Carbohydrate Research 404 (2015) 108–116

Contents lists available at ScienceDirect

Carbohydrate Research

journal homepage: www.elsevier.com/locate/carres

Structural analysis of the O-acetylated O-polysaccharide isolated from *Salmonella paratyphi* A and used for vaccine preparation

N. Ravenscroft [a], P. Cescutti [b], M. Gavini [c], G. Stefanetti [c], C. A. MacLennan [c], L. B. Martin [c], F. Micoli [c,*]

[a] Department of Chemistry, University of Cape Town, Rondebosch 7701, South Africa
[b] Department of Life Sciences, Blg. C11, Università di Trieste, via L. Giorgieri 1, 34127 Trieste, Italy
[c] Novartis Vaccines Institute for Global Health, Via Fiorentina 1, I-53100 Siena, Italy

ARTICLE INFO

Article history:
Received 23 October 2014
Received in revised form 10 December 2014
Accepted 13 December 2014
Available online 20 December 2014

Keywords:
Salmonella paratyphi A
O-Polysaccharide
Bacterial polysaccharide structure
O-Acetylation

ABSTRACT

Salmonella paratyphi A is increasingly recognized as a common cause of enteric fever cases and there are no licensed vaccines against this infection. Antibodies directed against the O-polysaccharide of the lipopolysaccharide of *Salmonella* are protective and conjugation of the O-polysaccharide to a carrier protein represents a promising strategy for vaccine development. O-Acetylation of *S. paratyphi* A O-polysaccharide is considered important for the immunogenicity of S. *paratyphi* A conjugate vaccines.

Here, as part of a programme to produce a bivalent conjugate vaccine against both *S. typhi* and *S. paratyphi* A diseases, we have fully elucidated the O-polysaccharide structure of *S. paratyphi* A by use of HPLC–SEC, HPAEC–PAD/CD, GLC, GLC–MS, 1D and 2D-NMR spectroscopy. In particular, chemical and NMR studies identified the presence of *O*-acetyl groups on C-2 and C-3 of rhamnose in the lipopolysaccharide repeating unit, at variance with previous reports of O-acetylation at a single position. Moreover HR-MAS NMR analysis performed directly on bacterial pellets from several strains of *S. paratyphi* A also showed O-acetylation on C-2 and C-3 of rhamnose, thus this pattern is common and not an artefact from O-polysaccharide purification. Conjugation of the O-polysaccharide to the carrier protein had little impact on O-acetylation and therefore should not adversely affect the immunogenicity of the vaccine.

1. Introduction

Salmonella enterica serovars are responsible for invasive disease in the developing world. Most cases of enteric fever are caused by *Salmonella enterica* serovar *typhi* (*S. typhi*). The annual global burden of disease due to typhoid fever was estimated at approximately 22 million illnesses in 2000 with a case-fatality rate of 1% resulting in more than 200,000 deaths.[1] The highest incidence was reported in South Asia,[1] and in pre-school and school-aged children.[2–4] An additional 5.4 million cases of enteric fever in 2000 were caused by *Salmonella enterica* serovar *paratyphi* A (*S. paratyphi* A), with children mainly affected.[1] A more recent revised estimate of the global burden of typhoid and paratyphoid fever has reported 26.9 million and 4.9 million adjusted cases of *S. typhi* and *S. paratyphi* A, respectively, in 2010.[5] *S. paratyphi* A is an increasingly common cause of enteric fever in the Indian subcontinent and Southeast Asia, with data suggesting it may account for up to 50% of enteric fever cases.[6]

There are licensed vaccines against *S. typhi*,[7] but they do not offer protection against enteric fever caused by *S. paratyphi* A. The increasing frequency of *S. paratyphi* A as a cause of enteric fever and the emergence of antibiotic resistance[8] are of great concern, particularly because there are no licensed vaccines against *S. paratyphi* A infection. Furthermore, enteric fever caused by *S. paratyphi* A and *S. typhi* is clinically indistinguishable. As a result, in areas with high incidence of paratyphoid fever there is the risk that even a highly effective *S. typhi* vaccine will be poorly efficacious against enteric fever as a whole. To address enteric fever more broadly, a combination vaccine that protects against both *S. typhi* and *S. paratyphi* A would be a valuable public health tool.

The O-polysaccharide of the lipopolysaccharide (LPS) of *S. paratyphi* A is a potentially protective antigen and therefore the target for vaccine development. Phase 1 and 2 clinical studies with an O-polysaccharide-TT vaccine[9] were conducted 14 years ago in Vietnamese adults and children and found to be safe and immunogenic.[10] The Chengdu Institutes of Biological Products in China together with the Lanzhou Institute are currently conducting

* Corresponding author. Tel.: +39 0577 539087; fax: +39 0577 243540.
E-mail address: francesca.micoli@novartis.com (F. Micoli).

http://dx.doi.org/10.1016/j.carres.2014.12.002
0008-6215/

a Phase 2 trial of this vaccine.[6,11] Other O-polysaccharide glycoconjugates using DT and CRM_{197} have been developed and tested in preclinical studies by the International Vaccine Institute (IVI)[6] and Novartis Vaccines Institute for Global Health (NVGH), respectively.[12] Both have been developed alongside Vi conjugate vaccines in order to be used in bivalent combinations and protect against both aetiologic agents of enteric fever.

Salmonella LPS consists of lipid A linked to the 3-deoxy-D-manno-octulosonic acid (Kdo) terminus of a conserved core region, which is linked to the variable O-polysaccharide chain. In *S. paratyphi* A, the chain repeating unit consists of a trisaccharide backbone composed of rhamnose (Rha), mannose (Man) and galactose (Gal), with a terminal paratose (Par) at C-3 of Man (which confers serogroup specificity: factor 2) and terminal glucose (Glc) at C-6 of Gal (Fig. 1).

A common feature shared by various O-polysaccharides is the presence of decorations such as *O*-acetyl groups. It has been reported that not only the presence but also the position of *O*-acetyl groups in the O-polysaccharide may influence the immune response.[13] The *S. paratyphi* A O-polysaccharide was reported to require *O*-acetyl groups in order to elicit serum LPS antibody with bactericidal activity in mice.[9] In the case of *S. paratyphi* A, a single site of O-acetylation has been reported, on C-3 of the Rha residue as part of the repeating O-polysaccharide chain[9,10,14,15] or on C-2 of the same monomer.[16] However, from analysis of the ^{13}C NMR spectrum, Konadu et al. suggested the presence of two different *O*-acetyl groups, the second one on the Par residue, akin to that found on C-2 of abequose for *S. typhimurium*.[17] The authors highlighted the need for further investigation given the important role that *O*-acetyl groups seem to play in immunogenicity.[9,10]

In the O-polysaccharide purification strategy adopted by NVGH, the labile linkage between the Kdo of the core and the lipid A is cleaved directly in the bacterial growth medium;[18] the O-polysaccharide, still linked to the core, is then easily purified from the cells and can be used for the production of conjugate vaccines. For conjugation to CRM_{197}, the *S. paratyphi* O-polysaccharide is linked to the carrier protein through the terminal Kdo, with the aim to preserve the O-polysaccharide antigen structure.[12] Here the O-polysaccharide structure has been elucidated by use of HPLC–SEC, HPAEC–PAD/CD, GLC, GLC–MS and 1D and 2D NMR studies. Heterogeneity due to non-stoichiometric glucosylation and O-acetylation resulted in complex NMR spectra and simpler structures were chemically derived to facilitate full assignment of the native O-polysaccharide structure. In particular, the presence of *O*-acetyl groups on both C-2 and C-3 of Rha was identified. The same pattern of O-acetylation was found in the O-polysaccharide-CRM_{197} conjugate and on the surface carbohydrate of *S. paratyphi* A bacteria. This confirmed that the purification and the conjugation processes do not alter the structure of the O-polysaccharide antigen, including retention of *O*-acetyl groups considered important for the immunogenicity of the *S. paratyphi* A vaccine.

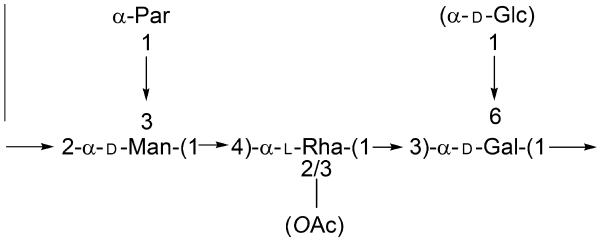

Figure 1. Structure of the repeating unit of *Salmonella paratyphi* A O-polysaccharide showing the positions of O-acetylation. The structural features that vary with the strain of origin are shown in curved brackets.

2. Results and discussion

2.1. Structural analysis of OPS

Purified O-polysaccharide (**OPS**) showed one main population with an average molecular mass of 40–45 kDa.[12,18] Sugar composition analysis of the **OPS** by GLC–MS analysis of the alditol acetate derivatives showed the presence of Rha, Man, Gal, Glc and a component eluting before Rha in the molar ratio 1.00:1.00:1.00:0.76:0.17, respectively. The early eluting component was identified by GLC–MS as a 3,6-dideoxy-hexose, ascribed to the expected Par, which was partially destroyed during acid hydrolysis. In the absence of a commercially available monomer standard, Par was quantified by ^1H NMR analysis which yielded the expected molar ratio of 1:1 with respect to Rha from the ratio of the corresponding H-6 resonances. Incomplete glucosylation is in agreement with HPAEC–PAD data (74%) previously reported.[18] GLC analysis of the chiral glycosides of the **OPS** showed that the hexoses were in the D absolute configuration and Rha in the L absolute configuration. As before, the absence of a commercial standard meant that analysis for Par could not be performed, and its D configuration was established by NMR glycosylation shifts. According to GLC and GLC–MS of the partially-methylated alditol acetate (PMAA) derivatives, **OPS** contains terminal non-reducing 3,6 dideoxyhexose, identified as Par by NMR (t-Par), 4-linked Rha (4-Rha), terminal non-reducing glucose (t-Glc), 3-linked Gal (3-Gal), 2,3-linked Man (2,3-Man) and 3,6-linked-Gal (3,6-Gal) in the molar ratio 0.16:∼0.94:∼0.71:0.30:1.00:0.67. These data are consistent with the **OPS** containing a mixture of ∼70% of the pentasaccharide repeating unit (containing Glc-(1→6)-Gal) and ∼30% of the tetrasaccharide repeating unit (without terminal glucose).[18] In addition to glucosylation, another source of the structural heterogeneity of the **OPS** is due to variable O-acetylation. Our earlier studies on several lots indicated a total *O*-acetyl content of 65–80% per repeating unit determined by ^1H NMR following in-tube de-O-acetylation using sodium deuteroxide. In this study, the total O-acetylation levels determined by HPAEC–CD were in agreement with values found by ^1H NMR and ranged from 65% to 75% per repeating unit for **OPS** lots evaluated.

The ^1H NMR spectrum of the **OPS** confirmed the presence of the two deoxy sugars with diagnostic signals for the H-3s of Par and H-6 of Rha (split) and Par, but the anomeric region could not be readily assigned due to overlap and the presence of multiple peaks. The complexity of the **OPS** spectrum was attributed to the twinning of some signals from incomplete glucosylation and the presence of both the pentasaccharide (major) and tetrasaccharide (minor) repeating units. This was further exacerbated by additional twinning of signals due to partial O-acetylation (shown by signals at 2.20 and 2.18 ppm) and the presence of small peaks from the core region and process residuals. Elucidation of these complex NMR spectra was achieved after full assignment of the ^1H and ^{13}C NMR spectra of simpler saccharides derived from **OPS**.

2.2. Characterization of OPS Smith degradation products

Complete periodate oxidation of **OPS**, followed by reduction with $NaBH_4$ and mild acid hydrolysis (Smith degradation) yielded an oligosaccharide (**OPS-S1**) as the major product. Full assignment of **OPS-S1** aided in the NMR interpretation of the polymer obtained after partial Smith degradation to reduce the level of glucosylation (**OPS-S2**), which in turn facilitated elucidation of the de-O-acetylated **OPS** (**OPS-D**) spectra and finally the spectra obtained for **OPS**.

The permethylated **OPS-S1** gave the mass spectrum (Fig. S1A) containing three ions at *m*/*z* 718.4, 723.4 and 739.4 corresponding to the parent adducts $[M+NH_4]^+$, $[M+Na]^+$, and $[M+K]^+$, respectively.

The MS^2 spectrum gave information on the sequence (Fig. S1B) as a trisaccharide Hex-[3,6-dideoxyHex]-Hex- attached to 1-deoxy-erythritol (1dEry-ol), derived from oxidation between C-2 and C-3 of the rhamnosyl ring. Detailed assignment of the ions is given in Table S1. Identification of **OPS-S1** as α-D-Gal*p*-(1→2)-[α-D-Par*p*-(1→3)]-α-D-Man*p*-(1→2-1dEry-ol followed from detailed 1D and 2D NMR analysis (Supplementary data). A second Smith degradation using partial periodate oxidation of **OPS** was performed in order to selectively remove the t-Glc residue and thereby increase the proportion of the tetrasaccharide repeating unit in the **OPS** preparation. Composition analysis by GLC of the alditol acetates derivatives gave Rha/Man/Gal/Glc in the relative molar ratios 0.66:1.00:1.03:0.44, thus indicating a successful partial removal of the t-Glc, but also some depletion of the Rha residue. As Rha is a backbone residue, its decrease would be associated with some in-chain cleavage and result in a lower molecular mass of the sample. Therefore, the partial Smith degradation product was separated on a Sephacryl S-300 size exclusion column and only the fractions with the higher molecular mass were pooled together to yield the polymeric **OPS-S2** sample containing lower levels of glucosylation.

2.3. Assignment of the NMR spectra of OPS-D

The ^{1}H NMR spectra of **OPS-S1**, **OPS-S2** and **OPS-D** are shown in Figure S2 and the corresponding ^{13}C NMR spectra in Figure S3. Elucidation of the sugar spin systems for the pentasaccharide repeating unit (major component, RU5, residues in upper case) and tetrasaccharide repeating unit (minor component, RU4, residues in lower case) for **OPS-D** was achieved using the 1D and 2D NMR experiments previously applied to **OPS-S1** (Supplementary data). The NMR study was hampered by the presence of heterogeneity leading to signal overlap, and although the terminal Glc and Par residues gave relatively sharp signals (and their spin systems were readily assigned), those from the backbone residues, particularly the 2,3-linked Man, gave broad signals and poor crosspeaks in the 2D NMR spectra. These problems were overcome by extensive use of hybrid experiments (HSQC-TOCSY and HSQC-NOESY) and by repeating the NMR study using a high field spectrometer (600 MHz).

Comparison of the 1D ^{1}H NMR spectra of **OPS-S2** and **OPS-D** (Fig. S2B and C) confirmed that **OPS-S2** contains a lower level of Glc compared to **OPS-D** (~25% vs ~80%). The largest visible effect of glucosylation was experienced by the anomeric signals of Man [5.30 ppm (C-1 at 100.3 ppm) in RU4 shifts to 5.18 ppm (C-1 at 100.7 ppm) in RU5] and this is the main source of spectral complexity in the anomeric region (see HSQC expansion, Fig. 2A). Once this was realized, the five major spin systems were identified by tracing correlations starting from H-1 (all), H-3 (Par) and H-6 (Par and Rha) resonances in COSY and TOCSY experiments and the corresponding carbons through use of HSQC, HSQC-TOCSY and HMBC experiments. The ^{1}H and ^{13}C NMR data are collected in Table 1. Ambiguities in tracing out the Man (**M**) spin systems were solved as for **OPS-S1**, by the HMBC crosspeaks from H-1 Man at 5.18 ppm to C-2 (small), C-3 and C-5 and the linkage shown by the crosspeak from H-1 to C-4 Rha (82.1 ppm). The spin system for the Man (**m**) from RU4 showed the same carbon correlations from H-1 at 5.30 ppm, indicating that the effect of glucosylation is mainly experienced by the anomeric position. The twinning of the H-6/C-6 crosspeaks for Rha due to glucosylation (1.32:17.7 ppm for **r** in RU4 and 1.34:17.9 ppm for **R** in RU5) provided a useful window for elucidation of these spin systems using the HSQC-TOCSY and HMBC experiments. Here the twinning of signals was detectable for H-5, H-4 and H-1/C-1 (Table 1 and Fig. 2). The α-L-Rha*p*-(1→3)-α-D-Gal*p* linkage (not present in **OPS-S1**) was shown by the HMBC crosspeak between H-1 Rha and C-3 of Gal at 77.9 ppm (partly overlapped with the H-1 Par/C-3 Man crosspeak at 77.8 ppm) and confirmed by the H-1/C-1 HSQC-NOESY

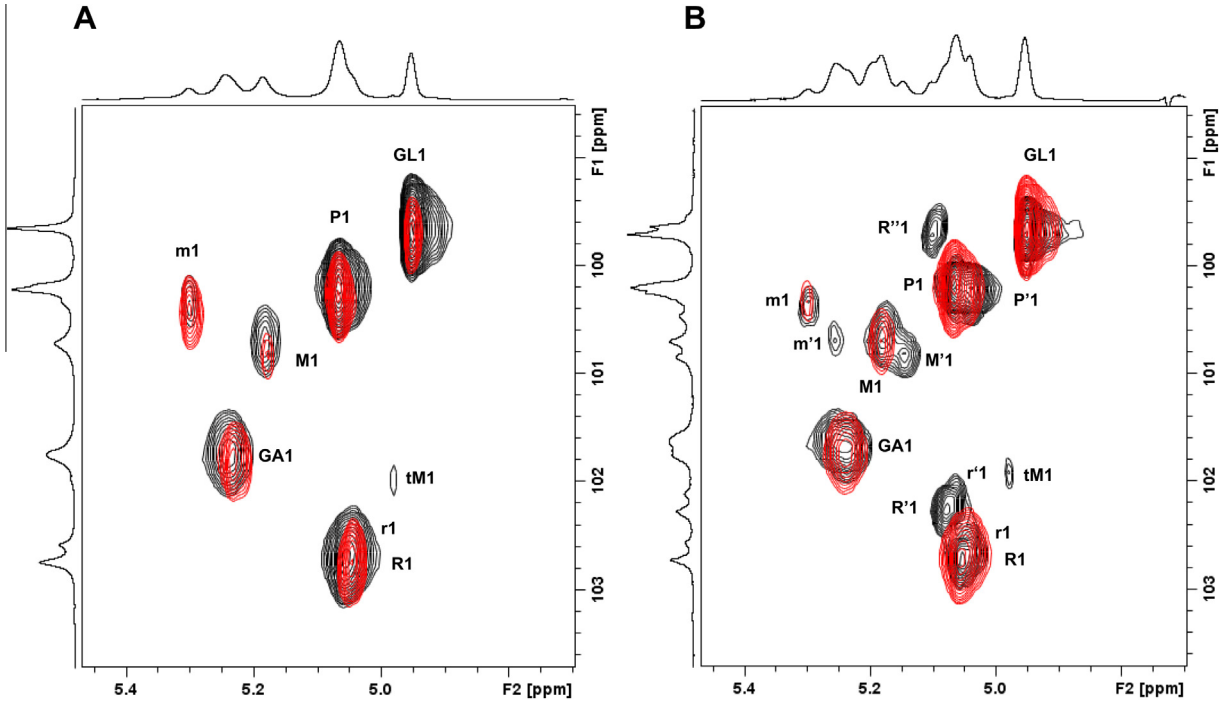

Figure 2. Overlays of the anomeric regions of the HSQC spectra of *S. paratyphi* A O-polysaccharide samples: (A) partial Smith degradation product **OPS-S2** (red) and de-O-acetylated **OPS-D** (black), (B) de-O-acetylated **OPS-D** (red) and partially O-acetylated **OPS** (black). Crosspeaks are labelled according to the proton/carbon pairs with upper case for the pentasaccharide repeating unit (P1 = H-1/C-1 of Par, etc.) and lower case for the tetrasaccharide repeating unit. The Rha O-acetylation is labelled with dashes (R1 = H-1/C-1 of Rha, R′1 = H-1/C-1 of Rha3Ac, R″1 = H-1/C-1 of Rha2Ac). Dashes are also used to label twinning of the other residues due to O-acetylation of Rha.

Table 1
NMR data of *S. paratyphi* A O-polysaccharide [de-OAc **OPS-D** (upper panel), part of the native **OPS** (lower panel)]

Residue	H-1 C-1	H-2 C-2	H-3 C-3	H-4 C-4	H-5 C-5	H-6 C-6	*O*-Acetyl Methyl	CO
3)-α-D-Gal*p*-(1→	5.24	3.95	3.96	4.16/4.05	4.23/4.07	3.79,3.69/3.71		
(GA/ga)	101.7	68.4	77.9	69.5	69.8/72.0	66.8/61.7		
α-D-Glc*p*-(1→	4.95	3.56	3.73	3.42	3.66	3.87,3.76		
(GL)	99.6	72.0	73.8	70.1	72.9	61.3		
4)-α-L-Rha*p*-(1→	5.06/5.04	4.08	3.97	3.60/3.55	3.96/3.93	1.34/1.32		
(R/r)	102.7/102.6	71.1	69.8	82.1	68.7	17.9/17.7		
2,3)-α-D-Man*p*-(1→	5.18/5.30	4.02/4.03	4.05	4.05	3.96	~3.83		
(M/m)	100.7/100.3	79.7	77.8	67.0	74.2	61.2		
α-D-Par*p*-(1→	5.07/5.04	3.83	2.13,1.71	3.38	3.72	1.24/1.23		
(P/p)	100.2	67.6	35.1	70.4	69.9	17.2		
4)-α-L-Rha*p*-(1→	5.06	4.07	3.96	3.60	3.96	1.34		
(R)	102.5	71.1	69.8	82.1	68.7	17.7		
4)-α-L-Rha*p*3Ac-(1→	5.09	4.21	5.19	3.88	4.06	1.36	2.20	
(R′)	102.3	68.9	73.4	78.6	68.9	18.1	21.6	174.1
4)-α-L-Rha*p*2Ac-(1→	5.10	5.20	4.19	3.63	4.02	1.37	2.18	
(R″)	99.6	73.5	68.2	81.5	68.8	18.0	21.1	173.9
2,3)-α-D-Man*p*-(1→	5.18	4.02/4.03	4.05	4.05	3.96	~3.83		
(M/M′ and M″)	100.7/100.3	79.7/79.4	77.8	67.0	74.2	61.2		

OPS-D: the residues of the major pentasaccharide repeating unit are indicated by upper case (e.g. Gal = **GA**) and the minor tetrasaccharide repeating unit by lower case (e.g. Gal = **GA**). Deshielded ring carbons (linkage positions) are underlined. The spectra were recorded at 303 K (600 MHz) and referenced to residual sodium acetate (present after de-O-acetylation). **OPS**: only the residues of the pentasaccharide repeating unit affected by O-acetylation are presented; minor peaks due to the presence of the tetrasaccharide repeating unit are not listed. The Rha, Rha3Ac and Rha2Ac spin systems are represented by R, R′ and R″, respectively, and the corresponding spin systems of Man by M, M′ and M″. The NMR spectra were recorded at 303 K (600 MHz) and referenced to DMSO.

crosspeak to H-3 of Gal at 3.96 ppm. The Gal spin systems were identified from examination of the 1D and 2D NMR spectra obtained for **OPS-S2** (mainly Gal **ga** in RU4) and **OPS-D** (mainly Gal **GA** in RU5). The HMBC experiments from H-1 of Gal gave crosspeaks to C-3 and C-5 of the residues and the crosspeak from H-1 to 79.7 ppm (C-2 Man) confirmed the α-D-Gal*p*-(1→2)-D-Man*p* linkage. The location of glucosylation on C-6 of Gal in RU5 was demonstrated by the deshielding of C-6 of Gal in RU4 (from 61.7 to 66.8 ppm) with concomitant shielding of C-5 (from 72.0 to 69.8 ppm) and deshielding of the attached H-5 (from 4.07 to 4.23 ppm). This was confirmed by HMBC crosspeaks between H-1 of Glc (4.95 ppm) and C-6 of Gal. Lastly, proof of the α-D-Par*p*-(1→3)-α-D-Man*p* linkage was shown by the H-1 Par/C-3 Man crosspeak in the HMBC spectrum thereby confirming the structure of the pentasaccharide repeating unit: →2)-[α-D-Par*p*-(1→3)]-α-D-Man*p*-(1→4)-α-L-Rha*p*-(1→3-[α-D-Glc*p*-(1→6)]-α-D-Gal*p*-(1→ (Fig. 1). Modelling studies were performed to account for the major chemical shift influence observed from glucosylation (H-1 Man and H-4 to H-6 Rha); the favored orientations of phi/psi dihedral angles for each of the glycosidic linkages in the pentasaccharide repeating unit bring Glc in close proximity to the Man and Rha backbone residues.[19] Additional crosspeaks were observed corresponding to small peaks in the 1D spectra derived from the core oligosaccharide as well as the non-reducing terminal of the **OPS** α-D-Par*p*-(1→3)-α-D-Man*p*-(1→ [with H-1, C-1 Par at 4.04 and 102.2 ppm and H-1, C-1 Man at 4.98 and 102.0 ppm, respectively]. Characterization of terminal residues of the *Salmonella typhimurium* O-polysaccharide chain has been reported;[20] it was also detected in our NMR studies performed on *Salmonella typhimurium*.[21]

2.4. Chemical and NMR characterization of OPS O-acetylation

The pattern of O-acetylation of **OPS** was investigated chemically and by use of NMR spectroscopy. **OPS** was permethylated following the method of Prehm,[22] which leaves the *O*-acetyl groups in their native positions, thus allowing the determination of their location. GLC chromatograms of the PMAA derivatives for the O-polysaccharide obtained without (A) and with (B) retention of the *O*-acetyl groups (Fig. 3) showed the presence of two new peaks

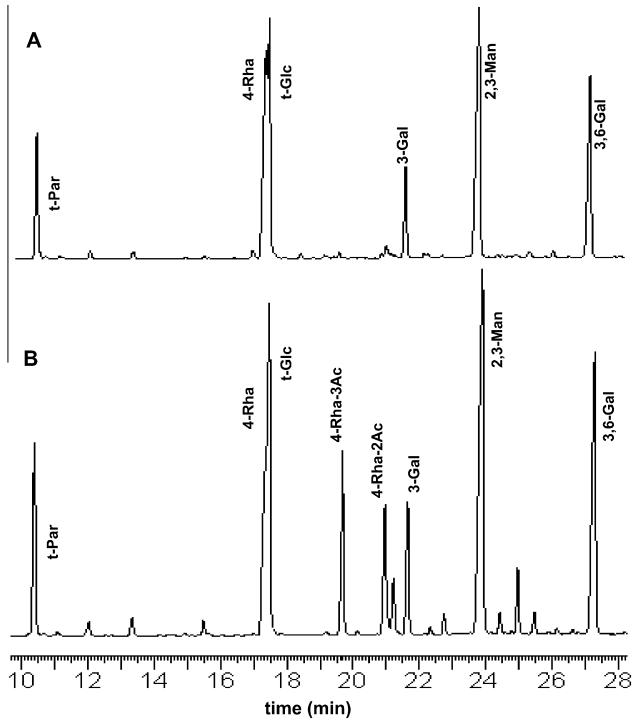

Figure 3. GLC chromatograms of the PMAA derivatives obtained from permethylation of *S. paratyphi* A O-polysaccharide: (A) without, (B) with retention of *O*-acetyl groups. Peak assignments are reported: t-Par = terminal non-reducing Par; 4-Rha-3-Ac = 4-linked rhamnose acetylated on C-3; 4-Rha-2-Ac = 4-linked rhamnose acetylated on C-2. The numbers indicate the position of hydroxyl functions engaged in linkages.

corresponding to 3,4-linked Rha and 2,4-linked Rha and a decrease in the peak intensity of 4-linked Rha. Although the peaks corresponding to 4-Rha and t-Glc overlap on the SP2330 column, integration of the signals corresponding to the new Rha derivatives indicated that the 34% of the 4-linked Rha was acetylated on C-3 and 25% on C-2 (Table 2). Subsequent GLC–MS of the samples on

Table 2
Determination of the glycosidic linkages in *S. paratyphi* A O-polysaccharide by GLC-MS of PMAA derivatives

Linkage	RRT[a]	I[b]	II[c]
t-Par[d]	0.435	0.26	0.26
4-Rha	0.726	0.93	0.40
t-Glc	0.731	0.66	0.72
3,4-Rha	0.824	—	0.34
2,4-Rha	0.877	—	0.25
3-Gal	0.906	0.24	0.23
2,3-Man	1.000	1.00	1.00
3,6-Gal	1.142	0.61	0.66

[a] Relative retention time.
[b] I = **OPS** methylated without retention of *O*-acetyl groups.
[c] II = **OPS** methylated with retention of *O*-acetyl groups.
[d] t-Par = the amount of this derivative is underestimated because it is acid labile.

a HP-1 column separated the 4-Rha and t-Glc peaks and integration (of the MS signal) gave similar results with 31% of the 4-linked Rha acetylated on C-3 and 21% on C-2, with the relative intensity of 4-Rha decreasing from 0.74 to 0.34.

The ^{1}H NMR spectrum of **OPS** contained two signals at 2.20 and 2.18 ppm (Fig. 4A), not present in **OPS-D** (Fig. S2C); these were attributed to the two *O*-acetyl groups detected chemically. An overlay of the ^{13}C NMR spectra of **OPS** and **OPS-D** (Fig. 5) showed a decrease in intensity of the Rha (**R**) signals and the appearance of several new signals attributed to Rha3Ac (**R′**) and Rha2Ac (**R″**). These spin systems were assigned using a full array of 2D experiments; a combination of the HSQC-NOESY (Fig. 6A) and HSQC-TOCSY (Fig. 6B) experiments were particularly useful in establishing connectivities from the *O*-acetyl signals (at 2.20 and 2.18 ppm) and the H-6/C-6 resonances, respectively, to the carbons of the Rha3Ac (R′) and Rha2Ac (R″) residues. Initial studies indicated a single position of O-acetylation manifest by a single HSQC crosspeak as recently reported.[15] However, repeating the NMR study at higher field (600 MHz) revealed the presence of overlapping crosspeaks due to Rha3Ac (major, R′ at 5.19/73.4 ppm) and Rha2Ac (minor, R″ at 5.20/73.5 ppm). The HMBC experiment permitted identification of the acetyl carbons (at 174.1 and 173.9 ppm) from the signals at 2.20 and 2.18 ppm, and the carbonyls gave small crosspeaks to the corresponding ring protons in the anomeric region. The NMR data are collected in Table 1. For Rha3Ac, the presence of the *O*-acetyl group on C-3 resulted in deshielding of H-3 (+1.23 ppm) and C-3 (+3.6 ppm) and shielding of the adjacent carbons C-2 (−2.2 ppm) and C-4 (−3.5 ppm), whereas for Rha2Ac, H-2 (+1.13 ppm) and C-2 (+2.4 ppm) were deshielded and the adjacent carbons C-1 (−2.9 ppm) and C-3 (−1.6 ppm) shielded. The effects of O-acetylation on Rha are comparable to those reported for →4)-α-Rha2/3Ac in *Salmonella typhimurium*[21] and a strain of *Aeromonas sobria*.[23] The major effect of O-acetylation on the rest of the sugars in the repeating unit was a slight shielding for positions 1 and 2 of the α-Man substituent at C-4 of the O-acetylated Rha and the anomeric positions of Gal and Par (see HSQC expansion, Fig. 2B).

Estimation of the relative intensities of the *O*-acetyl groups at 2.20 and 2.18 ppm compared to H-6 of Rha indicated a degree of O-acetylation of 30% and 20% at positions C-3 and C-2, respectively.

2.5. Little impact of conjugation on OPS O-acetylation pattern

Similar relative intensities of the *O*-acetyl signals were obtained for the derived conjugate **OPS-CRM$_{197}$** (previously[12] designated as O:2-CRM$_{197}$), proving that the conjugation process had little impact on the pattern of O-acetylation although there was a small decrease in total O-acetylation (Fig. 4B). A third smaller peak at 2.16 ppm (with CH_3 at 21.0 ppm and carbonyl at 173.8 ppm) was assumed to be a third position of O-acetylation, however,

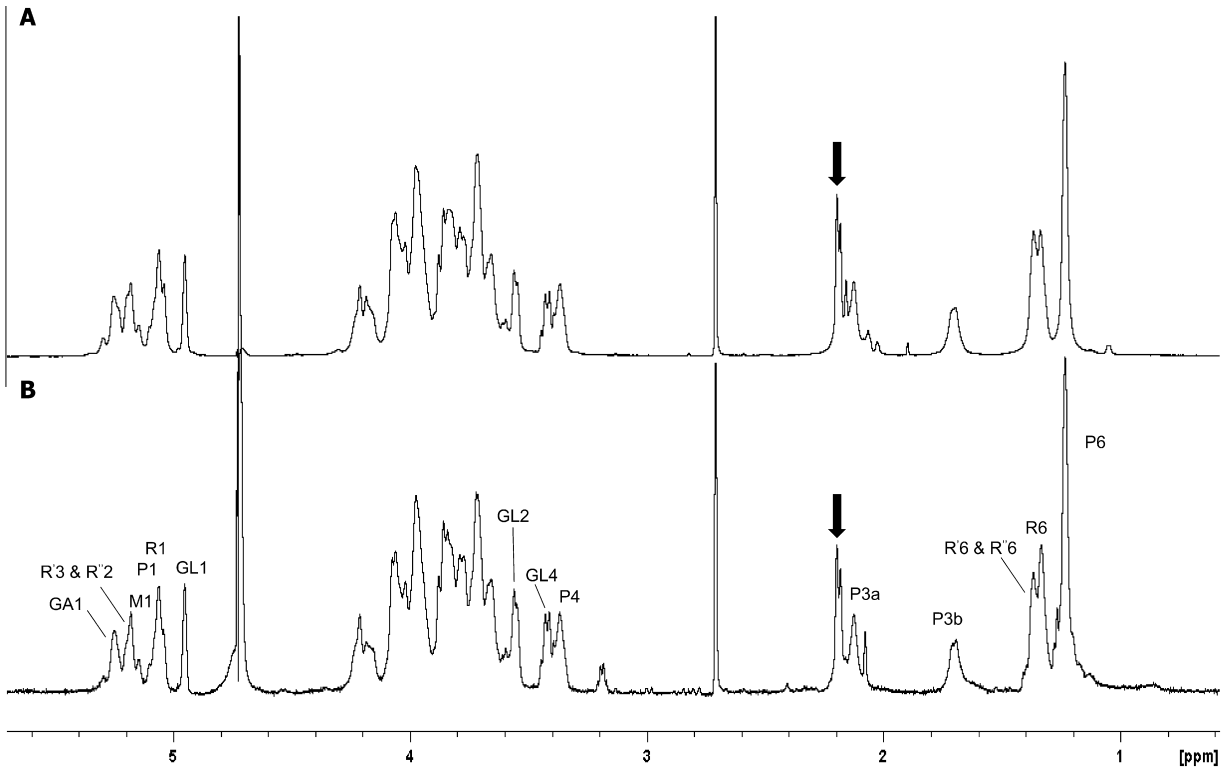

Figure 4. ^{1}H NMR spectra of *S. paratyphi* A samples: (A) partially O-acetylated O-polysaccharide **OPS**, (B) the purified O-polysaccharide conjugate **OPS-CRM$_{197}$**. Some assignments are labelled, upper case for the pentasaccharide repeating unit (P1 = H-1 of Par, etc.) and the Rha O-acetylation is labelled with dashes (R1 = H-1 of Rha, R′1 = H-1 of Rha3Ac, R″1 = H-1 of Rha2Ac). An arrow indicates *O*-acetyl signals.

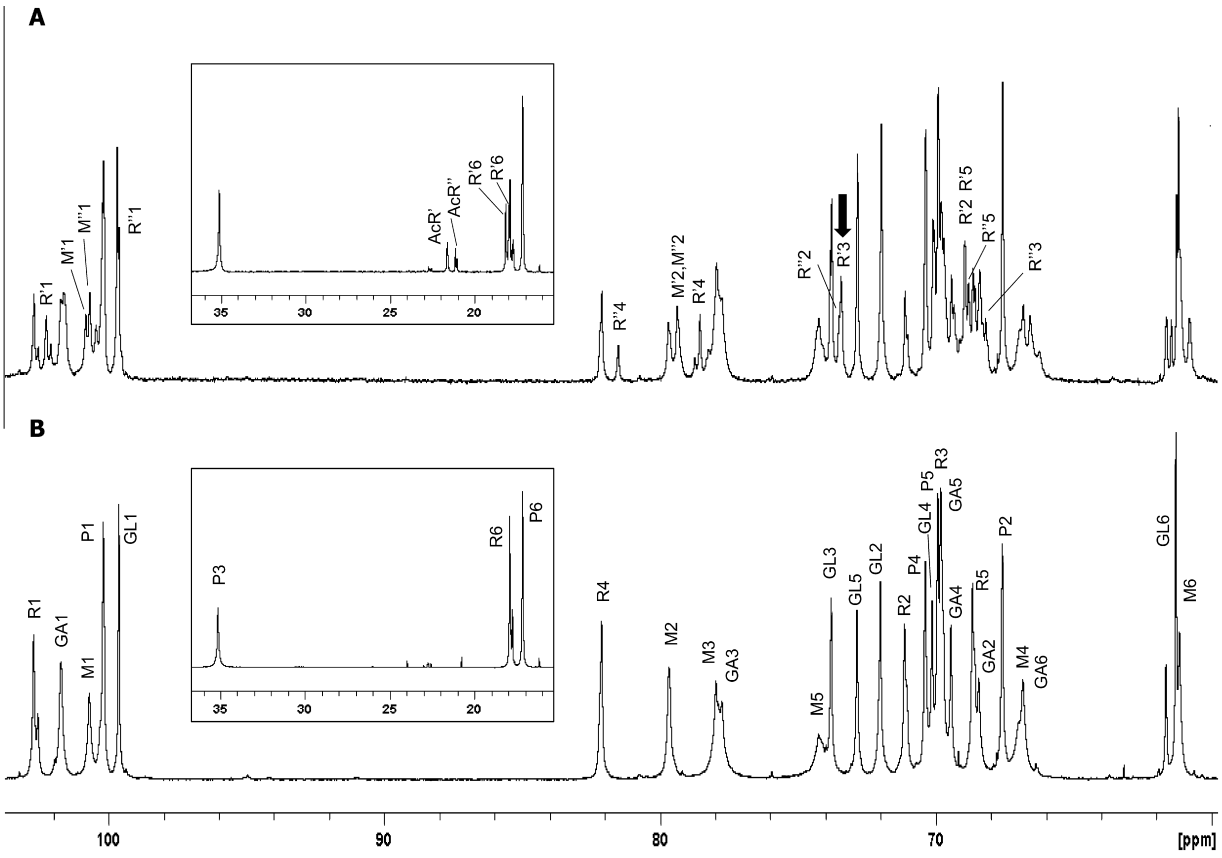

Figure 5. ^{13}C NMR spectra of *S. paratyphi* A O-polysaccharide samples: (A) partially O-acetylated O-polysaccharide **OPS**, (B) de-O-acetylated **OPS-D**. The inserts show the methyl and methylene region of the spectra. Some assignments are labelled, upper case for the pentasaccharide repeating unit (P1 = C-1 of Par, etc.) and the Rha O-acetylation is labelled with dashes (R1 = C-1 of Rha, R′1 = C-1 of Rha3Ac, R″1 = C-1 of Rha2Ac). Dashes are also used to label twinning of the other residues due to O-acetylation of Rha. An arrow indicates the O-acetylated carbons.

connectivity to the O-polysaccharide repeating unit could not be established and the peak is absent from the NMR spectrum of **OPS-CRM$_{197}$**. This peak could explain the higher total O-acetylation levels determined by NMR after in-tube de-O-acetylation or HPAEC–PAD compared to NMR analysis of the intact O-acetylated O-polysaccharide.

2.6. HR-MAS NMR analysis on bacterial pellets confirming O-acetylation pattern found on purified OPS

The O-acetylation pattern of O-polysaccharide expressed on surface of bacteria was also investigated in light of literature reports regarding a single position of O-acetylation and the possibility that the O-acetylation observed on C-2 and C-3 of Rha of **OPS** may arise from migration during O-polysaccharide preparation. HR-MAS NMR analysis, used to track the pattern of O-acetylation on live *Salmonella typhimurium* bacteria,[24] was performed on bacterial pellets from three clinical isolates and a genetically modified *S. paratyphi* A strain. The ^{1}H NMR spectra obtained (Fig. 7A–D) showed that all cells contain two peaks (~1:1 ratio) in the *O*-acetyl region assigned to O-acetylation on C-2 and C-3 of Rha. This proves that O-acetylation on both positions is common to different *S. paratyphi* A strains and it is not an artefact introduced during O-polysaccharide isolation and purification, corroborated by a comparison of the O-acetylation on the surface of bacteria (Fig. 7D) and on the O-polysaccharide purified from the homologous strain (Fig. 7E). Purified **OPS** has O-acetylation of 30% and 18% at positions C-3 and C-2 of Rha (compared to H-6 of Rha), suggesting that although some migration may occur during O-polysaccharide purification, it results in a consistent 3:2 ratio at positions C-3 and C-2 of Rha. The facile migration between positions C-2 and C-3 has been reported for methyl glycosides of O-acetylated α-L-Rhap.[25]

3. Conclusions

As part of a project to develop a bivalent conjugate vaccine against both *S. typhi* and *S. paratyphi* A infections, we have fully characterized the O-polysaccharide from *S. paratyphi* A. Sugar analysis by GLC-MS confirmed the composition of the tetrasaccharide repeating unit as Man, Gal, Rha and Par together with non-stoichiometric amounts (~75%) of glucosylation; this is in agreement with HPAEC–PAD data previously reported.[18] The structure of the repeating unit was confirmed by linkage analysis which showed that the terminal glucose present was linked 1→6 to Gal. Incomplete glucosylation and O-acetylation resulted in complex NMR spectra and simpler structures obtained by chemical modification of the native O-polysaccharide permitted assignment of the repeating unit backbone and determination of the location and extent of O-acetylation. The presence of *O*-acetyl groups on C-2 and C-3 of Rha was confirmed by linkage analysis and GLC-MS; this is in contrast to previous literature reports of a single position of O-acetylation.[9,10,14–16] The NMR assignments permitted evaluation of the O-acetylation pattern of the O-polysaccharide from a second *S. paratyphi* A strain as well as the surface of several different strains by use of HR-MAS NMR. All *S. paratyphi* A screened were O-acetylated on C-2 and C-3 of Rha suggesting that this is a common phenotype. Additionally, conjugation of the O-polysaccharide to a carrier protein had no impact on the position of the *O*-acetyl groups. Considering the importance of this decoration for the

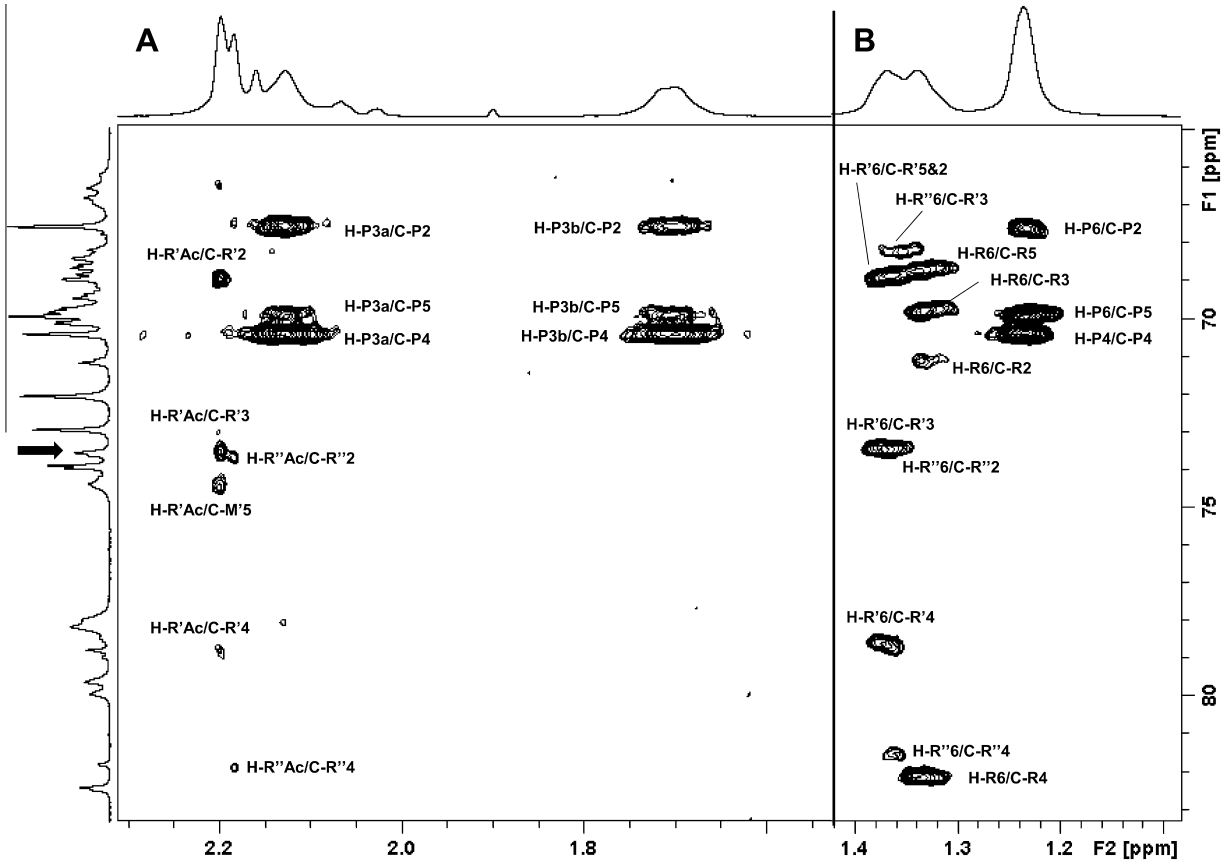

Figure 6. Expansions of the $^{1}H–^{13}C$ hybrid 2D NMR experiments performed on partially O-acetylated O-polysaccharide **OPS** (A) HSQC-NOESY spectrum (mixing time of 250 ms), (B) HSQC-TOCSY spectrum (mixing time of 120 ms). Key H/C correlations are labelled, upper case for the pentasaccharide repeating unit and the Rha O-acetylation is labelled with dashes (H-R1 = H-1 of Rha, H-R′1 = H-1 of Rha3Ac, H-R″1 = H-1 of Rha2Ac). Crosspeaks are labelled H/C (H-P6/C-P2 = H-6 Par/C-2 Par, etc.) The ^{1}H acetyl signals of Rha3Ac and Rha2Ac are labelled as H-R′Ac and H-R″Ac, respectively. An arrow indicates the O-acetylated carbons.

immunogenicity of vaccines against *S. paratyphi* A,[9,10] maintaining the structure of critical epitopes of the O-polysaccharide from the bacteria through to conjugate formation increases the likelihood of success of this much needed vaccine.

4. Experimental

4.1. Strains

S. paratyphi A used in this study were either clinical isolates, obtained from University of Oxford and Patan Hospital, Kathmandu Nepal (NVGH307, NVGH309, NVGH310) or genetically modified strains, prepared in house or supplied by the Center for Vaccine Development, University of Maryland (NVGH2041 and CVD 1901).

4.2. Samples

OPS was purified as previously described.[18] All preparations were characterized by protein content <1% (mg/mg OPS; by micro BCA), nucleic acid content <0.5% (mg/mg OPS; by absorption at 260 nm) and endotoxin level <0.1 UI/μg (by LAL).

De-O-acetylation of **OPS** was performed in the NMR tube, however the higher pH and ionic strength resulted in minor chemical shift differences with the parent **OPS** sample that hampered direct comparison of the spectra. Therefore the **OPS-D** sample was obtained by de-O-acetylation of the O-polysaccharide sample (200 mM NaOH at 37 °C for 2 h) followed by recovery of the sample by freeze-drying after dialysis against water.

The O-polysaccharide conjugate (**OPS-CRM$_{197}$**) was prepared as previously described.[12] Briefly the sugar chain was linked to adipic acid dihydrazide (ADH) by reductive amination through the ketone group of its terminal Kdo sugar. An additional linker, SIDEA, was then attached to ADH before conjugation to the carrier protein. The conjugate prepared for this structural analysis was characterized by a saccharide to CRM_{197} w/w ratio of 2.7, with no free CRM_{197} detectable (by HPLC–SEC)[12] and less than 10% free saccharide (following precipitation of the conjugate with deoxycholate).[26]

4.3. Smith degradation of OPS to yield OPS-S1 and OPS-S2

OPS was subjected to complete oxidation with $NaIO_4$,[27,28] by dissolving 180 mg of the sample in 50 mL of 30 mM $NaIO_4$ and incubating at 10 °C for 7 days in the dark. The reaction was stopped by the addition of glycerol and the products were reduced with $NaBH_4$. Addition of 50% CH_3COOH after 16 h destroyed the excess of reducing reagent, the sample was dialysed and the solution volume subsequently reduced to 10 mL. TFA was added in order to have a final concentration of 0.1 M and the hydrolysis was conducted at room temperature for 7 days. The solution was taken to dryness under reduced pressure, dissolved in water, its pH adjusted to neutrality, and the product was recovered by lyophilization. It was then separated on a Bio Gel P2 column (1.6 cm i.d. × 90 cm) equilibrated in 50 mM $NaNO_3$ which was also used as eluent. The flow rate was 6 mL/h and fractions were collected at 15 min intervals. Elution was monitored using a refractive index detector (WGE Dr. Bures, LabService Analitica) connected to a paper recorder and interfaced with a computer via picolog software. One major oligosaccharide, named **OPS-S1**, was obtained from the chromatographic separation and analysed by NMR and ESI-MS.

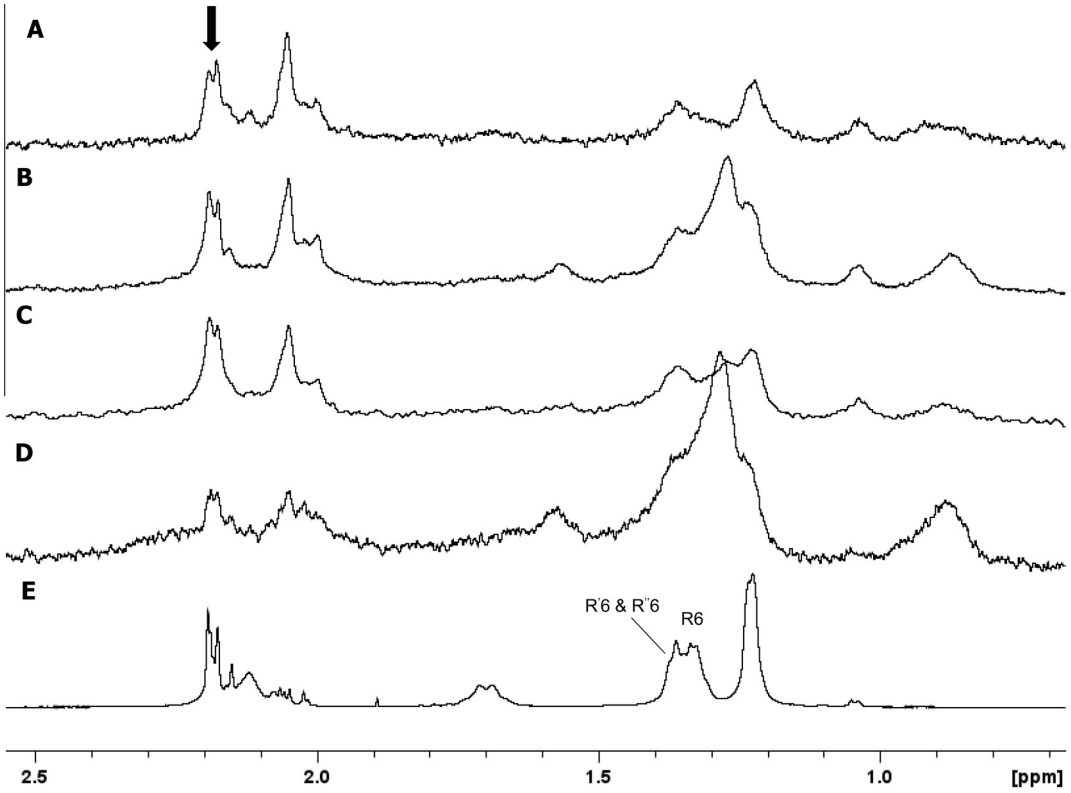

Figure 7. Expansions of the ^{1}H HR-MAS-NMR spectra of *S. paratyphi* A samples: (A) clinical isolate NVGH307, (B) clinical isolate NVGH309, (C) clinical isolate NVGH310, (D) genetically modified strain NVGH2041, and (E) the ^{1}H NMR spectrum the O-polysaccharide purified from strain NVGH2041. An arrow indicates the *O*-acetyl signals and some assignments are labelled with Rha O-acetylation indicated with dashes (R6 = H-6 of Rha, R′6 = H-6 of Rha3Ac, R″6 = H-6 of Rha2Ac).

OPS was also subjected to partial periodate oxidation in order to selectively remove the t-Glc residue: 21 mg of sample were dissolved in 7 mL of 16 mM $NaIO_4$ and stirred for 75 min at 10 °C in the dark. The oxidation was then stopped by addition of glycerol and the products were treated as reported above. Reduction of the aldehyde groups with $NaBH_4$ resulted in de-O-acetylation. After a mild hydrolysis for 7 days with 0.05 M TFA, the sample was taken to dryness under reduced pressure in order to remove most of the acid, then taken to neutral pH and recovered by lyophilization. The sample was subjected to composition analysis and then separated on a Sephacryl S-300 column (1.6 cm i.d. × 92 cm), equilibrated in 50 mM $NaNO_3$ and using the same solution as eluent. The flow rate was 6 mL/h and fractions were collected at 15 min intervals. Elution was monitored as reported above for the sample **OPS-S1**. The fractions with the highest molecular mass were pooled together, named **OPS-S2** and investigated by NMR spectroscopy.

4.4. Chemical analysis

Analytical GLC was performed on a Perkin Elmer Autosystem XL gas chromatograph equipped with a flame ionization detector and an SP2330 capillary column (Supelco, 30 m), using He as carrier gas. The following temperature programmes were used: for alditol acetates, 200–245 °C at 4 °C/min; for partially methylated alditol acetates, 150–250 °C at 4 °C/min. Separation of the trimethylsilylated (+)-2-butyl glycosides was obtained on a HP-1 column (Hewlett–Packard, 50 m), using the following temperature programme: 50 °C 1 min, 50–130 °C at 45 °C/min, hold 1 min, 130–240 °C at 1 °C/min, hold 20 min. GLC–MS analyses were carried out on an Agilent Technologies 7890A gas chromatograph coupled to an Agilent Technologies 5975C VL MSD. Hydrolysis of the samples was performed with 2 M TFA for 1 h at 125 °C. Alditol acetates were prepared as already described.[29] Permethylation of polysaccharides and oligosaccharides was achieved following the protocols by Harris[30] and Dell,[31] respectively. The method of Prehm was used to methylate **OPS** without removal of *O*-acetyl groups.[22]

Determination of the absolute configuration of the sugar residues was performed as described.[32]

Quantification of total O-acetylation was performed by use of high-performance anion-exchange chromatography with conductivity detection (HPAEC–CD). **OPS** samples (total volume of 225 μL) were added with 25 μL of 200 μg/mL sodium propionate in 100 mM NaOH. The samples were heated at 37 °C for 4 h. Samples were then filtered with 0.45 μm Acrodisc (PALL) filters before chromatographic analysis. Calibration curve was built using sodium acetate as standard in the range of 0.5–10 μg/mL. The standards were treated and analysed as described for the samples. HPAEC–CD was performed with a Dionex ICS3000 equipped with a Dionex IonPac AS11 (4 × 250 mm) coupled with IonPac AG11 Guard Column (4 × 50 mm) and quantified by a conductivity detector with ASRS suppressor under 2 mM NaOH elution.

4.5. ESI MS analysis

ESI mass spectra were recorded on a Bruker Esquire 4000 ion trap mass spectrometer connected to a syringe pump for the injection of the samples. The instrument was calibrated using a tune mixture provided by Bruker. Permethylated samples were dissolved in chloroform/methanol/0.75 M NH_4OAc 1:1:0.03 (v/v/v) at an appropriate concentration and injected at 180 μL/h. Detection was always performed in the positive ion mode.

4.6. NMR spectroscopy

Samples of **OPS** [90 mg (400 MHz study) and 20 mg (600 MHz study)], **OPS-D** (70 mg), **OPS-S1** (10 mg), **OPS-S2** (6 mg) and

OPS-CRM$_{197}$ (1.9 mg saccharide) were prepared for NMR analysis by repeatedly dissolving and drying the sample in D_2O. NMR spectra were acquired on samples dissolved in D_2O (0.6 mL, Aldrich) and transferred to 5 mm NMR tubes (Wilmad). Spectra were also recorded on an **OPS** sample after the addition of sodium deuteroxide to a final concentration of 200 mM in order to achieve de-O-acetylation in the NMR tube. 1D (^1H and ^{13}C) and 2D COSY, TOCSY, NOESY, HSQC, HMBC and hybrid HSQC-TOCSY and HSQC-NOESY NMR spectra were obtained using a Bruker Avance III 400 MHz spectrometer. Many of the experiments were repeated using a Bruker Avance III 600 MHz NMR spectrometer equipped with a BBO Prodigy cryoprobe. The probe temperature was set at 303 K and the spectra were acquired and processed using standard Bruker software (Topspin 3.2). 2D TOCSY experiments were performed using mixing times of 120 and 180 ms and the 1D variants using mixing times of 200 or 250 ms. The HSQC experiment was optimized for $J = 145$ Hz (for directly attached ^1H–^{13}C correlations), and the HMBC experiment optimized for a coupling constant of 8, 6 or 5 Hz (for long-range ^1H–^{13}C correlations). HSQC-TOCSY and HSQC-NOESY NMR spectra were recorded using mixing times of 120 and 250 ms, respectively. Spectra were referenced to residual acetate (^1H signal at 1.903 ppm and ^{13}C signal at 23.97 ppm, DMSO (^1H signal at 2.71 ppm and ^{13}C signal at 39.39 ppm[33] or acetone (^1H signal at 2.225 ppm and ^{13}C signal at 31.45 ppm).

Proton high-resolution magic angle spinning (HR-MAS) NMR spectra were recorded on bacterial pellets as previously described for *S. typhimurium*.[18] The spectra were recorded on a Bruker Avance III 400-MHz spectrometer using a Bruker 4-mm HR-MAS probe. Samples were spun at 4.5 kHz at 25 °C and the proton spectra acquired with a diffusion filter pulse sequence to remove sharp lines arising from low-molecular-mass species in solutin.

Acknowledgements

The authors thank S. Rondini, L. Lanzilao and the technical development team of NVGH for preparation of O:2 polysaccharide; O. Rossi for culturing *S. paratyphi* A strains; F. Berti for HR-MAS NMR spectra; Y. Herasimenka for GLC and GLC–MS analysis. The authors are also grateful to University of Oxford and Patan Hospital, Kathmandu Nepal for providing S. *paratyphi* A clinical isolates and University of Maryland for providing CVD 1901.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.carres.2014.12.002.

References

1. Crump, J. A.; Luby, S. P.; Mintz, E. D. *Bull. World Health Organ.* **2004**, *82*, 346–353.
2. Sinha, A.; Sazawal, S.; Kumar, R.; Sood, S.; Reddaiah, V. P.; Singh, B.; Rao, M.; Naficy, A.; Clemens, J. D.; Bhan, M. K. *Lancet* **1999**, *354*, 734–737.
3. Saha, S. K.; Baqui, A. H.; Hanif, M.; Darmstadt, G. L.; Ruhulamin, M.; Nagatake, T.; Santosham, M.; Black, R. E. *Pediatr. Infect. Dis. J.* **2001**, *20*, 521–524.
4. Brooks, W. A.; Hossain, A.; Goswami, D.; Nahar, K.; Alam, K.; Ahmed, N.; Naheed, A.; Nair, G. B.; Luby, S.; Breiman, R. F. *Emerg. Infect. Dis.* **2005**, *11*, 326–329.
5. Buckle, G. C.; Walker, C. L.; Black, R. E. *J. Global Health* **2012**, *2*, 010401.
6. Sahastrabuddhe, S.; Carbis, R.; Wierzba, T. F.; Ochiai, R. L. *Expert Rev. Vaccines* **2013**, *12*, 1021–1031.
7. Szu, S. C. *Expert Rev. Vaccines* **2013**, *12*, 1273–1286.
8. Kulkarni, S. V.; Narayan, A.; Indumathi, V. A.; Rao, T. S.; Kempegowda, P. *Asian J. Med. Sci.* **2011**, *2*, 14–17.
9. Konadu, E. Y.; Shiloach, J.; Bryla, D. A.; Robbins, J. B.; Szu, S. C. *Infect. Immun.* **1996**, *64*, 2709–2715.
10. Konadu, E. Y.; Lin, F. Y.; Ho, V. A.; Thuy, N. T.; Van, B. P.; Thanh, T. C.; Khiem, H. B.; Trach, D. D.; Karpas, A. B.; Li, J., et al. *Infect. Immun.* **2000**, *68*, 1529–1534.
11. http://who.int/immunization/research/meetings_workshops/pdvac/en/.
12. Micoli, F.; Rondini, S.; Gavini, M.; Lanzilao, L.; Medaglini, D.; Saul, A.; Martin, L. B. *PLoS One* **2012**, *7*, e47039.
13. Perepelov, A. V.; L'vov, V. L.; Liu, B.; Senchenkova, S. N.; Shekht, M. E.; Shashkov, A. S.; Feng, L.; Aparin, P. G.; Wang, L.; Knirel, Y. A. *Carbohydr. Res.* **2009**, *344*, 687–692.
14. Hellerqvist, C. G.; Lindberg, B.; Samuelsson, K.; Lindberg, A. A. *Acta Chem. Scand.* **1971**, *25*, 955–961.
15. Kothari, S.; Kim, J. A.; Kothari, N.; Jones, C.; Choe, W. S.; Carbis, R. *Vaccine* **2014**, *32*, 2457–2462.
16. Liu, B.; Knirel, Y. A.; Feng, L.; Perepelov, A. V.; Senchenkova, S. Y. N.; Reeves, P. R.; Wang, L. *FEMS Microbiol. Rev.* **2014**, *38*, 56–89.
17. Hellerqvist, C. G.; Lindberg, A. A. *Carbohydr. Res.* **1971**, *16*, 39–48.
18. Micoli, F.; Rondini, S.; Gavini, M.; Pisoni, I.; Lanzilao, L.; Colucci, A. M.; Giannelli, C.; Pippi, F.; Sollai, L.; Pinto, V.; Berti, F.; MacLennan, C. A.; Martin, L. B.; Saul, A. *Anal. Biochem.* **2013**, *434*, 136–145.
19. Kuttel, M. K. *Pers. commun.* **2014**.
20. De Castro, C.; Lanzetta, R.; Leone, S.; Parrilli, M.; Molinaro, A. *Carbohydr. Res.* **2013**, *370*, 9–12.
21. Micoli, F.; Ravenscroft, N.; Cescutti, P.; Stefanetti, G.; Londero, S.; Rondini, S.; MacLennan, C. A. *Carbohydr. Res.* **2014**, *385*, 1–8.
22. Prehm, P. *Carbohydr. Res.* **1980**, *78*, 372–374.
23. Turska-Szewczuk, A.; Pietras, H.; Duda, K. A.; Kozińska, A.; Pękala, A.; Holst, O. *Carbohydr. Res.* **2015**, *403*, 142–148.
24. Ilg, K.; Zandomeneghi, G.; Rugarabamu, G.; Meier, B. H.; Aebi, M. *Carbohydr. Res.* **2013**, *382*, 58–64.
25. Rönnols, J.; Pendrill, R.; Fontana, C.; Hamark, C.; d'Ortoli, T. A.; Engström, O.; Ståhle, J.; Zaccheus, M. V.; Säwén, E.; Hahn, L. E.; Iqbal, S.; Widmalm, G. *Carbohydr. Res.* **2013**, *380C*, 156–166.
26. Guo, Y. Y.; Anderson, R.; McIver, J.; Gupta, R. K.; Siber, G. R. *Biologicals* **1998**, *26*, 33–38.
27. Hay, G. W.; Lewis, B. A.; Smith, F. *Methods Carbohydr. Chem.* **1965**, *5*, 357–361.
28. Goldstein, I. J.; Hay, G. W.; Lewis, B. A.; Smith, F. *Methods Carbohydr. Chem.* **1965**, *5*, 361–370.
29. Albersheim, P.; Nevins, D. J.; English, P. D.; Karr, A. *Carbohydr. Res.* **1967**, *5*, 340–345.
30. Harris, P. J.; Henry, R. J.; Blakeney, A. B.; Stone, B. A. *Carbohydr. Res.* **1984**, *127*, 59–73.
31. Dell, A. *Methods Enzymol.* **1990**, *193*, 647–660.
32. Gerwig, G. J.; Kamerling, J. P.; Vliegenthart, J. F. G. *Carbohydr. Res.* **1978**, *62*, 349–357.
33. Gottlieb, H. E.; Kotlyar, V.; Nudelman, A. *J. Org. Chem.* **1997**, *62*, 7512–7515.