Structure of the high molecular weight exopolysaccharide produced by *Bifidobacterium animalis* subsp. *lactis* IPLA-R1 and sequence analysis of its putative *eps* cluster

Shaun Leivers[a], Claudio Hidalgo-Cantabrana[b], Glenn Robinson[a], Abelardo Margolles[b], Patricia Ruas-Madiedo[b] and Andrew P. Laws[a]*

[a] Department of Chemical and Biological Sciences, University of Huddersfield, Queensgate, Huddersfield, HD1 3DH, United Kingdom

[b] Instituto de Productos Lácteos de Asturias, Consejo Superior de Investigaciones Científicas (IPLA-CSIC), Carretera de Infiesto s/n, 33300 Villaviciosa, Asturias, Spain

*Corresponding Author: Tel+44(1484)472668; Fax +44(1484)472182;

E-mail address a.p.laws@hud.ac.uk

Abstract

The bile adapted strain *Bifidobacterium animalis* subsp. *lactis* IPLA-R1 secretes a high molecular weight exopolysaccharide (HMW-EPS) when grown on the surface of agar-MRSC. This EPS is composed of L-rhamnopyranosyl, D-glucopyranosyl, D-galactopyranosyl and D-galactofuranosyl residues in the ratio of 3:1:1:1. Linkage analysis and 1D and 2D-NMR spectroscopy were used to show that the EPS has a hexasaccharide repeating unit with the following structure:

α-D-Gal*f*-(1→2)-α-L-Rha*p*
1
↓
2
→4)-β-D-Glc*p*-(1→3)-α-L-Rha*p*-(1→2)-α-L-Rha*p*-(1→4)-β-D-Gal*p*-(1→

Treatment of the EPS with mild acid cleanly removed the terminal D-galactofuranosyl residue. The *eps* cluster sequenced for strain IPLA-R1 showed high genetic homology with putative *eps* clusters annotated in the genomes of strains from the same species. It is of note that several genes coding for rhamnose-precursors are present in the *eps* cluster, which could be correlated with the high percentage of rhamnose detected in its EPS repeated unit.

Keywords:

Bifidobacterium, exopolysaccharide, EPS structure, *eps* cluster, glycosyltransferase

1. Introduction

Bifidobacteria are Gram-positive non-spore forming, non motile, non filamentous rods which can display various shapes, the most typical one is bifurcated with spatulated extremities. They are strict anaerobes, with high G+C content belonging to the phylum *Actinobacteria* and they are normal constituents of a healthy gut microbiota of animals. Currently more than 30 species are included in the genus *Bifidobacterium*, whilst most abundant in the human gastrointestinal tract are *B. adolescentis, B. angulatum, B. bifidum, B. breve, B. catenulatum, B. longum* and *B. pseudocatenulatum.* Bifidobacteria are regarded as probiotic microorganisms and are increasingly being consumed as supplements in foods or in pharmaceutical formulations, to promote a healthy gut microbiota balance. The species most often found in functional dairy products is *Bifidobacterium animalis* subsp. *lactis*[1].

Probiotics have been defined as “live microorganisms, which when administered in adequate amounts confer a health benefit on the host”[2]. The degree of scientific evidence of probiotic effect in humans is scarce since most of these beneficial claims are based on the extrapolation of results of *in vitro* and animal model experiments. There are only a limited number of reports showing efficacy of a few specific probiotic strains in human intervention studies. Several meta-analyses collecting clinical evidence have demonstrated probiotic efficacy in alleviating lactose intolerance, antibiotic associated diarrhoea, atopic allergy in infants and some inflammatory bowel diseases[3,4]. One of the potential mechanisms by which probiotic bacteria can elicit their health benefits is through the surface molecules such as the exopolysaccharides (EPS). These biopolymers are exocellular carbohydrates that can be: covalently linked to bacterial surface forming a capsule; they can be non-

covalently associated with the surface or be totally secreted. Several health benefits have been *in vitro* attributed to EPS, such as cholesterol lowering capability, prebiotic effect and modulating the immune response[5]. Bifidobacteria isolated from human intestinal origin are able to synthesise EPS composed of more than one type of monosaccharide[6,7]. Regarding the putative role of EPS produced by bifidobacteria in the gut environment, it has been reported that bile salts induce their synthesis in some strains of *B. animalis* subsp. *animalis*[8]. Thereby, it seems that these polymers could have a protective role for the producing bacteria. This property is interesting for orally delivered strains since it could help bifidobacteria to survive the challenges, mainly acidic conditions and high concentration of bile salts that they will encounter on their transit from the mouth to the small intestine. In the large intestine, EPS-producing bifidobacteria will meet a complex ecosystem inhabited by a vast number of microorganisms. Salazar and co-workers have reported that a number EPS isolated from bifididobacteria have an ability to *in vitro* modulate the composition of human intestinal microbiota [9].A similar effect has been recently shown *in vivo* using rats fed with the EPS-producing strain *B. animalis* subsp. *lactis* IPLA-R1[10]. Additionally, the EPS produced by this strain was able to *in vitro* counteract the cytotoxic effect of bacterial toxins upon colonocyte-like Caco-2 cells. This EPS-fraction was analysed by size exclusion chromatography coupled with multi-angle laser light scattering detection (SEC-MALLS) and it was found a distribution of three molecular weight peaks differing in size[11].

Whilst there are a number of reports of EPS producing bifidobacteria, very little work has been undertaken to fully characterise the EPS that they produce, as well as their genetic determinants. The structures of the EPS produced by *B. bifidum* BIM B-465[12] and *B. longum* JBL05[13] both of human origin, have recently been

reported. But, as far as we know, the functional characterization of genes coding for enzymes involved in EPS-synthesis in *Bifidobacterium* has not been undertaken to date. A recent comparative analysis of bifidobacterial genomes shows that the presence of putative *eps* clusters seems to be an ubiquous character in this genus[14]. Thereby, the aim of this study was to analyse the structure of the EPS produced by *B. animals* subsp. *lactis* IPLA-R1, a promising strain with probiotic potential, and to analyse the sequence of the putative *eps* cluster coding for proteins involved in the synthesis of this polymer.

2. Results and Discussion

2.1. Structure of the HMW-EPS polymer synthesised by strain IPLA-R1

The crude-EPS sample purified from *B. animals* subsp. *lactis* IPLA-R1 had a protein content of 3.9% and the molecular weight distribution was similar to that previously reported by SEC-MALLS[11]: a high molecular weight fraction (HMW) with average molecular weight of 3.5 x 10^6 Da, a middle weight EPS (3.0 x 10^4 Da) and a low molecular weight EPS (4.9 x 10^3 Da). Dialysis of the crude-EPS sample against a 100 kDa cellulose acetate membrane separated the HWM-EPS, which was isolated in the retentate with reasonable purity. By SEC-MALLS separation, the average molecular weight of the HMW-EPS in the retentate was measured as 3.5 x 10^6 Da. The purity of the HMW-EPS was also determined by comparison of the anomeric region of the NMR spectra before (Fig 1a) and after dialysis (Fig1b), only a small amount of additional material (assumed to be middle weight EPS) was present. Six anomeric protons appear in the anomeric region of the ^{1}H NMR spectrum of the HMW-EPS suggesting that the repeating unit is a hexasaccharide; from this point

forward, the anomeric signals of the individual monomers are arbitrarily labelled as **A** to **F**, in decreasing order of their chemical shifts.

The results of monomer analysis and determination of the absolute configuration of the monomers indicate that the polysaccharide is composed of L-rhamnose, D-galactose and D-glucose in a molar ratio of 2.85:1.97:1. After performing linkage analysis, five unique methylated alditol acetates were obtained including: a 1,4,5-tri-*O*-acetyl-2,3,6 -tri-*O*-methylglucitol (from 1,4-Glc*p*); a 1,4,5-tri-*O*-acetyl-2,3,6-tri-*O*-methylgalactitol (from 1,4-Gal*p*); a 1,2,5-tri-*O*-acety-3,4-di-*O*-methylrhamnitol (from 1,2-Rha*p*); a 1,2,3,5-tetra-*O*-acetyl-4-*O*-methylrhamnitol (from 1,2,3- Rha*p*); and a 1,4-di-*O*-acetyl-2,3,5,6-tetra-*O*-methylgalactitol (from t-Gal*f*).

The structure of the HMW-EPS was determined using the results of the linkage analysis and by inspection of a range of 1D and 2D-NMR spectra including: ^{1}H-^{1}H COSY; ^{1}H-^{1}H TOCSY; ^{1}H-^{13}C HMBC; ^{1}H-^{13}C HSQC and ^{1}H-^{13}C HSQC-TOCSY spectra. The first thing to note is the anomeric configuration of each of the monomers, for monomers **A** to **D** this was determined by measurement of the magnitude of the $^{1}J_{C\text{-}H}$ coupling constant for the anomeric signals **A** (173 Hz)**, B** (175 Hz)**, C** (176 Hz) and **D** (175 Hz) these values are all greater than 170 Hz and identifies each as having α-linkages. The anomeric configuration of the two remaining monomers (**E** and **F)** was determined by measurement of the $^{3}J_{H1\text{-}H2}$ coupling constants which were both greater than 8 Hz identifying that **E** and **F** residues have β-linkages.

The position of the remaining proton resonances (H-2 to H-6) was determined using a combination of the COSY and TOCSY spectra. On the ^{1}H-^{1}H TOCSY spectrum (120 ms, data not shown) there are cross peaks from the anomeric protons of residues **A, B** and **C** to H-4 and, in the methyl region, from H-6 to H-3, identifying

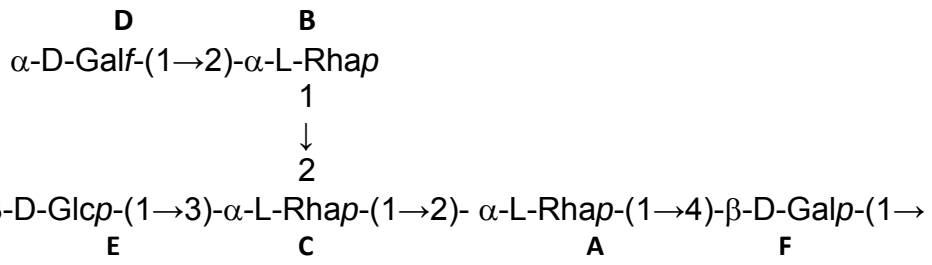
A, B and **C** as the rhamnose monomers. For residues **D, E** and **F** cross peaks on the COSY and TOCSY spectra identified the positions of H-1 through to H-4. The exact positions of H-5- and H-6 were not easily determined as there is poor transmission of coupling beyond H-4. The position of the remaining resonances was obtained from inspection of a ^{13}C Dept spectrum and from the HQSC spectrum. On the ^{13}C Dept spectrum, the C-6 resonances are located together at approximately 60 ppm. Once all the cross-peaks on the HSQC spectrum for C1/H1 to C4/H4 and for C6/6 had been assigned (Fig 2) the three remaining cross peaks must be those belonging to C/H-5s. Finally, these were assigned to individual monomers by cross reference to a HSQC-TOCSY spectrum. As the structure does not have any glycosidic links involving the hydroxyls at C6, our failure to assign the individual H-6 residues to **D**, **E** and **F** has no consequence for the characterisation of the EPS. For clarity, the resonance position for the ^{1}H signals and ^{13}C signals (H2 to H5 and C2 to C5) are indicated on the ^{1}H-^{13}C HSQC spectrum (Fig 2a, anomeric signals appear on the inset Fig 2b) the anomeric proton resonances are also indicated on the ^{1}H-^{1}H NOESY spectrum (Fig 3) and the combined chemical shift data for the complete repeating unit is presented in Table 1.

The very low field chemical shift values for H-2, H-3 and the H-4 resonances of residue **D,** combined with the loss of this residue under mildly acidic conditions (see discussion below), is evidence for residue **D** being the terminal Gal*f*. This would leave residues **E** and **F** as the 1→4 linked hexoses. One way of differentiating between galacto- and glucopyranoses is from inspection of the chemical shifts for their H-4 resonances: for a galactose H-4 is shifted substantially to lower field than that of a glucose regardless of the anomeric configuration and linkage. Data collected from assignments for lactic acid bacteria (LAB) EPS structures show that

the H-4 resonances for galactose lie in the range 4.30-3.85 δ whilst those for glucose lie in the range 3.45 -3.75δ[15]. The chemical shift for **E** H-4 (3.61) and that for **F** H-4 (4.39) implies that **E** is the glucopyranose sugar and **F** is the galactopyranose.

Information regarding the sequence of the sugar residues in the repeating unit was obtained from examination of the NOESY spectrum (Fig 3) and the anomeric region of the HMBC spectrum (not shown). On the NOESY spectrum there are strong inter residue NOEs between: **A** H-1 and **F** H-4, **B** H-1 and **C** H-2, **D** H-1 and **B** H-2, **E** H-1 and **C** H-3, and **F** H-1 and **E** H-4; identifying **A**(1→4)**F** , **B**(1→2)**C**, **D**(1→2)**B, E**(1→3)**C** and **F**(1→4)**E** linkages. On the HMBC spectrum inter-residue couplings are observed between: **A** H-1 and **F** H-4, **B** H-1 and **C** H-2, and **D** H-1 and **B** H-2, matching the NOE signals and, additionally, a cross peak is observed between **C** H-1 and **A** H-2.

Using a combination of the results for the linkage analysis and the NMR identifies the structure for the repeating unit as:



This is a novel structure and differs to those of the EPS structures that have been reported for *B. bifidum* BIM B-465[12] and *B. longum* JBL05[13] and is also different to the EPSs that have been isolated and characterised from LAB[15].

Mild acid catalysed hydrolysis of the HMW EPS

It is well known that glycosidic links to Gal*f* residues are hydrolysed in acidic solution[16]. Treatment of the HMW-EPS with a dilute solution of trifluroacetic acid in

193 an NMR tube caused the loss of the Gal*f* residue which we were able to monitor over
194 time (Fig 4). Within a period of 8 hours the anomeric signal, from residue **D**, reduced
195 in intensity and ultimately merged with the spectral noise. It seems that the EPS
196 IPLA-R1 could have a protective role during the transit of the producing bacteria
197 through the upper part of the intestinal tract[10]. However, in the *in vivo* situation this
198 Gal*f* residue could be either partially or totally lost. Thereby, care will be needed in
199 attempting to correlate EPS structure with biological activity measured *in vitro*, since
200 the passage through the gastrointestinal tract could modify its composition.
201
202 ***2.2 Putative* eps *cluster of* B. animalis *subps.* lactis *IPLA-R1***

203 A fragment of 54,259 bp containing the putative *eps* cluster of the strain *B.*
204 *animalis* subsp. *lactis* IPLA-R1 was sequenced and the putative function of coded
205 proteins has been studied by homology comparison with sequences held in the
206 GenBank database. A high genetic homology was detected among the putative *eps*
207 cluster of our strain IPLA-R1 and those of the five *B. animalis* subsp. *lactis* whose
208 genomes are currently available (data not shown). As expected, the functional maps
209 were almost identical among the six strains; this was not surprising since it has been
210 shown that there is a scarce inter-strain genetic variability within this *Bifidobacterium*
211 species[17].

212 The best characterised *eps* clusters, either by homology comparison or by
213 functional studies, are those of LAB and, in fact, a functional structure has been
214 found among these clusters[18]. Taking into account this functional structure, several
215 proteins involved in the synthesis of EPS have been found in the *eps* cluster of *B.*
216 *animalis* subp. *lactis* IPLA-R1 (Figure 5). This is the case of glycosyltransferases

(GTF), priming-GTF (p-GTF), genes involved in export of repeated unit, its polymerization and chain length determination, as well as mobile elements (transposase and insertion sequences (IS)). However, no gene regulators have been found although this function could be played by proteins with unknown function within the cluster or others outside it. It is also surprising that a number of membrane proteins are present in the bifidobacterial *eps* cluster. Another remarkable feature in the IPLA-R1 *eps* cluster is the presence of two p-GTF, as previously denoted for *B. longum* subsp. *longum* NCC2705[19] and for the other *B. animalis* subsp. *lactis* strains[14]. This enzyme catalyses the transfer of an activated sugar to the lipid carrier C55, being the first step in the synthesis of the repeated unit that build the polymer. In *B. animalis* subsp. *lactis* IPLA-R1, one of the p-GTF was located in the 5´ end (annotated as "undecaprenyl-phosphate sugar phosphotraferase") and the second one was located down-stream (annotated as galactosyl transferase CpsD). The nucleotide sequences of the two p-GTFs of the strain IPLA-R1 were different to the corresponding homologues of the type-strain DSM10140. In addition, in the strain IPLA-R1 the change in the nucleotide sequence of *cpsD*, but no that of the undecaprenyl-phosphate sugar phosphotraferase gene, leaded to a modification in the translated amino acid. However, we do not know if this different amino acid residue could modify the function of this p-GTF and thereby, influence the synthesis of the EPS in IPLA-R1 strain. Variations in the nucleotide sequence were also detected in the transposase IS204/IS1001/IS1096/IS1165 located at the 3´end, probably due to gain or lose of nucleotides in this mobile element during each transposition.

In relation to the structure of the HMW-EPS IPLA-R1 determined in this study, it is worth mentioning the presence of genes coding for proteins involved in the

biosynthesis of rhamnose precursors in the putative *eps* gene cluster of this strain, which could correlate with the high rhamnose content (50%) of its hexasaccharide repeated unit. Normally, in LAB strains, the content of rhamnose does not exceed that of the other two common EPS monosaccharides: glucose and galactose. An exception is the strain *Lactobacillus rhamnosus* RW9595M whose glucose: galactose: rhamnose ratio is 2:1:4[20] and it also presents in its *eps* cluster rhamnose-precursor biosynthseis genes[21]. Additionally, in EPS isolated from intestinal strains the content of rhamnose was higher (52%) than in those isolated from foods (28%)[7]. In this regard, the *in silico* comparative analysis of five bifidobacteria species shows that rhamnose-precursor biosynthesis genes are present in three out of the five species, showing a high degree of protein homology (data not shown). These facts suggest that the high presence of rhamnose in intestinal EPS could play a role in this ecological niche, which deserves future investigation.

3. Experimental

3.1. Bacterial growth and purification of the *HMW-EPS*

The growth conditions and purification of the crude-EPS from strain *B. animalis* subsp. *lactis* IPLA-R1 (previously named A1dOxR) have previously been described[11]. In short, bacterial biomass grown in an anaerobic chamber [MG500 (Down Whitley Scientific, West Yorkshire, UK): 80% (v/v) N_2, 10% CO_2, 10% H_2] at 37ºC for 5 days on the surface of agar-MRSC [MRS (Biokar Diagnostics, Beauvais, France) + 0.25% L-cysteine (Sigma Chemical Co. St. Louis, MO, USA)] was collected with water. The bacterial suspension was mixed with 1 volume of 2M NaOH and the crude-EPS from the cell-free supernatant was precipitated with 2 volumes of chilled-absolute ethanol for 48 h at 4ºC. The precipitated fraction was

resuspended in ultra-pure water, dialysed in 12-14 kDa MWCO cellulose membranes (Sigma) against water for 3 days at 4ºC and finally freeze-dried to obtain the EPS-crude fraction.

A pure sample of the high molecular weight (HMW)-EPS from the crude-EPS fraction isolated from *B. animals* subsp. *lactis* IPLA-R1 was obtained by dialysis as follows. A crude-EPS sample (25 mg) was dissolved in distilled water (10 mL) with gentle heating (less than 50°C) and the HMW material was isolated in the retentate (about 10 mL), after dialysis (Spectra/Por Float-A-Lyser 100 KDa MWCO, Sigma) for 72 hours at 4°C against three changes of distilled water per day. The content of the dialysis membrane were freeze dried to provide HMW-EPS.

The purity of the HMW-EPS was determined by SEC-MALLS and NMR analysis. For SEC-MALLS, solutions of EPS in deionised water (1 mg mL^{-1}) were prepared and left for 24 h to completely dissolve. Samples (100 µl) were injected onto an analytical SEC system comprising three columns Aquagel-OH 40, 50 and 60 (15 µm particle size, 25 cm x 4 mm, Varian, Oxford, UK) connected in series. The neutral analytes were eluted with deionised water flowing at 1 mL min^{-1}. The concentration of the EPS fractions eluting from the column were determined by a differential refractometer (Optilab rEX, Wyatt technology, Santa Barbara, USA) and the weight average molecular weight was measured using a Dawn-EOS MALLS operating with a 690 nm laser (Wyatt technology, Santa Barbara).

3.2. Monomer composition and linkages of the HMW-EPS

The monomer composition of the HMW-EPS was determined after acid hydrolysis by HPAEC-PAD. The HMW-EPS (1 mg mL^{-1}) was treated with 2M TFA (120°C for 2 h) and the identity of the released monomers was determined using

high performance anion exchange chromatography (HPAEC) on a Dionex BioLC system (Sunnyvale CA, USA) equipped with a CarboPac PA20 column (150 mm x 3 mm). Monomers were eluted using a sodium 10 mM hydroxide mobile phase at a flow rate of 0.5 mL min^{-1} and detected using a pulsed amperometric detector (PAD) ED50 (Dionex) operating with a dual potential waveform. The ratio of monomers was determined by comparison of the detector response to calibration standards of the individual monomers (galactose, glucose and rhamnose, 5-100 ppm). The absolute configurations of monosaccharides were determined by conversion to their butyl glycosides using the procedure described by Gerwig et al[22].

For linkage analysis, the HMW-EPS was permethylated using the procedures described by Stellner et al[23]. The methylated-polysaccharide was hydrolysed by treatment with 2M TFA (120 °C for 2 h) and the methylated monosaccharides converted to their corresponding methylated alditol acetates. The identity of the variously methylated alditol acetates was determined by GLC-MS and by analysis of the individual fragmentation patterns observed in the MS. GLC-MS analyses were performed on an Agilent 7890A GC system (Santa Clara, CA, USA) coupled to an Agilent 5675c quadrupole MS. The samples were eluted from a HP-5 column (30 m x 0.25 mm-id, 0.25 µm film) using helium as carrier (9 psi, flow rate 1 mL min^{-1}) and using the following temperature programme: start temperature 155°C, hold time 1 min, and a final column temperature of 195°C reached via a rising gradient of 0.75°C min^{-1}.

3.3. NMR structure of the HMW-EPS

NMR spectra were recorded for HWM-EPS samples that were dissolved (10 mg mL^{-1}) directly in D_2O (Goss Scientific Instruments Ltd., Essex, UK). NMR spectra were recorded at a probe temperature of 70°C. The elevated temperature was

315 initially chosen as it shifted the HOD signal to higher field, into a clear region of the
316 spectrum. The higher temperature also increased spectral resolution by reducing the
317 sample viscosity. All of the NMR spectra were recorded on a Bruker Avance 500.13
318 MHz ^{1}H (125.75 MHz ^{13}C) spectrometer (Bruker-biospin, Coventry, UK) operating
319 with Z-field gradients where appropriate, and using Bruker's pulse programs.
320 Chemical shifts are expressed in ppm relative to either internal or external acetone; δ
321 2.225 for ^{1}H and δ 31.55 for ^{13}C. The 2D gs-DQF-COSY spectrum was recorded in
322 magnitude mode at 70°C. TOCSY experiments were recorded with variable mixing
323 times (60, 90, 120 ms). The 2D-heteronuclear ^{1}H-^{13}C HSQC, and phase sensitive
324 HSQC-TOCSY were recorded using Bruker pulse sequences and 256 experiments
325 of 1024 data points. The NOESY spectrum was recorded using a Bruker pulse
326 sequence and 256 experiments of 1024 data points were recorded using a mixing
327 time of 200 ms. For the majority of spectra, time-domain data were multiplied by
328 phase-shifted (squared-) sine-bell functions. After applying zero-filling and Fourier
329 transformation, data sets of 1024-1024 points were obtained.

330 A mild acid hydrolysis treatment of the HMW-EPS was carried out as follows:
331 a solution of EPS (10 mg mL^{-1}) was mixed with 20 µl of TFA in an NMR tube and
332 kept at 70ºC for 24 h. Hydrolysis was monitored by ^{1}H NMR, spectra were recorded
333 every hour for first 8 h and then after 24 h.

334 ***3.4. Sequencing of the putative* eps *cluster of* B. animalis *subsp.* lactis *IPLA-R1***

335 Strain *B. animalis* subsp. *lactis* IPLA-R1 was grown for 24 h in 10 mL of
336 MRSC broth to isolate DNA using the "GenElute Bacterial Genomic DNA" kit (Sigma)
337 following the manufacturer instructions, but including a previous step of bacterial
338 lysis with mutanolysin and lysozyme[8]. For sequencing the putative *eps* cluster, 54

pair of PCR primers were designed taking into account the sequence of the type strain *B. animalis* subsp. *lactis* DSM10140, whose genome is publicly available (GenBank accession number CP001606)[24]. Primers, synthesised by Thermo-Fisher Scientific GmbH (Ulm, Germany), amplified regions of about 1,000 bp. The PCR reaction mixture in a final volume of 50 µl was: 1 µl chromosomal DNA, 0.20 µM of each primer, 200 µM dNTPs (Amersham Bioscience, Upsala, Sweden) and 2.5 U Taq DNA-polymerase (Eppendorf, Hamburg, Germany). The PCR thermal conditions were an initial denaturalisation cycle 95ºC for 5 min, 30 amplification cycles of: 95ºC for 1 min, 52 or 56ºC (variable according to the pair of primers) for 50 s and 68ºC for 2 min, and a final extension step of 68ºC for 10 min. Amplification was done in a UnoCycler thermal cycler (VWR International Eurolab S.L., Barcelona, Spain). PCR amplified products were visualised under UV in 1% agarose gels, after staining with ethidium bromide. Purification and sequencing of each amplicons (both strands) was performed by Macrogen (Seoul, Korea). Sequences obtained were processed with the free Chromas 1.45 software (Technelysium Pty Ltd., Australia) and used for comparison with those held in the GenBank database (http://www.ncbi.nlm.nih.gov/genbank) using the BLASTn and BLASTp tools.

Acknowledgements

This work was funded by the Spanish “Plan Nacional I+D+I” from the “Ministerio de Ciencia e Innovación” (MICINN, FEDER funds from European Union) through the project AGL2009-09445 and by the University of Huddersfield. C. Hidalgo-Cantabrana acknowledges his FPI fellowship to MICINN.

References

(1) Margolles, A.; Mayo, B.; Ruas-Madiedo, P. Screening, identifiation and characterization of *Lactobacillus* and *Bifidobacterium* strains. In *Handbook of probiotics and prebiotics,* 2nd ed; Nomoto, K., Salmine, S., Lee, Y.K., Eds; John Willey & Sons Inc.: New Jersey, **2009**; pp 4-43.

(2) FAO/WHO. Probiotics in foods. Health and nutritional properties and guidelines for evaluation. In *FAO food and nutrition paper* No. 85 (ISBN 92-5-105513-0). **2006**; pp 1-50.

(3) Mayo, B.; Delgado, S., Rodríguez, J.M.; Gueimonde, M.; *CAB Rev*., **2008**, No. 055 (Online ISSN 1749-8848); pp 1-17.

(4) WGO. In *World Gastroenterology Organisation Practice Guidelines: Probiotics and Prebiotics* (http://www.worldgastroenterology.org/assets/downloads/en/pdf/guidelines/19_probiotics_prebiotics.pdf) **2008**; pp 1-22.

(5) Ruas-Madiedo, P.; Abraham, A.; Mozzi, F.; de los Reyes-Gavilán, C.G. Functionality of exopolysaccharides produced by lactic acid bacteria. In *Molecular aspects of lactic acid bacteria for traditional and new applications*, Mayo, B., López, P., Pérez-Martínez, G., Eds.; Research Signpost Publishers: Kerala, India, **2008**; pp 137-166.

(6) Ruas-Madiedo, P.; Moreno, J.A.; Salazar, N.; Delgado, S.; Mayo, B.; Margolles, A.; de los Reyes-Gavilán, C.G. *Appl. Environ. Microbiol*., **2007**, *73*, 4385-4388.

(7) Salazar, N.; Prieto, A.; Leal, J.A.; Mayo, B.; Bada-Gancedo, J.C.; de los Reyes-Gavilán, C.G.; Ruas-Madiedo, P., *J. Dairy Sci*., **2009**, *92*, 4158-4168.

(8) Ruas-Madiedo, P.; Gueimonde, M.; Arigoni, F.; de los Reyes-Gavilán, C.G.; Margolles, A. *Appl. Environ. Microbiol*., **2009**, *75*, 1204-1207.

(9) Salazar, N.; Gueimonde, M.; Hernández-Barranco, A.M.; Ruas-Madiedo, P.; de los Reyes-Gavilán, C.G. *Appl. Environ. Microbiol*., **2008**, *74*, 4737-4745.

(10) Salazar, N.; Binetti, A.; Gueimonde, M.; Alonso, A.; Garrido, P.; González del Rey, C.; González, C.; Ruas-Madiedo, P.; de los Reyes-Gavilán, C.G., *Int. J. Food Microbiol.*, **2011**, *144*, 342-351.

(11) Ruas-Madiedo, P.; Medrano, M., Salazar, N.; de los Reyes-Gavilán, C.G.; Pérez, P.F.; Abraham, A. *J. Appl. Microbiol*., **2010**, *109*, 2079-2086.

(12) Zdorovenko, E.L.; Kachala, V.V.; Sidarenka, A.V.; Izhik, A.V.; Kisileva E.P.; Shashkov, A.S.; Novik, G.I.; Knirel, Y.A. *Carbohyd. Res*., **2009**, *344*, 2417-2420.

(13) Kohno, M.; Suzuki, S.; Kanaya, T.; Yoshino, T.; Matsuura, Y.; Asada, M.; Kitamura, S., *Carbohyd. Polym.*, **2009**, *77*, 351-357.

(14) Lee, J.H.; O´Sullivan, J., *Microbiol. Mol. Biol. Rev*., **2010**, *74*, 378-416.

(15) Harding, L. P.; Marshall, V. M.; Hernandez, Y.; Gu, Y. C.; Maqsood, M.; McLay, N.; Laws, A. P. *Carbohydr. Res.,* **2005**, *340*, 1107-1111.

(16) Carpita, N.C.; Shea, E.M. Linkage Structure Of Carbohydrates By Gas-Chromatography - Mass Spectrometry Of Partially Methylated Alditol

406 Acetates. In *Analysis of Carbohydrates by GLC and MS.* Bierman C.J.;
407 McGinnis G.D. Eds.; CRC Press: Florida, USA, **1988**; pp 157-216.

408 (17) Briczinski, E.P.; Loquasto, J.R.; Barrangou, R.; Dudley, E.G.; Roberts, A.M.;
409 Roberts, R.F., *Appl. Environ. Microbiol*., **2009**, *75*, 7501-7508.

410 (18) Ruas-Madiedo, P.; Salazar, N.; de los Reyes-Gavilán, C.G. Biosynthesis and
411 chemical composition of exopolysaccharides produced by lactic acid bacteria.
412 In *Bacterial polysaccharides: current innovations and future trends.* Ullrich, M.,
413 Ed.; Caister Academic Press: Norfolk, UK, **2009**; pp 279-310.

414 (19) Schell, M.A., Karmirantzou, M., Snell, B.; Vilanova, D.; Berger, B.; Pessim G.;
415 Zwahlen, M.C.; Desiere, F.; Bork, P.; Delley, M.; Pridmore, R.D.; Arigoni, F.
416 *PNAS,* **2002**, *99*, 14422-14427.

417 (20) van Calsteren, M.R.; Pau-Roblot, C.; Begin, A.; Roy, D., *Biochem. J*., **2002**, *363*,
418 7-17.

419 (21) Péant, B.; LaPointe, G.; Gilbert, C.; Atlan, D.; Ward, P.; Roy, D., *Microbiology*,
420 **2005**, *151*, 1839-1851.

421 (22) Gerwig, G. J.; Kamerling, J. P.; Vliegenthart, J. F. G. *Carbohydr. Res.* **1978**, *62*,
422 349-357.

423 (23) Stellner, K.; Saito, H.; Hakomori, S. I. *Arch. Biochem. Biophys.***1973**, *155*, 464-
424 472.

425 (24) Barrangou, R.; Briczinski, E.P.; Traeger, L.L.; Loquasto, J.R.; Richards, M.;
426 Hovath, P.; Coûte-;Monvoisin, A.C.; Leyer, G.; Rendulic, S.; Steele, J.L.;

Broadbent, J.R.; Oberg, T.; Dudley, E.G.; Schuster, S.; Romero, D.A.; Roberts, R.F., *J. Bacteriol*., **2009**, *191*, 4144-4151.

Figure Legends

Figure 1 500 MHz ^{1}H NMR spectrum of the HMW-EPS obtained after separation of crude-EPS by dialysis in a 100 kDa cellulose acetate membrane; spectra recorded in D_2O at 70^{o}C. Inset shows an expanded plot of the anomeric region.

Figure 2 (a): 500-MHz ^{1}H-^{13}C HSQC spectrum of a selected region of the HMW-EPS from *B. animalis* subsp. *lactis* IPLA-R1 recorded in D_2O at 70^{o}C. The identity of the cross peaks is noted by the sugar residue (**A-F**) and by identifying the location of hydrogens/carbons within the ring as **2-5**. (b): anomeric region of the 500-MHz ^{1}H-^{13}C HSQC spectrum.

Figure 3: 500-MHz ^{1}H-1H NOESY spectrum of a selected region of the HMW-EPS from *B. animalis* subsp. *lactis* IPLA-R1 recorded in D_2O at 70^{o}C. The identity of the cross peaks is noted by the sugar residue (**A-F**) and by identifying the location of hydrogens within the ring as **1-5**. Intra-residue couplings are highlighted in red and inter-residue couplings are highlighted in yellow. For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.

Figure 4: 500-MHz 1H–13C HMBC spectrum of a selected region of the HMW-EPS from *B. animalis* subsp. *lactis* IPLA-R1 recorded in D2O at 70°C. The identity of the

cross peaks is noted by the sugar residue (A–D) and by identifying the location of hydrogens within the ring as 1–5. Intra-residue couplings are highlighted in red and inter-residue couplings are highlighted in yellow. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

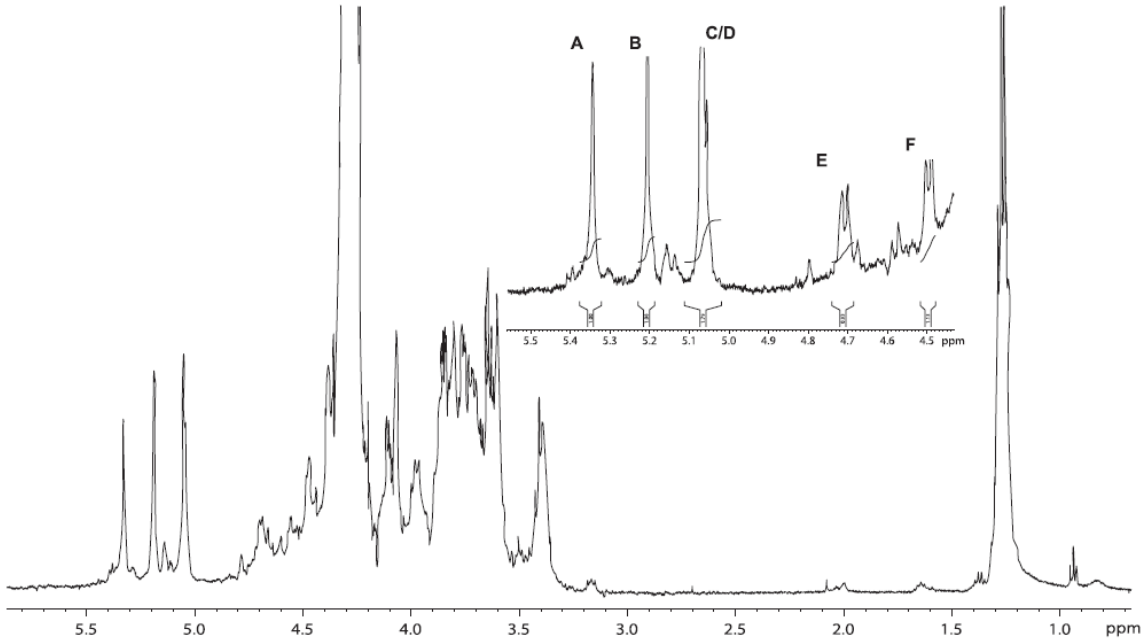
Figure 5: 500 MHz ^{1}H NMR spectra of the anomeric region following the acid catalysed hydrolysis of the HMW-EPS from *B. animalis* subsp. *lactis* IPLA-R1 as a function of time; spectra were recorded in D_2O at 70^{o}C and the sample was maintained at 70^{o}C for the full reaction period. Sugar residues in the native HMW-EPS are identified at the bottom (**A to F**), whereas those of the hydrolysed sample are identified at the top (**A' to F'**).

Figure 6: Physical map of the putative *eps* cluster (54,259 bp) of *B. animalis* subsp. *lactis* IPLA-R1. Predicted protein functions are categorised as follows: **GTF**, glycosyltransferase; **p-GTF**, priming-GTF/ undecaprenyl-phosphate sugar phosphotransferase / galactosyltransferase (CpsD); **Rh-B**, rhamnose biosynthesis precursors; **P-ChL**, polymerization (polymerase) – chain length determination; **P-E**, polymerization – export (Wzx and Wzz flippases); **PB**, polysaccharide biosynthesis; **AcS**, acyl-synthetase; **Ph**, phosphorilase; **MP**, membrane protein; **T-IS**, transposase – IS mobile elements; White arrows without label indicate hypothetical proteins.

Table 1: ^{1}H and ^{13}C NMR chemical shifts of the HMW-EPS from *B. animalis* subsp. *lactis* IPLA-R1 recorded in D_2O at 70^{o}C

476 **Figure 1**

477



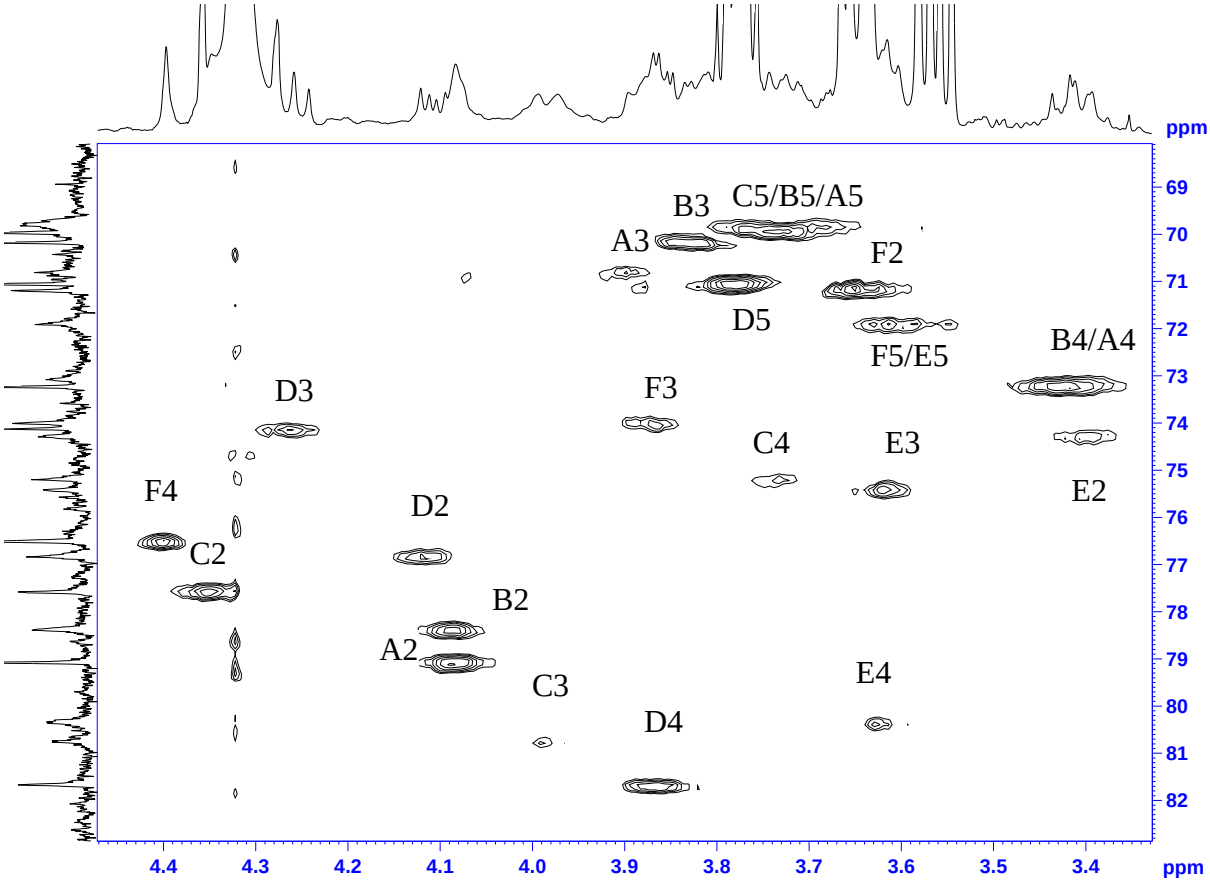
478

479

480

Figure 2

a


ppm
69
70
71
72
73
74
75
76
77
78
79
80
81
82
B3
C5/B5/A5
A3
F2
D5
F5/E5
B4/A4
D3
F3
C4
E3
F4
D2
E2
C2
B2
A2
C3
E4
D4
4.4
4.3
4.2
4.1
4.0
3.9
3.8
3.7
3.6
3.5
3.4
ppm

b

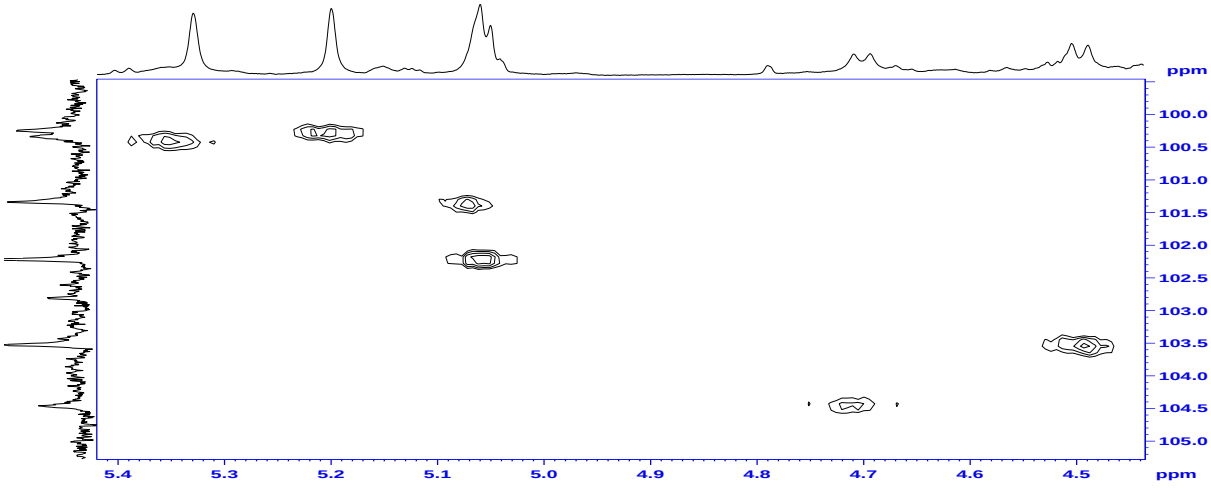
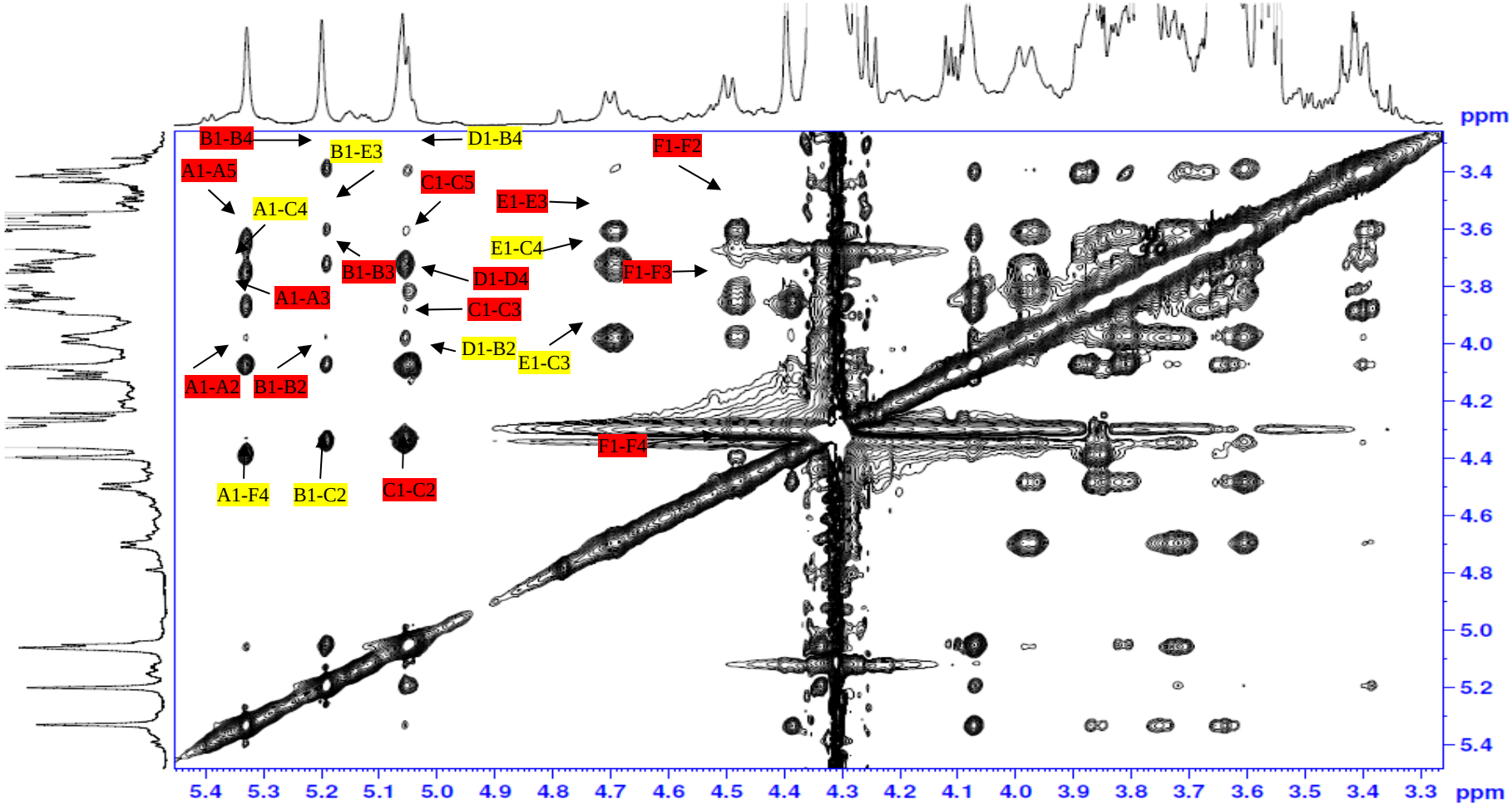

ppm
100.0
100.5
101.0
101.5
102.0
102.5
103.0
103.5
104.0
104.5
105.0
5.4
5.3
5.2
5.1
5.0
4.9
4.8
4.7
4.6
4.5
ppm

Figure 3


B1-B4
B1-E3
D1-B4
F1-F2
A1-A5
C1-C5
E1-E3
A1-C4
E1-C4
B1-B3
F1-F3
D1-D4
A1-A3
C1-C3
D1-B2
E1-C3
A1-A2
B1-B2
F1-F4
A1-F4
B1-C2
C1-C2
ppm
3.4
3.6
3.8
4.0
4.2
4.4
4.6
4.8
5.0
5.2
5.4
5.4 5.3 5.2 5.1 5.0 4.9 4.8 4.7 4.6 4.5 4.4 4.3 4.2 4.1 4.0 3.9 3.8 3.7 3.6 3.5 3.4 3.3
ppm

..

Figure 4

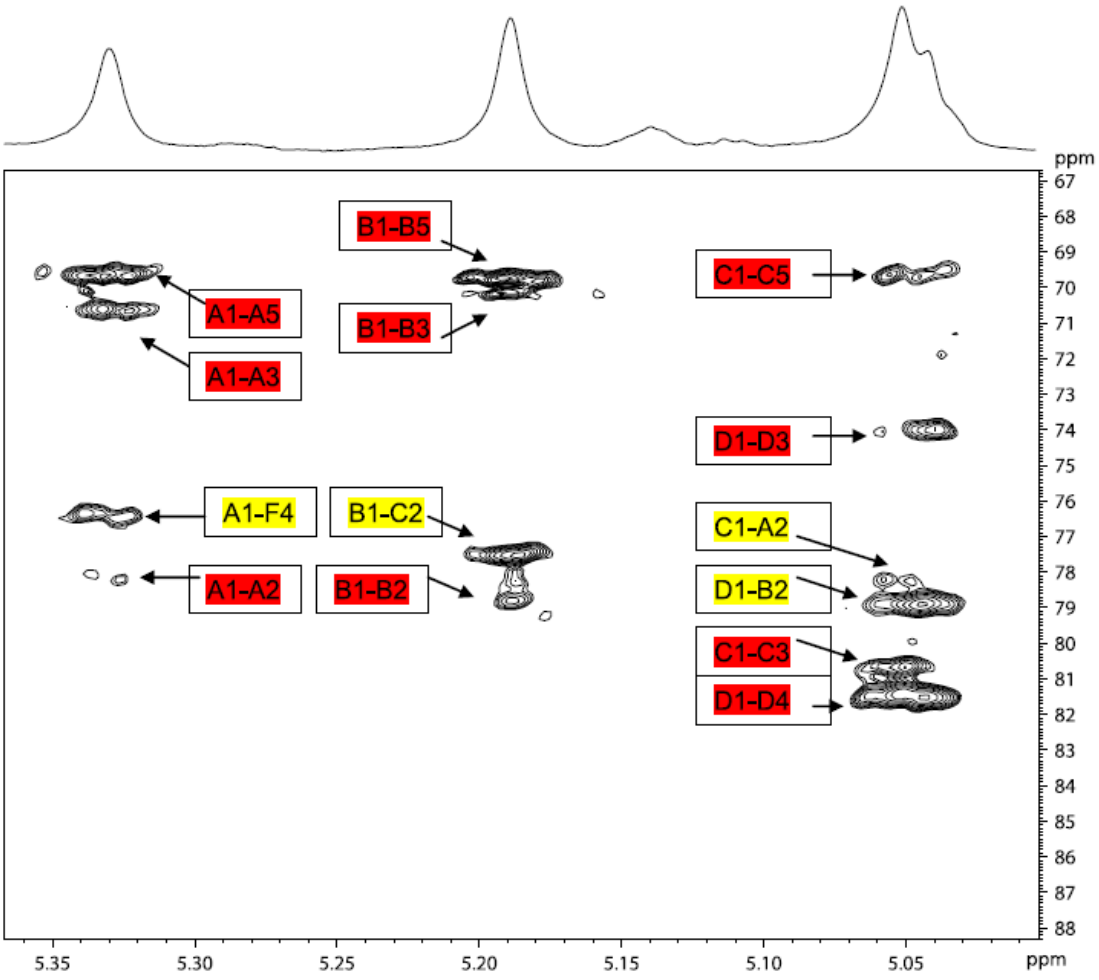

B1-B5
A1-A5
B1-B3
A1-A3
C1-C5
D1-D3
A1-F4
B1-C2
C1-A2
A1-A2
B1-B2
D1-B2
C1-C3
D1-D4
ppm
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
5.35
5.30
5.25
5.20
5.15
5.10
5.05
ppm

Figure 5

Figure 4

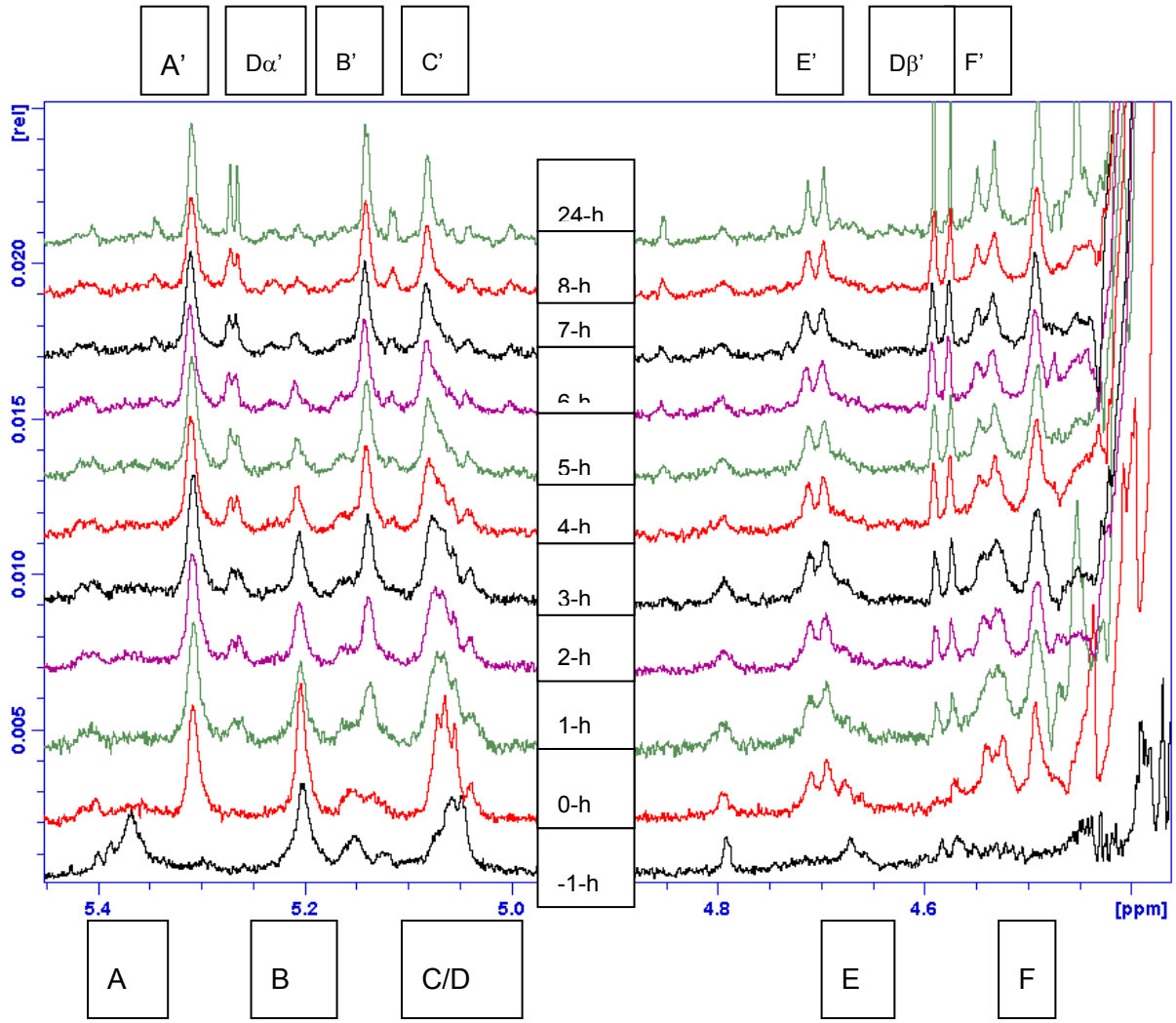


Figure 6

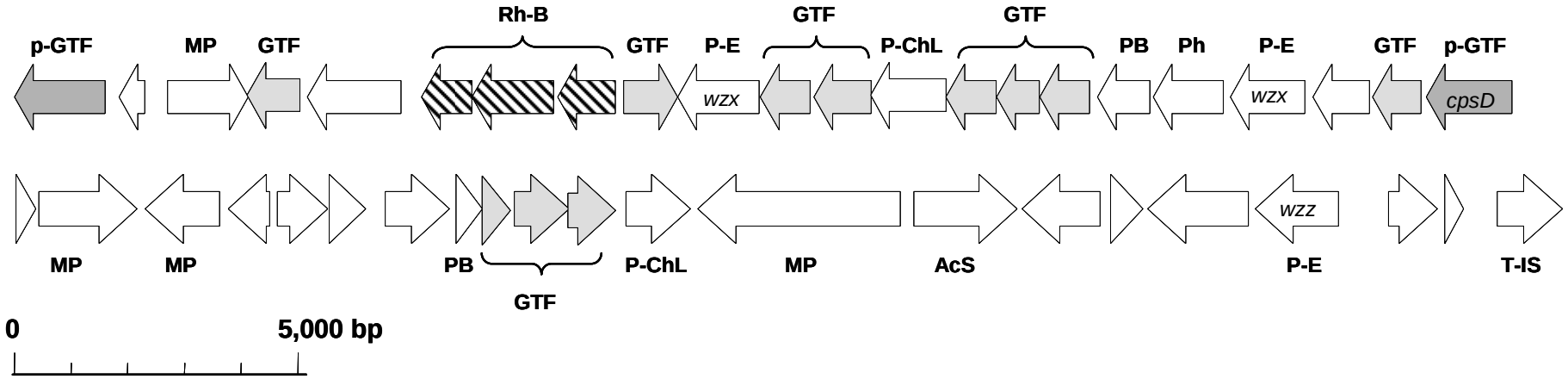

Rh-B
GTF
GTF
p-GTF
MP
GTF
GTF
P-E
P-ChL
PB
Ph
P-E
GTF
p-GTF
wzx
wzx
cpsD
MP
MP
PB
P-ChL
MP
AcS
wzz
P-E
T-IS
GTF
0
5,000 bp

Table 1

Residue	^{1}H Chemical Shift (ppm)							^{13}C Chemical Shift (ppm)					
	H1	**H2**	**H3**	**H4**	**H5**	**H6**	**H6'**	**C1**	**C2**	**C3**	**C4**	**C5**	**C6**
A	5.43	4.08	3.89	3.40	3.67	1.25	1.25	100.35	79.09	70.84	73.09	69.80	17.40
B	5.20	4.08	3.83	3.42	3.72	1.27	1.27	100.28	78.40	70.17	73.25	69.98	17.23
C	5.06	4.32	3.98	3.62	3.77	1.29	1.29	101.36	77.53	80.74	80.37	69.81	17.44
D	5.05	4.11	3.25	3.86	3.78	3.68*	3.90*	102.22	76.84	74.13	81.70	71.09	61.61
E	4.70	3.39	3.73	3.61	3.64	3.68*	3.90*	104.47	74.29	75.43	75.22	71.86	61.41
F	4.50	3.64	3.86	4.39	3.64	3.68*	3.90*	103.55	71.21	74.00	76.47	71.92	61.22

*Represents partially overlapping H6 and H6' resonances