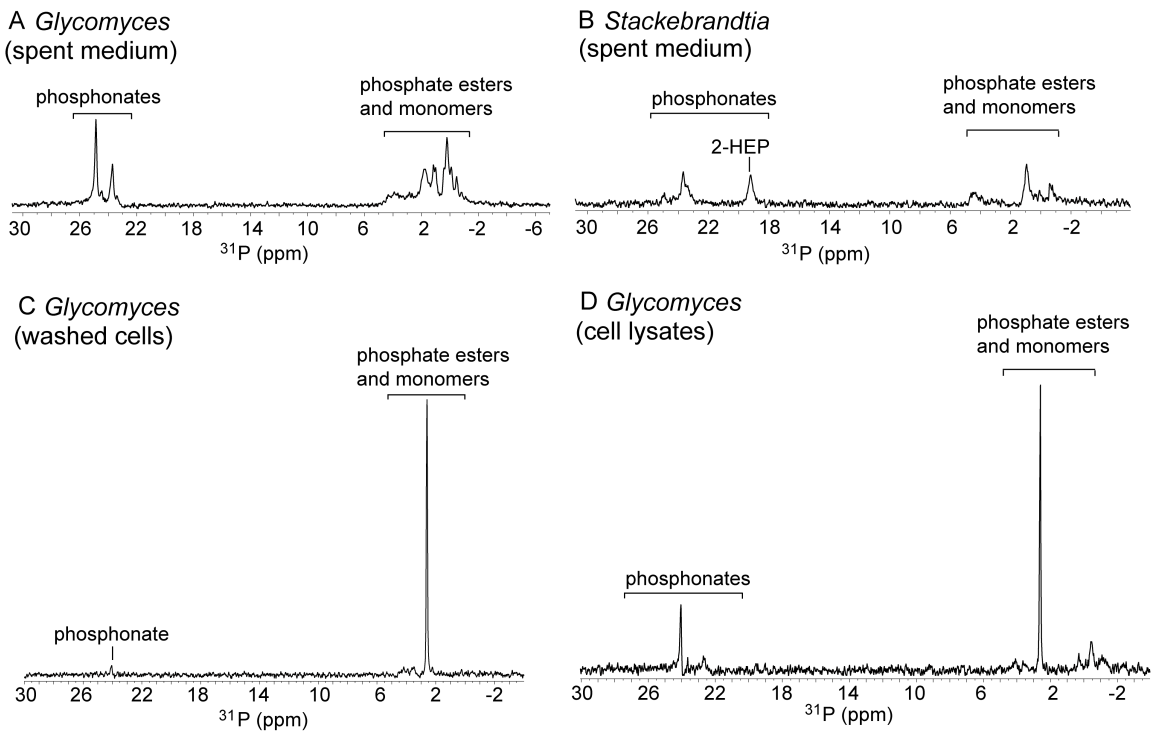

Fig. S1. ^{31}P NMR spectra of phosphonates from *Glycomyces* and *Stackebrandtia*. (A) Concentrated culture extract of *Glycomyces* shows the presence of phosphonates. (B) Concentrated culture extract of *Stackebrandtia* shows the presence of phosphonates. (C) Washed cells of *Glycomyces* contain a small amount of phosphonates. (D) Lysozyme treatment of the sample in (C) releases cell-bound phosphonates. Concentrated culture extracts were supplemented with 20% D_2O for ^{31}P NMR analyses, as shown in (A) and (B). *Glycomyces* cells (750 mg wet weight) were harvested and washed three times with 10 mM Tris buffer (pH=7.0). Washed buffers were combined, dried and redissolved in D_2O for ^{31}P NMR analysis, as shown in (C). Washed cells were resuspended in 10 mM Tris buffer (pH=7.0) and treated with DNase, RNase and proteinase K. Enzymes were inactivated by successive extractions with phenol:chloroform:isoamyl (25:24:1) and chloroform:isoamyl (24:1) and centrifugations. The aqueous phase collected was filtered using a 10 kDa Amicon filter. The retentate on the filter was washed 10 times with water, lyophilized, resuspended in 20% D_2O for ^{31}P NMR analysis, as shown in (D). P peaks with chemical shifts between 5 and 40 ppm are usually indicative of phosphonates and phosphinates whereas P peaks between +5 to -20 ppm range are usually phosphate monomers, esters and pyrophosphates.

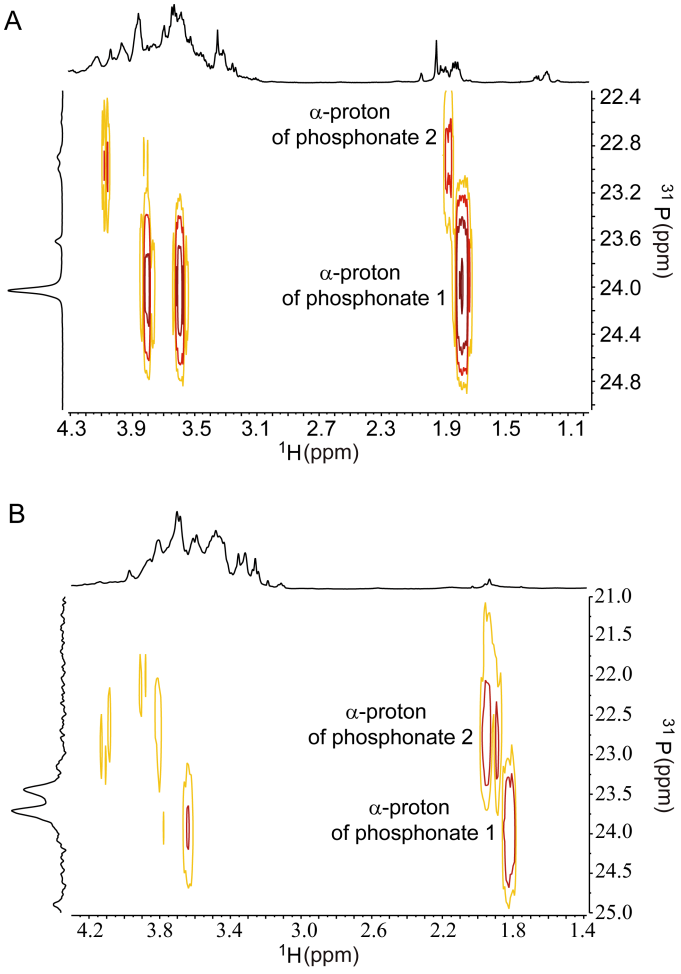

Fig. S2. ^{1}H-^{31}P HMBC spectra of purified phosphonoglycans from *Glycomyces* (A) and *Stackebrandtia* (B). In the ^{1}H-^{31}P HMBC spectrum of purified phosphonoglycans from *Glycomyces* (A), the major ^{31}P signal at 24.03 ppm had cross-peaks with three proton signals at 1.78, 3.59 and 3.80 ppm. The proton signal at 1.78 and 3.59 ppm was assigned to CH_2P and CH_2OH of a phosphonate head group, respectively, by ^{1}H-^{1}H COSY correlations of H-1 with H-2. The proton signal at 3.80 ppm may originate from a phosphonate ester. There were two minor ^{31}P signals at 22.89 and 22.99 ppm. One of them (or both) had cross-peaks with three proton signals at 1.86 (assigned to CH_2P of a phosphonate), 3.81 and 4.07 ppm. ^{31}P signal overlapping in this region prevented assignments of other proton signals. In the ^{1}H-^{31}P HMBC spectrum of purified phosphonoglycans from *Stackebrandtia* (B), the ^{31}P signal at 23.71 ppm had cross-peaks with three protons signals at 1.81, 3.64 and 3.78 ppm. The proton signals at 1.81 and 3.64 ppm were assigned to CH_2P and CH_2OH of a phosphonate head group, respectively, by ^{1}H-^{1}H COSY correlations of H-1 with H-2. The proton signal at 3.78 ppm may originate from a phosphonate ester. The ^{31}P signal at 23.45 ppm had a cross-peak with the proton signal at 1.94 ppm (assigned to CH_2P of a phosphonate) and possibly also had cross-peaks with proton signals at 3.81 and 4.11 ppm. The proton signals at 3.81 and 4.11 ppm could not

be confidently assigned due to low signal intensities.

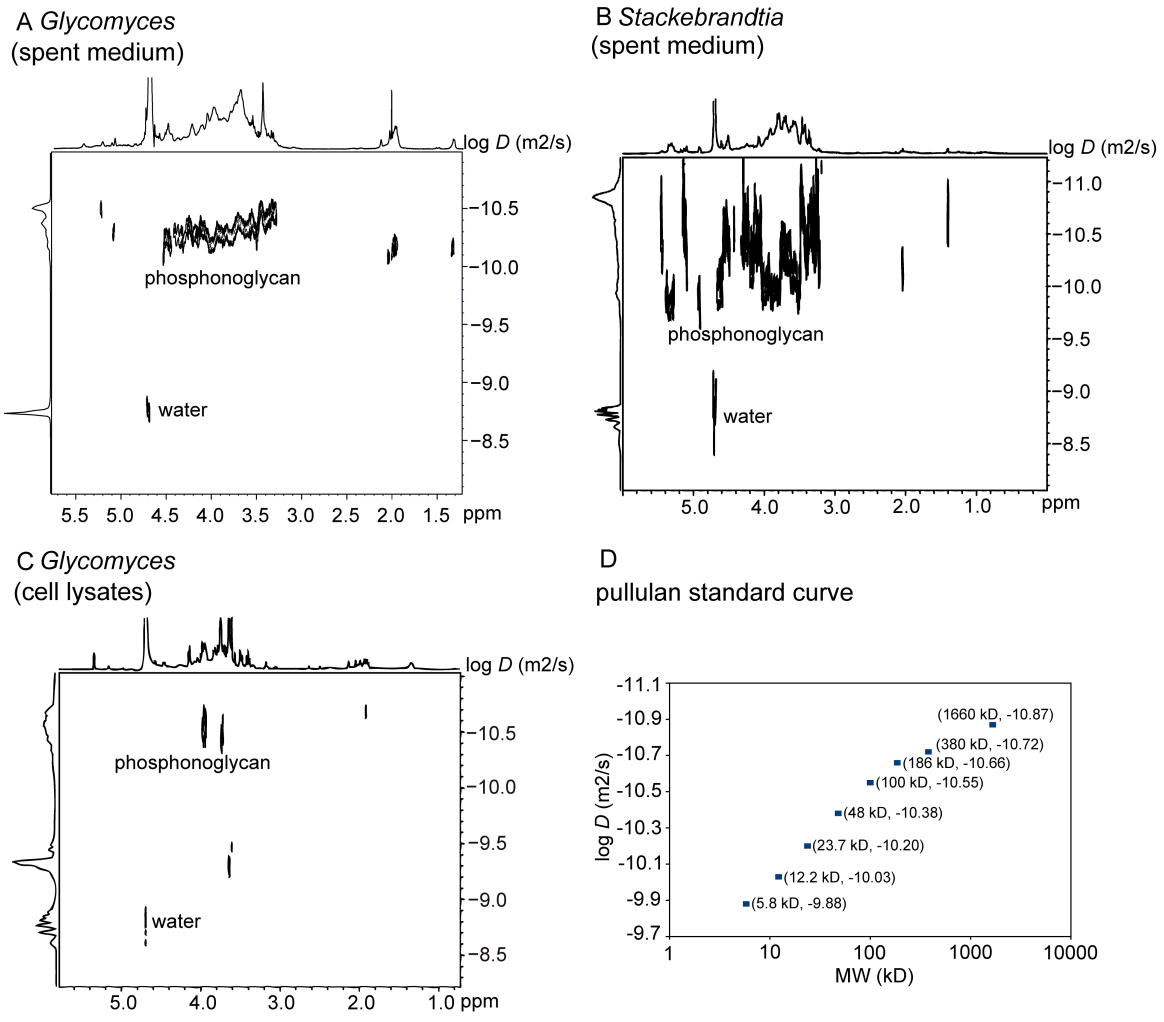

Fig. S3. ^{1}H DOSY NMR spectra of phosphonoglycans from *Glycomyces* and *Stackebrandtia*. (A) ^{1}H DOSY NMR spectrum of *Glycomyces* phosphonoglycans purified from the culture extract. (B) ^{1}H DOSY NMR spectrum of *Stackebrandtia* phosphonoglycans purified from the culture extract. (C) ^{1}H DOSY NMR spectrum of *Glycomyces* cell-bound phosphonoglycans. (D) A standard curve to plot log MW v.s. log *D* of a series of Shodex pullulan standards (in the range of 5.8 and 1660 kDa) in D_2O. The x axis in the ^{1}H DOSY NMR spectrum shows proton chemical shifts whereas the y axis shows the diffusion coefficiencies (log *D*) of compounds. The standard curve can be used to determine the approximate molecular weights of phosphonoglycans. The proton signal at 4.6 ppm in (A), (B) and (C) originates from water.

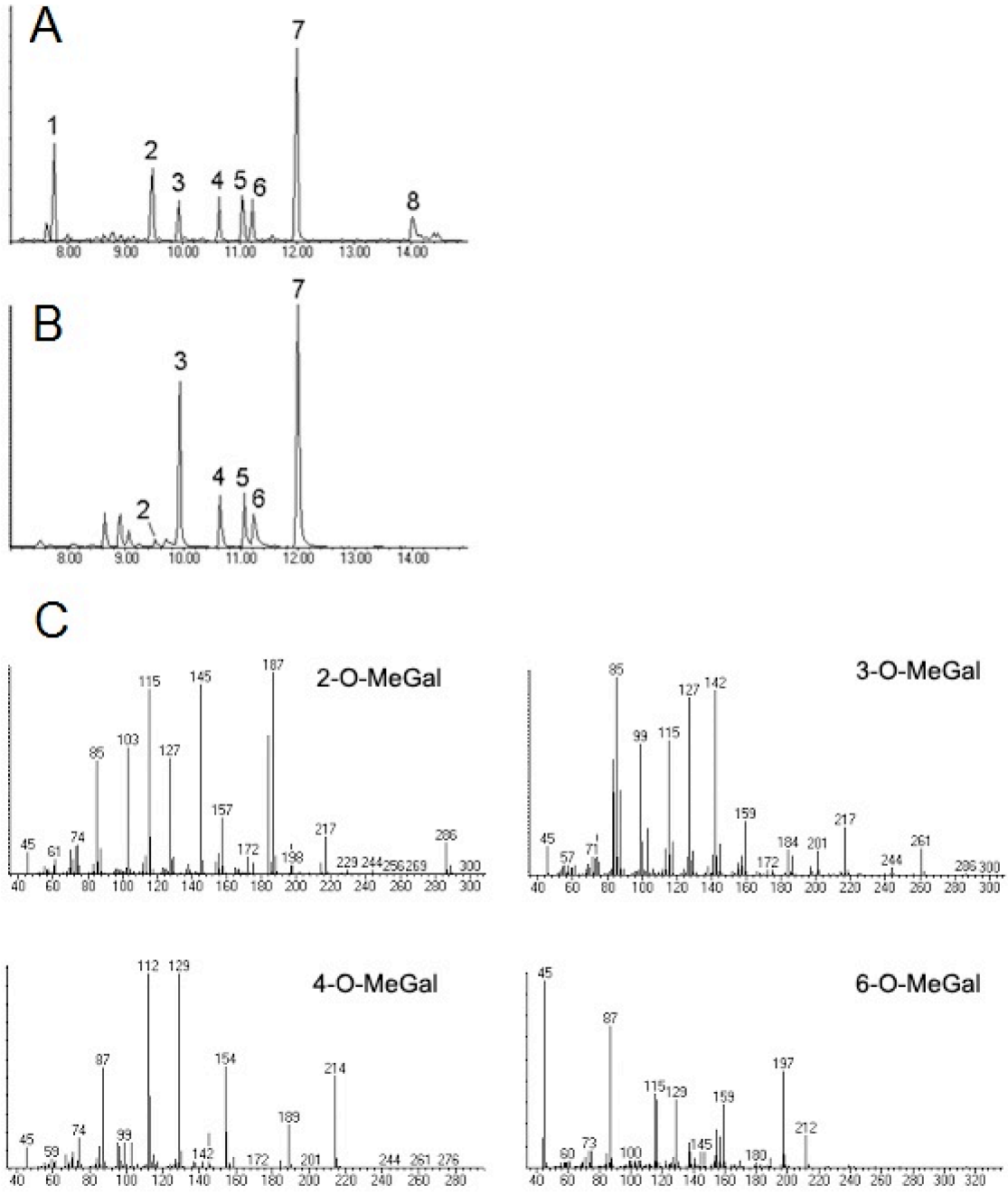

Fig. S4. GC chromatographs of acid labile monosaccharide composition from the major phosphonoglycan of *Glycomyces* (A) and partially *O*-methylated galactose standards (B). (C) EI-MS spectra for GC peaks 3, 4, 5 and 6. 1. Xyl; 2. 2,3-diMeGal; 3. 6-MeGal; 4. 2-MeGal; 5. 3-MeGal; 6. 4-MeGal; 7. Gal; 8. GlcNAc.

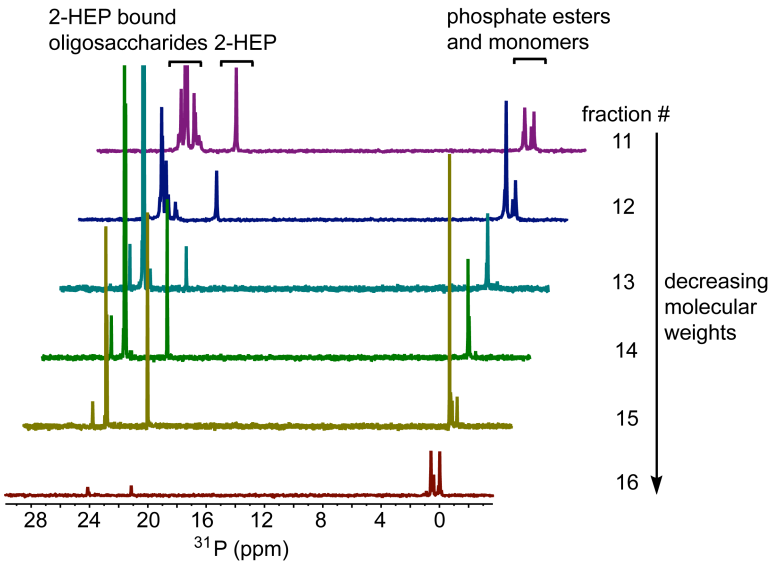

Fig. S5. Partial acid hydrolysis of *Glycomyces* phosphonoglycans produced various polymeric fragments containing different amounts of 2-HEP bound oligosaccharides and unbound 2-HEP. Fractions #11 to #16 were phosphonate-containing fractions eluted from Bio-Gel P2 column (Bio-Rad). The identity of the peak with the chemical shift of 21.2 ppm was confirmed to be 2-HEP by spiking the sample with an authentic 2-HEP standard. Fractions #14 to #16 were pooled for subsequent purification and characterization.

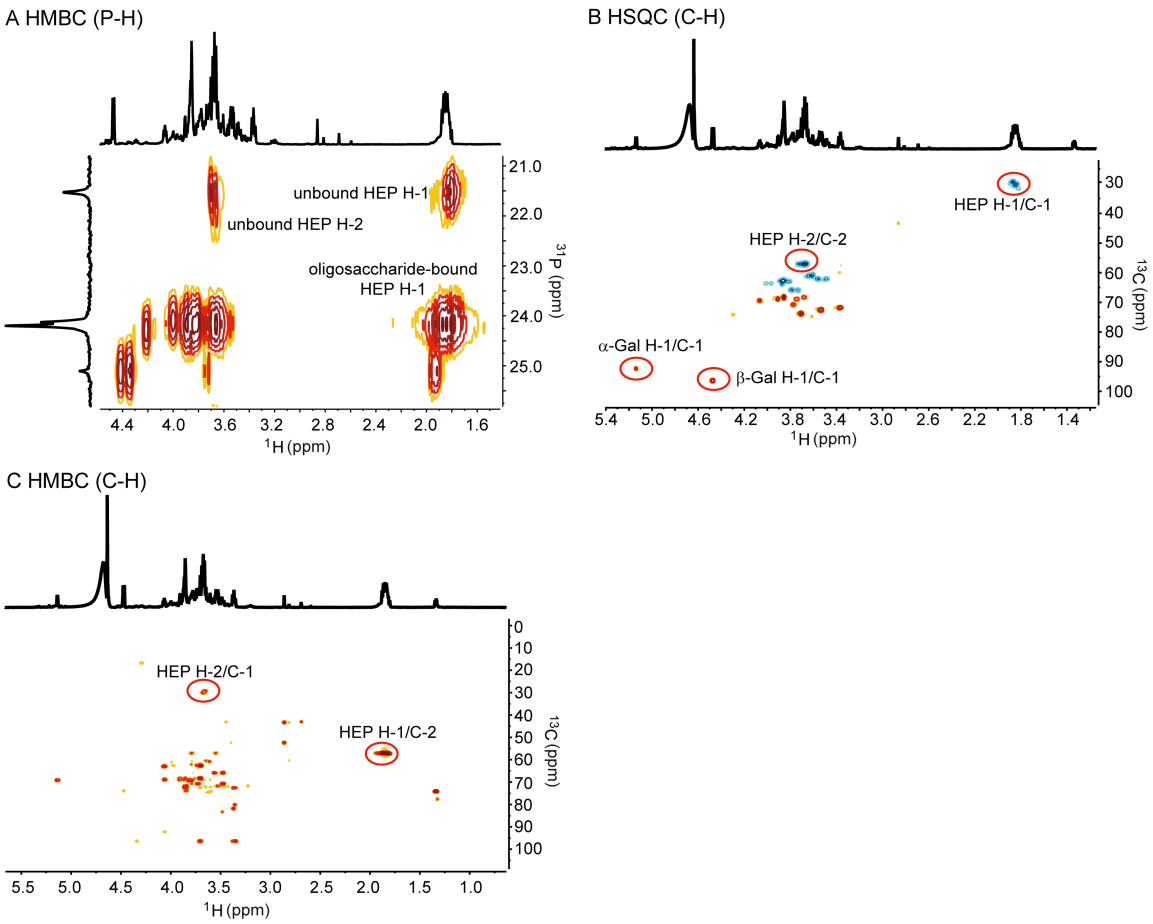

Fig. S6. 2D NMR analyses of retentate (MWCO= 0.5-1.0 kDa) after partial acid hydrolysis of *Glycomyces* phosphonoglycans. (A) ^{1}H-^{31}P HMBC spectrum. (B) ^{1}H-^{13}C HSQC spectrum. (C) ^{1}H-^{13}C HMBC spectrum. In the ^{1}H-^{31}P HMBC spectrum, the proton signal at 1.86 ppm had a cross-peak with one of the ^{31}P signals (or both) at 24.14 and 24.20 ppm. This was assigned to H-1 of the 2-HEP moiety, which was likely bound to oligosaccharides. C-1 (30.1 ppm) of the 2-HEP moiety, determined by the ^{1}H-^{13}C HSQC NMR experiment, had only one correlation with a proton signal at 3.67 ppm in the ^{1}H-^{13}C HMBC spectrum. This proton was assigned to H-2 of the 2-HEP moiety, which also showed the correlation with the ^{31}P signal at 24.14 (or 24.20) ppm in the ^{1}H-^{31}P HMBC spectrum. H-2 of 2-HEP only had one correlation with C-1 of 2-HEP but not with any other carbons in the ^{1}H-^{13}C HMBC spectrum, excluding the possibility of an ether linkage between 2-HEP and sugars.

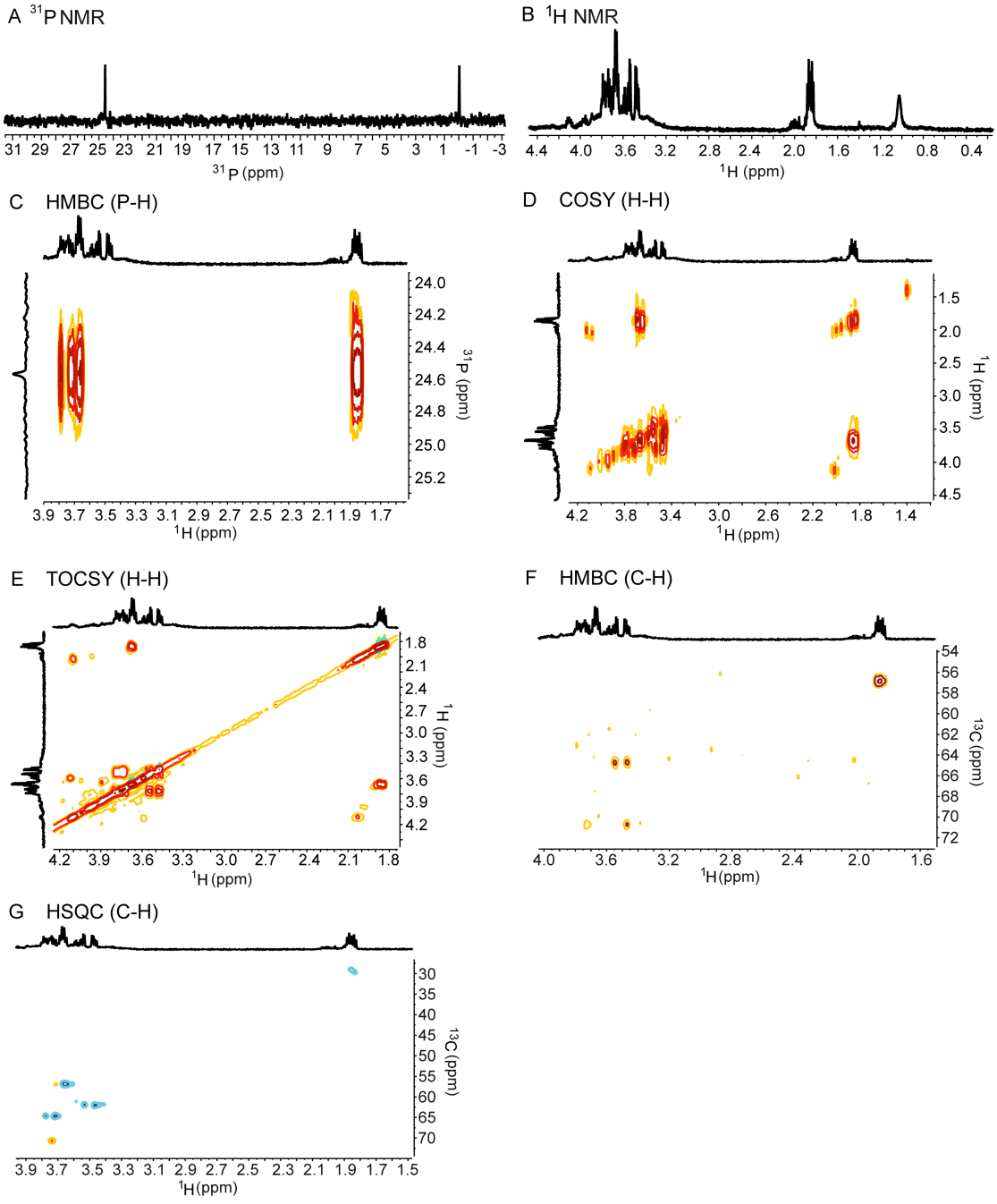

Fig. S7. 1D and 2D NMR analyses of 2-HEP mono(2,3-dihydroxypropyl) ester after partial acid hydrolysis of *Stackebrandtia* phosphonoglycans. (A) ^{31}P NMR spectrum. (B) ^{1}H NMR spectrum. (C) ^{1}H-^{31}P HMBC spectrum. (D) ^{1}H-^{1}H COSY NMR spectrum. (E) ^{1}H-^{1}H TOCSY NMR spectrum. (F) ^{1}H-^{13}C HMBC spectrum. (G) ^{1}H-^{13}C HSQC spectrum. In the ^{1}H-^{31}P HMBC spectrum, four proton signals at 1.85, 3.67, 3.72 and 3.78 ppm, had cross-peaks with the ^{31}P signal at 24.5 ppm. The proton signals at 1.85 and 3.67 ppm were assigned to CH_2P and CH_2OH of the 2-HEP moiety, respectively, by ^{1}H-^{1}H COSY and ^{1}H-^{1}H TOCSY correlations of

H-1 with H-2 and ^{1}H-^{13}C HMBC correlations of H-1 with C-2. Proton signals at 3.72 and 3.78 ppm are both from C-1′, as evident from ^{1}H-^{13}C HMBC correlations of H-1′ with C-2′ but not with C-1 or C-2. Additionally, two cross-peaks were found between the proton signal at 3.47 ppm and the carbon signals at 64.8 (C-1′) and 70.7 (C-2′) ppm. This proton signal was assigned to H-3′, consistent with ^{1}H-^{1}H TOCSY correlations of H-2′ with H-3′. Together, these data allow to assign the structure as 2-HEP mono(2,3-dihydroxypropyl) ester.

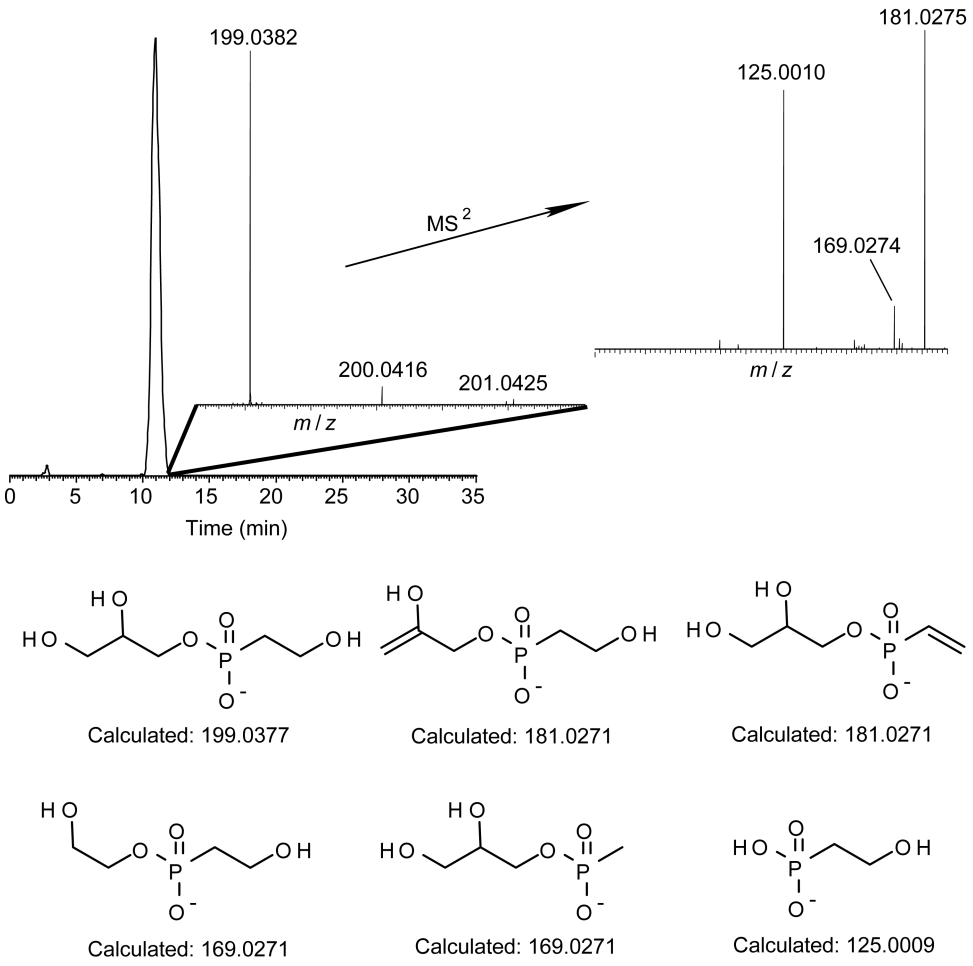

Fig. S8. High resolution LC-MS analysis of 2-HEP mono(2,3-dihydroxypropyl) ester after partial acid hydrolysis of *Stackebrandtia* phosphonoglycans. The precursor ion at *m*/*z* 199.0382 indicates HEP-linked glycerol, namely, 2-HEP mono(2,3-dihydroxypropyl) ester. This ion was selected for MS^2. Proposed ion structures were shown at the bottom. The peak labels in the spectrum indicate observed values whereas the numbers below the proposed ion structures indicate calculated monoisotopic values. LC-MS analysis was performed on a custom 11T linear ion-trap Fourier-transform mass spectrometer (LTQ-FT, Thermo Fisher Scientific) equipped with a 1200 HPLC system (Agilent). The mass spectral analysis consisted of a full scan at resolution 100,000 (*m/z* 100-1000), a source fragmentation scan (85V, *m*/*z* 50-110) detected in the ion trap and a targeted CID MS2 scan with FT detection to obtain tandem mass spectra of target compounds.

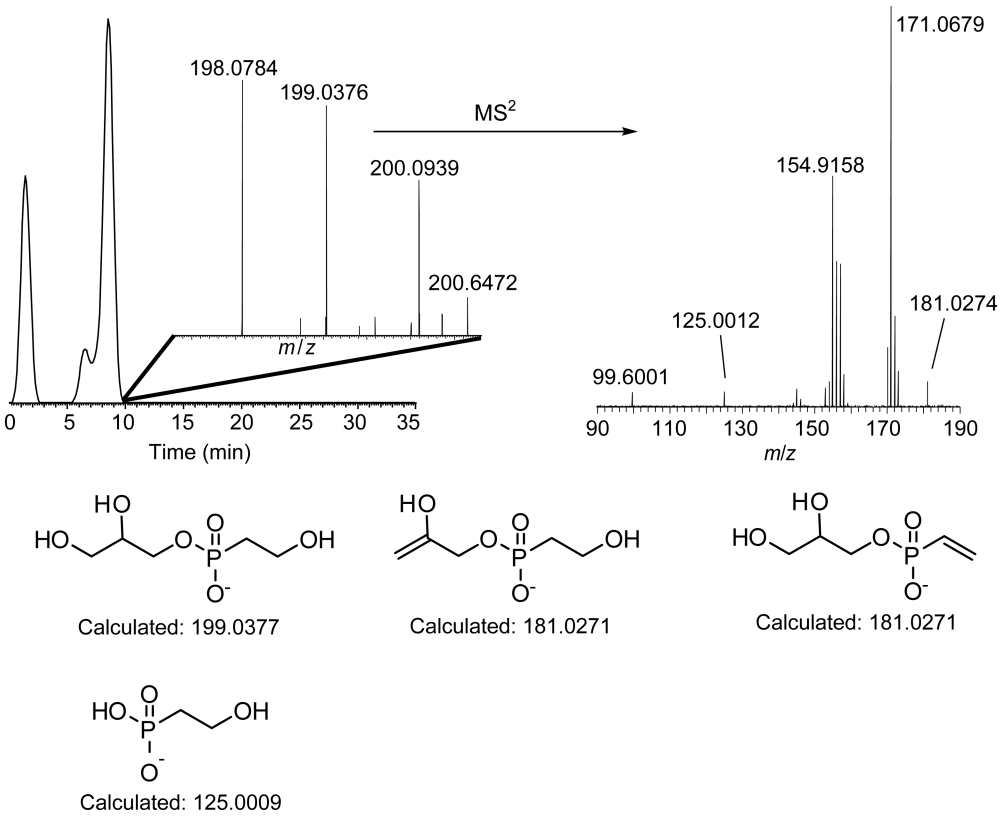

Fig. S9. High resolution LC-MS analysis of 2-HEP mono(2,3-dihydroxypropyl) ester after partial acid hydrolysis of *Glycomyces* phosphonoglycans. The precursor ion at *m*/*z* 199.0376 indicates HEP-linked glycerol, namely, 2-HEP mono(2,3-dihydroxypropyl) ester. This ion was selected for MS^2. Proposed ion structures were shown at the bottom. The peak labels in the spectrum indicate observed values whereas the numbers below the proposed ion structures indicate calculated monoisotopic values. LC-MS analysis was performed on a custom 11T linear ion-trap Fourier-transform mass spectrometer (LTQ-FT, Thermo Fisher Scientific) equipped with a 1200 HPLC system (Agilent). The mass spectral analysis consisted of a full scan at resolution 100,000 (*m/z* 100-1000), a source fragmentation scan (85V, *m*/*z* 50-110) detected in the ion trap and a targeted CID MS2 scan with FT detection to obtain tandem mass spectra of target compounds.

235 **Table S1.** Summary of ORFs in putative phosphonoglycan biosynthetic locus of

236 *Stackebrandtia*

ORF	No. of amino acids	Predicted function (GenBank accession)
1′	116	ABC transporter (YP_003514402)
2′	474	nucleotide sugar dehydrogenase (YP_003514401)
3′	310	endonuclease/exonuclease/phosphatase (YP_003514400)
4′	193	hypothetical protein Snas_5676 (YP_003514399)
5′	704	LmbE family protein (YP_003514398)
6′	762	glycosyl transferase group 1 (YP_003514397)
7′	549	N-acetylglucosamine phosphotransferase (YP_003514396)
8′	593	N-acetylglucosamine phosphotransferase (YP_003514395)
9′	685	CDP-alcohol phosphatidyltransferase (YP_003514394)
10′	253	ABC transporter related protein (YP_003514393)
11′	291	CDP-alcohol phosphatidyltransferase (YP_003514392)
12′	239	2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase (YP_003514391)
13′	551	hypothetical protein Snas_5667 (YP_003514390)
14′	569	hypothetical protein Snas_5666 (YP_003514389)
15′	437	PEP mutase (YP_003514388)

16′	373	phosphonopyruvate decarboxylase (YP_003514387)
17′	386	iron-containing alcohol dehydrogenase (YP_003514386)
18′	238	hypothetical protein Snas_5662 (YP_003514385)
19′	138	hypothetical protein Snas_5661 (YP_003514384)
20′	106	hypothetical protein Snas_5660 (YP_003514383)
21′	237	hypothetical protein Snas_5659 (YP_003514382)
22′	196	hypothetical protein Snas_5658 (YP_003514381)
23′	102	hypothetical protein Snas_5657 (YP_003514380)
24′	662	major facilitator superfamily protein (YP_003514379)
25′	239	flavoprotein (YP_003514378)
26′	449	radical SAM protein (YP_003514377)
27′	222	NAD-dependent epimerase/dehydratase (YP_003514376)
28′	143	glyoxalase/bleomycin resistance protein/dioxygenase (YP_003514375)
29′	219	GntR family transcriptional regulator (YP_003514374)