**Supplementary Data: Monosaccharide Composition and Linkage Analysis of RPS**

Methods:

Glycosyl composition analysis was done by gas chromatography-mass spectrometry (GC-MS) of the monosaccharide alditol acetates at the Complex Carbohydrate Research Center at the Univ. of Georgia. After hydrolysis of the RPS in 4N TFA at 100$^{o}$C (4 hr), resulting monosaccharides were reduced with $NaBH_4$, acetylated in pyridine/acetic anhydride and analyzed on a DB-1 capillary column. Linkage analysis was done by GC-MS of the partially methylated alditol acetates. RPS samples were permethylated by the method of Ciucanu and Kerek (1984). The partially methylated RPS was hydrolyzed, reduced and acetylated as described above and analyzed by GC-MS on a DB-1 column. Aqueous HF was used to cleave the phosphate from selected samples prior to the analyses. RPS was treated in 48% aqueous HF at 4$^{o}$C for two days, then lyophilized and used for composition and linkage analysis.

Results:

Composition analyses are given in Table S1. The RPS of *S. gordonii* 38, whose structure has been previously determined (Reddy *et al.*, 1994), was used as a control. Comparison of the composition with the reported structure shows the GalNAc measured composition to be low by one residue since the 6-phosphate linkage is not cleaved under the hydrolysis conditions. Rhamnose composition is low for unknown reasons. In this sample as well as most of the other samples, GlcNAc is found in variable amounts apparently from contaminating polysaccharides that co-purify with the RPS. GC-MS in the linkage analysis shows terminal rhamnose and 2,3-linked rhamnose, 4-linked glucose, 3-linked galactopyranoside, 6-linked galactofuranoside and 6-Linked GalNAc in agreement with the known structure. Additional minor peaks are seen for 2-linked rhamnose, 3-linked rhamnose, 3-linked galactofuranoside, 3-linked GlcNAc, terminal glucose and terminal GalNAc. These levels of impurities are typical of the other samples and are indicative of the reliability of the analysis.

For strain GC27 RPS, the gene structure and the NMR spectra imply a structure identical to that of previously reported strain MC2 RPS (Yoshida *et al.*, 2006). The composition of Table S1 agrees with the proposed structure considering that one residue of GalNAc is not detected due to the 6-linked phosphate. GC-MS of the PMAA shows 3-linked rhamnose, 4-linked glucose, 6-linked galactofuranoside, 6-linked galactopyranoside and 3-linked GalNAc in agreement with the proposed structure. Minor peaks for terminal rhamnose, 2-linked rhamnose and 2,3-linked rhamnose were detected in addition to terminal GlcNAc and 3-linked GlcNAc.

For strain GC37 RPS, the composition in Table S1 shows galactose and GalNAc to be low apparently due to the 6-phosphate on galactose and to the resistance of galactosamine to acid hydrolysis. GC-MS of the PMAA shows 3-linked rhamnose, 4-linked glucose, 6-linked galactopyranoside, 3-linked GalNAc and 3-linked galactofuranoside rather than the 6-linked galactofuranoside found in the RPS for strain GC-27. Minor peaks, similar to those seen for GC27 RPS, were also detected from strain GC37 RPS indicating similar contaminants.

The composition of strain GC32 RPS in Table S1 agrees with the proposed structure. The reduction of galactose by one residue due to the 6-phosphate is partially

corrected by pretreatment of the sample with HF. In the GC-MS of the PMAA, 3-linked rhamnose, 4-linked glucose, 6-linked galactopyranoside, 6-linked galactofuranoside and 3-linked GalNAc are detected along with peaks for terminal rhamnose, 2-linked rhamnose and 2,3-linked rhamnose presumably from contaminating polysaccharides. In the analysis of PMAA from HF-treated strain GC32 RPS, the same peaks are detected plus terminal galactose from the non-reducing terminal of the hexasaccharide and 3-linked GalNAc furanoside from anomerization at the reducing terminus of the hexasaccharide.

The composition analysis of strain GC35 RPS agrees with the proposed structure considering that the yield of galactose is reduced by one residue due to the 6-phosphate. Treatment with HF removes the phosphate leading to better agreement. The GC-MS of the PMAA shows 3-linked rhamnose, 4-linked glucose, 3-linked galactofuranoside, 6-linked galactopyranoside and 3-linked GalNAc in agreement with the proposed structure along with the same contaminating peaks seen in the analysis of strain GC32 RPS. After HF-treatment, terminal galactose from the non-reducing end of the hexasaccharide is detected along with 3-linked GalNAc-furanoside from the reducing end.

The composition analysis of strain GC39 RPS in Table S1 is in agreement with the proposed structure; HF-treatment of this polysaccharide raised the yield of galactosamine to match the expected composition. GC-MS of the PMAA shows peaks agreeing with the proposed structure along with the same contaminating peaks observed from strain GC32 and GC35 RPS. After HF-pretreatment, PMAA are also detected for terminal GalNAc from the non-reducing end of the hexasaccharide and 3-linked GalNAc-furanoside from the reducing end. We previously observed (McIntire *et al.*,1988, Cisar *et al.*, 1995) that HF-treatment of similar polysaccharides cleaved not only the phosphate linkage, but also in some circumstances, partially cleaved the galactofuranoside linkage in the oligosaccharide. For the case of strain GC39 RPS, the GC-MS of the PMAA for the HF-treated sample shows terminal galactose and 3-linked galactofuranoside because of this effect.

For strain GC51 RPS, composition analysis in Table S1 shows the rhamnose yield to be low, similar to that found for wild type strain 38 RPS. The structures of these polysaccharides differ only at residue A, which is GalNAc in strain 38 and Gal in strain GC51. GC-MS of the PMAA shows terminal rhamnose and 2,3-linked rhamnose, 4-linked glucose, 3-linked galactose, 6-linked galactofuranose and 6-linked GalNAc. HF-pretreatment produces terminal galactose and 3-linked galactofuranoside from the heptasaccharide.

References:

Cisar, J.O., Sandberg, A.L., Abeygunawardana, C., Reddy, G.P., and Bush, C.A. (1995) *Glycobiology* **5,** 655-662

Ciucanu I. and Kerek F. (1984) *Carbohydr. Res.* **131**:209-217

McIntire, F.C., Crosby, L.K., Vatter, A.E., Cisar, J.O., McNeil, M.R., Bush, C.A., Tjoa, S.S., and Fennessey, P.V. 1988. *J. Bacteriol.* **170**: 2229-2235

Reddy, G.P., Abeygunawardana, C., Bush, C.A., and Cisar, J.O. 1994. *Glycobiology* **4**: 183-192

Yoshida, Y., Ganguly, S., Bush, C. A., and Cisar, J. O. 2006. *J. Bacteriol.* **188**, 4125-4130

TABLE S1. Monosaccharide composition (%) of RPS from different *S. gordonii* strains as determined by GC-MS of the alditol acetates.

| Strain | Rha | Glc | Gal | GlcNAc | GalNAc |
|---|---|---|---|---|---|
| 38 (wild type) | 15.9 | 20.5 | 42.0 | 1.1 | 20.5 |
| GC27 | 16.8 | 18.0 | 41.3 | 9.0 | 14.9 |
| GC37 | 19.2 | 16.7 | 55.1 | 1.3 | 7.7 |
| GC32 | 19.0 | 16.5 | 35.5 | 7.7 | 21.3 |
| GC32 (HF)[a] | 19.7 | 15.3 | 39.5 | 8.0 | 17.5 |
| GC35 | 14.3 | 19.7 | 39.7 | 6.4 | 19.9 |
| GC35 (HF) | 16.2 | 16.8 | 48.3 | 4.2 | 14.5 |
| GC39 | 12.6 | 19.2 | 38.7 | 3.6 | 25.9 |
| GC39 (HF) | 18.4 | 14.6 | 34.9 | 1.7 | 30.4 |
| GC51 | 16.7 | 19.7 | 40.1 | 1.1 | 22.4 |
| GC51 (HF) | 17.4 | 15.2 | 32.3 | 0.9 | 34.2 |

[a](HF) indicates that the RPS was treated aqueous HF prior to analysis.

**Supplementary Data: NMR Spectra**

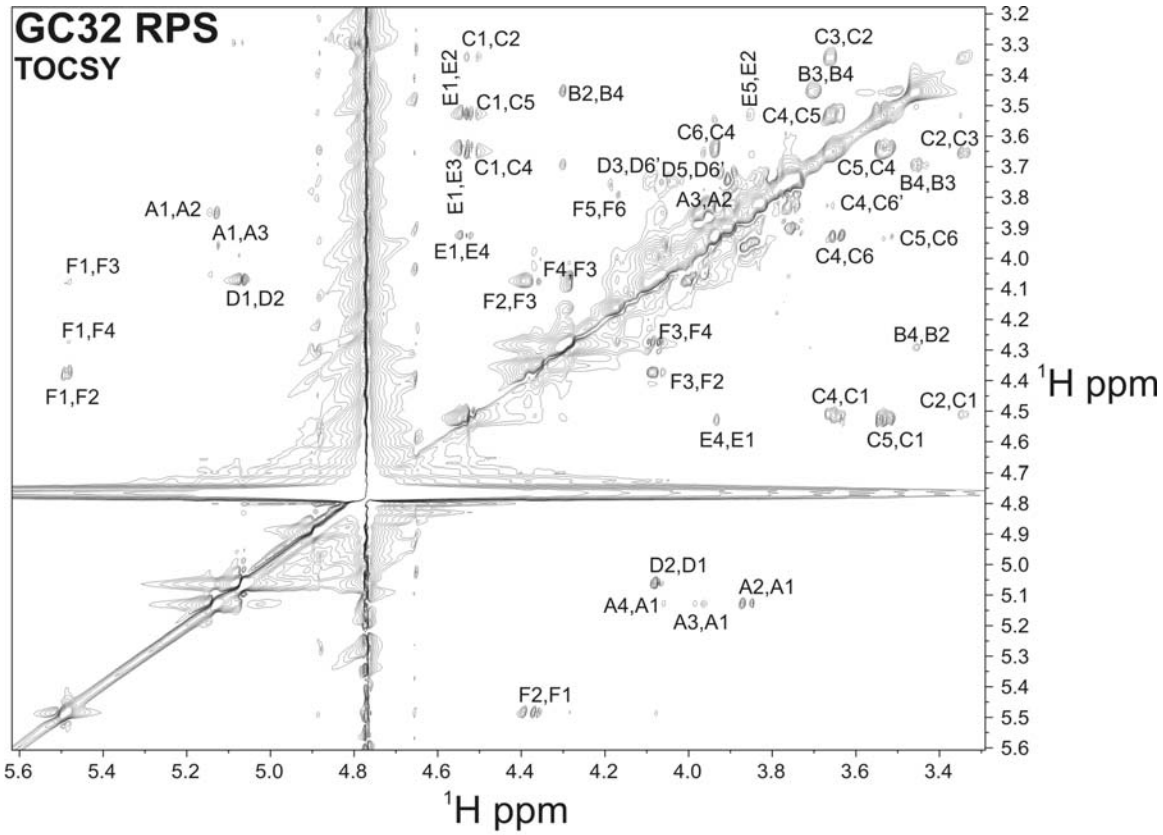


FIGURE S1. TOCSY spectrum (tm= 90 msec) of *S. gordonii* GC32 RPS

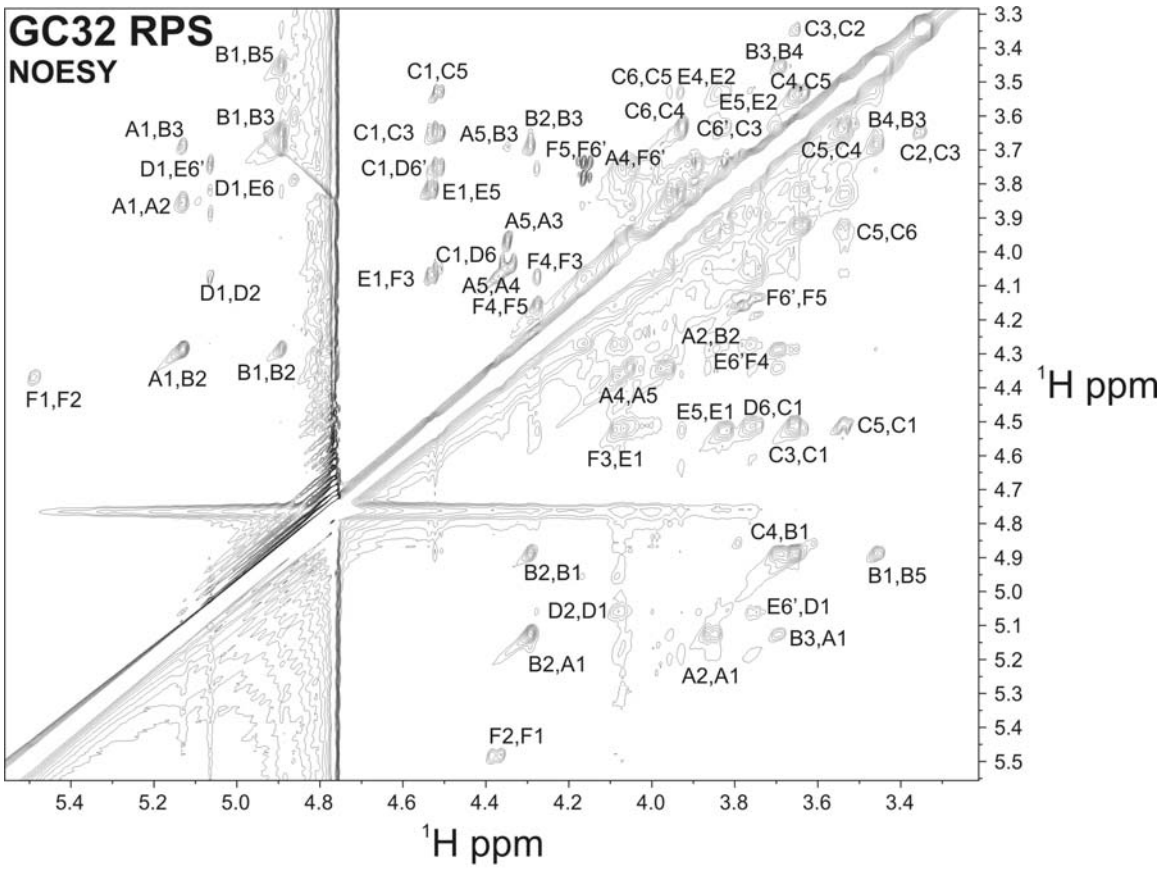


FIGURE S2 NOESY spectrum (tm=300 msec) of *S. gordonii* GC32 RPS

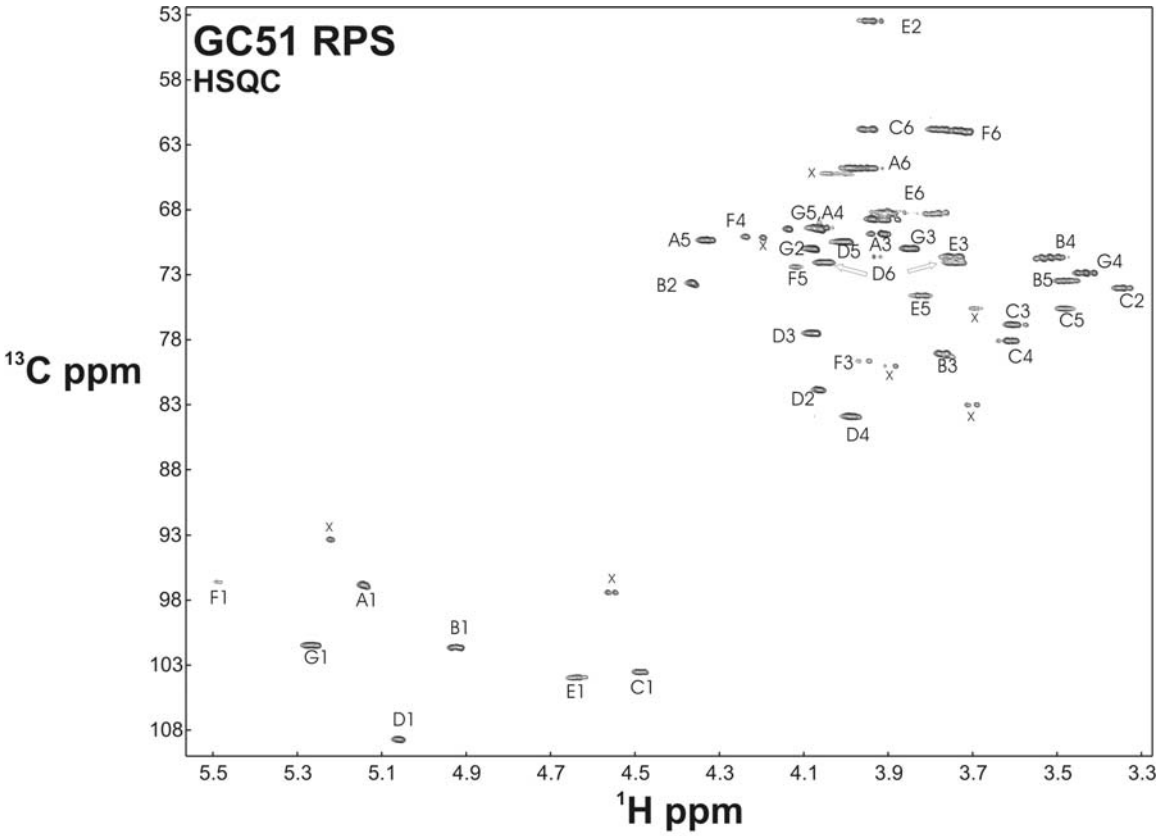


FIGURE S3. HSQC spectrum of S. gordonii GC51 RPS. Signals marked with X result from unassigned impurities. The intensity of the signal from F1 appears diminished due to heteronuclear coupling from $^{31}$P.

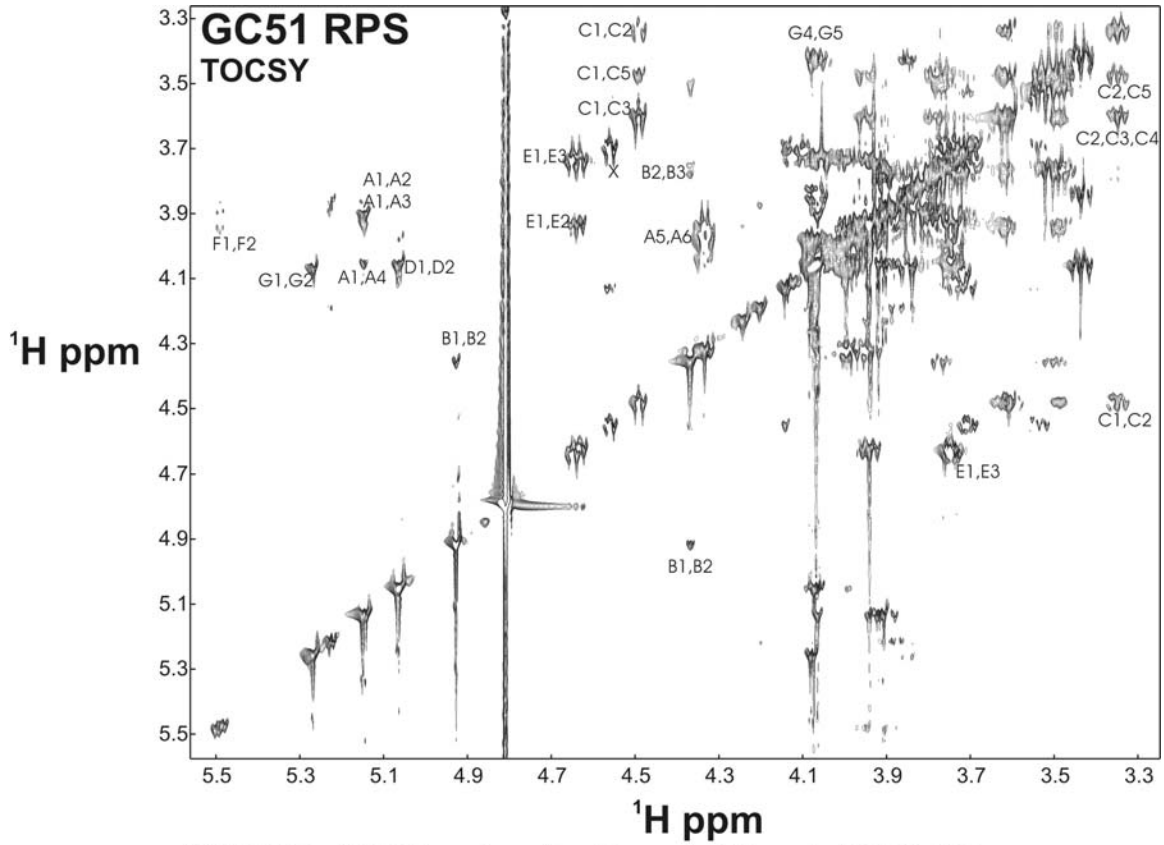


FIGURE S4. TOCSY spectrum (tm= 70 msec) of *S. gordonii* GC51 RPS

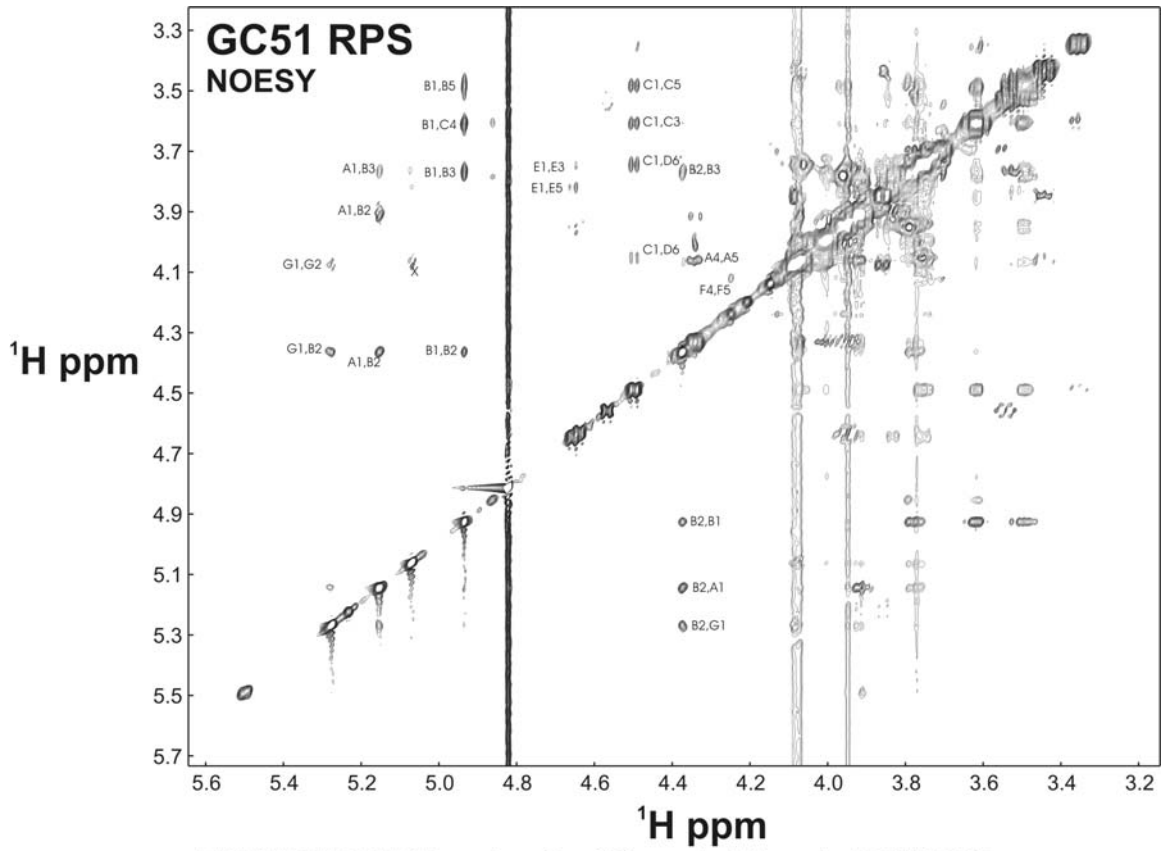


FIGURE S5. NOESY spectrum (tm=300 msec) of *S. gordonii* GC51 RPS

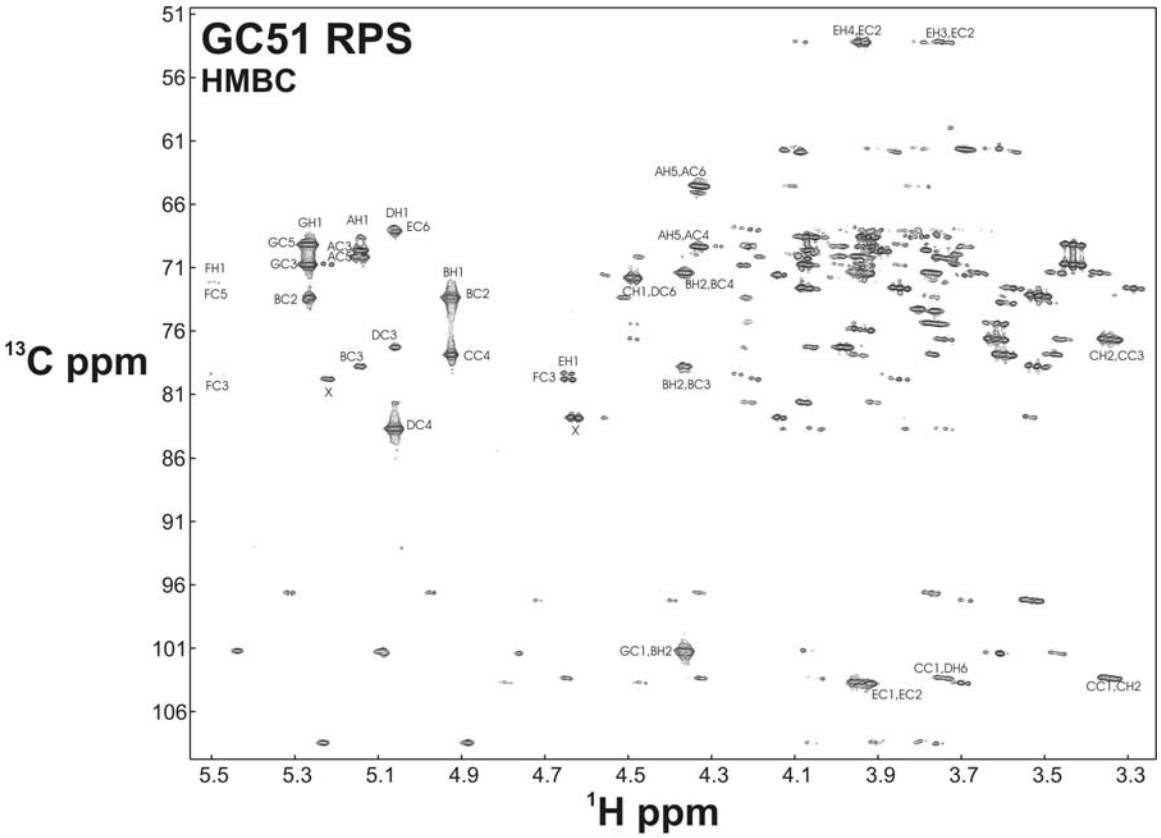


FIGURE S6. HMBC spectrum (evolution delay= 60 msec) of *S. gordonii* GC51 RPS

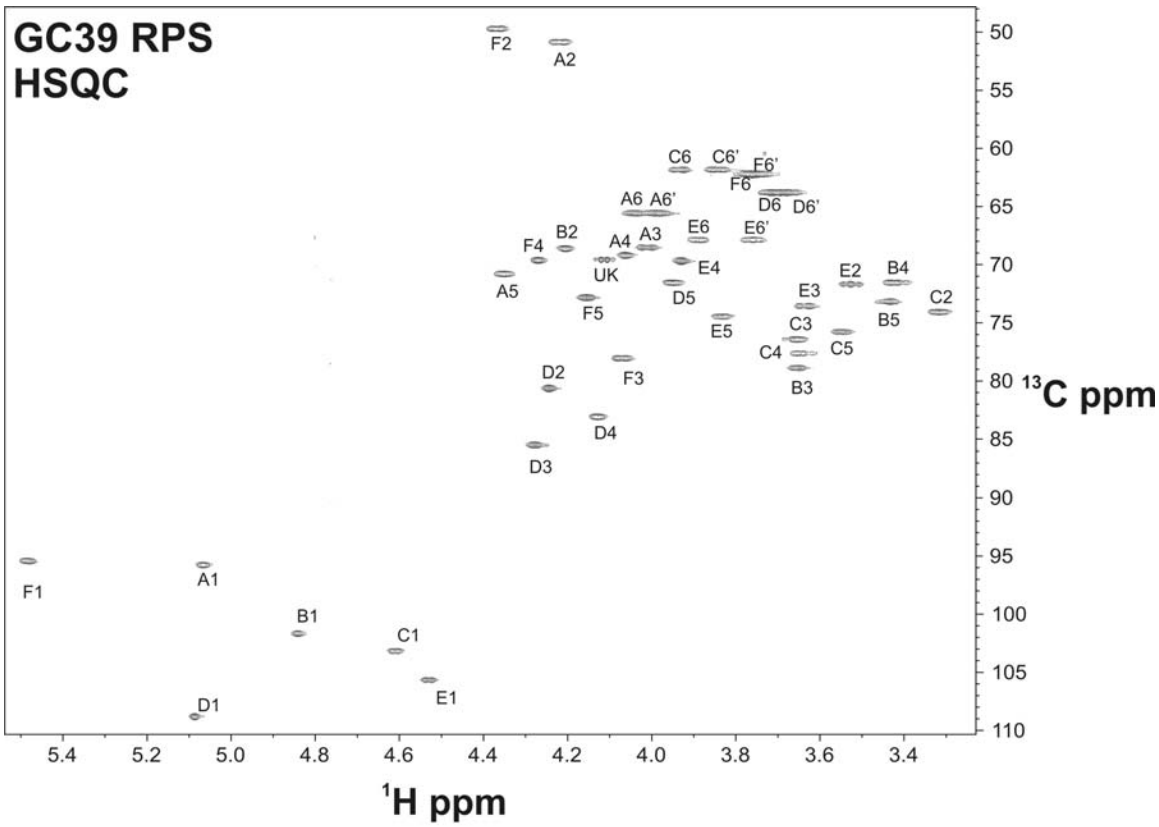


FIGURE S7. HSQC spectrum of *S. gordonii* GC39 RPS

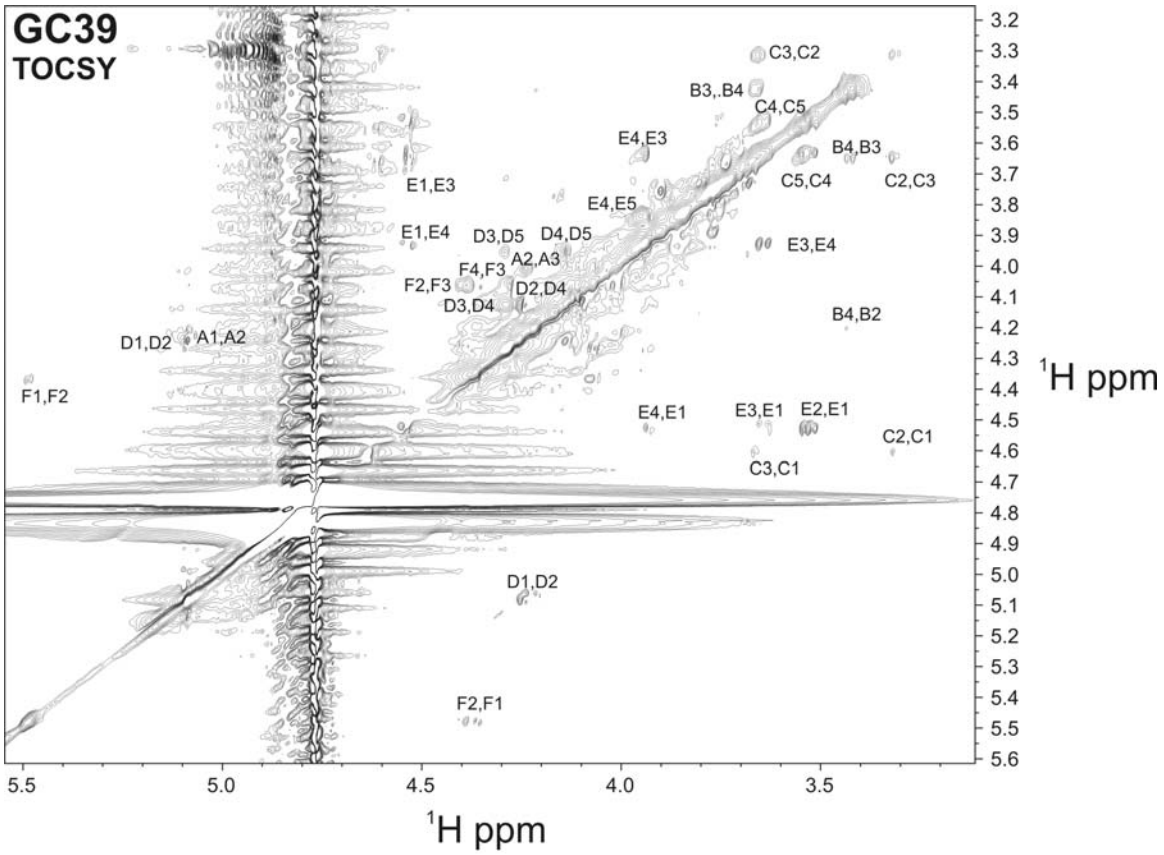


FIGURE S8. TOCSY spectrum (tm= 90 msec) of *S. gordonii* GC39 RPS

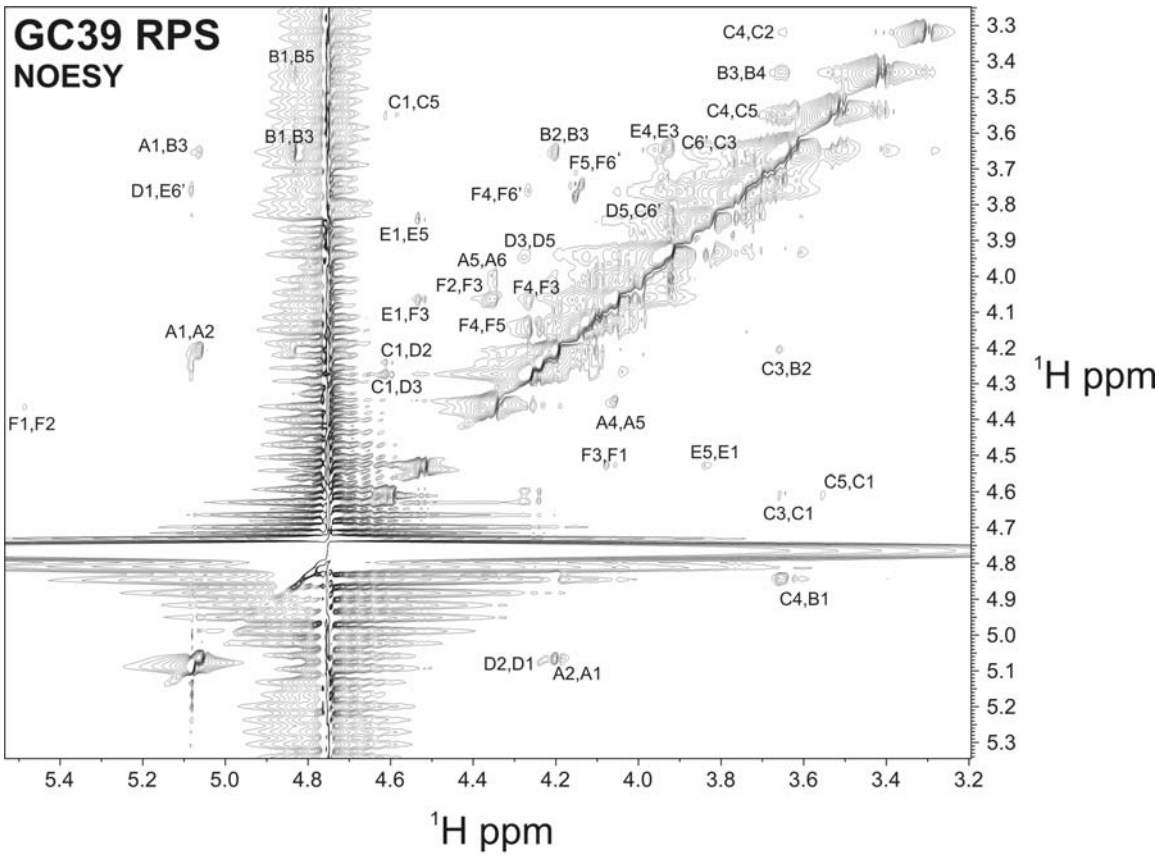


FIGURE S9. NOESY spectrum (tm= 300 msec) of *S. gordonii* GC39 RPS