1 **Structure of the high molecular weight exopolysaccharide**

2 **produced by *Bifidobacterium animalis* subsp. *lactis* IPLA-R1 and**

3 **sequence analysis of its putative *eps* cluster**

4

5 Shaun Leivers[a], Claudio Hidalgo-Cantabrana[b], Glenn Robinson[a], Abelardo

6 Margolles[b], Patricia Ruas-Madiedo[b] and Andrew P. Laws[a]*

7

8

9 [a] Department of Chemical and Biological Sciences, University of Huddersfield,

10 Queensgate, Huddersfield, HD1 3DH, United Kingdom

11 [b] Instituto de Productos Lácteos de Asturias, Consejo Superior de Investigaciones

12 Científicas (IPLA-CSIC), Carretera de Infiesto s/n, 33300 Villaviciosa, Asturias,
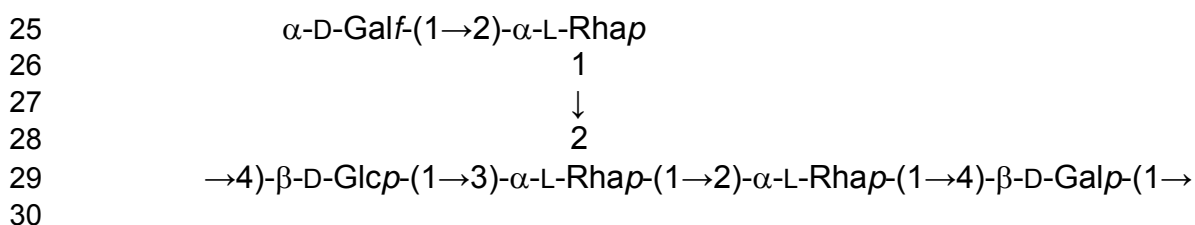
13 Spain

14

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

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

17

18   **Abstract**

19   The bile adapted strain *Bifidobacterium animalis* subsp. *lactis* IPLA-R1

20   secretes a high molecular weight exopolysaccharide (HMW-EPS) when grown on

21   the surface of agar-MRSC. This EPS is composed of L-rhamnopyranosyl, D-

22   glucopyranosyl, D-galactopyranosyl and D-galactofuranosyl residues in the ratio of

23   3:1:1:1. Linkage analysis and 1D and 2D-NMR spectroscopy were used to show that

24   the EPS has a hexasaccharide repeating unit with the following structure:

25   α-D-Gal*f*-(1→2)-α-L-Rha*p*
26   1
27   ↓
28   2
29   →4)-β-D-Glc*p*-(1→3)-α-L-Rha*p*-(1→2)-α-L-Rha*p*-(1→4)-β-D-Gal*p*-(1→
30

31   Treatment of the EPS with mild acid cleanly removed the terminal D-

32   galactofuranosyl residue. The *eps* cluster sequenced for strain IPLA-R1 showed high

33   genetic homology with putative *eps* clusters annotated in the genomes of strains

34   from the same species. It is of note that several genes coding for rhamnose-

35   precursors are present in the *eps* cluster, which could be correlated with the high

36   percentage of rhamnose detected in its EPS repeated unit.

37

40

1. Introduction

41

42   Bifidobacteria are Gram-positive non-spore forming, non motile, non

43 filamentous rods which can display various shapes, the most typical one is bifurcated

44 with spatulated extremities. They are strict anaerobes, with high G+C content

45 belonging to the phylum *Actinobacteria* and they are normal constituents of a healthy

46 gut microbiota of animals. Currently more than 30 species are included in the genus

47 *Bifidobacterium*, whilst most abundant in the human gastrointestinal tract are *B.*

48 *adolescentis, B. angulatum, B. bifidum, B. breve, B. catenulatum, B. longum* and *B.*

49 *pseudocatenulatum.* Bifidobacteria are regarded as probiotic microorganisms and

50 are increasingly being consumed as supplements in foods or in pharmaceutical

51 formulations, to promote a healthy gut microbiota balance. The species most often

52 found in functional dairy products is *Bifidobacterium animalis* subsp. *lactis*[1].

53   Probiotics have been defined as "live microorganisms, which when

54 administered in adequate amounts confer a health benefit on the host"[2]. The degree

55 of scientific evidence of probiotic effect in humans is scarce since most of these

56 beneficial claims are based on the extrapolation of results of *in vitro* and animal

57 model experiments. There are only a limited number of reports showing efficacy of a

58 few specific probiotic strains in human intervention studies. Several meta-analyses

59 collecting clinical evidence have demonstrated probiotic efficacy in alleviating lactose

60 intolerance, antibiotic associated diarrhoea, atopic allergy in infants and some

61 inflammatory bowel diseases[3,4]. One of the potential mechanisms by which probiotic

62 bacteria can elicit their health benefits is through the surface molecules such as the

63 exopolysaccharides (EPS). These biopolymers are exocellular carbohydrates that

64 can be: covalently linked to bacterial surface forming a capsule; they can be non-

65   covalently associated with the surface or be totally secreted. Several health benefits

66   have been *in vitro* attributed to EPS, such as cholesterol lowering capability,

67   prebiotic effect and modulating the immune response[5]. Bifidobacteria isolated from

68   human intestinal origin are able to synthesise EPS composed of more than one type

69   of monosaccharide[6,7]. Regarding the putative role of EPS produced by bifidobacteria

70   in the gut environment, it has been reported that bile salts induce their synthesis in

71   some strains of *B. animalis* subsp. *animalis*[8]. Thereby, it seems that these polymers

72   could have a protective role for the producing bacteria. This property is interesting for

73   orally delivered strains since it could help bifidobacteria to survive the challenges,

74   mainly acidic conditions and high concentration of bile salts that they will encounter

75   on their transit from the mouth to the small intestine. In the large intestine, EPS-

76   producing bifidobacteria will meet a complex ecosystem inhabited by a vast number

77   of microorganisms. Salazar and co-workers have reported that a number EPS

78   isolated from bifididobacteria have an ability to *in vitro* modulate the composition of

79   human intestinal microbiota [9].A similar effect has been recently shown *in vivo* using

80   rats fed with the EPS-producing strain *B. animalis* subsp. *lactis* IPLA-R1[10].

81   Additionally, the EPS produced by this strain was able to *in vitro* counteract the

82   cytotoxic effect of bacterial toxins upon colonocyte-like Caco-2 cells. This EPS-

83   fraction was analysed by size exclusion chromatography coupled with multi-angle

84   laser light scattering detection (SEC-MALLS) and it was found a distribution of three

85   molecular weight peaks differing in size[11].

86   Whilst there are a number of reports of EPS producing bifidobacteria, very

87   little work has been undertaken to fully characterise the EPS that they produce, as

88   well as their genetic determinants. The structures of the EPS produced by *B. bifidum*

89   BIM B-465[12] and *B. longum* JBL05[13] both of human origin, have recently been

90   reported. But, as far as we know, the functional characterization of genes coding for

91   enzymes involved in EPS-synthesis in *Bifidobacterium* has not been undertaken to

92   date. A recent comparative analysis of bifidobacterial genomes shows that the

93   presence of putative *eps* clusters seems to be an ubiquous character in this genus[14].

94   Thereby, the aim of this study was to analyse the structure of the EPS produced by

95   *B. animals* subsp. *lactis* IPLA-R1, a promising strain with probiotic potential, and to

96   analyse the sequence of the putative *eps* cluster coding for proteins involved in the

97   synthesis of this polymer.

98

99   **2. Results and Discussion**

100   ***2.1. Structure of the HMW-EPS polymer synthesised by strain IPLA-R1***

101   The crude-EPS sample purified from *B. animals* subsp. *lactis* IPLA-R1 had a

102   protein content of 3.9% and the molecular weight distribution was similar to that

103   previously reported by SEC-MALLS[11]: a high molecular weight fraction (HMW) with

104   average molecular weight of 3.5×10^6 Da, a middle weight EPS (3.0×10^4 Da) and a

105   low molecular weight EPS (4.9×10^3 Da). Dialysis of the crude-EPS sample against

106   a 100 kDa cellulose acetate membrane separated the HWM-EPS, which was

107   isolated in the retentate with reasonable purity. By SEC-MALLS separation, the

108   average molecular weight of the HMW-EPS in the retentate was measured as 3.5 x

109   10^6 Da. The purity of the HMW-EPS was also determined by comparison of the

110   anomeric region of the NMR spectra before (Fig 1a) and after dialysis (Fig1b), only a

111   small amount of additional material (assumed to be middle weight EPS) was present.

112   Six anomeric protons appear in the anomeric region of the ^{1}H NMR spectrum of the

113   HMW-EPS suggesting that the repeating unit is a hexasaccharide; from this point

114   forward, the anomeric signals of the individual monomers are arbitrarily labelled as **A**

115   to **F**, in decreasing order of their chemical shifts.

116   The results of monomer analysis and determination of the absolute

117   configuration of the monomers indicate that the polysaccharide is composed of L-

118   rhamnose, D-galactose and D-glucose in a molar ratio of 2.85:1.97:1. After

119   performing linkage analysis, five unique methylated alditol acetates were obtained

120   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*-

121   acetyl-2,3,6-tri-*O*-methylgalactitol (from 1,4-Gal*p*); a 1,2,5-tri-*O*-acety-3,4-di-*O*-

122   methylrhamnitol (from 1,2-Rha*p*); a 1,2,3,5-tetra-*O*-acetyl-4-*O*-methylrhamnitol (from

123   1,2,3- Rha*p*); and a 1,4-di-*O*-acetyl-2,3,5,6-tetra-*O*-methylgalactitol (from t-Gal*f*).

124   The structure of the HMW-EPS was determined using the results of the

125   linkage analysis and by inspection of a range of 1D and 2D-NMR spectra including:

126   ^{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-

127   TOCSY spectra. The first thing to note is the anomeric configuration of each of the

128   monomers, for monomers **A** to **D** this was determined by measurement of the

129   magnitude of the $^{1}J_{\text{C-H}}$ coupling constant for the anomeric signals **A** (173 Hz)**, B** (175

130   Hz)**, C** (176 Hz) and **D** (175 Hz) these values are all greater than 170 Hz and

131   identifies each as having α-linkages. The anomeric configuration of the two

132   remaining monomers (**E** and **F)** was determined by measurement of the $^{3}J_{\text{H1-H2}}$

133   coupling constants which were both greater than 8 Hz identifying that **E** and **F**

134   residues have β-linkages.

135   The position of the remaining proton resonances (H-2 to H-6) was determined

136   using a combination of the COSY and TOCSY spectra. On the ^{1}H-^{1}H TOCSY

137   spectrum (120 ms, data not shown) there are cross peaks from the anomeric protons

138   of residues **A, B** and **C** to H-4 and, in the methyl region, from H-6 to H-3, identifying
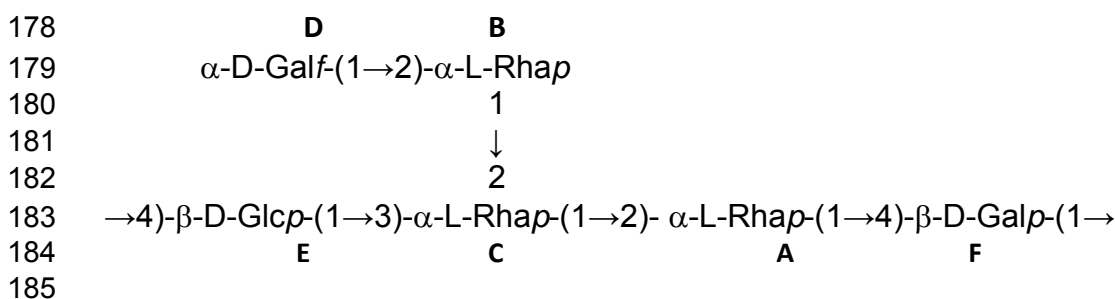
139 **A, B** and **C** as the rhamnose monomers. For residues **D, E** and **F** cross peaks on the

140 COSY and TOCSY spectra identified the positions of H-1 through to H-4. The exact

141 positions of H-5- and H-6 were not easily determined as there is poor transmission of

142 coupling beyond H-4. The position of the remaining resonances was obtained from

143 inspection of a ^{13}C Dept spectrum and from the HQSC spectrum. On the ^{13}C Dept

144 spectrum, the C-6 resonances are located together at approximately 60 ppm. Once

145 all the cross-peaks on the HSQC spectrum for C1/H1 to C4/H4 and for C6/6 had

146 been assigned (Fig 2) the three remaining cross peaks must be those belonging to

147 C/H-5s. Finally, these were assigned to individual monomers by cross reference to a

148 HSQC-TOCSY spectrum. As the structure does not have any glycosidic links

149 involving the hydroxyls at C6, our failure to assign the individual H-6 residues to **D**, **E**

150 and **F** has no consequence for the characterisation of the EPS. For clarity, the

151 resonance position for the ^{1}H signals and ^{13}C signals (H2 to H5 and C2 to C5) are

152 indicated on the ^{1}H-^{13}C HSQC spectrum (Fig 2a, anomeric signals appear on the

153 inset Fig 2b) the anomeric proton resonances are also indicated on the ^{1}H-^{1}H

154 NOESY spectrum (Fig 3) and the combined chemical shift data for the complete

155 repeating unit is presented in Table 1.

156   The very low field chemical shift values for H-2, H-3 and the H-4 resonances

157 of residue **D,** combined with the loss of this residue under mildly acidic conditions

158 (see discussion below), is evidence for residue **D** being the terminal Gal*f*. This would

159 leave residues **E** and **F** as the 1→4 linked hexoses. One way of differentiating

160 between galacto- and glucopyranoses is from inspection of the chemical shifts for

161 their H-4 resonances: for a galactose H-4 is shifted substantially to lower field than

162 that of a glucose regardless of the anomeric configuration and linkage. Data

163 collected from assignments for lactic acid bacteria (LAB) EPS structures show that

164   the H-4 resonances for galactose lie in the range 4.30-3.85 δ whilst those for glucose

165   lie in the range 3.45 -3.75δ[15]. The chemical shift for **E** H-4 (3.61) and that for **F** H-4

166   (4.39) implies that **E** is the glucopyranose sugar and **F** is the galactopyranose.

167   Information regarding the sequence of the sugar residues in the repeating unit

168   was obtained from examination of the NOESY spectrum (Fig 3) and the anomeric

169   region of the HMBC spectrum (not shown). On the NOESY spectrum there are

170   strong inter residue NOEs between: **A** H-1 and **F** H-4, **B** H-1 and **C** H-2, **D** H-1 and **B**

171   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**,

172   **D**(1→2)**B, E**(1→3)**C** and **F**(1→4)**E** linkages. On the HMBC spectrum inter-residue

173   couplings are observed between: **A** H-1 and **F** H-4, **B** H-1 and **C** H-2, and **D** H-1 and

174   **B** H-2, matching the NOE signals and, additionally, a cross peak is observed

175   between **C** H-1 and **A** H-2.

176   Using a combination of the results for the linkage analysis and the NMR

177   identifies the structure for the repeating unit as:

```
178              D                B
179        α-D-Galf-(1→2)-α-L-Rhap
180                               1
181                               ↓
182                               2
183  →4)-β-D-Glcp-(1→3)-α-L-Rhap-(1→2)- α-L-Rhap-(1→4)-β-D-Galp-(1→
184          E              C                A              F
185
```

186   This is a novel structure and differs to those of the EPS structures that have been

187   reported for *B. bifidum* BIM B-465[12] and *B. longum* JBL05[13] and is also different to

188   the EPSs that have been isolated and characterised from LAB[15].

189

190   ***Mild acid catalysed hydrolysis of the HMW EPS***

191   It is well known that glycosidic links to Gal*f* residues are hydrolysed in acidic

192   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

217 (GTF), priming-GTF (p-GTF), genes involved in export of repeated unit, its

218 polymerization and chain length determination, as well as mobile elements

219 (transposase and insertion sequences (IS)). However, no gene regulators have been

220 found although this function could be played by proteins with unknown function

221 within the cluster or others outside it. It is also surprising that a number of membrane

222 proteins are present in the bifidobacterial *eps* cluster. Another remarkable feature in

223 the IPLA-R1 *eps* cluster is the presence of two p-GTF, as previously denoted for *B.*

224 *longum* subsp. *longum* NCC2705[19] and for the other *B. animalis* subsp. *lactis*

225 strains[14]. This enzyme catalyses the transfer of an activated sugar to the lipid carrier

226 C55, being the first step in the synthesis of the repeated unit that build the polymer.

227 In *B. animalis* subsp. *lactis* IPLA-R1, one of the p-GTF was located in the 5´ end

228 (annotated as "undecaprenyl-phosphate sugar phosphotraferase") and the second

229 one was located down-stream (annotated as galactosyl transferase CpsD). The

230 nucleotide sequences of the two p-GTFs of the strain IPLA-R1 were different to the

231 corresponding homologues of the type-strain DSM10140. In addition, in the strain

232 IPLA-R1 the change in the nucleotide sequence of *cpsD*, but no that of the

233 undecaprenyl-phosphate sugar phosphotraferase gene, leaded to a modification in

234 the translated amino acid. However, we do not know if this different amino acid

235 residue could modify the function of this p-GTF and thereby, influence the synthesis

236 of the EPS in IPLA-R1 strain. Variations in the nucleotide sequence were also

237 detected in the transposase IS204/IS1001/IS1096/IS1165 located at the 3´end,

238 probably due to gain or lose of nucleotides in this mobile element during each

239 transposition.

240   In relation to the structure of the HMW-EPS IPLA-R1 determined in this study,

241 it is worth mentioning the presence of genes coding for proteins involved in the

242 biosynthesis of rhamnose precursors in the putative *eps* gene cluster of this strain,

243 which could correlate with the high rhamnose content (50%) of its hexasaccharide

244 repeated unit. Normally, in LAB strains, the content of rhamnose does not exceed

245 that of the other two common EPS monosaccharides: glucose and galactose. An

246 exception is the strain *Lactobacillus rhamnosus* RW9595M whose glucose:

247 galactose: rhamnose ratio is 2:1:4[20] and it also presents in its *eps* cluster rhamnose-

248 precursor biosynthseis genes[21]. Additionally, in EPS isolated from intestinal strains

249 the content of rhamnose was higher (52%) than in those isolated from foods (28%)[7].

250 In this regard, the *in silico* comparative analysis of five bifidobacteria species shows

251 that rhamnose-precursor biosynthesis genes are present in three out of the five

252 species, showing a high degree of protein homology (data not shown). These facts

253 suggest that the high presence of rhamnose in intestinal EPS could play a role in this

254 ecological niche, which deserves future investigation.

255 **3. Experimental**

256 *3.1. Bacterial growth and purification* of the *HMW-EPS*

257 The growth conditions and purification of the crude-EPS from strain *B. animalis*

258 subsp. *lactis* IPLA-R1 (previously named A1dOxR) have previously been

259 described[11]. In short, bacterial biomass grown in an anaerobic chamber [MG500

260 (Down Whitley Scientific, West Yorkshire, UK): 80% (v/v) N_2, 10% CO_2, 10% H_2] at

261 37ºC for 5 days on the surface of agar-MRSC [MRS (Biokar Diagnostics, Beauvais,

262 France) + 0.25% L-cysteine (Sigma Chemical Co. St. Louis, MO, USA)] was

263 collected with water. The bacterial suspension was mixed with 1 volume of 2M

264 NaOH and the crude-EPS from the cell-free supernatant was precipitated with 2

265 volumes of chilled-absolute ethanol for 48 h at 4ºC. The precipitated fraction was

266 resuspended in ultra-pure water, dialysed in 12-14 kDa MWCO cellulose membranes

267 (Sigma) against water for 3 days at 4ºC and finally freeze-dried to obtain the EPS-

268 crude fraction.

269   A pure sample of the high molecular weight (HMW)-EPS from the crude-EPS

270 fraction isolated from *B. animals* subsp. *lactis* IPLA-R1 was obtained by dialysis as

271 follows. A crude-EPS sample (25 mg) was dissolved in distilled water (10 mL) with

272 gentle heating (less than 50°C) and the HMW material was isolated in the retentate

273 (about 10 mL), after dialysis (Spectra/Por Float-A-Lyser 100 KDa MWCO, Sigma) for

274 72 hours at 4°C against three changes of distilled water per day. The content of the

275 dialysis membrane were freeze dried to provide HMW-EPS.

276   The purity of the HMW-EPS was determined by SEC-MALLS and NMR analysis.

277 For SEC-MALLS, solutions of EPS in deionised water (1 mg mL^{-1}) were prepared

278 and left for 24 h to completely dissolve. Samples (100 µl) were injected onto an

279 analytical SEC system comprising three columns Aquagel-OH 40, 50 and 60 (15 µm

280 particle size, 25 cm x 4 mm, Varian, Oxford, UK) connected in series. The neutral

281 analytes were eluted with deionised water flowing at 1 mL min^{-1}. The concentration

282 of the EPS fractions eluting from the column were determined by a differential

283 refractometer (Optilab rEX, Wyatt technology, Santa Barbara, USA) and the weight

284 average molecular weight was measured using a Dawn-EOS MALLS operating with

285 a 690 nm laser (Wyatt technology, Santa Barbara).

286 ***3.2. Monomer composition and linkages of the HMW-EPS***

287   The monomer composition of the HMW-EPS was determined after acid

288 hydrolysis by HPAEC-PAD. The HMW-EPS (1 mg mL^{-1}) was treated with 2M TFA

289 (120°C for 2 h) and the identity of the released monomers was determined using

290   high performance anion exchange chromatography (HPAEC) on a Dionex BioLC

291   system (Sunnyvale CA, USA) equipped with a CarboPac PA20 column (150 mm x 3

292   mm). Monomers were eluted using a sodium 10 mM hydroxide mobile phase at a

293   flow rate of 0.5 mL min^{-1} and detected using a pulsed amperometric detector (PAD)

294   ED50 (Dionex) operating with a dual potential waveform. The ratio of monomers was

295   determined by comparison of the detector response to calibration standards of the

296   individual monomers (galactose, glucose and rhamnose, 5-100 ppm). The absolute

297   configurations of monosaccharides were determined by conversion to their butyl

298   glycosides using the procedure described by Gerwig et al[22].

299   For linkage analysis, the HMW-EPS was permethylated using the procedures

300   described by Stellner et al[23]. The methylated-polysaccharide was hydrolysed by

301   treatment with 2M TFA (120 °C for 2 h) and the methylated monosaccharides

302   converted to their corresponding methylated alditol acetates. The identity of the

303   variously methylated alditol acetates was determined by GLC-MS and by analysis of

304   the individual fragmentation patterns observed in the MS. GLC-MS analyses were

305   performed on an Agilent 7890A GC system (Santa Clara, CA, USA) coupled to an

306   Agilent 5675c quadrupole MS. The samples were eluted from a HP-5 column (30 m

307   x 0.25 mm-id, 0.25 μm film) using helium as carrier (9 psi, flow rate 1 mL min^{-1}) and

308   using the following temperature programme: start temperature 155°C, hold time 1

309   min, and a final column temperature of 195°C reached via a rising gradient of 0.75°C

310   min^{-1}.

311   ***3.3. NMR structure of the HMW-EPS***

312   NMR spectra were recorded for HWM-EPS samples that were dissolved (10

313   mg mL^{-1}) directly in D_2O (Goss Scientific Instruments Ltd., Essex, UK). NMR spectra

314   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 70°C 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

339 pair of PCR primers were designed taking into account the sequence of the type

340 strain *B. animalis* subsp. *lactis* DSM10140, whose genome is publicly available

341 (GenBank accession number CP001606)[24]. Primers, synthesised by Thermo-Fisher

342 Scientific GmbH (Ulm, Germany), amplified regions of about 1,000 bp. The PCR

343 reaction mixture in a final volume of 50 µl was: 1 µl chromosomal DNA, 0.20 µM of

344 each primer, 200 µM dNTPs (Amersham Bioscience, Upsala, Sweden) and 2.5 U

345 Taq DNA-polymerase (Eppendorf, Hamburg, Germany). The PCR thermal conditions

346 were an initial denaturalisation cycle 95ºC for 5 min, 30 amplification cycles of: 95ºC

347 for 1 min, 52 or 56ºC (variable according to the pair of primers) for 50 s and 68ºC for

348 2 min, and a final extension step of 68ºC for 10 min. Amplification was done in a

349 UnoCycler thermal cycler (VWR International Eurolab S.L., Barcelona, Spain). PCR

350 amplified products were visualised under UV in 1% agarose gels, after staining with

351 ethidium bromide. Purification and sequencing of each amplicons (both strands) was

352 performed by Macrogen (Seoul, Korea). Sequences obtained were processed with

353 the free Chromas 1.45 software (Technelysium Pty Ltd., Australia) and used for

354 comparison with those held in the GenBank database

355 (http://www.ncbi.nlm.nih.gov/genbank) using the BLASTn and BLASTp tools.

356

362

363 **References**

364 (1) Margolles, A.; Mayo, B.; Ruas-Madiedo, P. Screening, identifiation and

365   characterization of *Lactobacillus* and *Bifidobacterium* strains. In *Handbook of*

366   *probiotics and prebiotics,* 2nd ed; Nomoto, K., Salmine, S., Lee, Y.K., Eds;

367   John Willey & Sons Inc.: New Jersey, **2009**; pp 4-43.

368 (2) FAO/WHO. Probiotics in foods. Health and nutritional properties and guidelines

369   for evaluation. In *FAO food and nutrition paper* No. 85 (ISBN 92-5-105513-0).

370   **2006**; pp 1-50.

371 (3) Mayo, B.; Delgado, S., Rodríguez, J.M.; Gueimonde, M.; *CAB Rev.*, **2008**, No.

372   055 (Online ISSN 1749-8848); pp 1-17.

373 (4) WGO. In *World Gastroenterology Organisation Practice Guidelines: Probiotics*

374   *and Prebiotics*

375   (http://www.worldgastroenterology.org/assets/downloads/en/pdf/guidelines/19_probio

376   tics_prebiotics.pdf) **2008**; pp 1-22.

377 (5) Ruas-Madiedo, P.; Abraham, A.; Mozzi, F.; de los Reyes-Gavilán, C.G.

378   Functionality of exopolysaccharides produced by lactic acid bacteria. In

379   *Molecular aspects of lactic acid bacteria for traditional and new applications*,

380   Mayo, B., López, P., Pérez-Martínez, G., Eds.; Research Signpost

381   Publishers: Kerala, India, **2008**; pp 137-166.

382 (6) Ruas-Madiedo, P.; Moreno, J.A.; Salazar, N.; Delgado, S.; Mayo, B.; Margolles,

383   A.; de los Reyes-Gavilán, C.G. *Appl. Environ. Microbiol.*, **2007**, *73*, 4385-

384   4388.

385   (7) Salazar, N.; Prieto, A.; Leal, J.A.; Mayo, B.; Bada-Gancedo, J.C.; de los Reyes-

386   Gavilán, C.G.; Ruas-Madiedo, P., *J. Dairy Sci.*, **2009**, *92*, 4158-4168.

387   (8) Ruas-Madiedo, P.; Gueimonde, M.; Arigoni, F.; de los Reyes-Gavilán, C.G.;

388   Margolles, A. *Appl. Environ. Microbiol.*, **2009**, *75*, 1204-1207.

389   (9) Salazar, N.; Gueimonde, M.; Hernández-Barranco, A.M.; Ruas-Madiedo, P.; de

390   los Reyes-Gavilán, C.G. *Appl. Environ. Microbiol.*, **2008**, *74*, 4737-4745.

391   (10) Salazar, N.; Binetti, A.; Gueimonde, M.; Alonso, A.; Garrido, P.; González del

392   Rey, C.; González, C.; Ruas-Madiedo, P.; de los Reyes-Gavilán, C.G., *Int. J.*

393   *Food Microbiol.*, **2011**, *144*, 342-351.

394   (11) Ruas-Madiedo, P.; Medrano, M., Salazar, N.; de los Reyes-Gavilán, C.G.;

395   Pérez, P.F.; Abraham, A. *J. Appl. Microbiol.*, **2010**, *109*, 2079-2086.

396   (12) Zdorovenko, E.L.; Kachala, V.V.; Sidarenka, A.V.; Izhik, A.V.; Kisileva E.P.;

397   Shashkov, A.S.; Novik, G.I.; Knirel, Y.A. *Carbohyd. Res.*, **2009**, *344*, 2417-

398   2420.

399   (13) Kohno, M.; Suzuki, S.; Kanaya, T.; Yoshino, T.; Matsuura, Y.; Asada, M.;

400   Kitamura, S., *Carbohyd. Polym.*, **2009**, *77*, 351-357.

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

402   (15) Harding, L. P.; Marshall, V. M.; Hernandez, Y.; Gu, Y. C.; Maqsood, M.; McLay,

403   N.; Laws, A. P. *Carbohydr. Res.,* **2005**, *340*, 1107-1111.

404   (16) Carpita, N.C.; Shea, E.M. Linkage Structure Of Carbohydrates By Gas-

405   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.;

427   Broadbent, J.R.; Oberg, T.; Dudley, E.G.; Schuster, S.; Romero, D.A.;

428   Roberts, R.F., *J. Bacteriol.*, **2009**, *191*, 4144-4151.

429

430

431   **Figure Legends**

432

433   **Figure 1** 500 MHz ^{1}H NMR spectrum of the HMW-EPS obtained after separation of

434   crude-EPS by dialysis in a 100 kDa cellulose acetate membrane; spectra recorded in

435   D_2O at 70°C. Inset shows an expanded plot of the anomeric region.

436

437   **Figure 2** (a): 500-MHz ^{1}H-^{13}C HSQC spectrum of a selected region of the HMW-

438   EPS from *B. animalis* subsp. *lactis* IPLA-R1 recorded in D_2O at 70°C. The identity of

439   the cross peaks is noted by the sugar residue (**A-F**) and by identifying the location of

440   hydrogens/carbons within the ring as **2-5**. (b): anomeric region of the 500-MHz [1]H-

441   ^{13}C HSQC spectrum.

442

443   **Figure 3**: 500-MHz [1]H-1H NOESY spectrum of a selected region of the HMW-EPS

444   from *B. animalis* subsp. *lactis* IPLA-R1 recorded in D_2O at 70°C. The identity of the

445   cross peaks is noted by the sugar residue (**A-F**) and by identifying the location of

446   hydrogens within the ring as **1-5**. Intra-residue couplings are highlighted in red and

447   inter-residue couplings are highlighted in yellow. For interpretation of the references

448   to colour in this figure legend, the reader is referred to the web version of this article.

449

450   **Figure 4**: 500-MHz 1H–13C HMBC spectrum of a selected region of the HMW-EPS

451   from *B. animalis* subsp. *lactis* IPLA-R1 recorded in D2O at 70°C. The identity of the

452 cross peaks is noted by the sugar residue (A–D) and by identifying the location of

453 hydrogens within the ring as 1–5. Intra-residue couplings are highlighted in red and

454 inter-residue couplings are highlighted in yellow. (For interpretation of the references

455 to colour in this figure legend, the reader is referred to the web version of this article).

456

457 **Figure 5**: 500 MHz ^{1}H NMR spectra of the anomeric region following the acid

458 catalysed hydrolysis of the HMW-EPS from *B. animalis* subsp. *lactis* IPLA-R1 as a

459 function of time; spectra were recorded in D_2O at 70°C and the sample was

460 maintained at 70°C for the full reaction period. Sugar residues in the native HMW-

461 EPS are identified at the bottom (**A to F**), whereas those of the hydrolysed sample

462 are identified at the top (**A' to F'**).

463

464 **Figure 6**: Physical map of the putative *eps* cluster (54,259 bp) of *B. animalis* subsp.

465 *lactis* IPLA-R1. Predicted protein functions are categorised as follows: **GTF**,

466 glycosyltransferase; **p-GTF**, priming-GTF/ undecaprenyl-phosphate sugar

467 phosphotransferase / galactosyltransferase (CpsD); **Rh-B**, rhamnose biosynthesis

468 precursors; **P-ChL**, polymerization (polymerase) – chain length determination; **P-E**,

469 polymerization – export (Wzx and Wzz flippases); **PB**, polysaccharide biosynthesis;

470 **AcS**, acyl-synthetase; **Ph**, phosphorilase; **MP**, membrane protein; **T-IS**, transposase

471 – IS mobile elements; White arrows without label indicate hypothetical proteins.
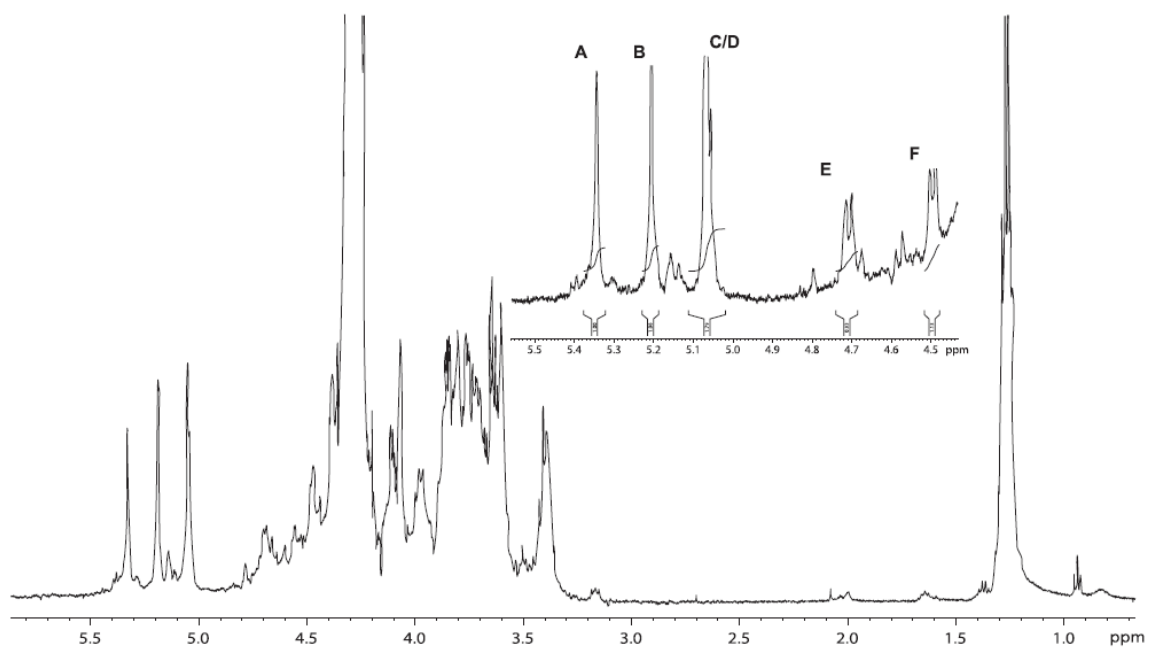
472

473 **Table 1**: ^{1}H and ^{13}C NMR chemical shifts of the HMW-EPS from *B. animalis* subsp.

474 *lactis* IPLA-R1 recorded in D_2O at 70°C
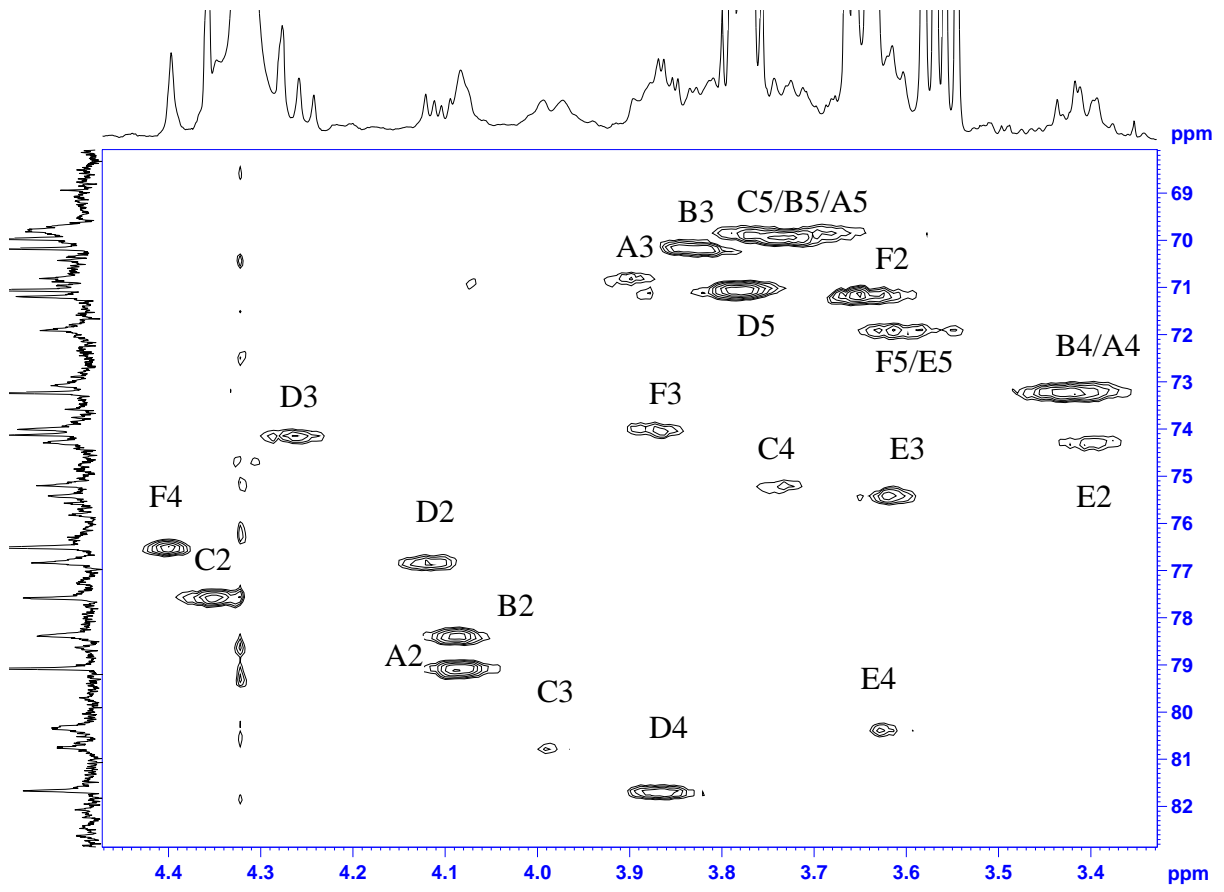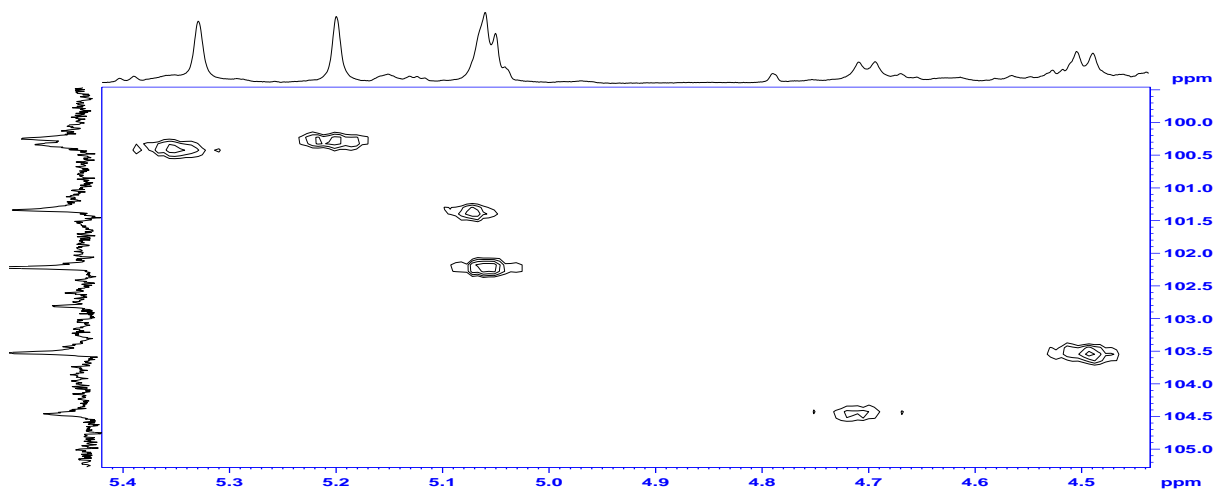
475

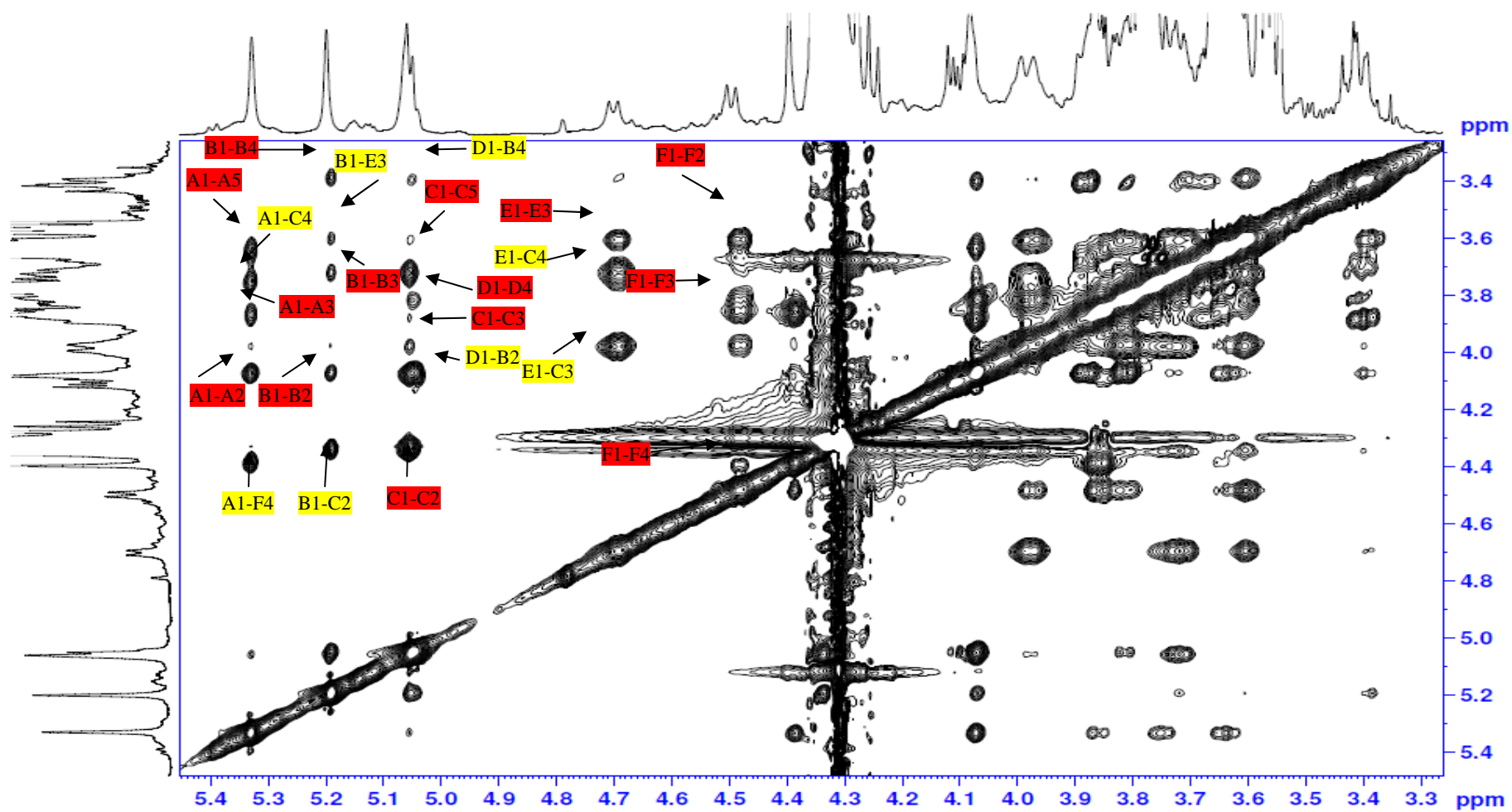476   **Figure 1**

477

478

479

480

Figure 2

a

b

487   **Figure 3**

488

489   ..
490

Figure 4
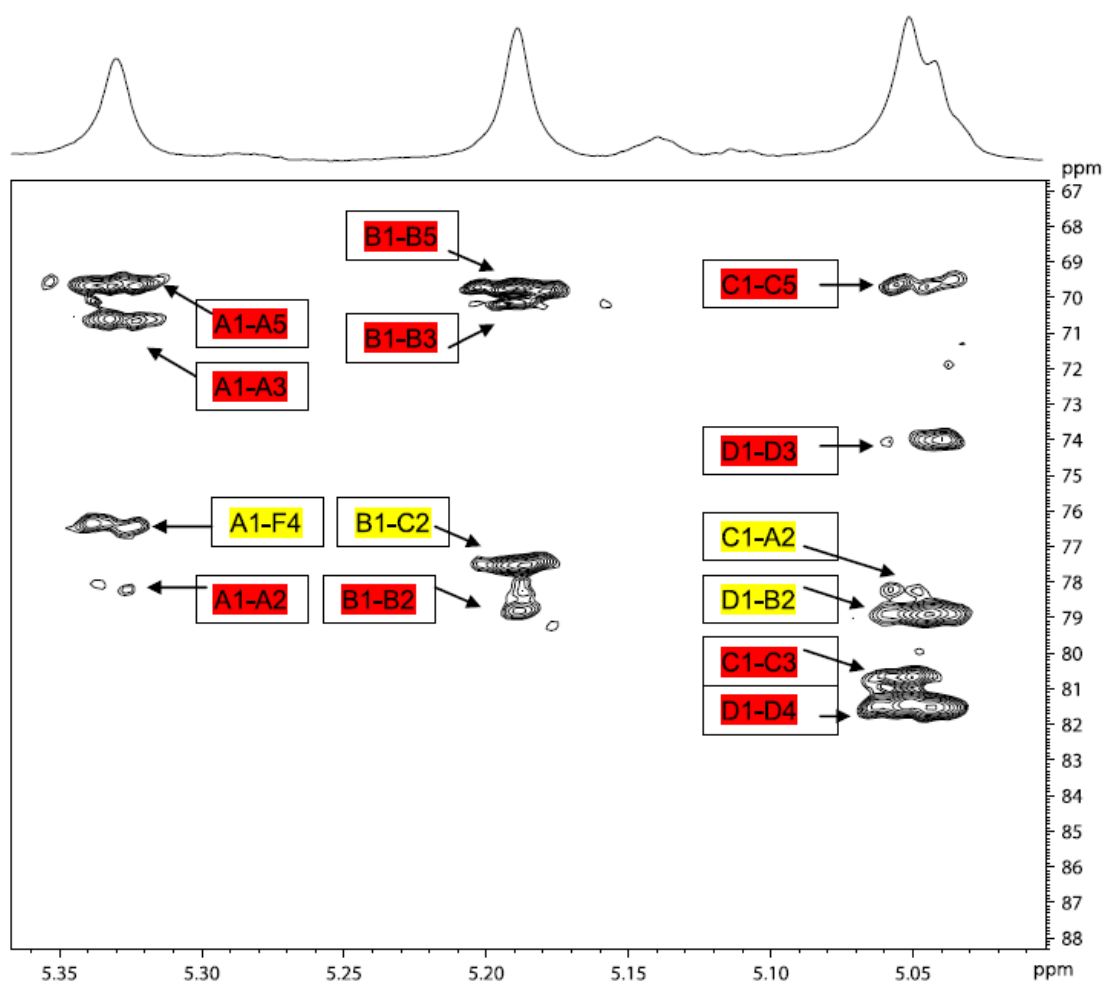
Figure 5

Figure 4

Figure 6

Top row (left to right): p-GTF, MP, GTF, Rh-B, GTF, P-E, GTF, P-ChL, GTF, PB, Ph, P-E, GTF, p-GTF, wzx, wzx, cpsD

Bottom row (left to right): MP, MP, PB, GTF, P-ChL, MP, AcS, P-E, wzz, T-IS

0   5,000 bp
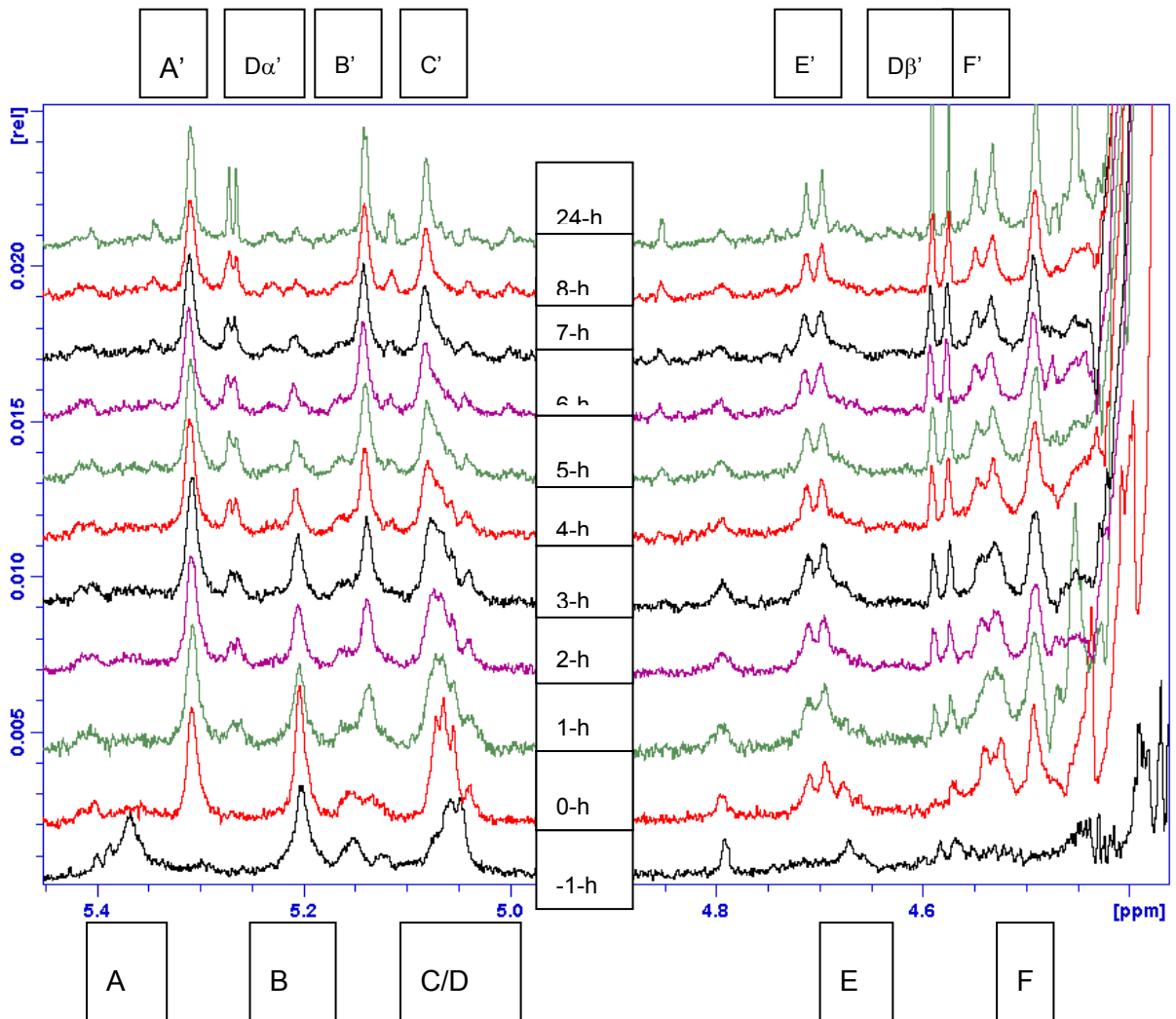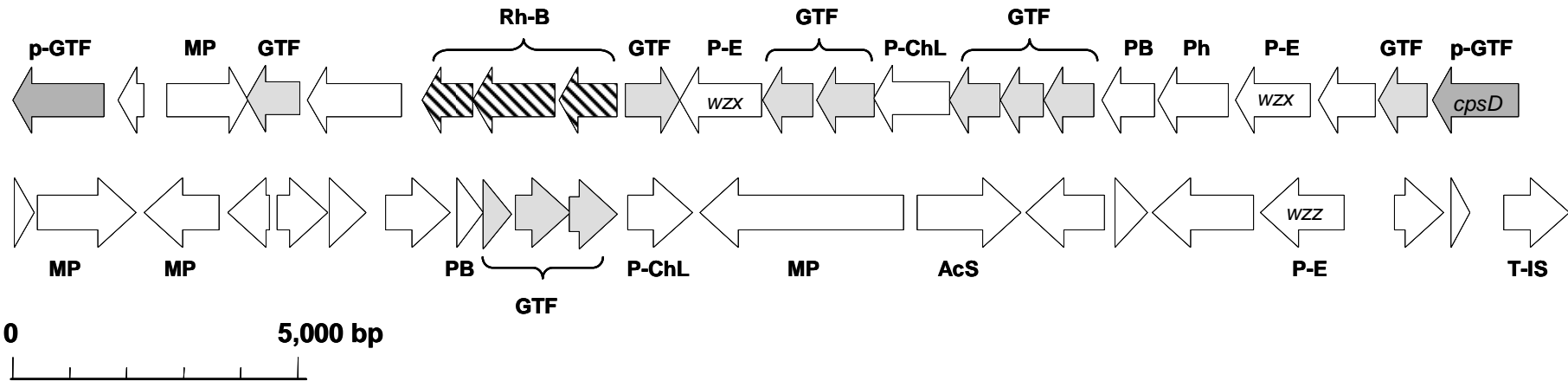
1 **Table 1**

2

Residue	¹H Chemical Shift (ppm)							¹³C 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

3 *Represents partially overlapping H6 and H6' resonances

4