us in the *P. shigelloides* strain 302-73 encoded by WapG. For this reason, we decided to blastx the *P. shigelloides* 302-73 WapG [16] against the *E. tarda* strain EIB 202 genome. We found a clear unique candidate, the gene encoding ETAE\_1955, which showed 58% identity and 74% similarity to WapG. We suggest that it could be responsible for  $\beta$ -Glc-residue linked to HepII-[(1 $\rightarrow$ 2)- $\alpha$ -L-HepII].

## 4. Materials and Methods

### 4.1. Growth Conditions and Isolation of the Lipopolysaccharide and the Polysaccharide

Bacteria *E. tarda* EIB 202 was obtained from the Y. Zhang laboratory [18]. The bacteria were grown and harvested as described previously [21]. The LPS was extracted from bacterial cells of *E. tarda* EIB 202 by the hot phenol/water method [22]. LPS (200 mg) was degraded by treatment with 1.5% acetic acid at 100 °C for 45 min. The supernatant was fractionated on a column (1.6  $\times$  100 cm) of Bio-Gel P-10, equilibrated with 0.05 M pyridine/acetic acid buffer, pH 5.4. Eluates were monitored with a Knauer differential refractometer, and all fractions were checked by NMR spectroscopy and mass spectrometry (MALDI-TOF and ESI MS<sup>n</sup>).

### 4.2. Chemical Methods

Methylation of oligosaccharide fractions was performed according to the method described by Ciucanu and Kerek [23]. The absolute configurations of the monosaccharides were determined as described by Gerwig et al. [24]. Alditol acetates and partially methylated alditol acetates were analyzed by gas chromatography GC-MS with the Thermo Scientific TSQ system using an RX5 fused-silica capillary column (0.2 mm by 30 m) and a temperature program of 150  $\rightarrow$  270 °C at 12 °C/min.

### 4.3. Instrumental Methods

All NMR spectra were recorded on a Bruker Avance III 600 MHz spectrometer equipped with a 5 mm QCI cryoprobe with z-gradient. The measurements were performed at 303 K without sample spinning and using the acetone signal ( $\delta_{\text{H}}/\delta_{\text{C}}$  2.225/31.05 ppm) as an internal reference. The signals were assigned by one- and two-dimensional experiments: <sup>1</sup>H-<sup>1</sup>H COSY (correlation spectroscopy), TOCSY (total correlated spectroscopy), NOESY, <sup>1</sup>H-<sup>13</sup>C HSQC-DEPT, HSQC-TOCSY, and HMBC. In the TOCSY experiments, the mixing times were 30, 60, and 100 ms. The NOESY experiment was performed with the mixing time of 200 ms, and HMBC experiment with a delay of 80 ms. For observation of phosphorus atoms, one-dimensional <sup>31</sup>P NMR spectra were recorded. The data were acquired and processed using standard Bruker software. The processed spectra were assigned with the help of the SPARKY program [25].



Core oligosaccharide (1 mg/mL in mQ) fractions were analyzed using a MALDI-TOF Ultraflextreme (Bruker, Germany) instrument. The MALDI-TOF MS spectra were obtained in a positive ion mode. 2,5-Dihydroxybenzoic acid (10 mg/mL in 1:1 AcN/0.2 M citric acid [*v/v*]) was used as a matrix for analyses.

Negative-ion electrospray ionization mass spectra (ESI-MS<sup>n</sup>) were recorded using an Amazon SL (Bruker Daltonics, Bremen, Germany) ion trap mass spectrometer. The samples were dissolved in a 50 µg/mL acetonitrile-water-formic acid solution (50:50:0.5 [*v/v/v*]).

#### 4.4. Comparative Genomics

For each analyzed genome we gathered all coding sequence (CDS) and pseudo-CDS information by parsing NCBI GenBank records. When we obtained the UniProt Knowledge Base records for these loci using the cross-reference with Entrez GeneIDs and parsed them for gene names, functional annotations, and associated COG, PFAM, and TIGRFAM protein domains were studied. To annotate orthologs, we wrote custom scripts to analyze reference sequence alignments made to subject genomes with blastn and tblastn via NCBI's web application programming interface.

**Acknowledgments:** This work was partially funded by BIO2016-80329-P from the Spanish Ministerio de Economía y Competitividad and by the Generalitat de Catalunya (Centre de Referència en Biotecnologia). We thank Maite Polo for her technical assistance and the Servicios Científico-Técnicos from the University of Barcelona, Spain. Publication supported by Wrocław Centre of Biotechnology, programme The Leading National Research Centre (KNOW) for years 2014-2018.

**Author Contributions:** Marta Kaszowska, Elena de Mendoza-Barberá, Susana Merino and Juan M. Tomás designed the research. Marta Kaszowska, Anna Maciejewska, and Elena de Mendoza-Barberá performed the experiments. Marta Kaszowska, Czesław Lugowski, Susana Merino and Juan M. Tomás wrote the manuscript. All authors analyzed the data.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Matsuyama, T.; Kamaishi, T.; Ooseko, N.; Kurohara, K.; Iida, T. Pathogenicity of motile and non-motile *Edwardsiella tarda* to some marine fish. *Fish Pathol.* **2005**, *40*, 133–135. [[CrossRef](#)]
2. Mohanty, B.R.; Sahoo, P.K. Edwardsiellosis in fish: A brief review. *J. Biosci.* **2007**, *32*, 1331–1344. [[CrossRef](#)] [[PubMed](#)]
3. Janda, J.M.; Abbott, S.L. Infections associated with the genus *Edwardsiella*: The role of *Edwardsiella tarda* in human disease. *Clin. Infect. Dis.* **1993**, *17*, 742–748. [[PubMed](#)]
4. Nelson, J.J.; Nelson, C.A.; Carter, J.E. Extraintestinal manifestations of *Edwardsiella tarda* infection: A 10-year retrospective review. *J. La. State Med. Soc.* **2009**, *161*, 103–106. [[PubMed](#)]
5. Leung, K.Y.; Siame, B.A.; Tenkink, B.J.; Noort, R.J.; Mok, Y.K. *Edwardsiella tarda* virulence mechanisms of an emerging gastroenteritis pathogen. *Microbes Infect.* **2012**, *14*, 26–34. [[CrossRef](#)] [[PubMed](#)]
6. Jin, R.P.; Hu, Y.H.; Sun, B.G.; Zhang, X.H.; Sun, L. *Edwardsiella tarda* sialidase: Pathogenicity involvement and vaccine potential. *Fish Shellfish Immunol.* **2012**, *33*, 514–521. [[CrossRef](#)] [[PubMed](#)]
7. Aquilini, E.; Tomás, J.M. (Eds.) *Lipopolysaccharides (Endotoxins)*; Reference Module in Biomedical Sciences; Elsevier: Amsterdam, The Netherlands, 2015. [[CrossRef](#)]
8. Katzenellenbogen, E.; Kocharova, N.A.; Shashkov, A.S.; Gorska-Fraczek, S.; Gamian, A.; Knirel, Y.A. Structure of the O-polysaccharide of *Edwardsiella tarda* PCM 1156. *Carbohydr. Res.* **2013**, *477*, 45–48. [[CrossRef](#)] [[PubMed](#)]
9. Katzenellenbogen, E.; Kocharova, N.A.; Shashkov, A.S.; Gorska-Fraczek, S.; Bogulska, M.; Gamian, A.; Knirel, Y.A. Structure of the O-polysaccharide of *Edwardsiella tarda* PCM 1150 containing an amide of D-glucuronic acid with L-alanine. *Carbohydr. Res.* **2013**, *347*, 84–88. [[CrossRef](#)] [[PubMed](#)]
10. Katzenellenbogen, E.; Kocharova, N.A.; Toukach, P.V.; Gorska, S.; Bogulska, M.; Gamian, A.; Knirel, Y.A. Structures of a unique O-polysaccharide of *Edwardsiella tarda* PCM 1153 containing an amide of galacturonic acid with 2-aminopropane-1,3-diol and an abequose-containing O-polysaccharide shared by *E. tarda* PCM 1145, PCM 1151 and PCM 1158. *Carbohydr. Res.* **2012**, *355*, 56–62. [[CrossRef](#)] [[PubMed](#)]

11. Vinogradov, E.; Nossova, L.; Perry, M.B.; Kay, W.W. Structural characterization of the O-polysaccharide antigen of *Edwardsiella tarda* MT 108. *Carbohydr. Res.* **2005**, *340*, 85–90. [[CrossRef](#)] [[PubMed](#)]
12. Heinrichs, D.E.; Yethon, J.A.; Whitfield, C. Molecular basis for structural diversity in the core regions of the lipopolysaccharides of *Escherichia coli* and *Salmonella enterica*. *Mol. Microbiol.* **1998**, *30*, 221–232. [[CrossRef](#)] [[PubMed](#)]
13. Regué, M.; Climent, N.; Abitiu, N.; Coderch, N.; Merino, S.; Izquierdo, L.; Altarriba, M.; Tomás, J.M. Genetic characterization of the *Klebsiella pneumoniae waa* gene cluster, involved in core lipopolysaccharide biosynthesis. *J. Bacteriol.* **2001**, *183*, 3564–3573. [[CrossRef](#)] [[PubMed](#)]
14. Knirel, Y.A.; Dentovskaya, S.V.; Bystrova, O.V.; Kocharova, N.A.; Senchenkova, S.N.; Shaikhutdinova, R.Z.; Titareva, G.M.; Bakhteeva, I.V.; Lindner, B.; Pier, G.B.; et al. Relationship of the lipopolysaccharide structure of *Yersinia pestis* to resistance to antimicrobial factors. *Adv. Exp. Med. Biol.* **2007**, *603*, 88–96. [[PubMed](#)]
15. Aquilini, E.; Azevedo, J.; Jimenez, N.; Bouamama, L.; Tomás, J.M.; Regué, M. Functional Identification of the *Proteus mirabilis* core lipopolysaccharide biosynthetic genes. *J. Bacteriol.* **2010**, *192*, 4413–4424. [[CrossRef](#)] [[PubMed](#)]
16. Aquilini, E.; Merino, S.; Regué, M.; Tomás, J.M. Genomic and proteomic studies of *Plesiomonas shigelloides* lipopolysaccharide core biosynthesis. *J. Bacteriol.* **2014**, *196*, 556–567. [[CrossRef](#)] [[PubMed](#)]
17. Xu, L.; Wang, Q.; Xiao, J.; Liu, Q.; Wang, X.; Chen, T.; Zhang, Y. Characterization of *Edwardsiella tarda waaL*: Roles in lipopolysaccharide biosynthesis, stress adaptation, and virulence toward fish. *Arch. Microbiol.* **2010**, *192*, 1039–1047. [[CrossRef](#)] [[PubMed](#)]
18. Xiao, J.F.; Wang, Q.Y.; Liu, Q.; Wang, X.; Liu, H.; Zhang, Y.X. Isolation and identification of fish pathogen *Edwardsiella tarda* from mariculture in China. *Aquac. Res.* **2008**, *40*, 13–17. [[CrossRef](#)]
19. Wang, Q.; Yang, M.; Xiao, J.; Wu, H.; Wang, X.; Lu, Y.; Xu, L.; Zheng, H.; Wang, S.; Zhao, G.; et al. Genome sequence of the versatile fish pathogen *Edwardsiella tarda* provides insights into its adaptation to broad host ranges and intracellular niches. *PLoS ONE* **2009**, *4*, e7646. [[CrossRef](#)] [[PubMed](#)]
20. Geerlof, A.; Lexendon, A.; Shaw, V. Purification and characterization of phosphopantetheine adenylyltransferase from *Escherichia coli*. *J. Biol. Chem.* **1999**, *274*, 27105–27111. [[CrossRef](#)] [[PubMed](#)]
21. Kaszowska, M.; Jachymek, W.; Jachymek, W.; Lukaszewicz, J.; Niedziala, T.; Kenne, L.; Lugowski, C. The unique structure of complete lipopolysaccharide isolated from semi-rough *Plesiomonas shigelloides* O37 (strain CNCTC 39/89) containing (2S)-O-(4-oxopentanoic acid)- $\alpha$ -D-Glcp ( $\alpha$ -D-lenose). *Carbohydr. Res.* **2013**, *378*, 98–107. [[CrossRef](#)] [[PubMed](#)]
22. Westphal, O.; Jann, K. Bacterial lipopolysaccharide: Extraction with phenol-water and further application of the procedure. *Methods Carbohydr. Chem.* **1965**, *5*, 83–89.
23. Ciucanu, I.; Kerek, F. A simple and rapid method for the permethylation of carbohydrates. *Carbohydr. Res.* **1984**, *131*, 209–217. [[CrossRef](#)]
24. Gerwig, G.J.; Kamerling, J.P.; Vliegthart, J.F.G. Determination of the D and L configuration of neutral monosaccharides by high-resolution capillary g.l.c. *Carbohydr. Res.* **1978**, *62*, 349–357. [[CrossRef](#)]
25. Goddard, T.D.; Kneller, D.G. SPARKY 3; University of California: San Francisco, CA, USA, 2001. Available online: <https://www.cgl.ucsf.edu/home/sparky/> (accessed on 30 May 2017).

