

Supplementary Materials for

Switchgrass metabolomics reveals striking genotypic and developmental differences in specialized metabolic phenotypes

Xingxing Li^{a, b}, Saurav J. Sarma^{c, f}, Lloyd W. Sumner^{d, e, f, g}, A. Daniel Jones^{a, b} and Robert L. Last^{a, b, c *}

^aDepartment of Biochemistry and Molecular Biology, Michigan State University, MI 48824

^b DOE Great Lakes Bioenergy Research Center, Michigan State University, MI 48824

^cDepartment of Plant Biology, Michigan State University, MI 48824

^dDepartment of Biochemistry, University of Missouri, Columbia, MO 65211

^eBond Life Sciences Center, University of Missouri, Columbia, MO 65211

^fMU Metabolomics Center, University of Missouri, Columbia, MO 65211

^gInterdisciplinary Plant Group, University of Missouri, Columbia, MO 65211

*Corresponding author: lastr@msu.edu

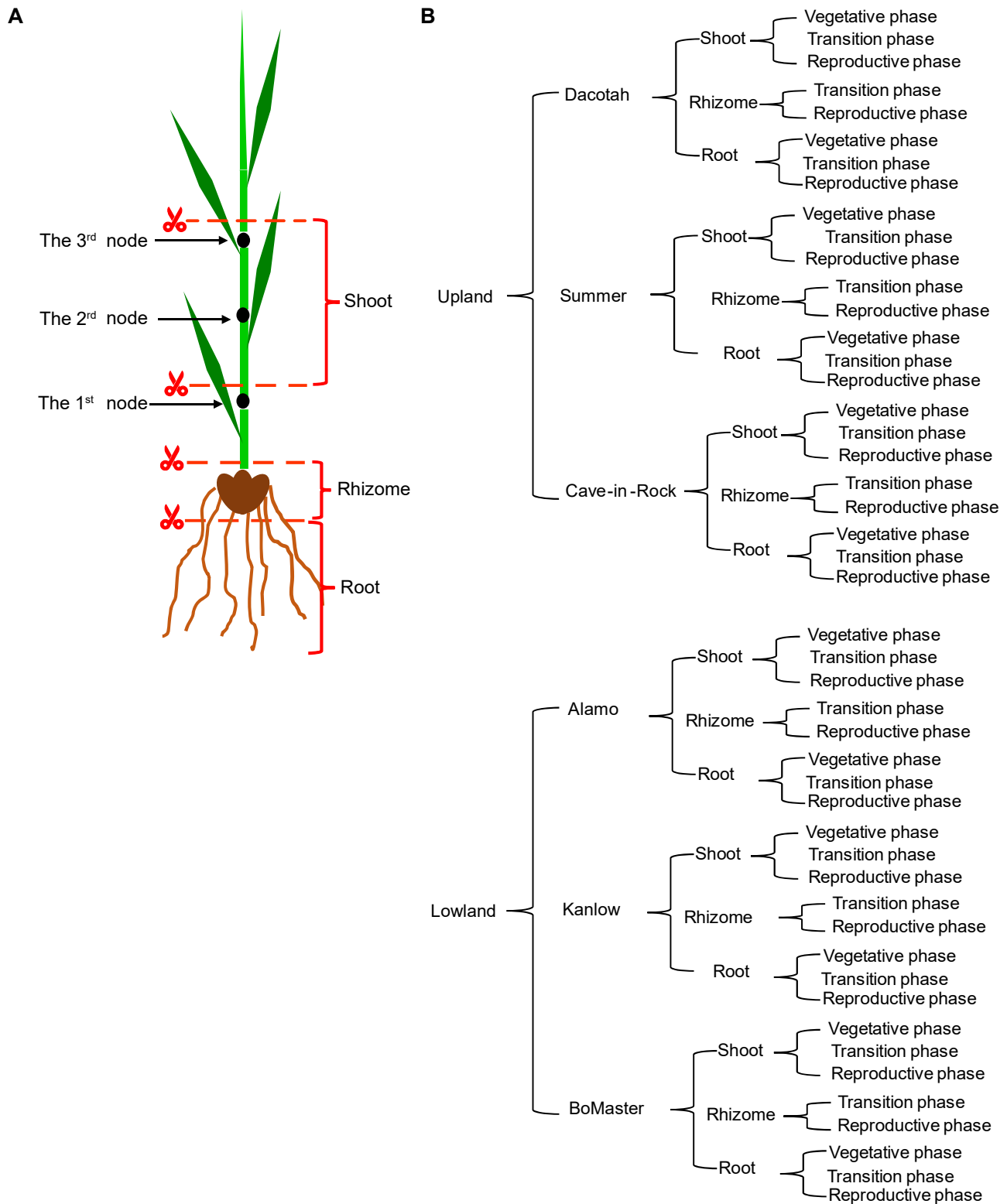


Figure S1. The method for sample collection and sample panel involved in this study. (A) Developmental stage identifications and sample collections were performed according to Hardin et al. (2013). The diagram shows how the shoot, rhizome and root tissues were collected from the plants at the transition- or reproductive-phase. The entire above- or below-ground parts were collected for the plants at the vegetative-phase. (B) The sample panel in this study contains six switchgrass cultivars, three tissue types and three developmental stages. Rhizomes were not collected for the plants at the vegetative-phase, due to the small size. In total, there are 48 samples groups in this study.

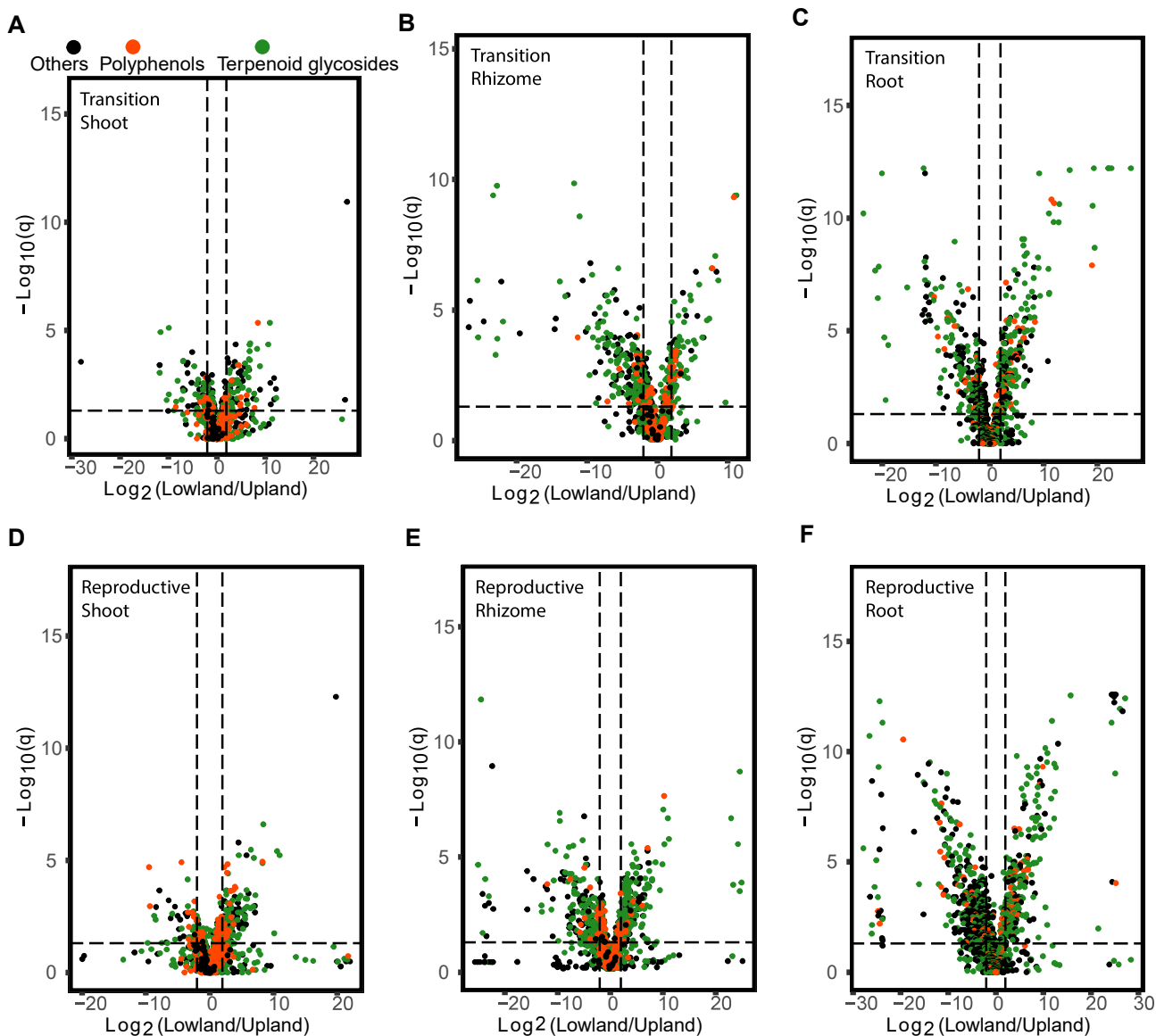


Figure S2. Significance analysis for the differentially accumulated features (DAFs) between the upland and lowland switchgrass ecotypes for each given *tissue type x development stage* combination. (A) *Shoot x Transition phase*. (B) *Rhizome x Transition phase*. (C) *Root x Transition phase*. (D) *Shoot x Reproductive phase*. (E) *Rhizome x Reproductive phase*. (F) *Root x Reproductive*. The horizontal dash line stands for the adjusted p -value = 0.05. Vertical dash lines stand for the \log_2 value of the fold-change = 2. Metabolite symbols are color-coded based on relative mass defect (RMD) analysis.

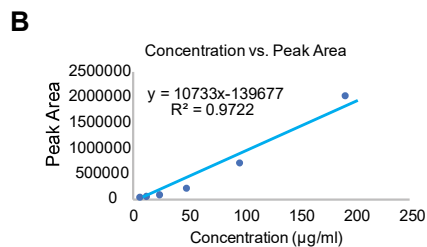
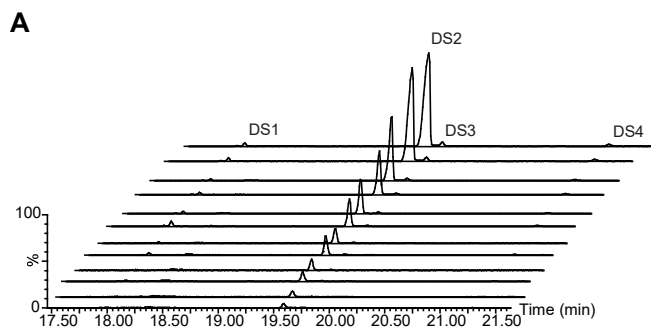


Figure S3. The six-point instrument response curve over a range of 6 – 192 mg/ml by serial dilution of the diosgenin standard. (A) Extracted ion (m/z 139) chromatograms of duplicate standard injections from the standard concentration series (6, 12, 24, 48, 96 and 192 $\mu\text{g/ml}$). One major (DS2) and three minor (DS1, DS3 and DS4) diosgenin standard derived peaks (due to the sample treatment, **Materials and Methods**) were identified. Areas for the four peaks were combined when plotted against the concentration. (B) Reproducibility and linearity ($R^2 > 0.97$) of the method.

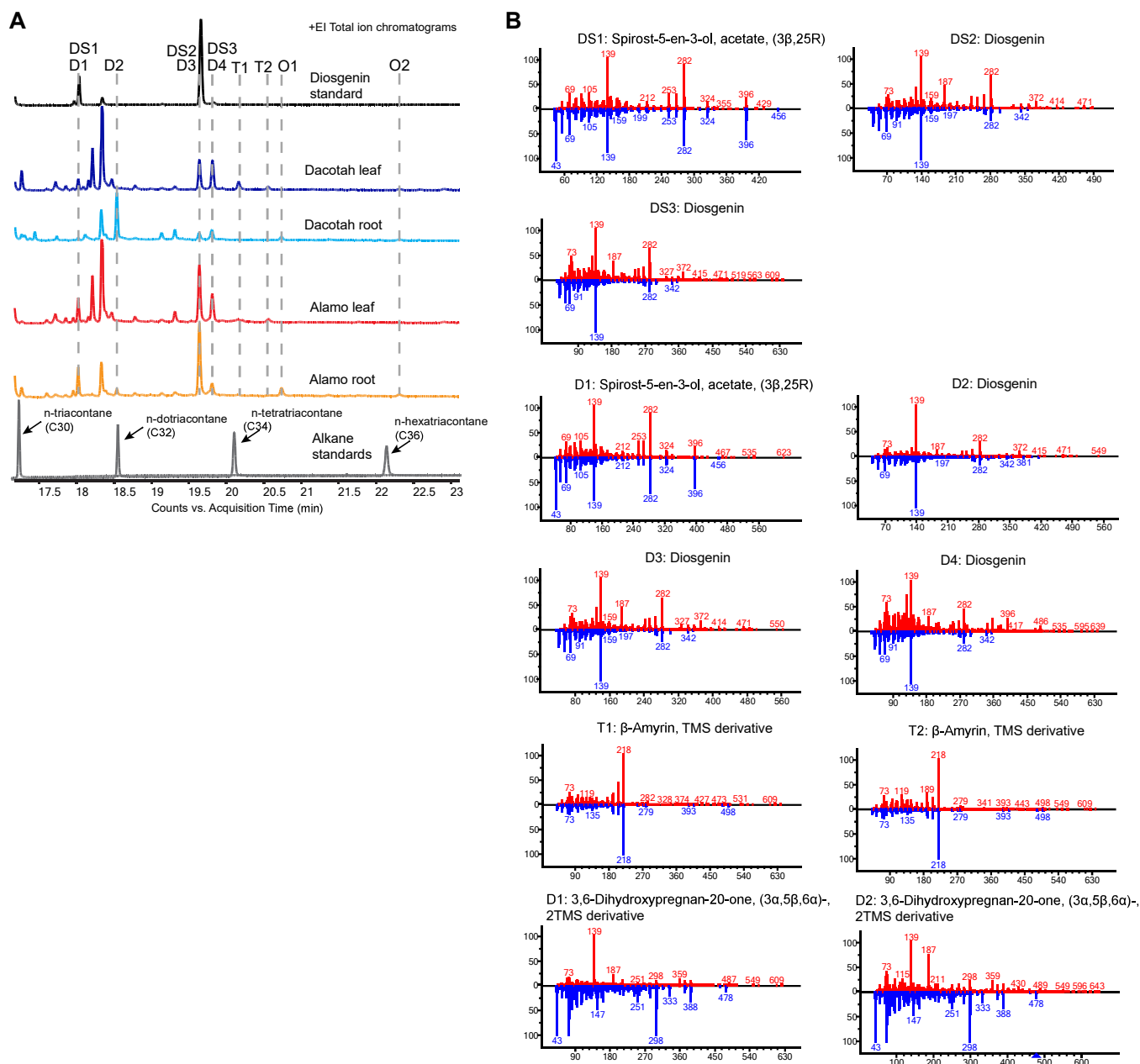


Figure S4. Sapogenin aglycone peaks identified in switchgrass extracts by (EI) GC-MS. (A) +EI GC-MS total ion chromatograms (TICs) of the extracts of Dacotah leaf (dark blue), Dacotah root (blue), Alamo leaf (red), Alamo root (orange) as well as a diosgenin standard sample (black) and an n-alkane standard mixture (gray) containing n-triacontane (C30), n-dotriacontane (C32), n-tetratriacontane (C34), and n-hexatriacontane (C36) that was used for calculating the RIs. The identified potential sapogenin peaks in the switchgrass and standard samples are indicated and aligned by the dashed lines. Note that the D1, D3 and D4 peaks detected in switchgrass samples have slightly different RTs from the DS1, DS2 and DS3 peaks detected in the diosgenin standard sample (black). (B) The (EI) GC-MS spectral pattern (upper/red) for each potential sapogenin peak. The abundant fragment ions for each peak are listed in **Table 2**. The best NIST library matches (lower/blue) for the EI spectral patterns of the detected sapogenin peaks, based on the highest reverse match scores, are shown.

Figure S5. Biplot of the PCA in Fig. 3C showing differential accumulations of the specialized metabolites among the four sample groups (upland shoot, upland root, lowland shoot and lowland root). For an easier visualization, the biplot was separated into an upper panel for sesquiterpenoids, diterpenoids and flavonoid glycosides (A) and a lower panel for saponins (B). The saponins from different classes were colored to match with the color schemes in **Fig 4A** and **Table S21** (Red, D-415-SCG; Purple, D-415; Dark blue, D-431-SCG; Blue, D-431; Orange, D-417-SCG; Green, D-457; Yellow, D-473) For the class names, the letter 'D' stands for 'diosgenin'; numerical number reflects *m/z* value of aglycone fragment ion detected by positive mode MS; 'SCG' stands for 'side chain glycosylation'. The metabolites further away from the center are more preferentially accumulated in certain tissue types and/or ecotypes.

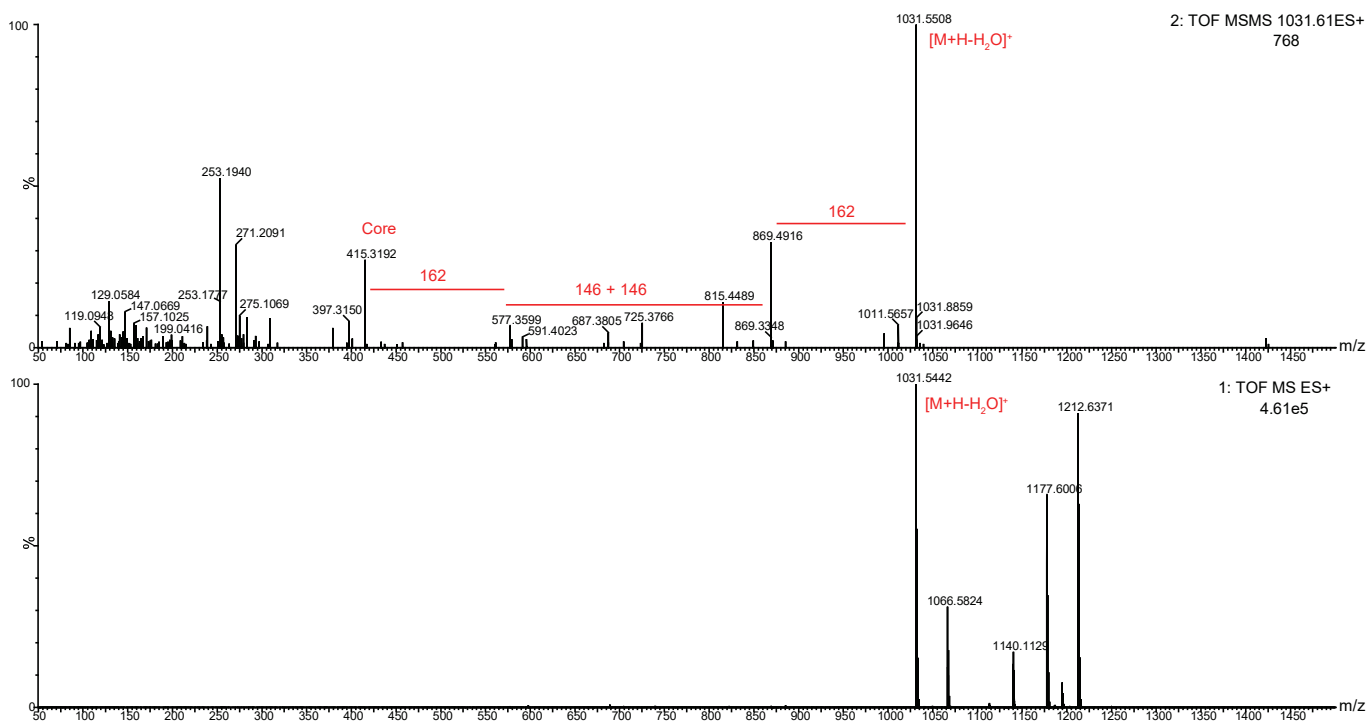


Figure S6. Mass spectra of the switchgrass saponin SS1031 analyzed by high resolution ES+ LC-MS using DDA mode. MS function: 1. survey scan; 2. DDA scan (MS/MS). The integer values of the MS neutral mass losses are labeled (18, loss of one molecule of water; 146, loss of one molecule of deoxyhexose; 162, loss of one molecule of hexose). See **Table S14** for additional details.

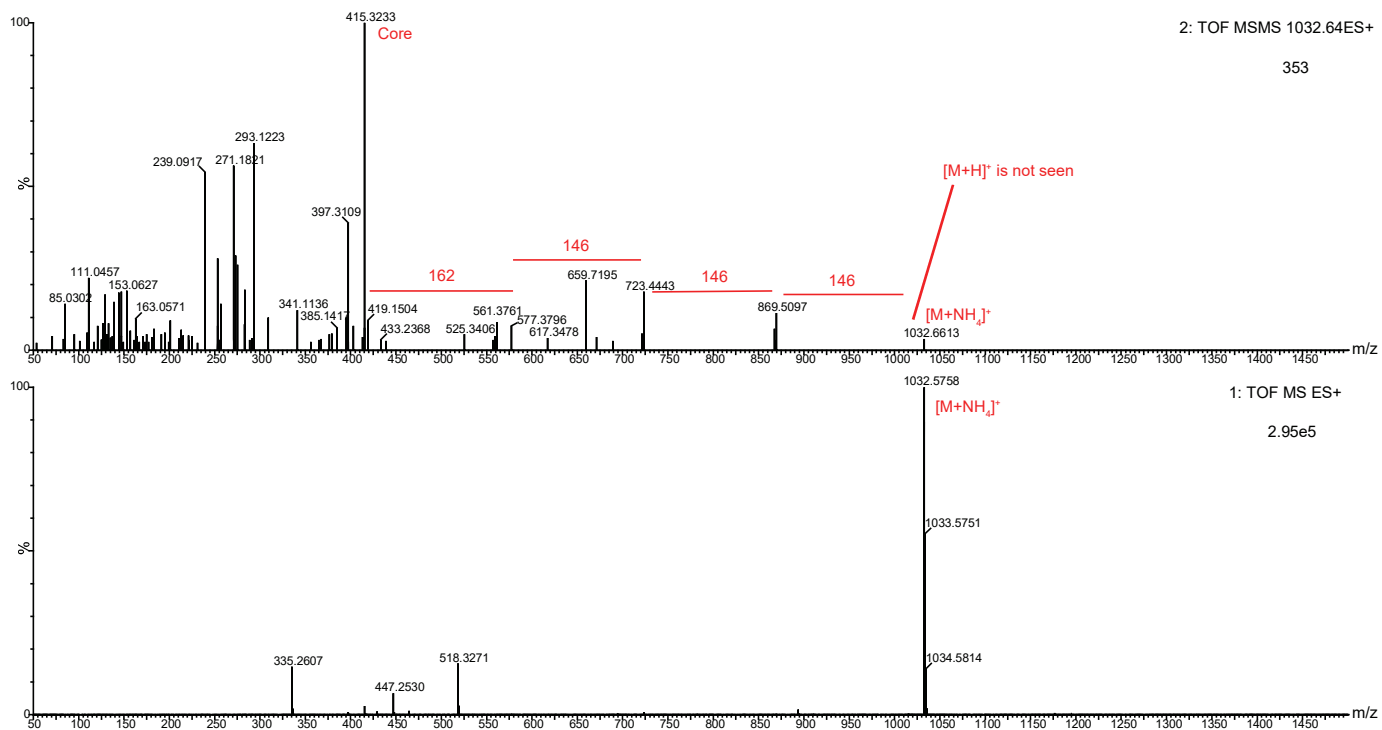


Figure S7. Mass spectra of the switchgrass saponin SS1032 analyzed by high resolution ES+ LC-MS using DDA mode. MS function: 1. survey scan; 2. DDA scan (MS/MS). The integer values of the MS neutral mass losses are labeled (18, loss of one molecule of water; 146, loss of one molecule of deoxyhexose; 162, loss of one molecule of hexose). See **Table S14** for additional details.

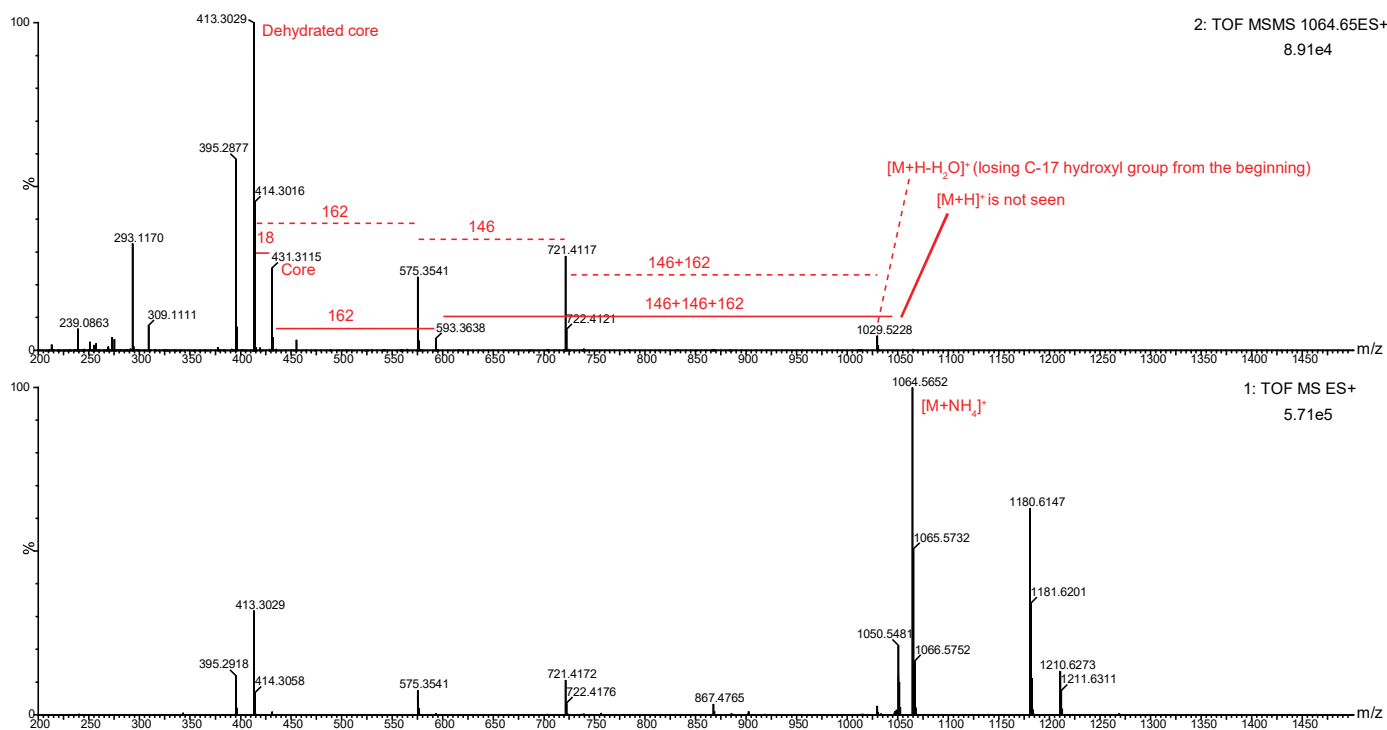


Figure S8. Mass spectra of the switchgrass saponin SS1064 analyzed by high resolution ES+ LC-MS using DDA mode. MS function: 1. survey scan; 2. DDA scan (MS/MS). The integer values of the MS neutral mass losses are labeled (18, loss of one molecule of water; 146, loss of one molecule of deoxyhexose; 162, loss of one molecule of hexose). See **Table S14** for additional details.

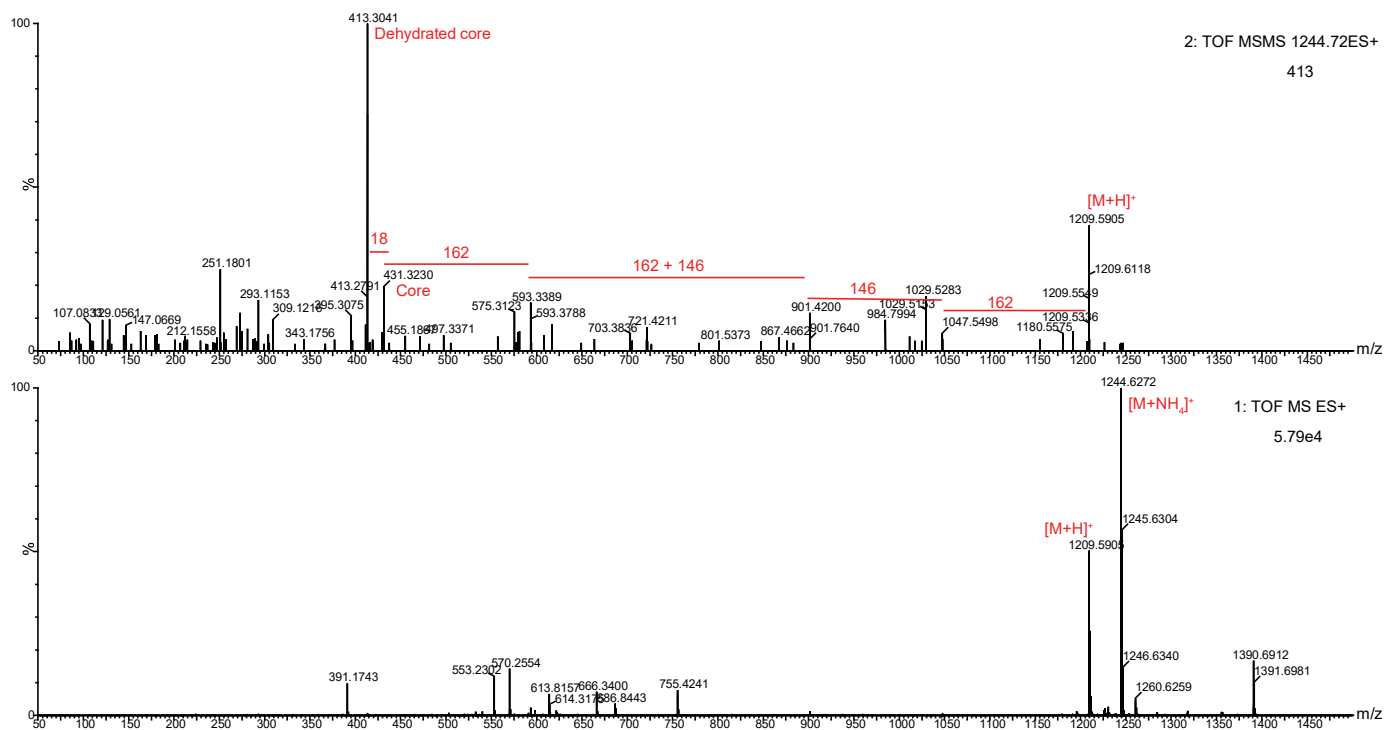


Figure S9. Mass spectra of the switchgrass saponin SS10244 analyzed by high resolution ES+ LC-MS using DDA mode. MS function: 1. survey scan; 2. DDA scan (MS/MS). The integer values of the MS neutral mass losses are labeled (18, loss of one molecule of water; 146, loss of one molecule of deoxyhexose; 162, loss of one molecule of hexose). See **Table S14** for additional details.

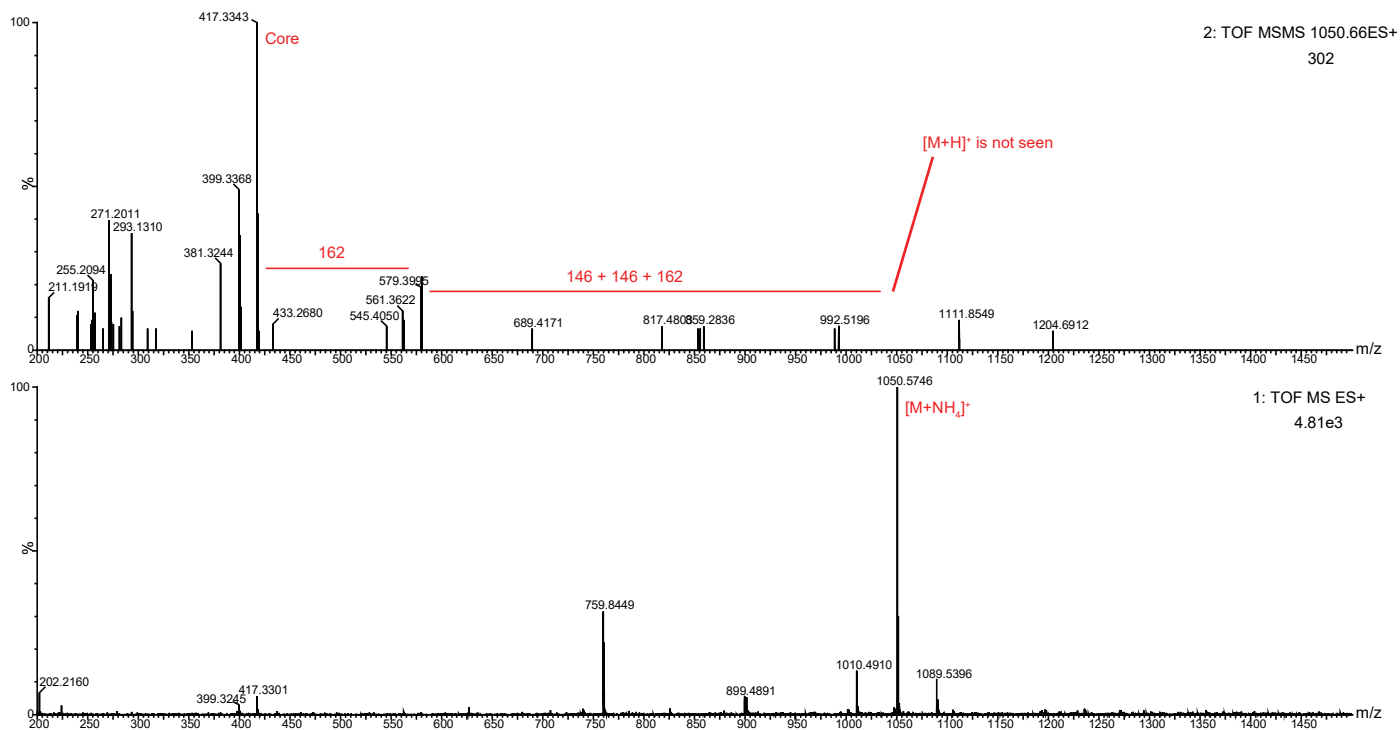


Figure S10. Mass spectra of the switchgrass saponin SS1050 analyzed by high resolution ES+ LC-MS using DDA mode. MS function: 1. survey scan; 2. DDA scan (MS/MS). The integer values of the MS neutral mass losses are labeled (18, loss of one molecule of water; 146, loss of one molecule of deoxyhexose; 162, loss of one molecule of hexose). See **Table S14** for additional details.

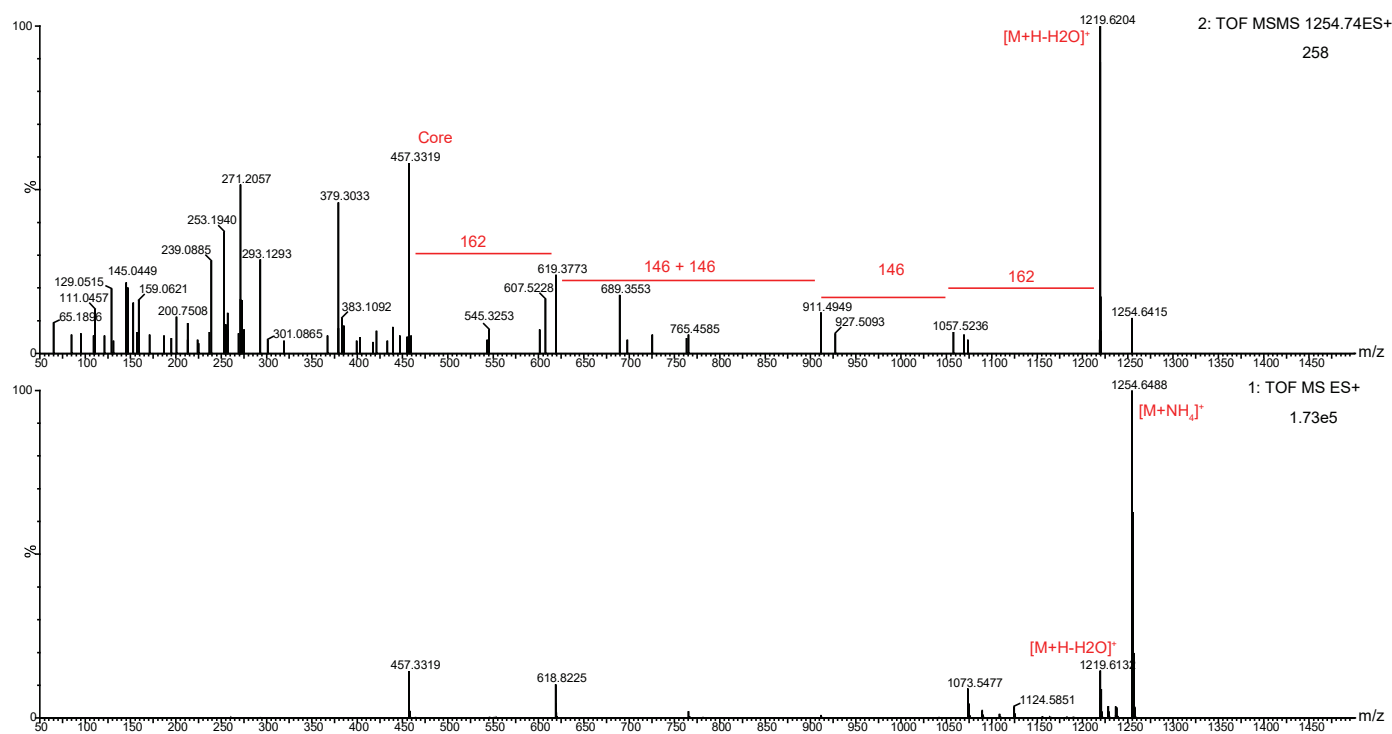


Figure S11. Mass spectra of the switchgrass saponin SS1254 analyzed by high resolution ES+ LC-MS using DDA mode. MS function: 1. survey scan; 2. DDA scan (MS/MS). The integer values of the MS neutral mass losses are labeled (18, loss of one molecule of water; 146, loss of one molecule of deoxyhexose; 162, loss of one molecule of hexose). See **Table S14** for additional details.

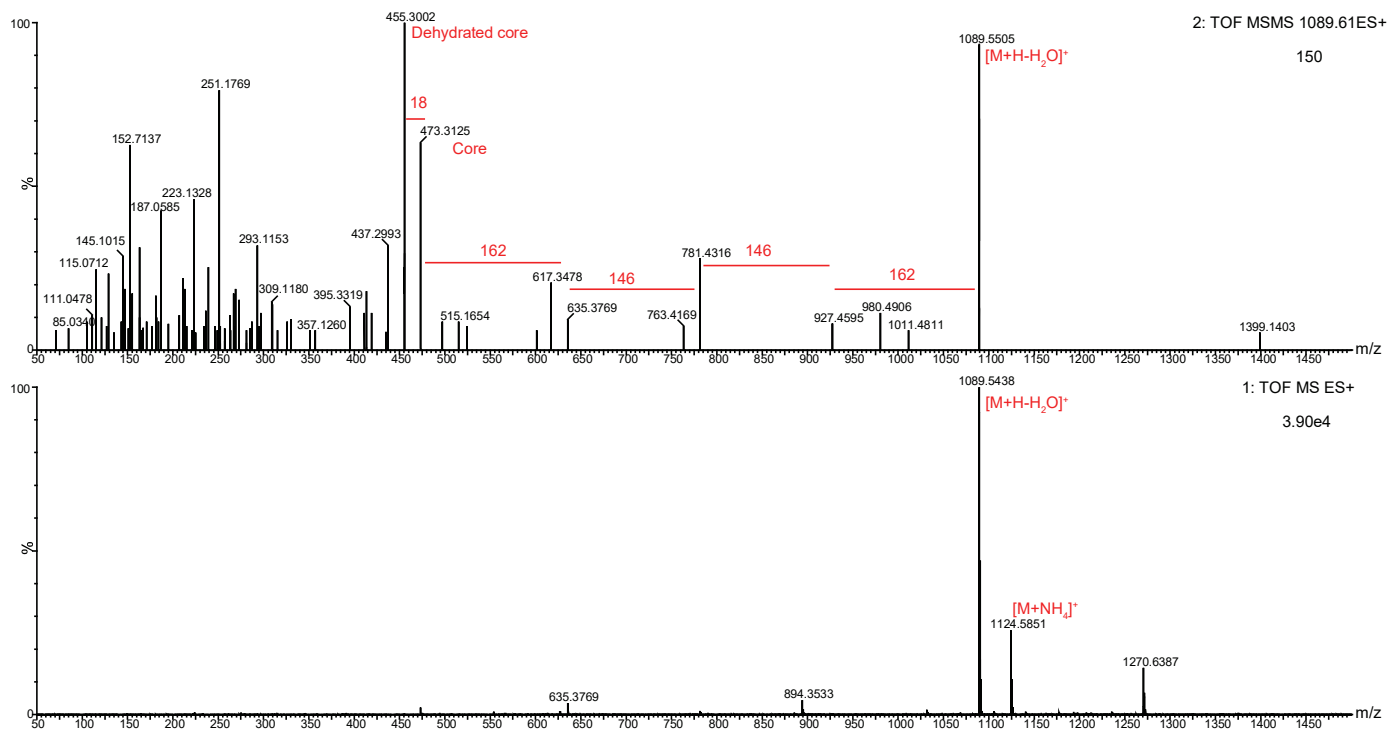


Figure S12. Mass spectra of the switchgrass saponin SS1089 analyzed by high resolution ES+ LC-MS using DDA mode. MS function: 1. survey scan; 2. DDA scan (MS/MS). The integer values of the MS neutral mass losses are labeled (18, loss of one molecule of water; 146, loss of one molecule of deoxyhexose; 162, loss of one molecule of hexose). See **Table S14** for additional details.

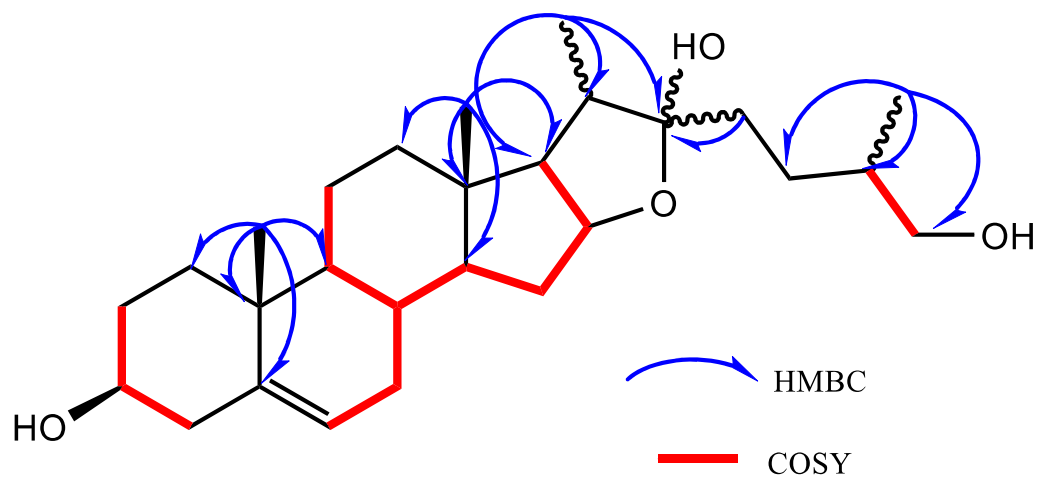


Figure S13. The key COSY (red, bold) and HMBC (blue arrow) correlations for the aglycone of the switchgrass saponin SS1031.

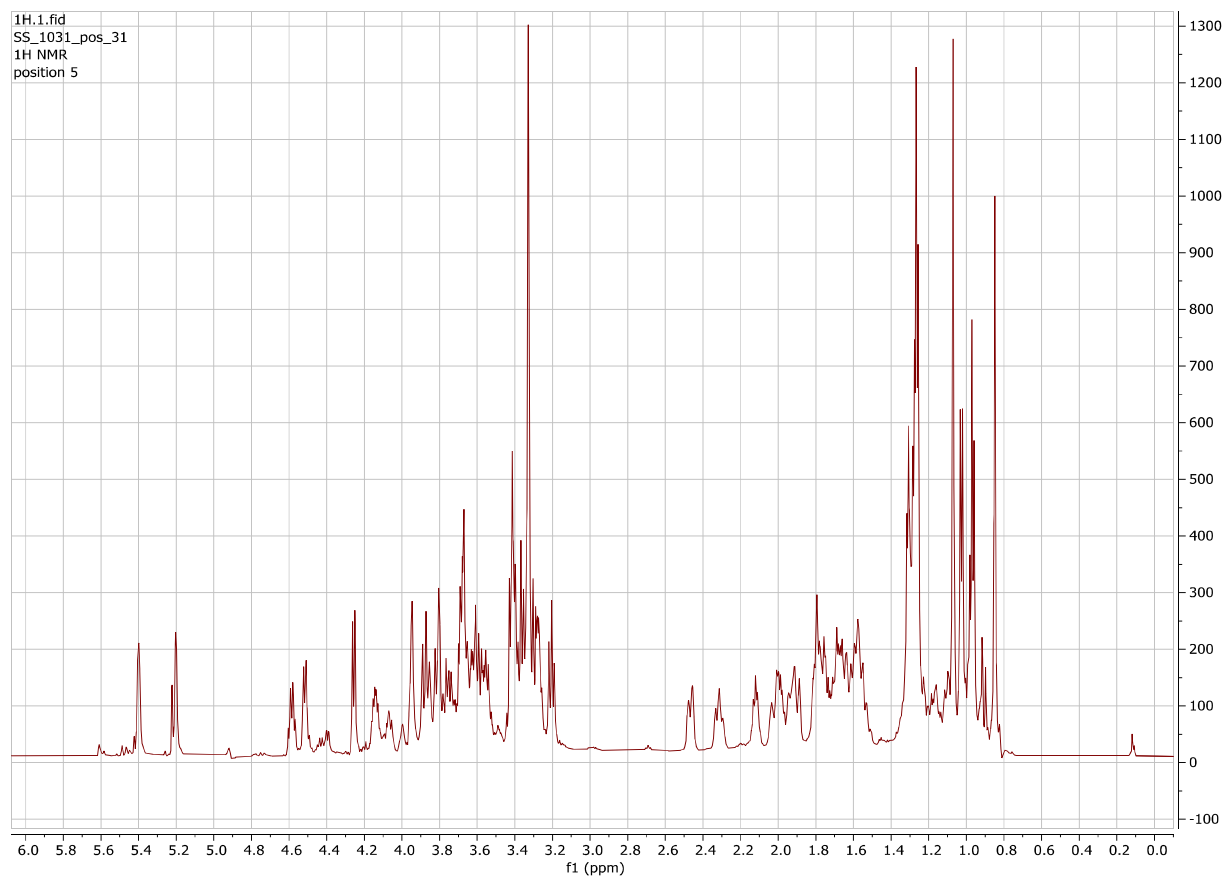


Figure S14. ¹H NMR spectrum for the intact molecule of the switchgrass saponin SS1031.

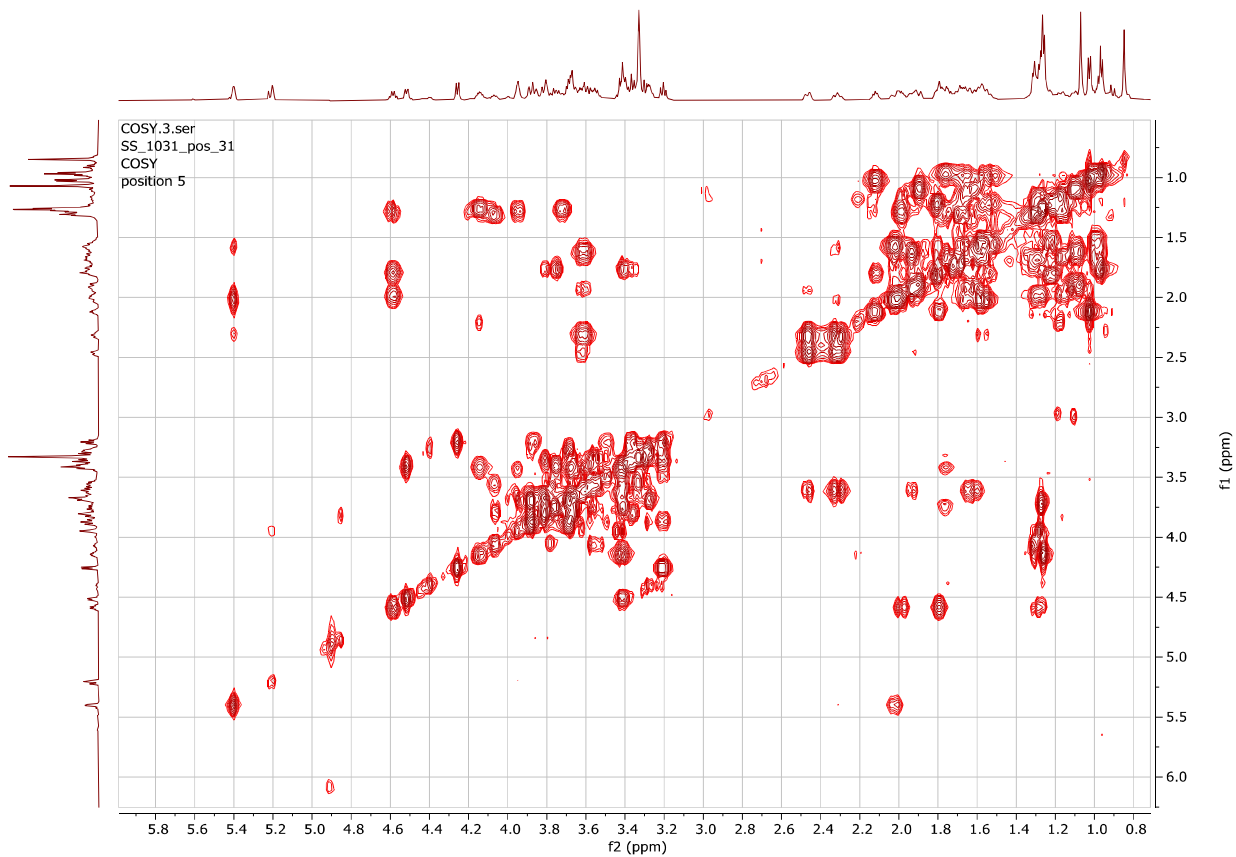


Figure S15. gCOSY ^1H - ^1H NMR spectrum for the intact molecule of the switchgrass saponin SS1031.



Figure S16. gHMBCAD NMR spectrum for the intact molecule of the switchgrass saponin SS1031.

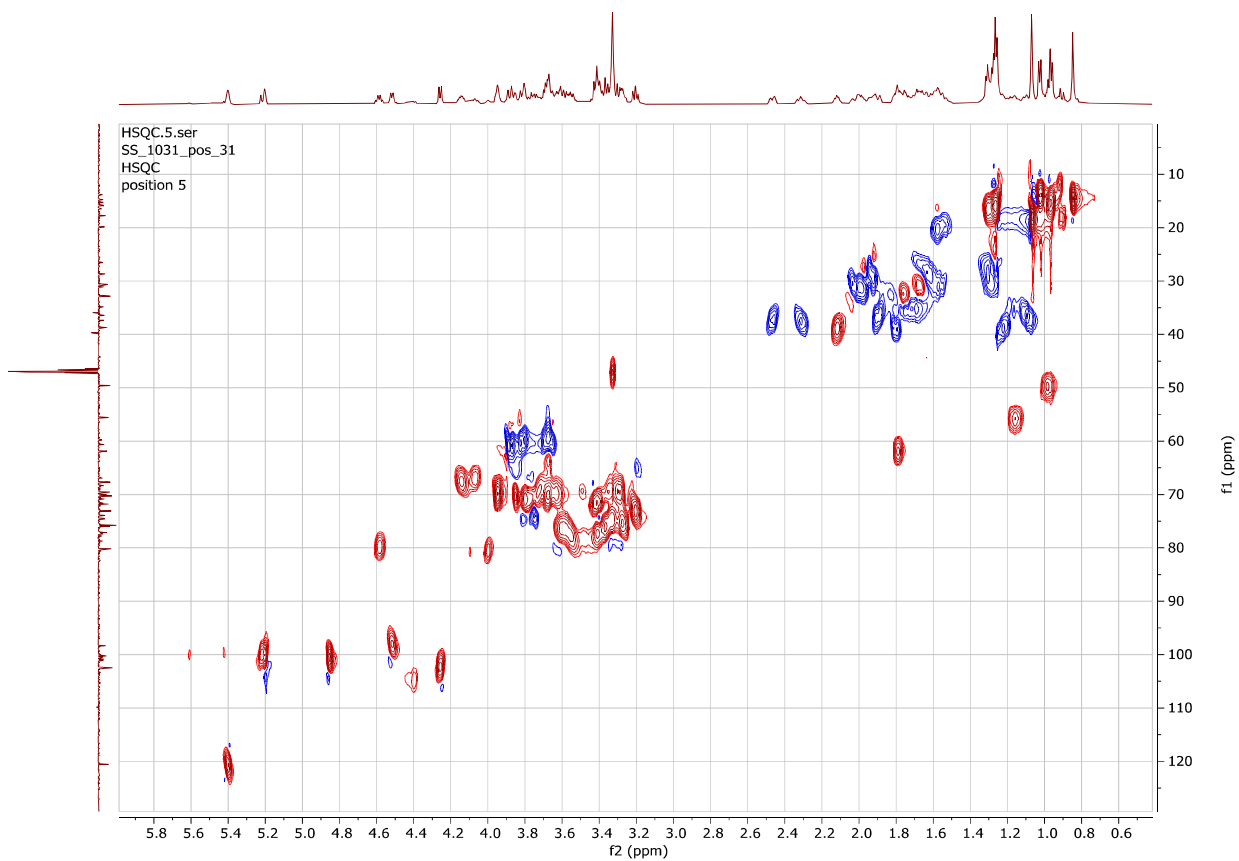


Figure S17. gHSQCAD NMR spectrum for the intact molecule of the switchgrass saponin SS1031.



Figure S18. gTOCSY NMR spectrum for the intact molecule of the switchgrass saponin SS1031.

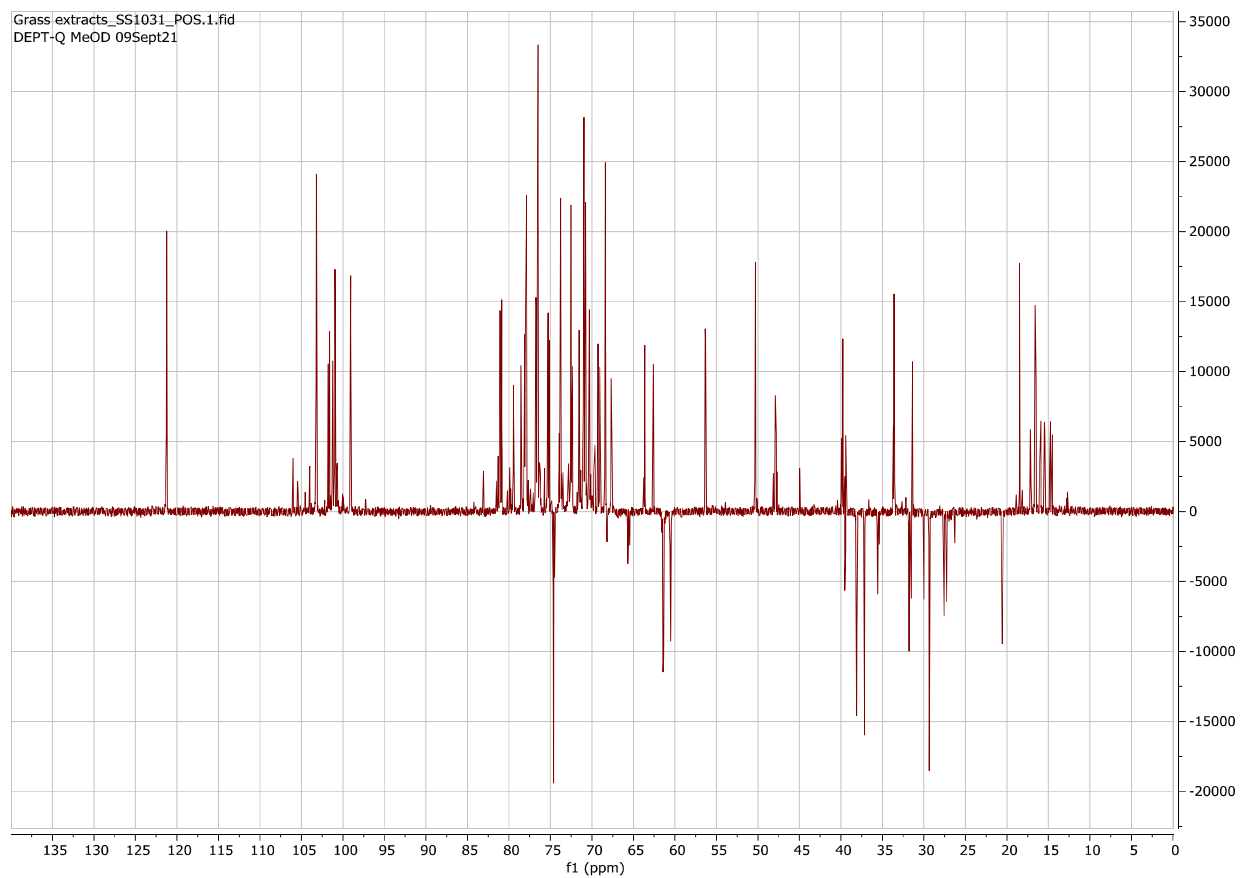


Figure S19. DEPT-Q (^{13}C) NMR spectrum for the intact molecule of the switchgrass saponin SS1031.

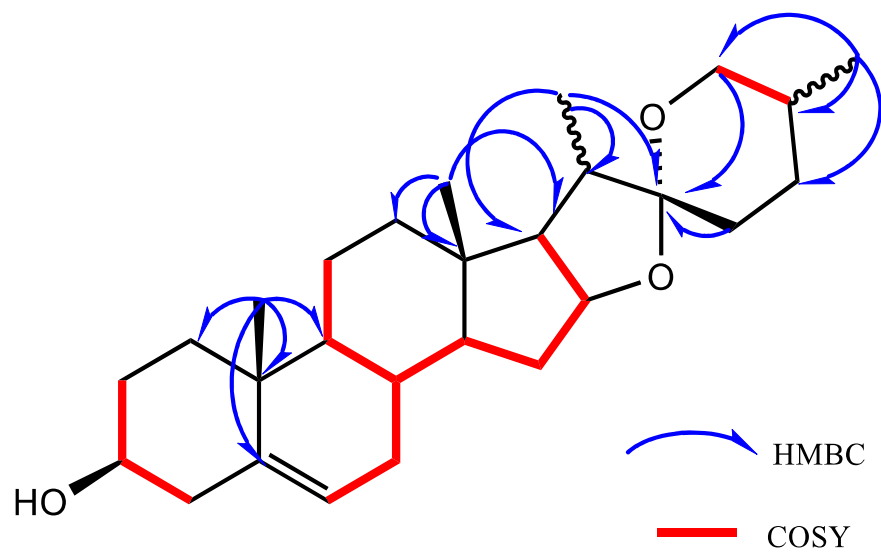


Figure S20. The key COSY (red, bold) and HMBC (blue arrow) correlations for the aglycone of the switchgrass saponin SS1032.

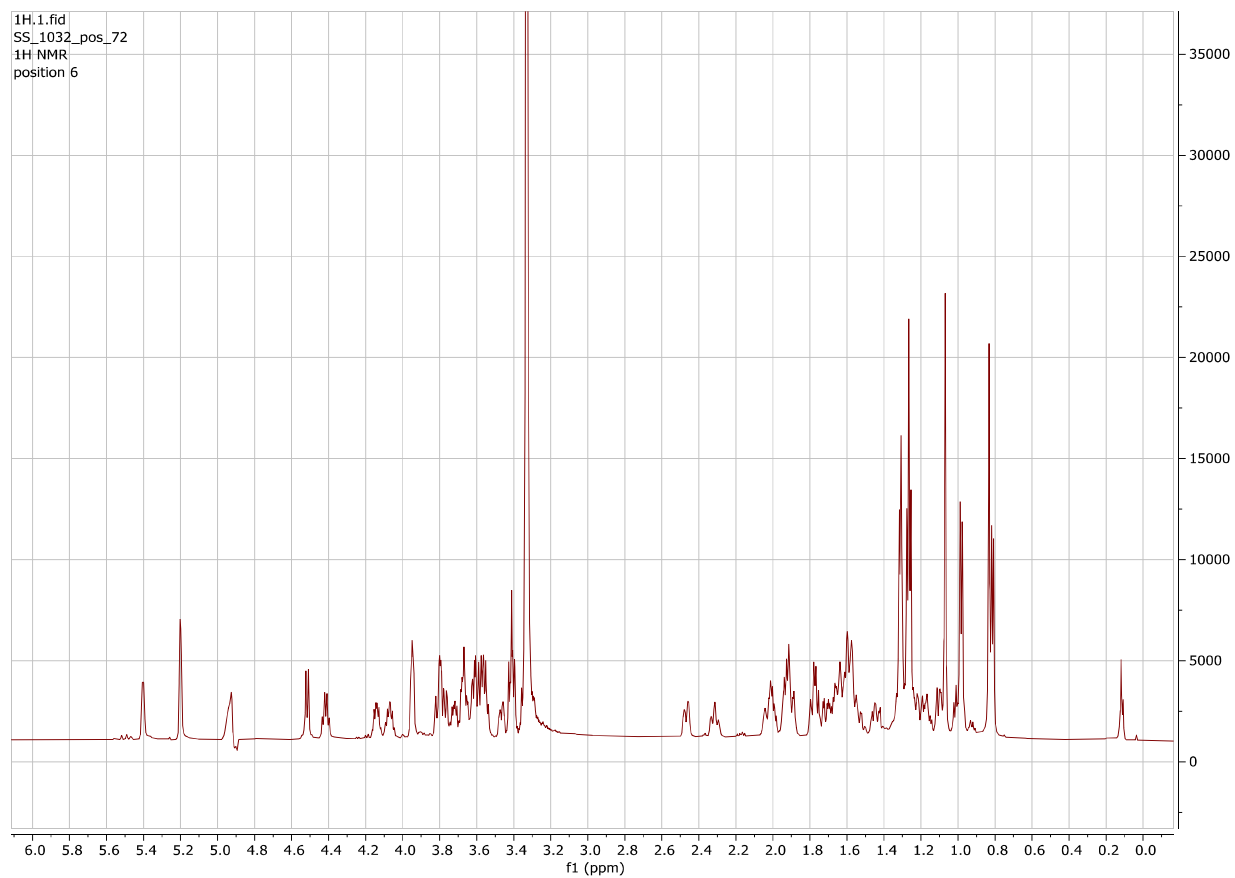


Figure S21. ¹H NMR spectrum for the intact molecule of the switchgrass saponin SS1032.

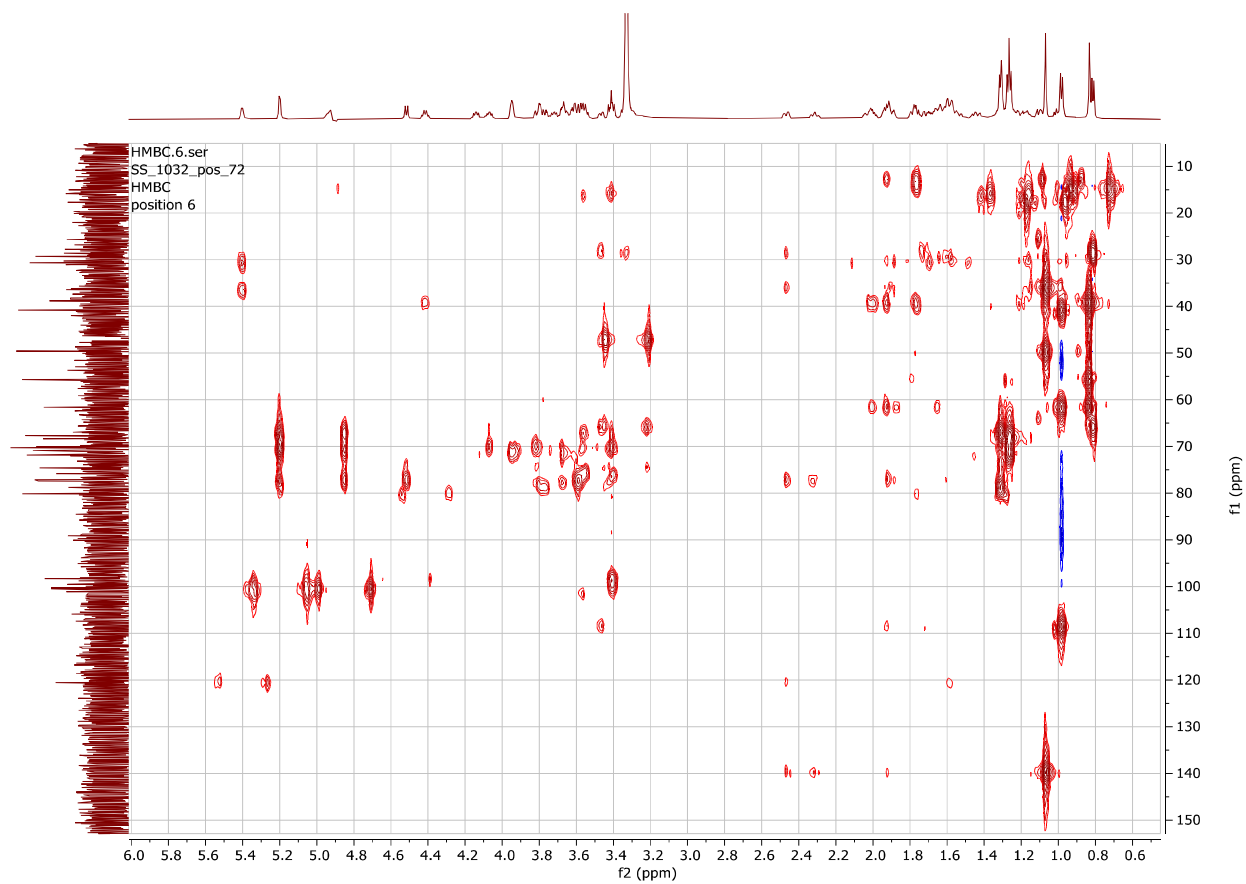


Figure S23. gHMBCAD NMR spectrum for the intact molecule of the switchgrass saponin SS1032.



Figure S24. gHSQCAD NMR spectrum for the intact molecule of the switchgrass saponin SS1032.

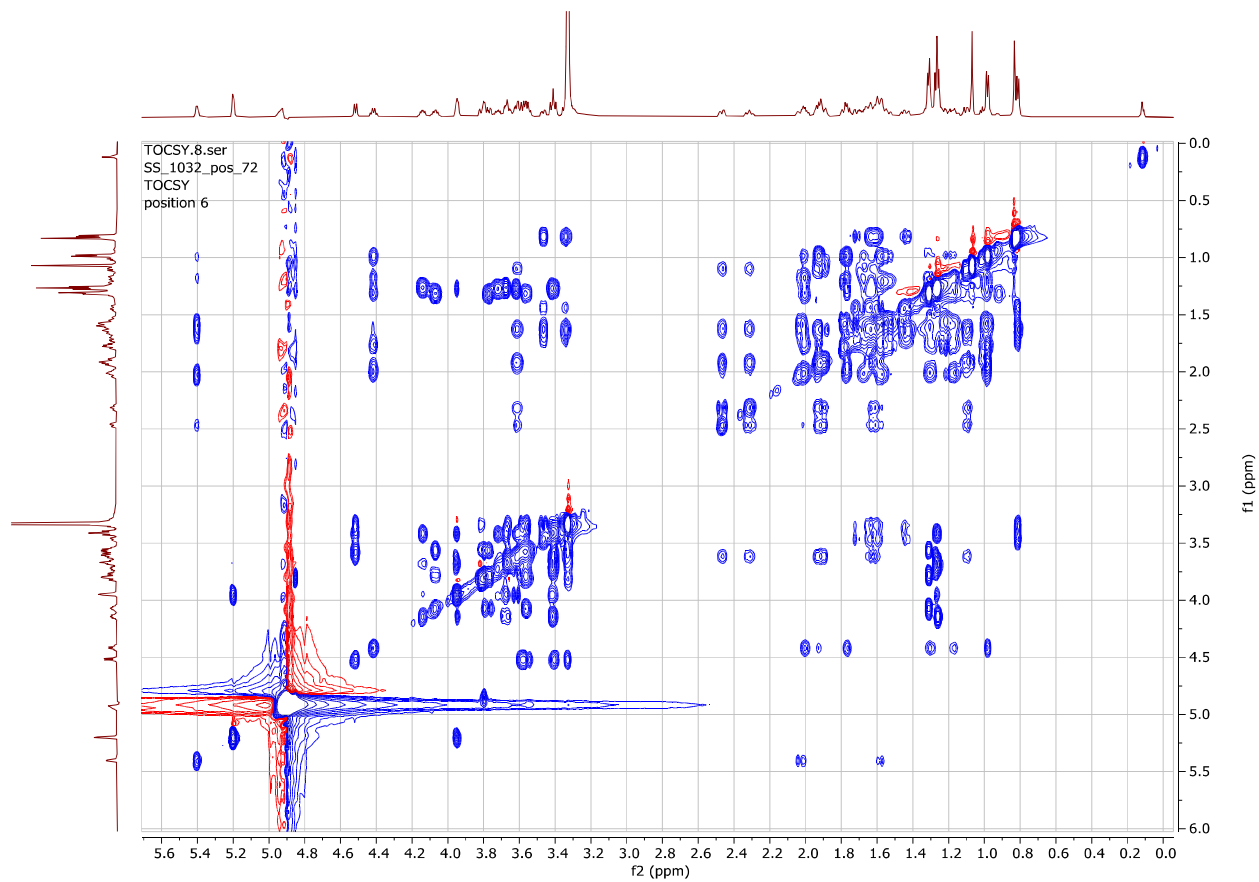


Figure S25. gTOCSY NMR spectrum for the intact molecule of the switchgrass saponin SS1032.

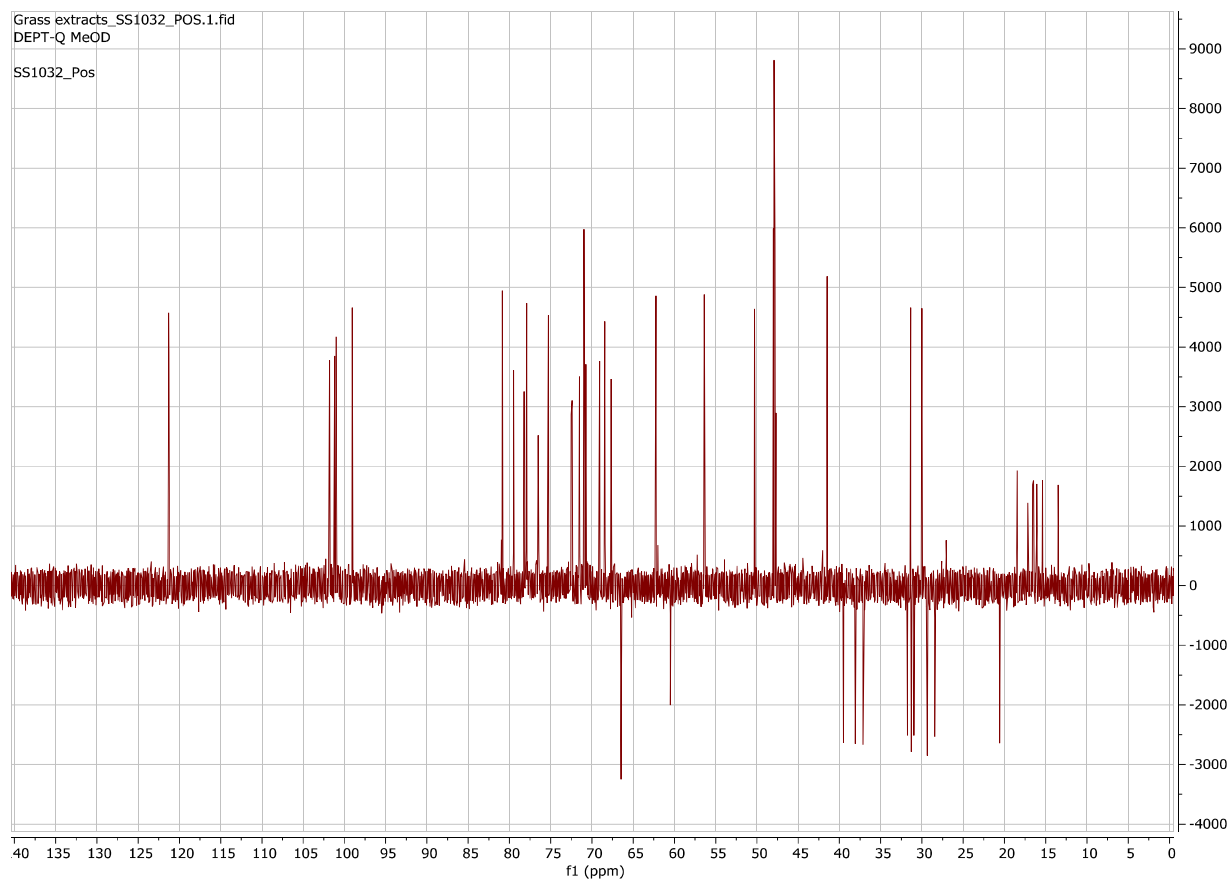


Figure S26. DEPT-Q (^{13}C) NMR spectrum for the intact molecule of the switchgrass saponin SS1032.

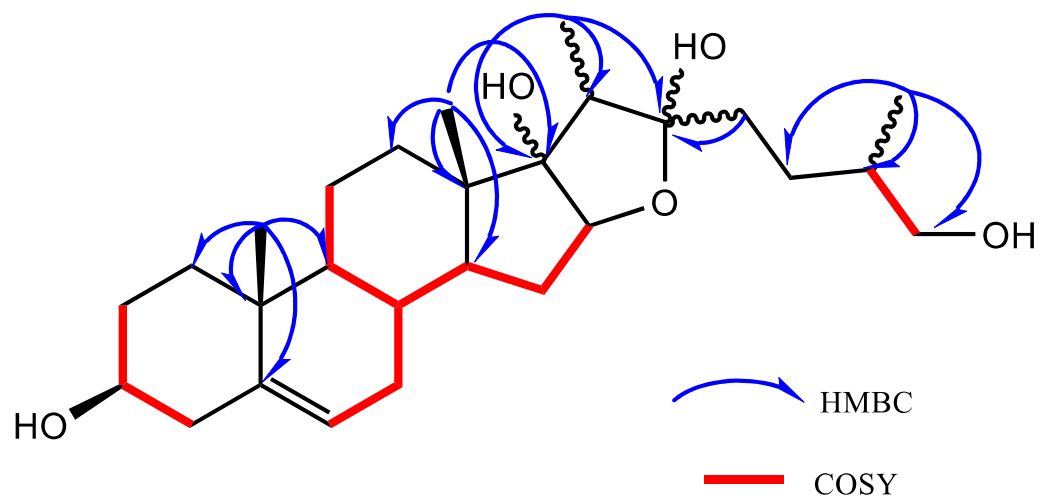


Figure S27. The key COSY (red, bold) and HMBC (blue arrow) correlations for the aglycone of the switchgrass saponin SS1244.

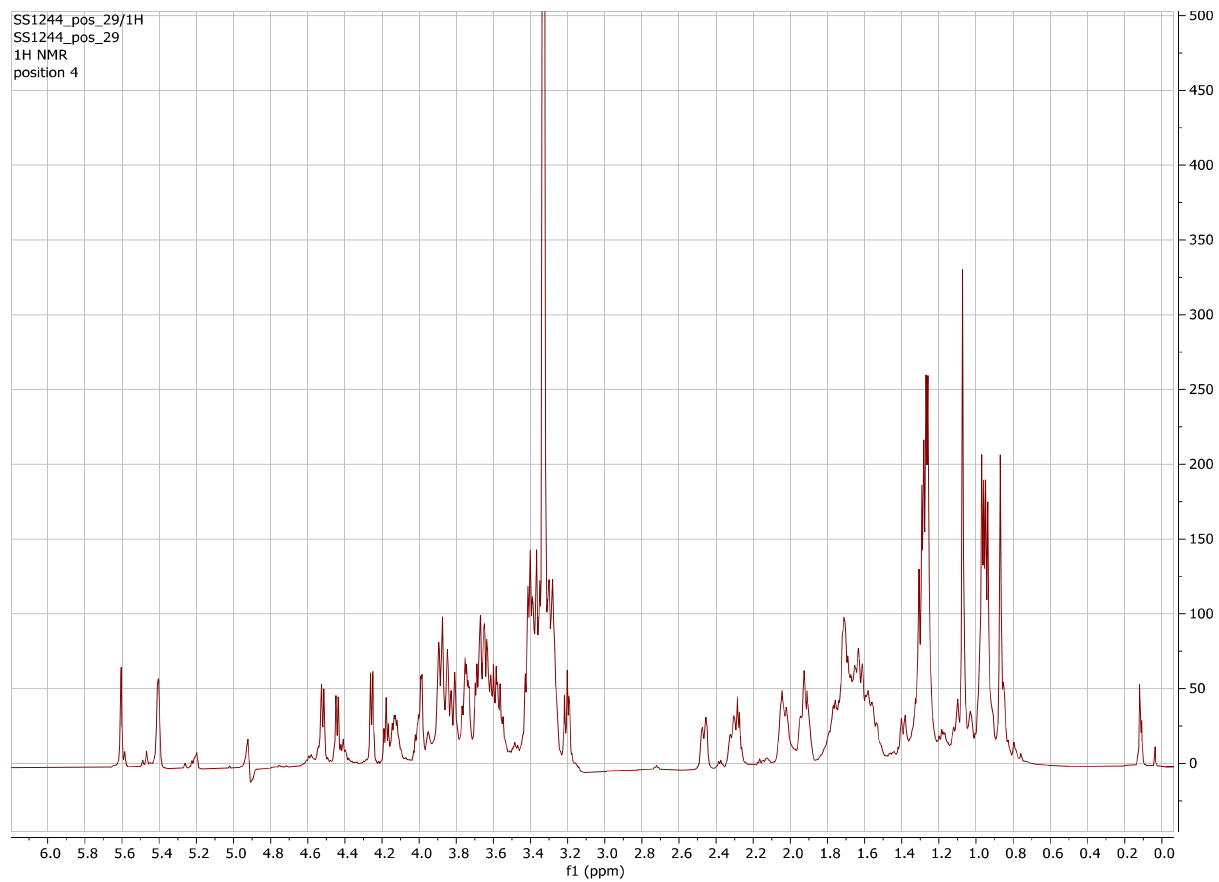


Figure S28. ¹H NMR spectrum for the intact molecule of the switchgrass saponin SS1244.



Figure S29. gCOSY ^1H - ^1H NMR spectrum for the intact molecule of the switchgrass saponin SS1244.

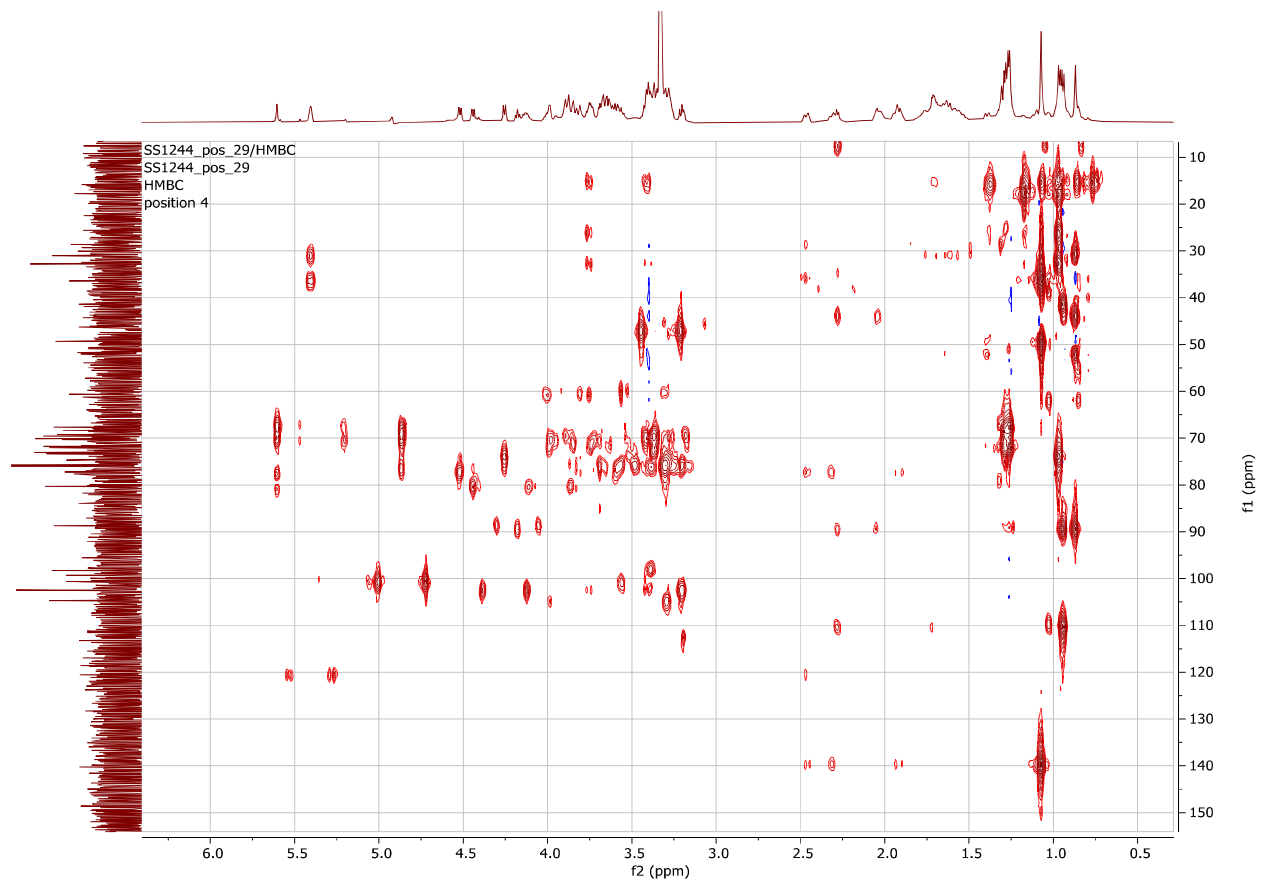


Figure S30. gHMBCAD NMR spectrum for the intact molecule of the switchgrass saponin SS1244.



Figure S31. gHSQCAD NMR spectrum for the intact molecule of the switchgrass saponin SS1244.

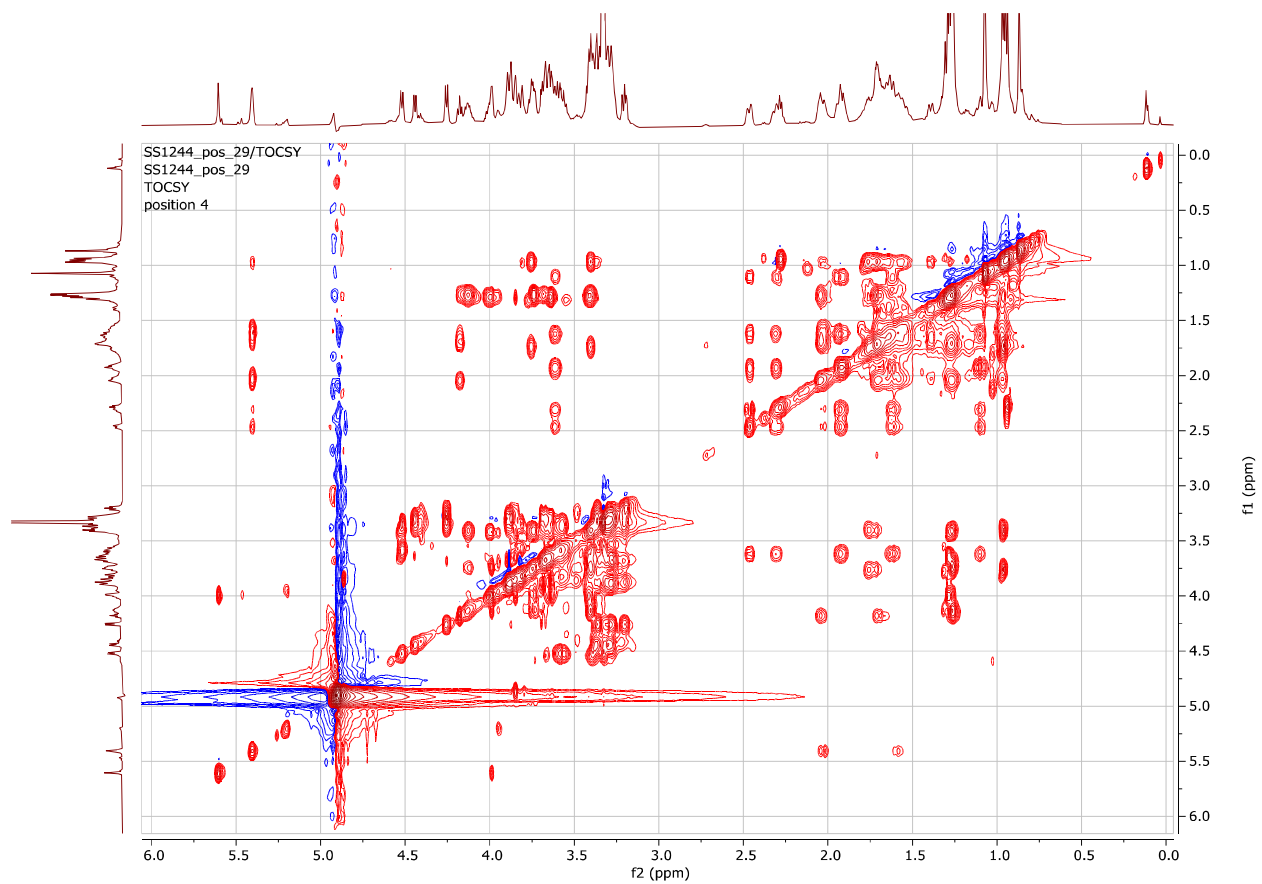


Figure S32. gTOCSY NMR spectrum for the intact molecule of the switchgrass saponin SS1244.

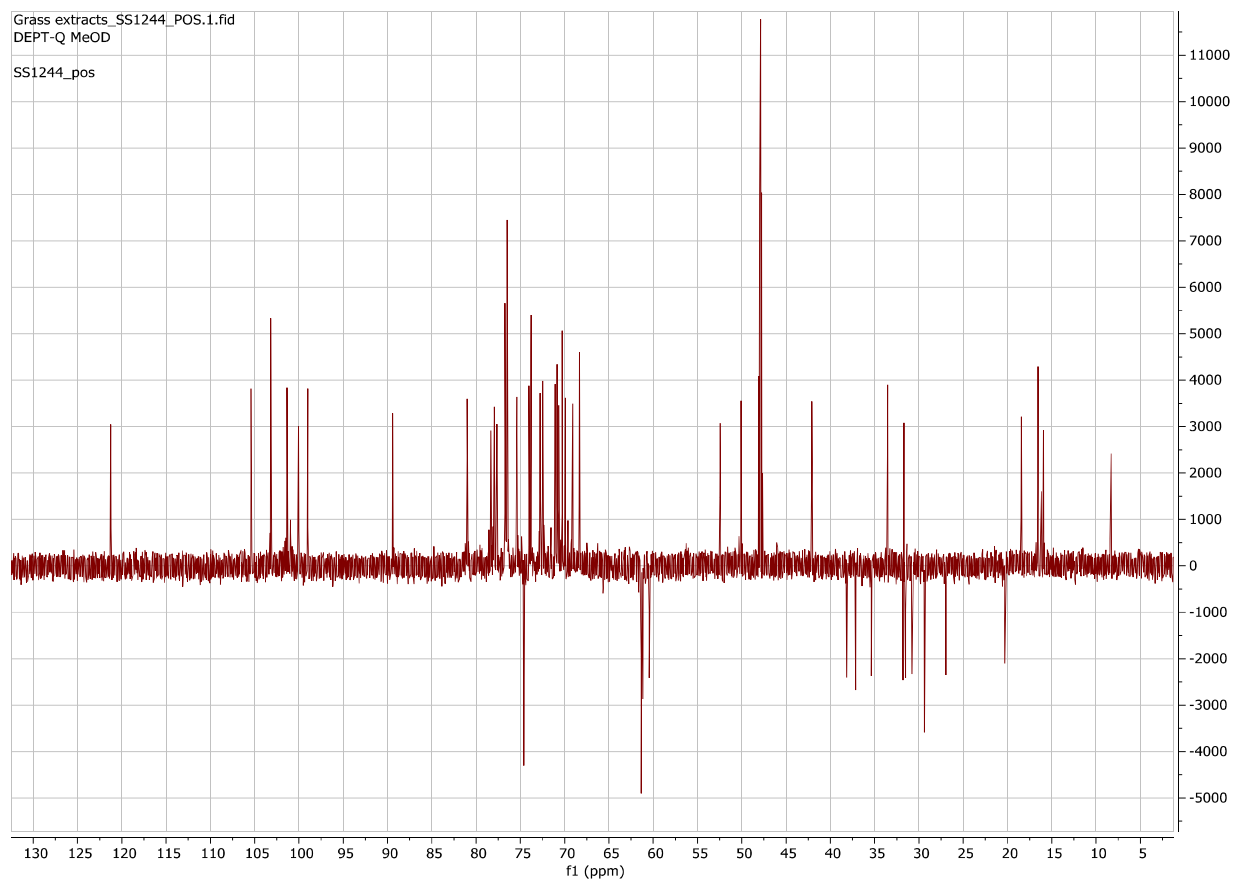


Figure S33. DEPT-Q (^{13}C) NMR spectrum for the intact molecule of the switchgrass saponin SS1244.

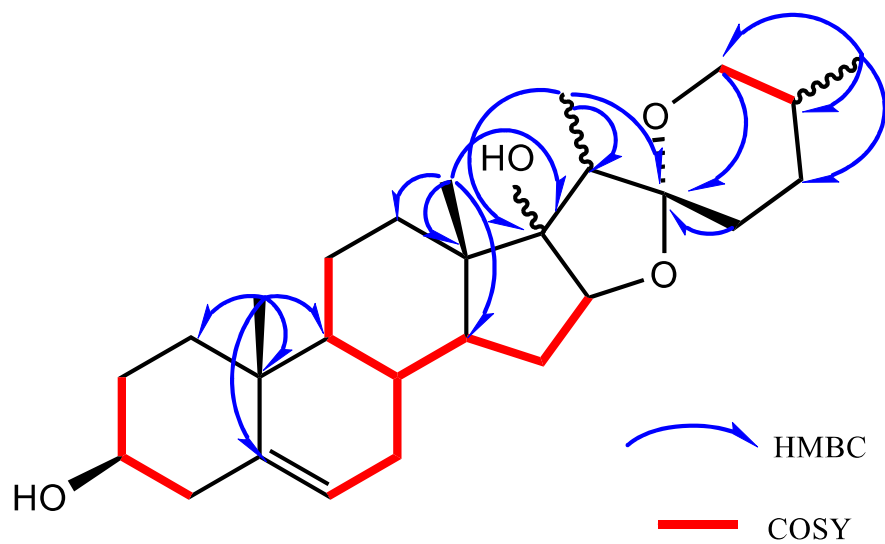


Figure S34. The key COSY (red, bold) and HMBC (blue arrow) correlations for the aglycone of the switchgrass saponin SS1064.

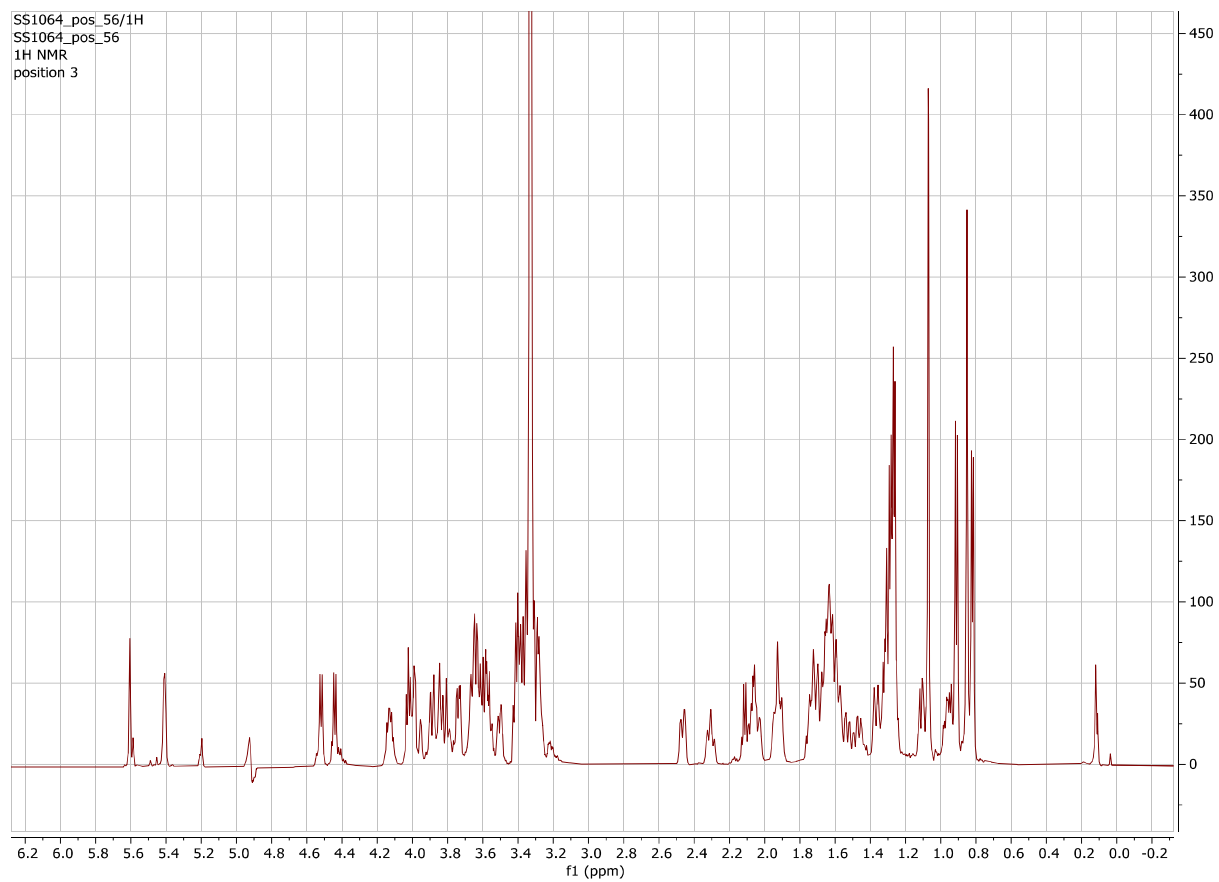


Figure S35. ^1H NMR spectrum for the intact molecule of the switchgrass saponin SS1064.

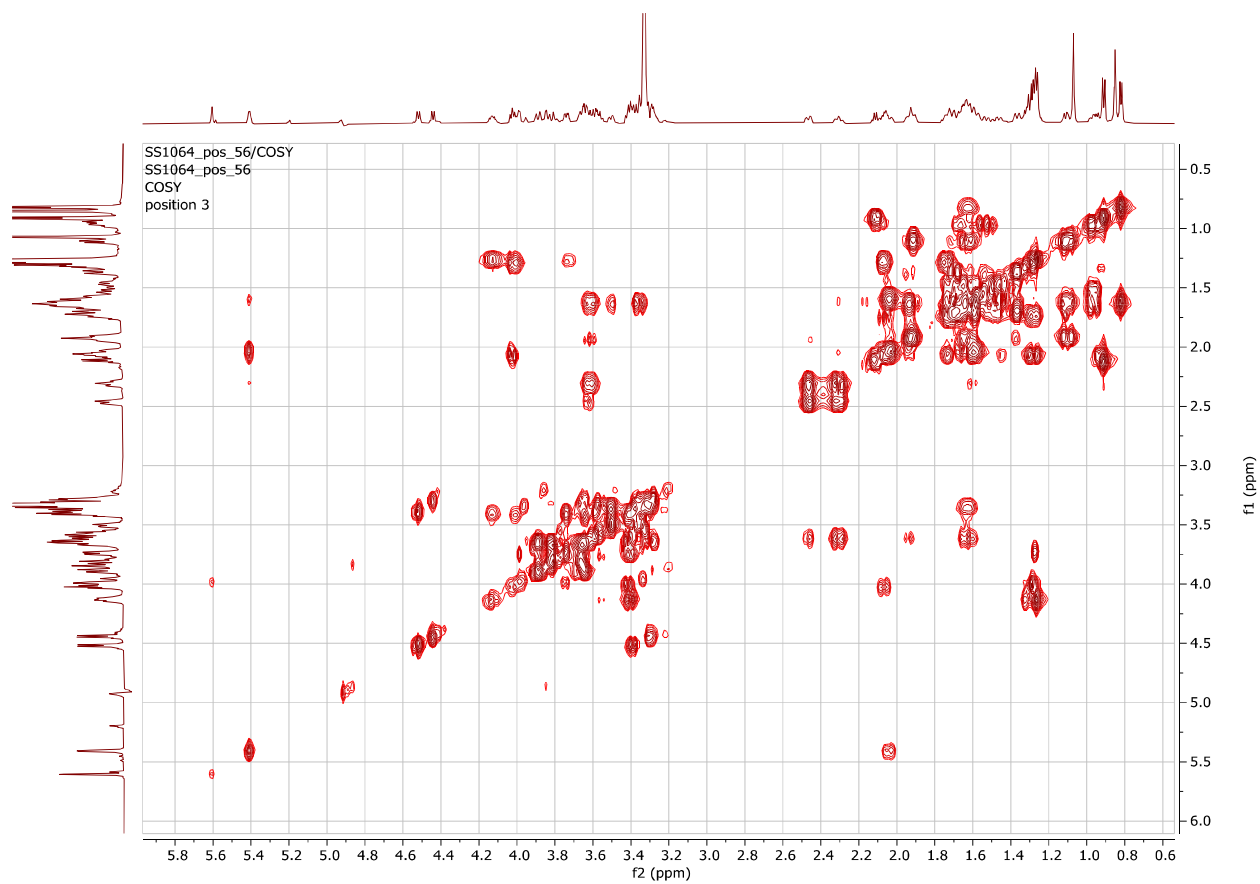


Figure S36. gCOSY NMR spectrum for the intact molecule of the switchgrass saponin SS1064.



Figure S37. gHMBCAD NMR spectrum for the intact molecule of the switchgrass saponin SS1064.

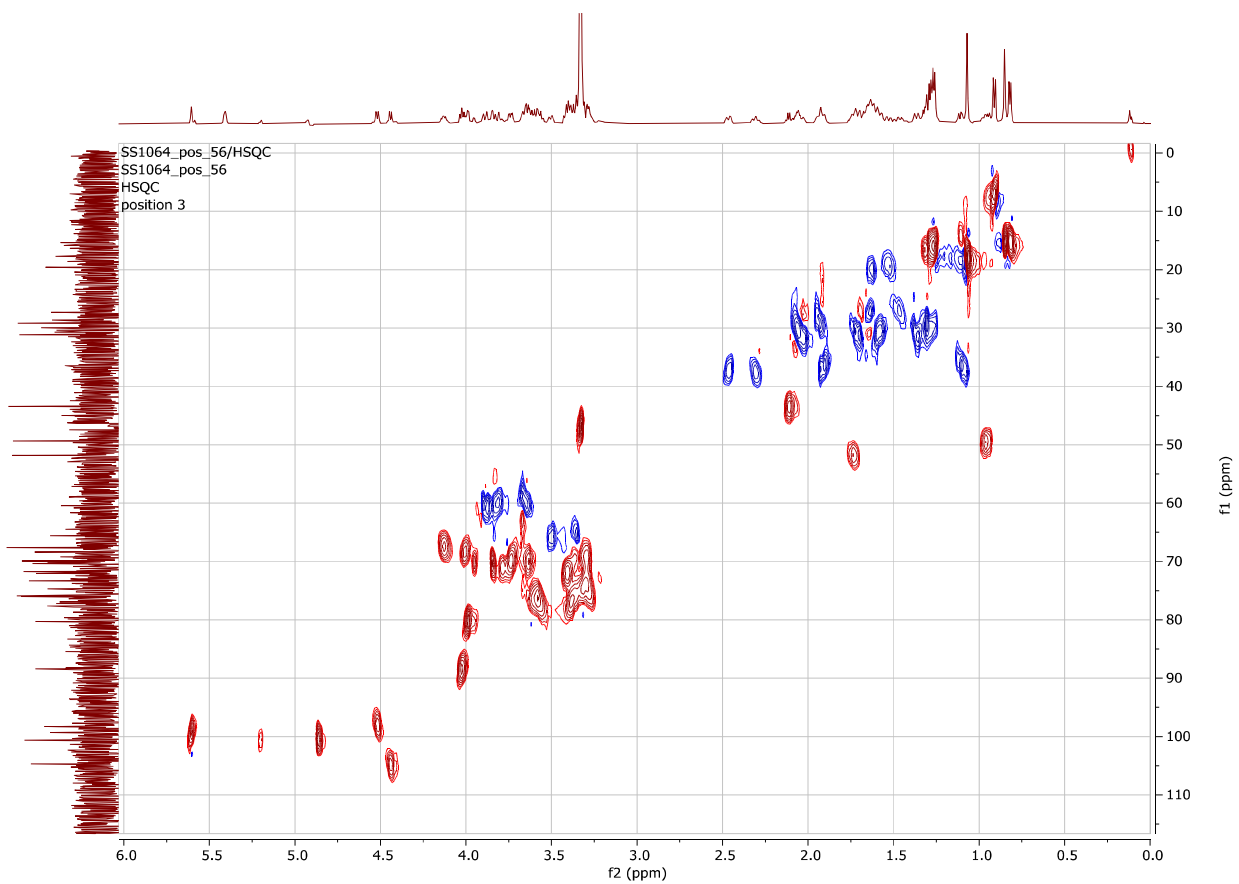


Figure S38. gHSQCAD NMR spectrum for the intact molecule of the switchgrass saponin SS1064.



Figure S39. gTOCSY NMR spectrum for the intact molecule of the switchgrass saponin SS1064.

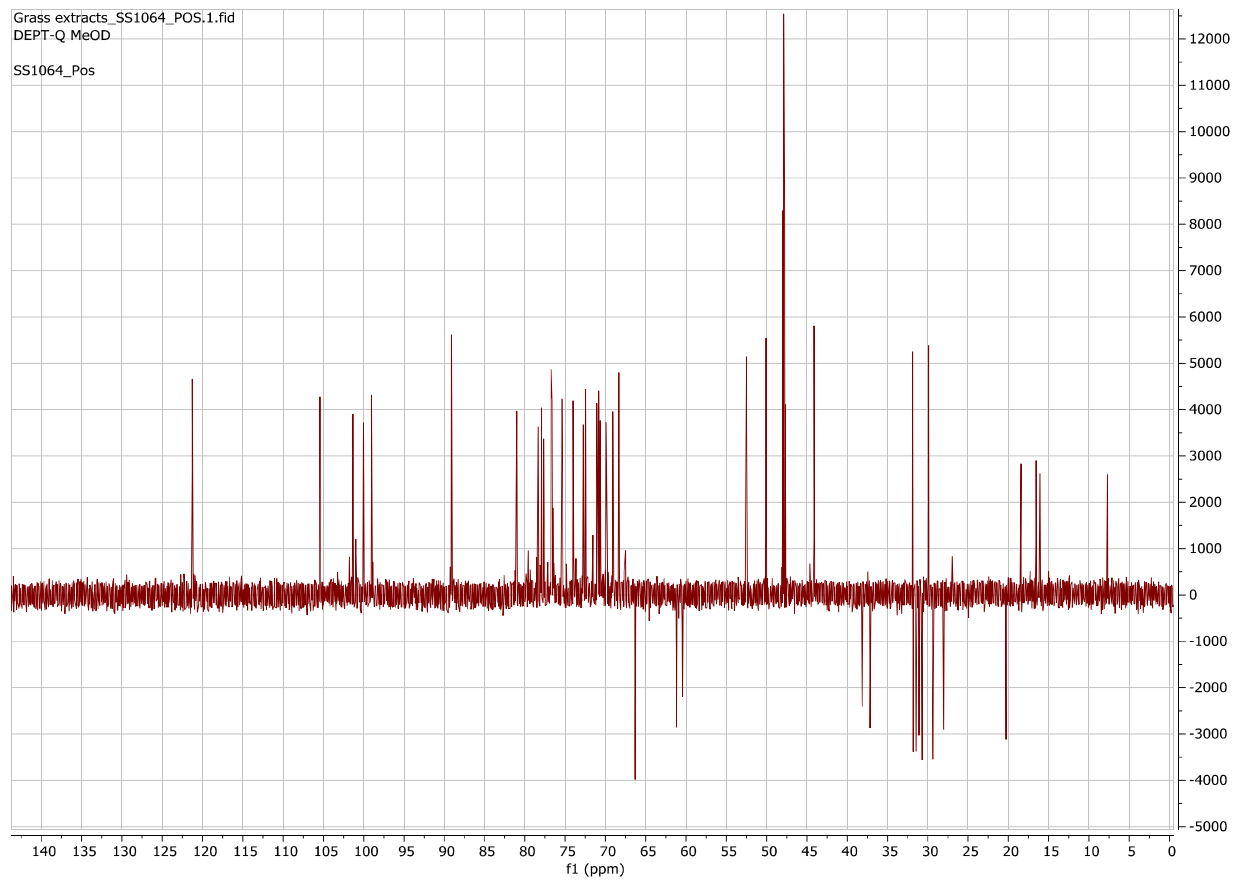


Figure S40. DEPT-Q (^{13}C) NMR spectrum for the intact molecule of the switchgrass saponin SS1064.

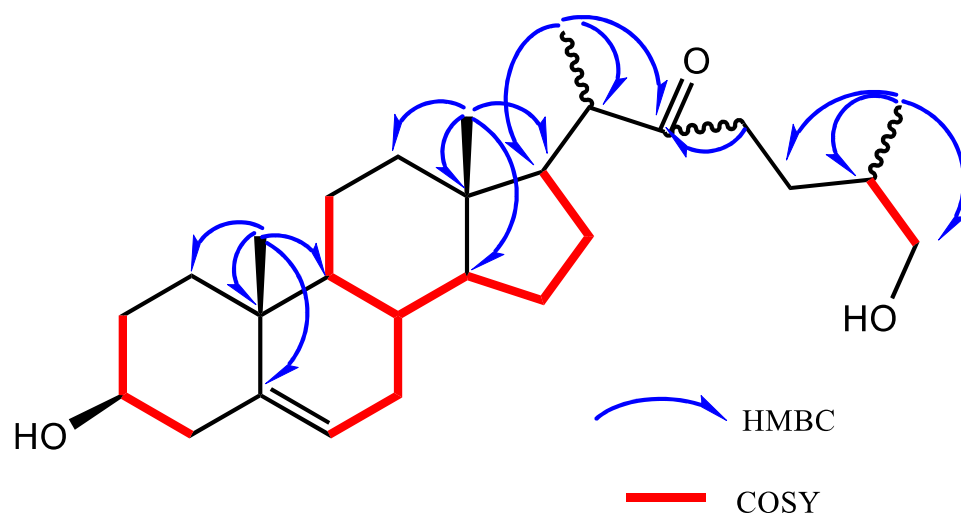


Figure S41. The key COSY (red, bold) and HMBC (blue arrow) correlations for the aglycone of the switchgrass saponin SS1050.

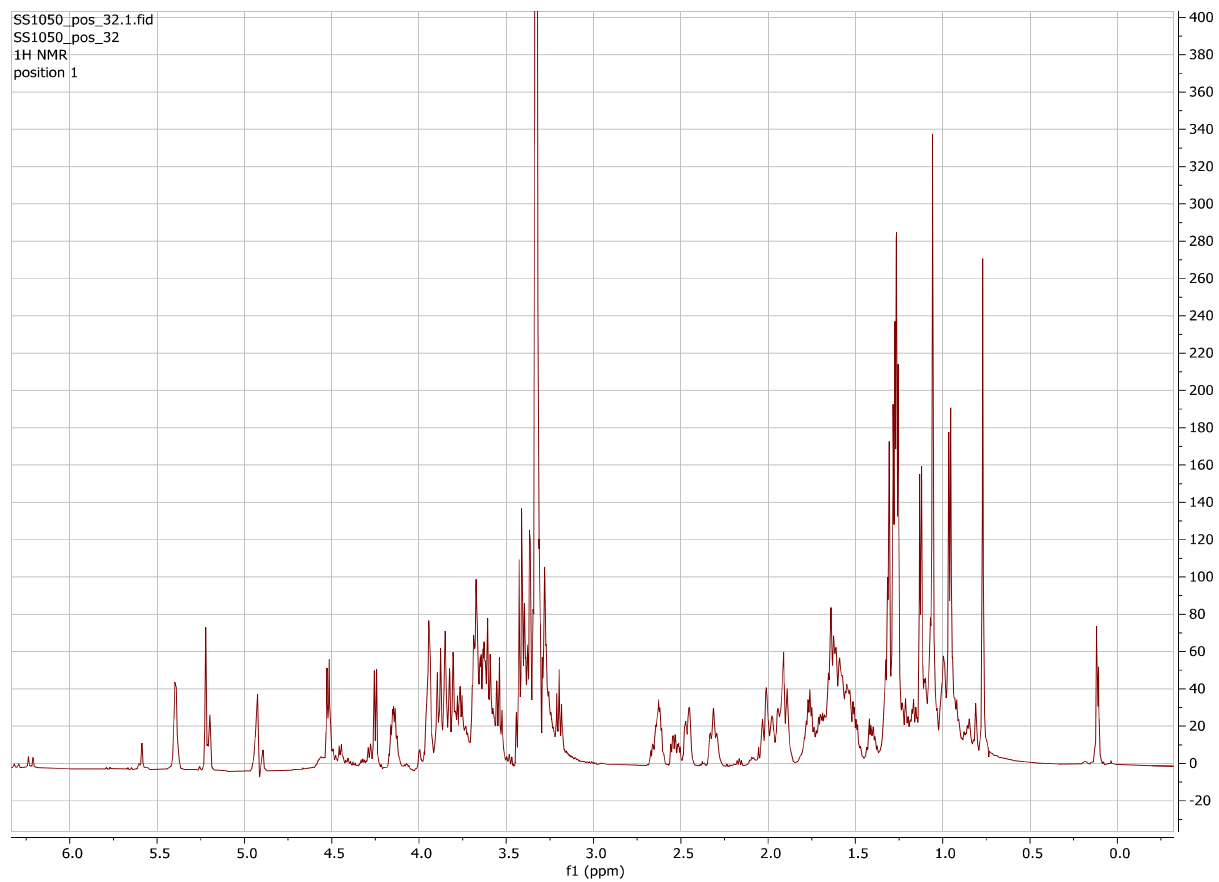


Figure S42. ¹H NMR spectrum for the intact molecule of the switchgrass saponin SS1050.

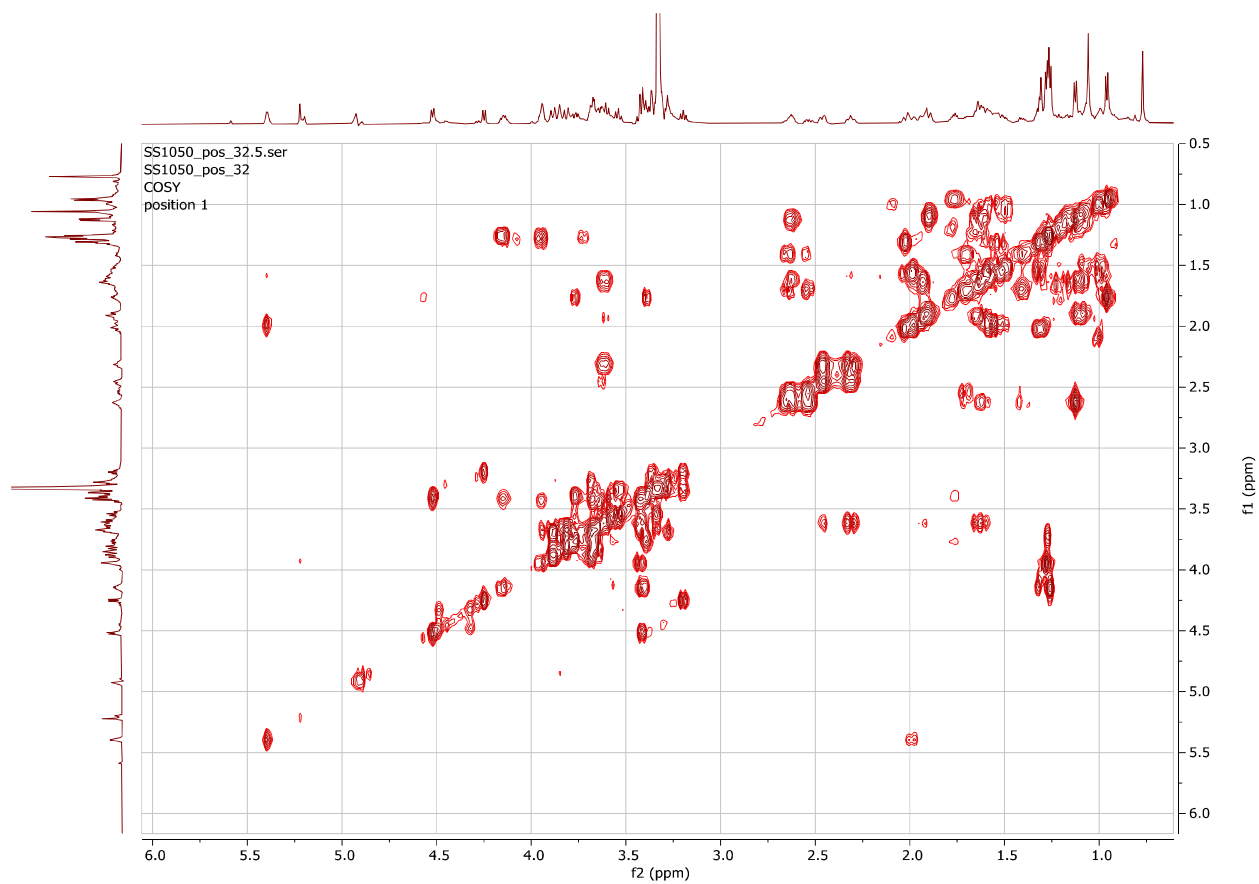


Figure S43. gCOSY NMR spectrum for the intact molecule of the switchgrass saponin SS1050.



Figure S44. gHMBCAD NMR spectrum for the intact molecule of the switchgrass saponin SS1050.



Figure S45. gHSQCAD NMR spectrum for the intact molecule of the switchgrass saponin SS1050.

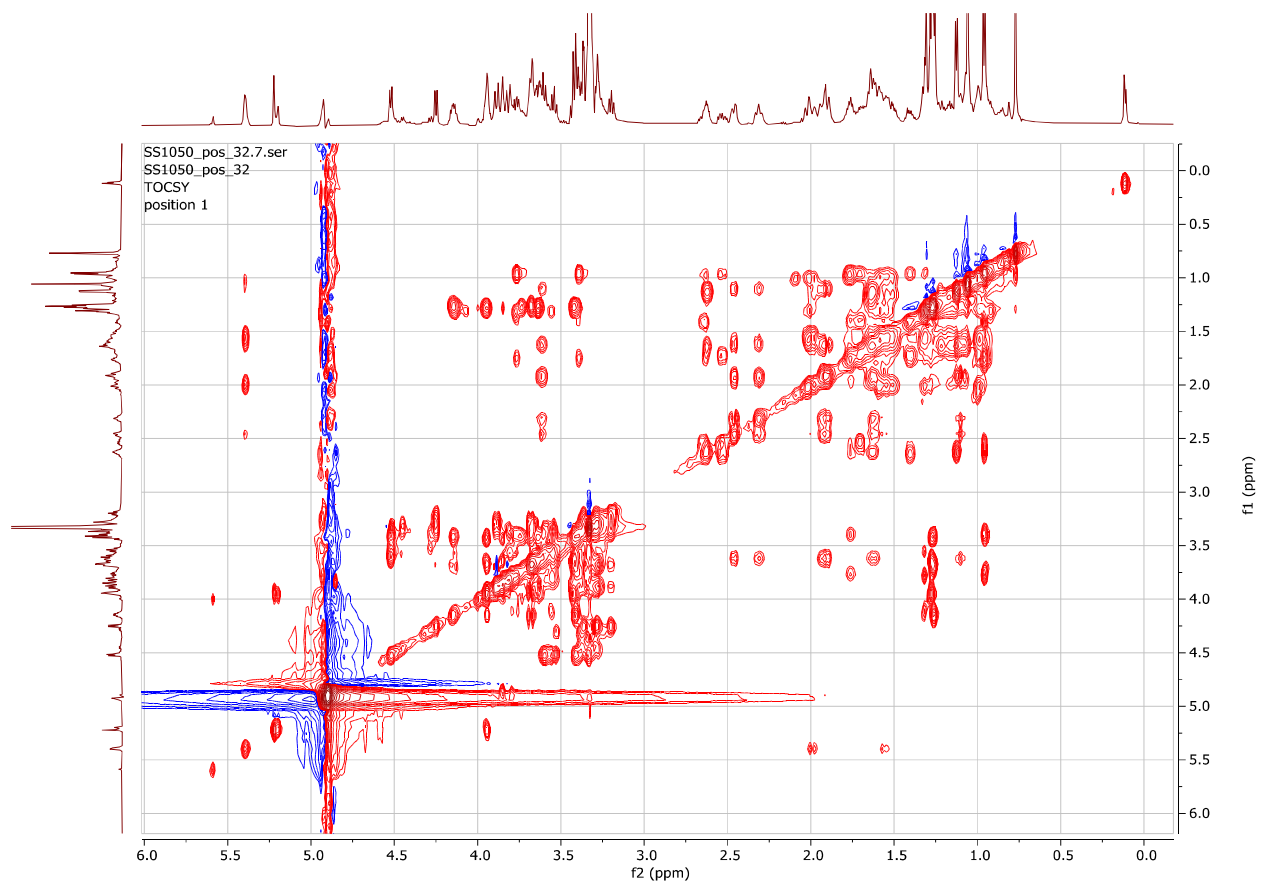


Figure S46. gTOCSY NMR spectrum for the intact molecule of the switchgrass saponin SS1050.

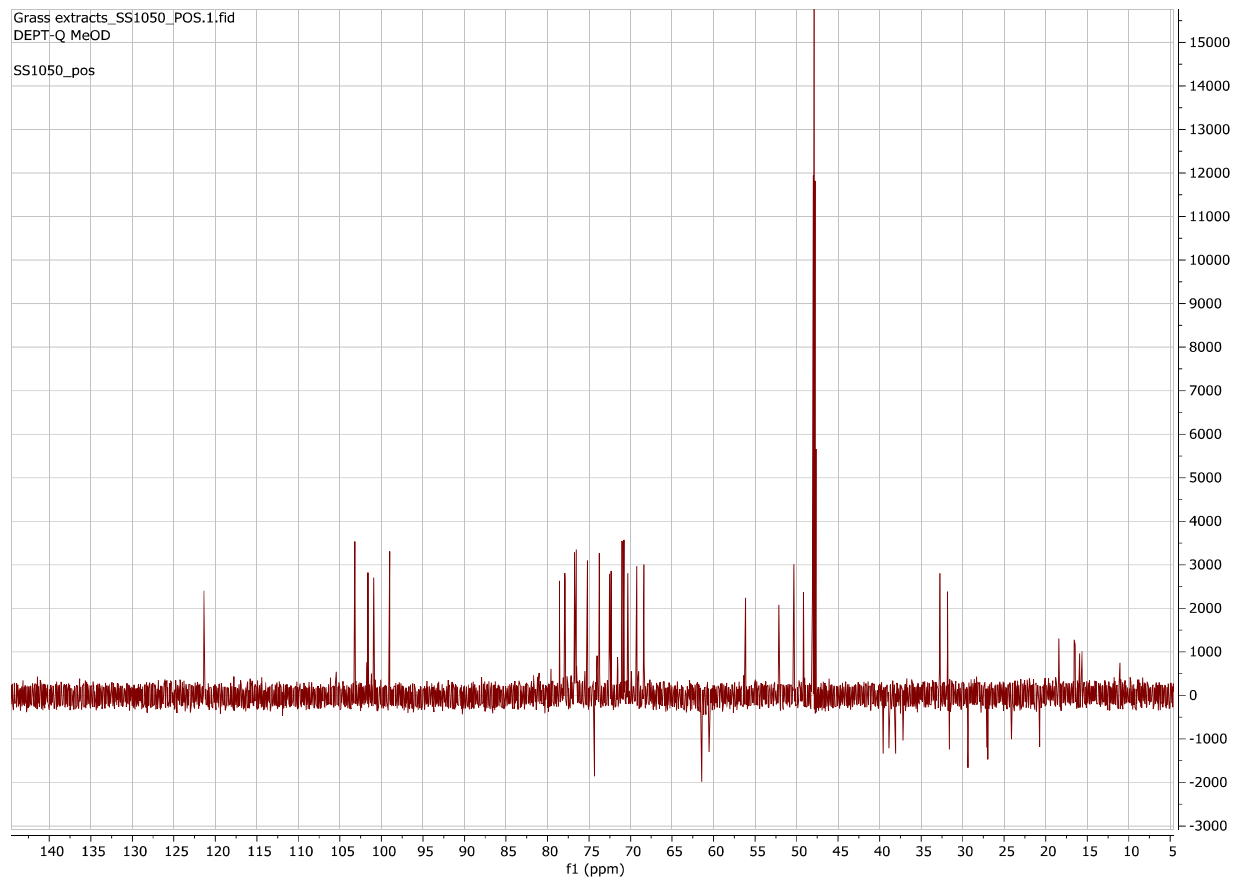


Figure S47. DEPT-Q (^{13}C) NMR spectrum for the intact molecule of the switchgrass saponin SS1050.

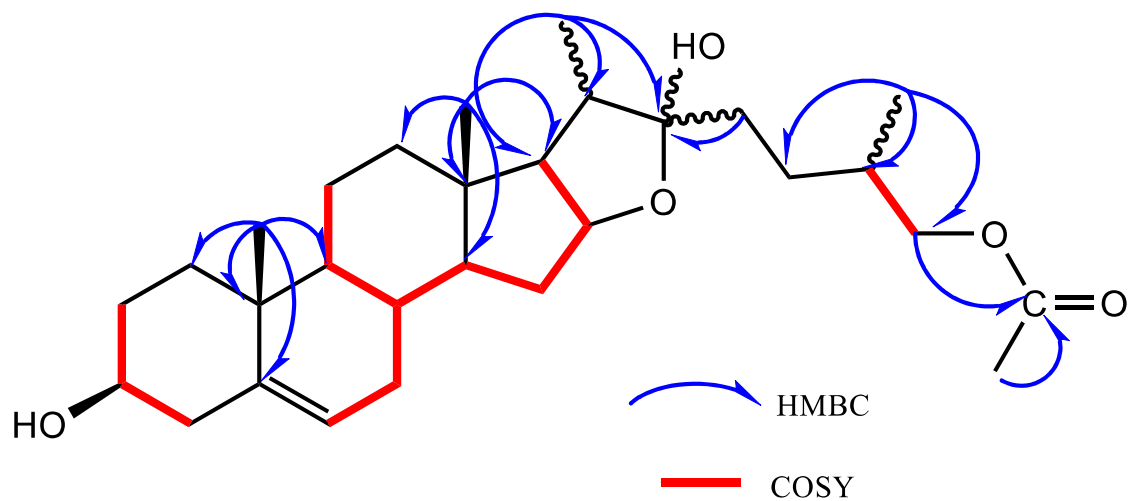


Figure S48. The key COSY (red, bold) and HMBC (blue arrow) correlations for the aglycone of the switchgrass saponin SS1254.

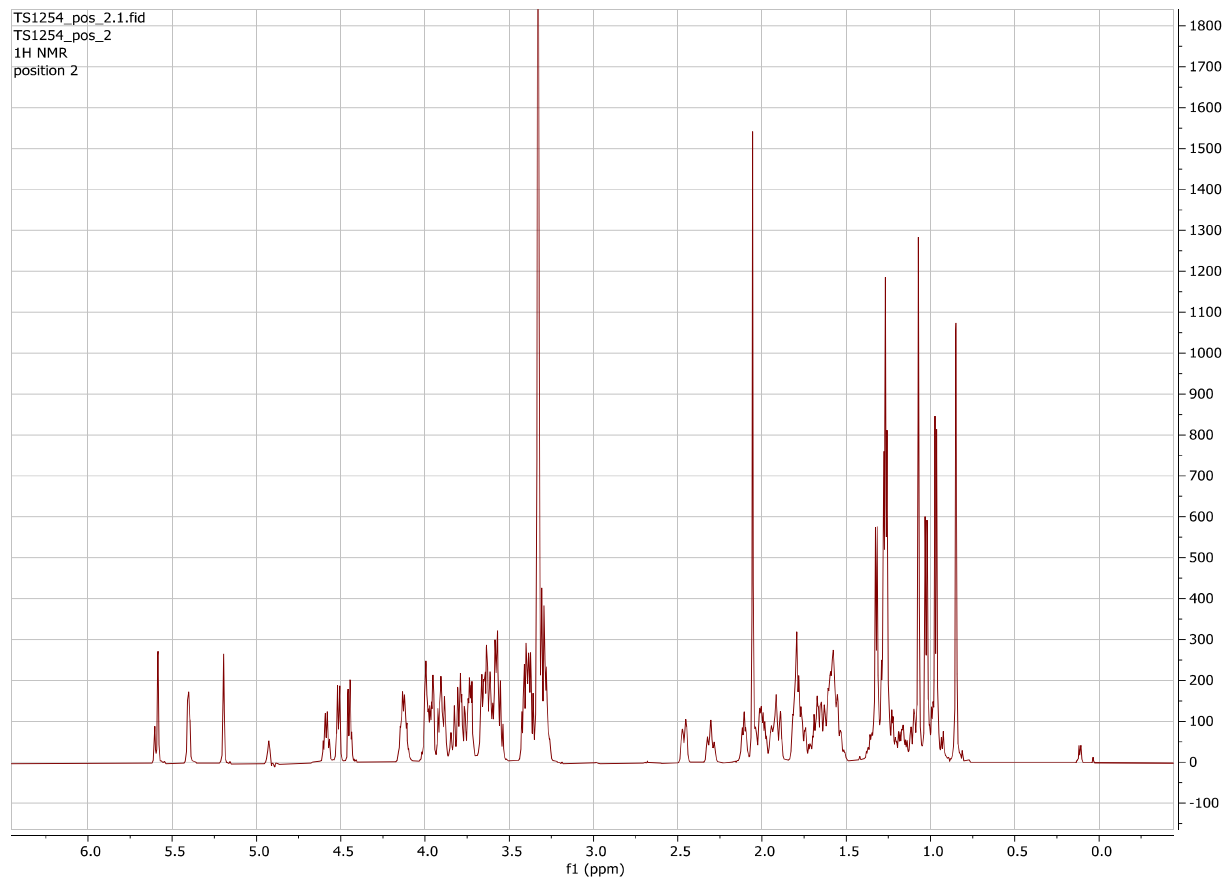


Figure S49. ¹H NMR spectrum for the intact molecule of the switchgrass saponin SS1254.

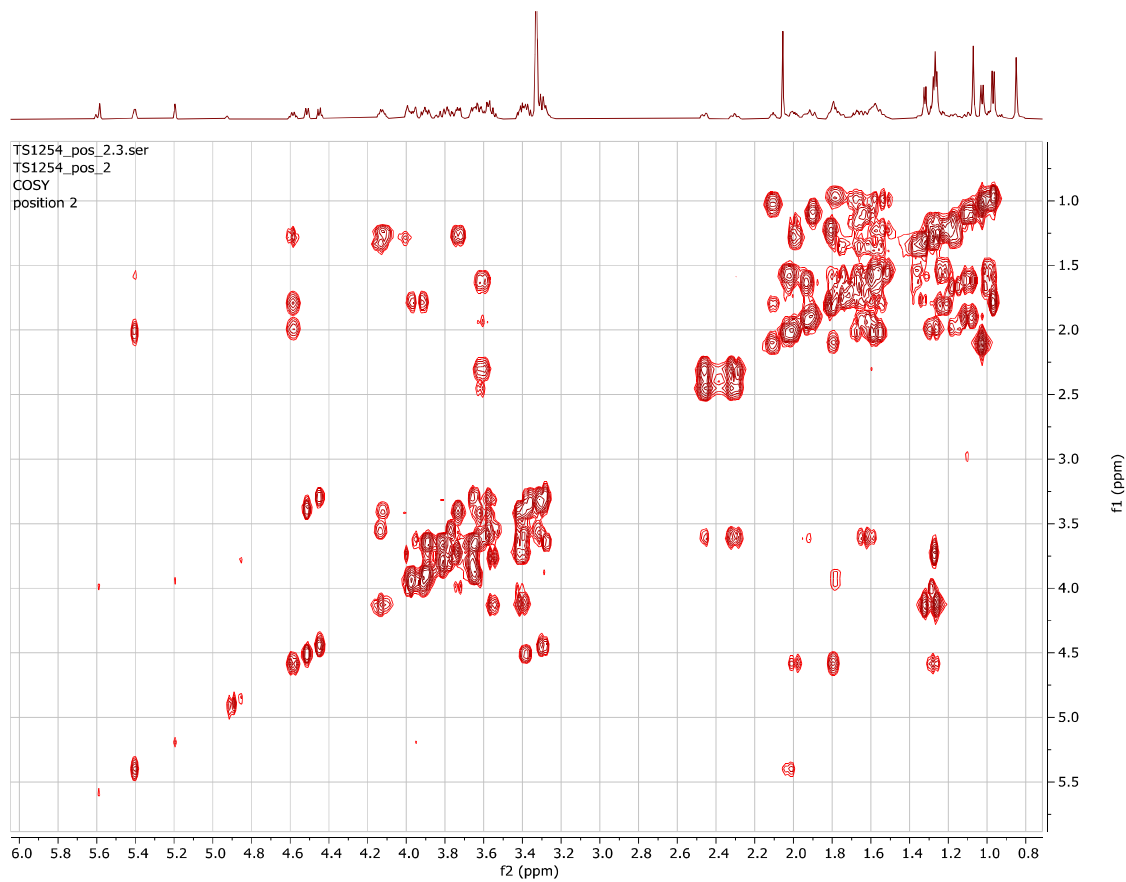


Figure S50. gCOSY NMR spectrum for the intact molecule of the switchgrass saponin SS1254.

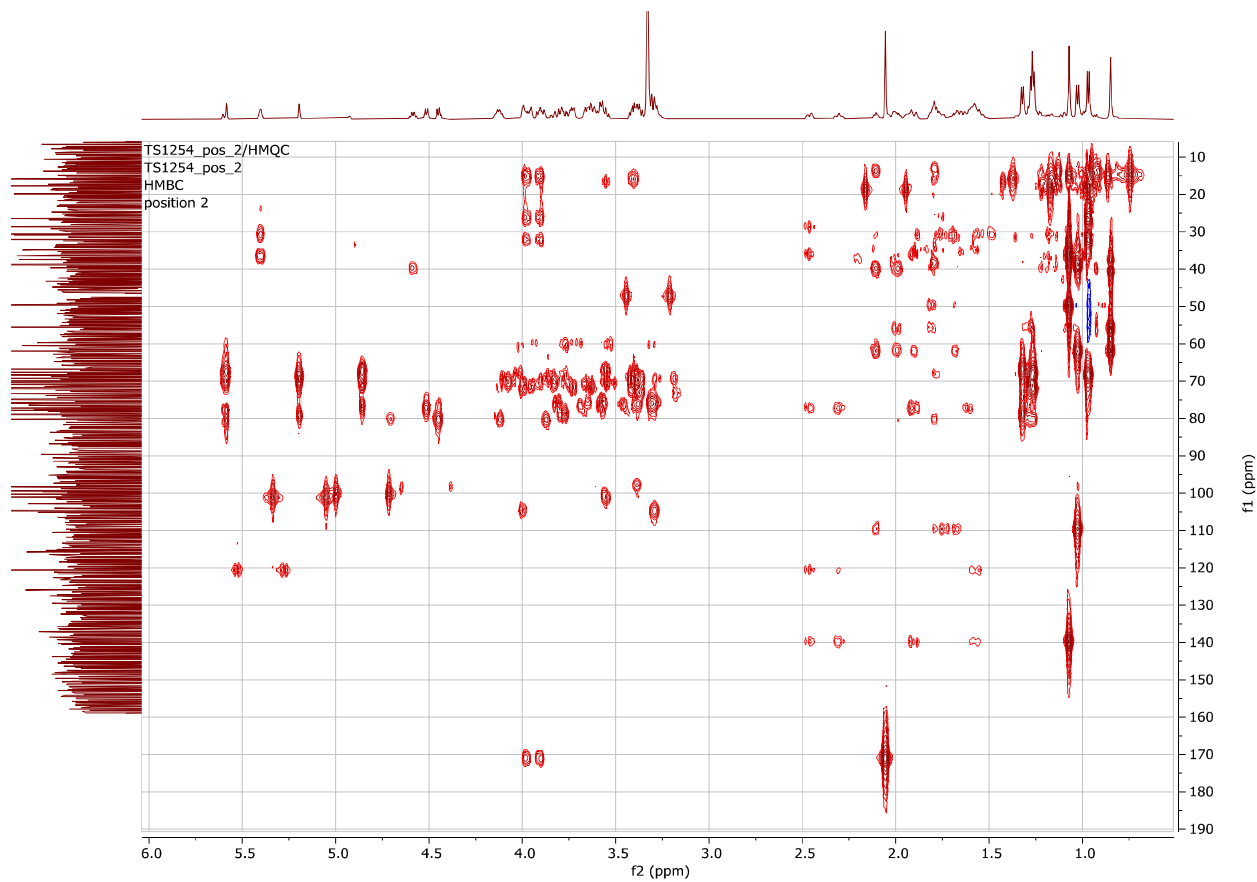


Figure S51. gHMBCAD NMR spectrum for the intact molecule of the switchgrass saponin SS1254.



Figure S52. gHSQCAD NMR spectrum for the intact molecule of the switchgrass saponin SS1254.

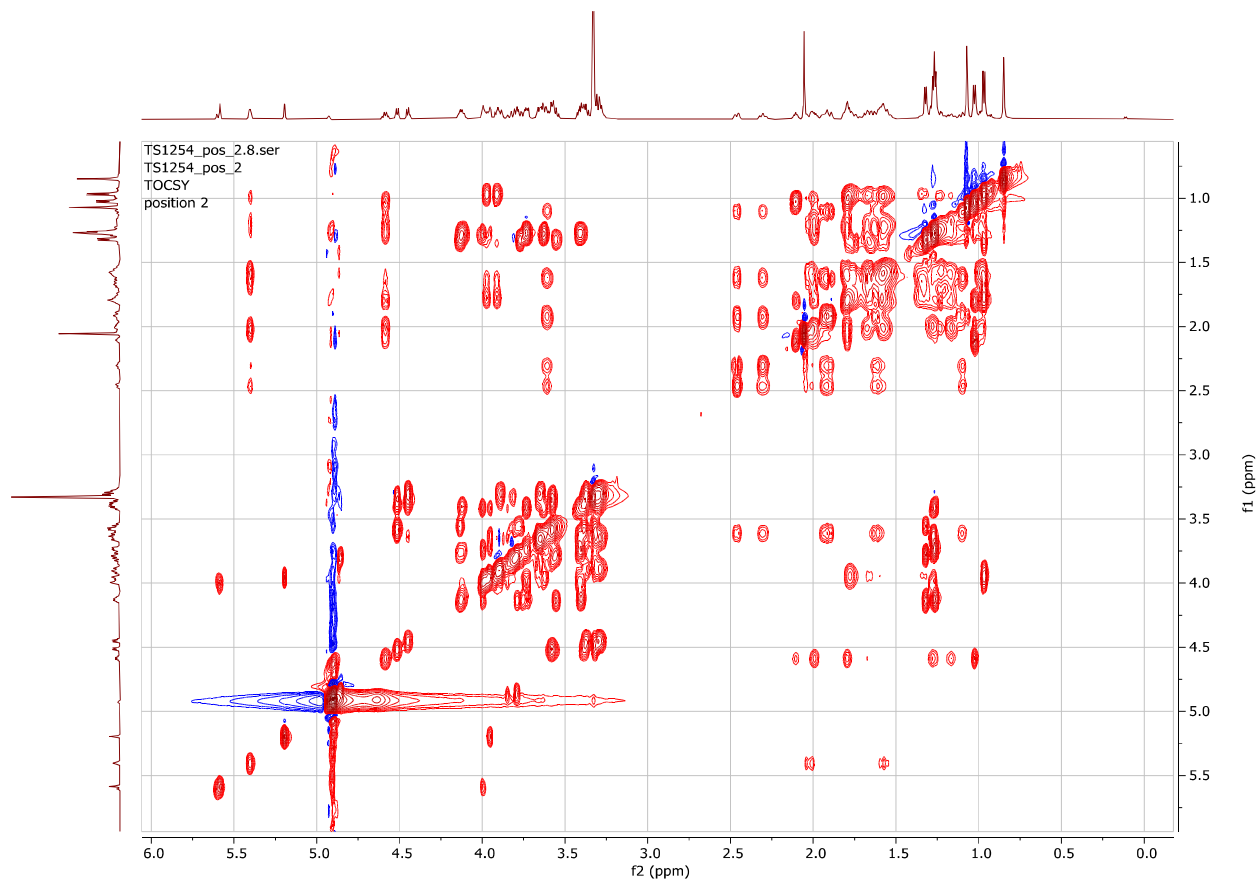


Figure S53. gTOCSY NMR spectrum for the intact molecule of the switchgrass saponin SS1254.

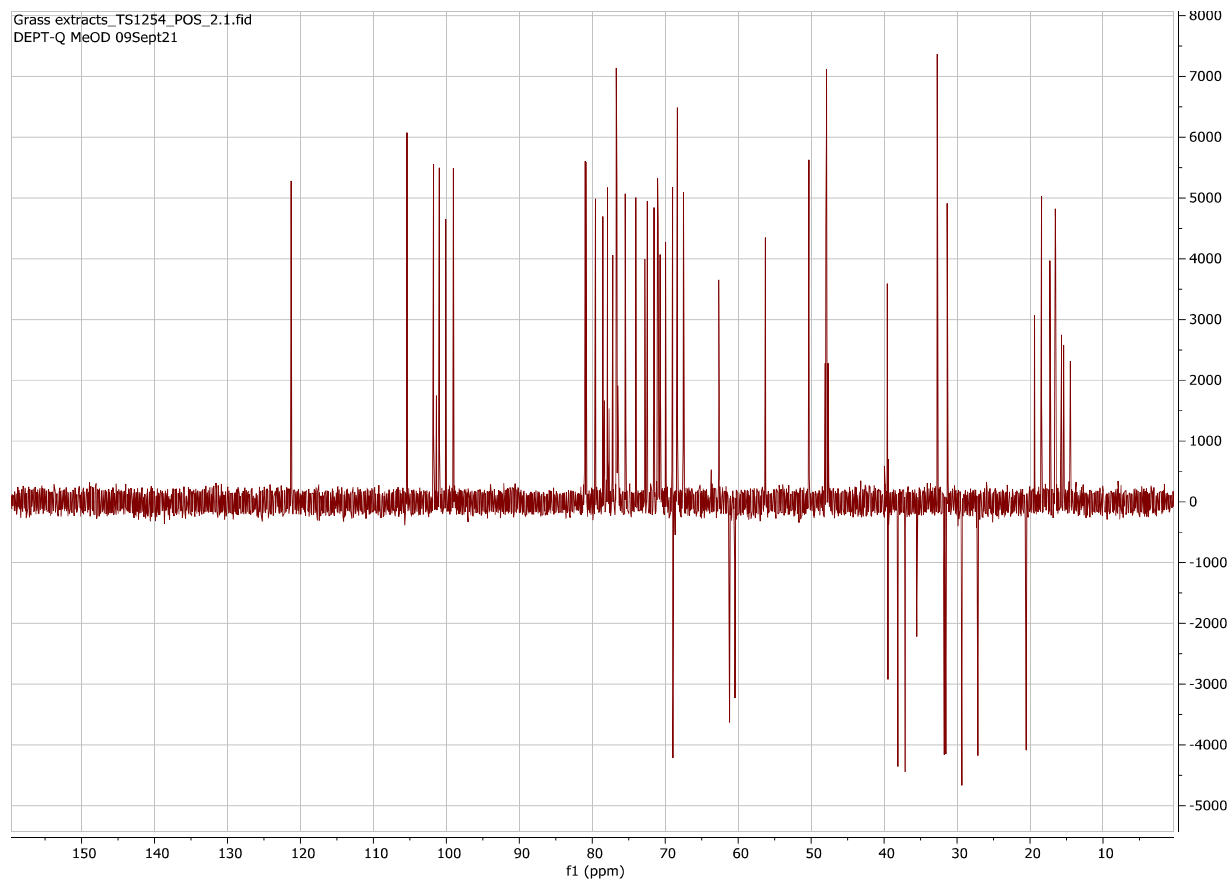


Figure S54. DEPT-Q (^{13}C) NMR spectrum for the intact molecule of the switchgrass saponin SS1254.

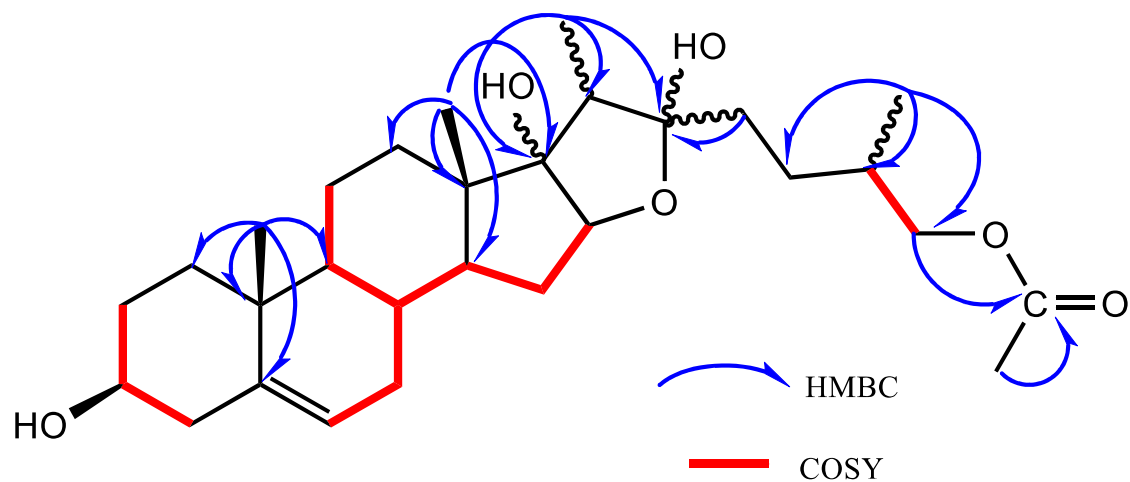


Figure S55. The key COSY (red, bold) and HMBC (blue arrow) correlations for the aglycone of the switchgrass saponin SS1089.

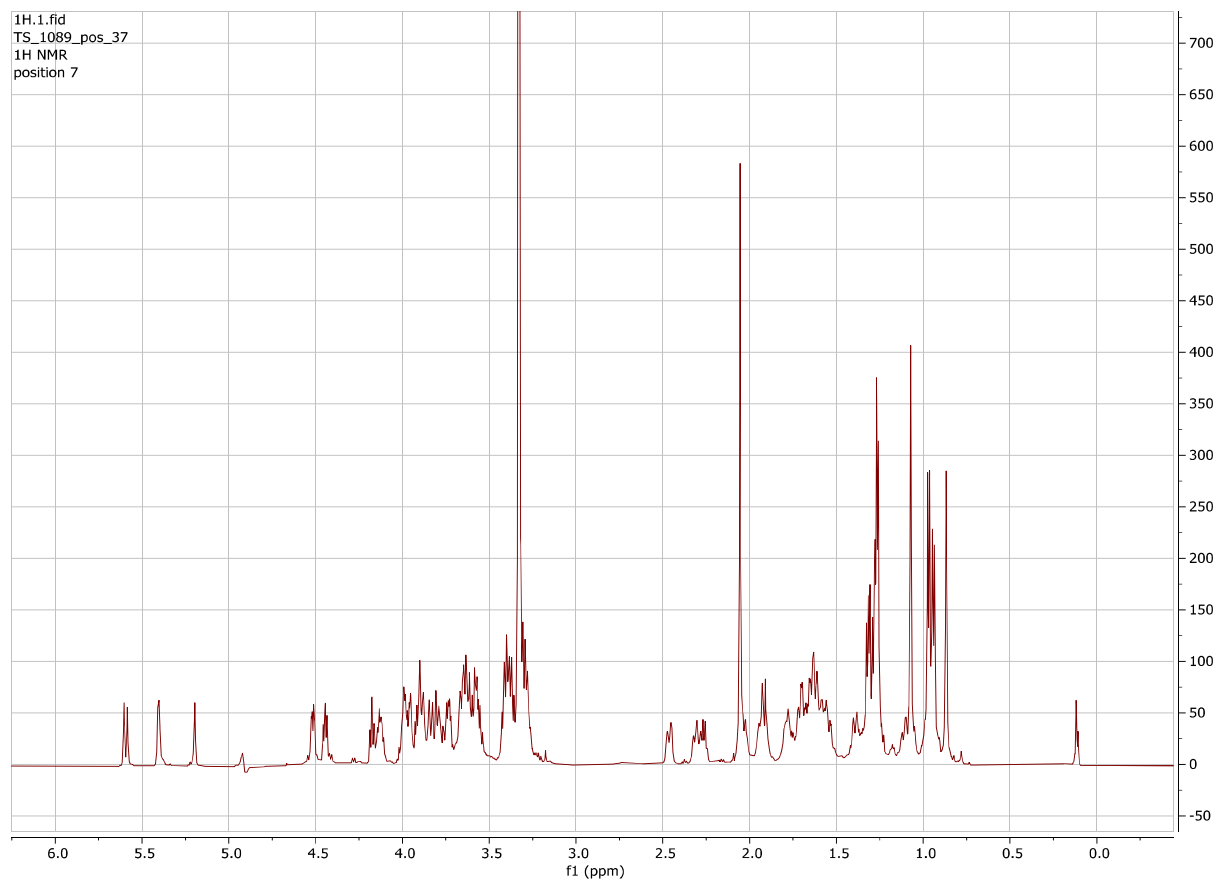


Figure S56. ¹H NMR spectrum for the intact molecule of the switchgrass saponin SS1089.

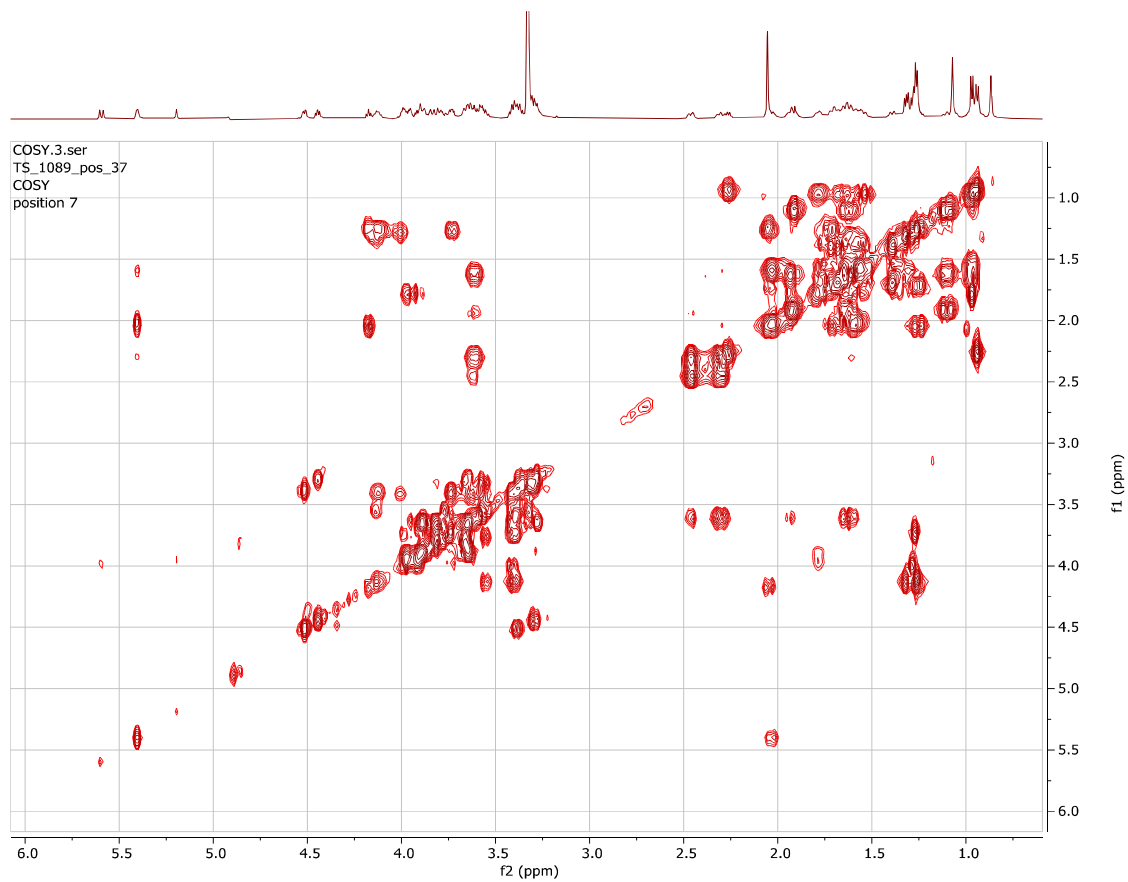


Figure S57. gCOSY NMR spectrum for the intact molecule of the switchgrass saponin SS1089.

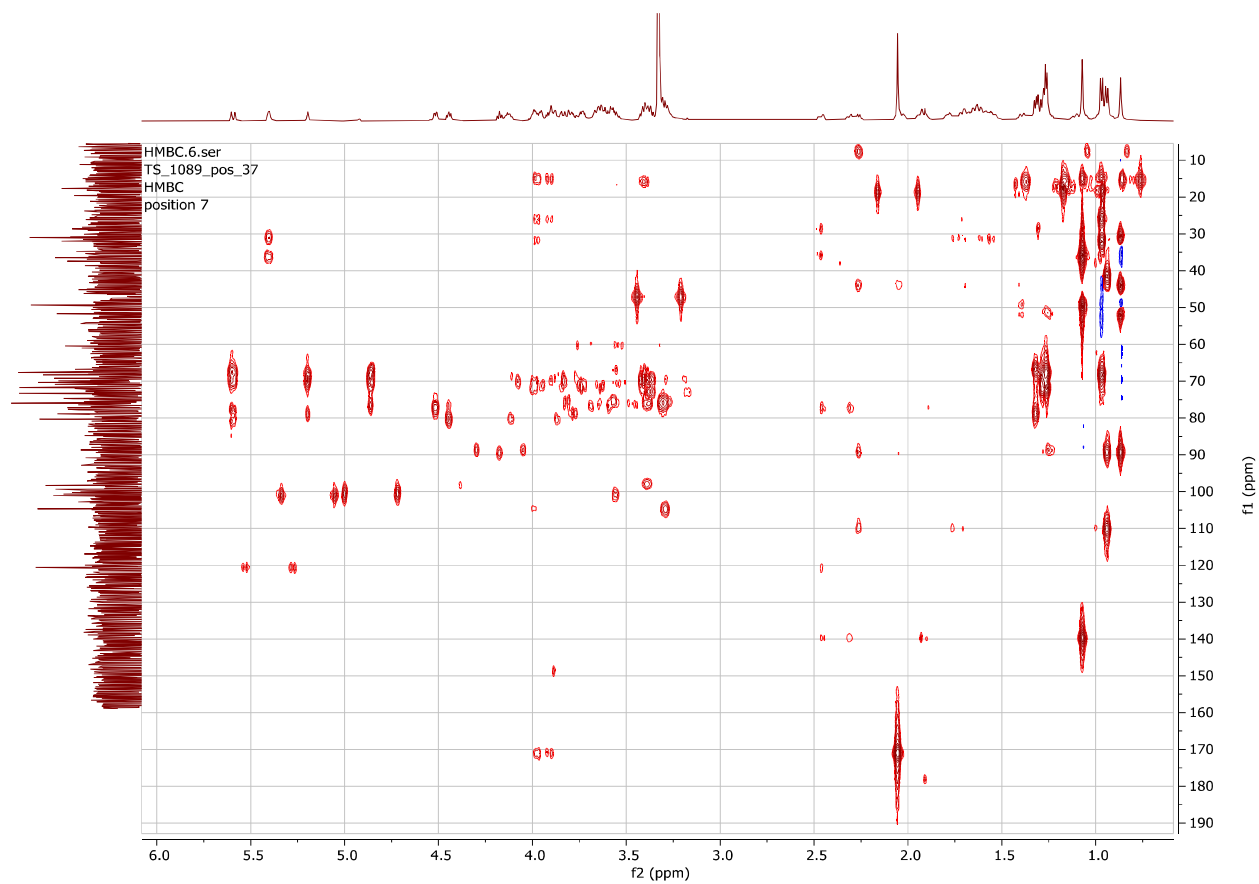


Figure S58. gHMBCAD NMR spectrum for the intact molecule of the switchgrass saponin SS1089.



Figure S59. gHSQCAD NMR spectrum for the intact molecule of the switchgrass saponin SS1089.

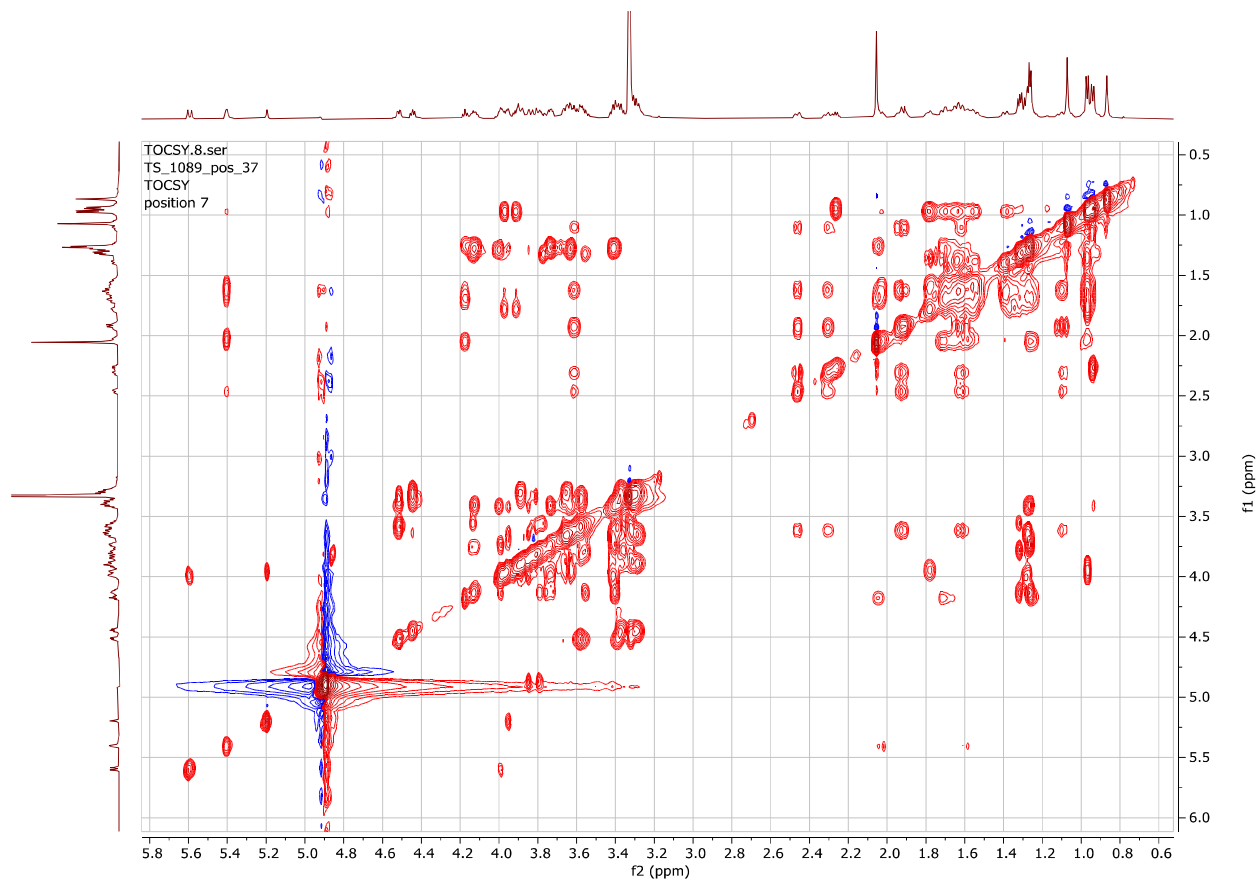


Figure S60. gTOCSY NMR spectrum for the intact molecule of the switchgrass saponin SS1089.

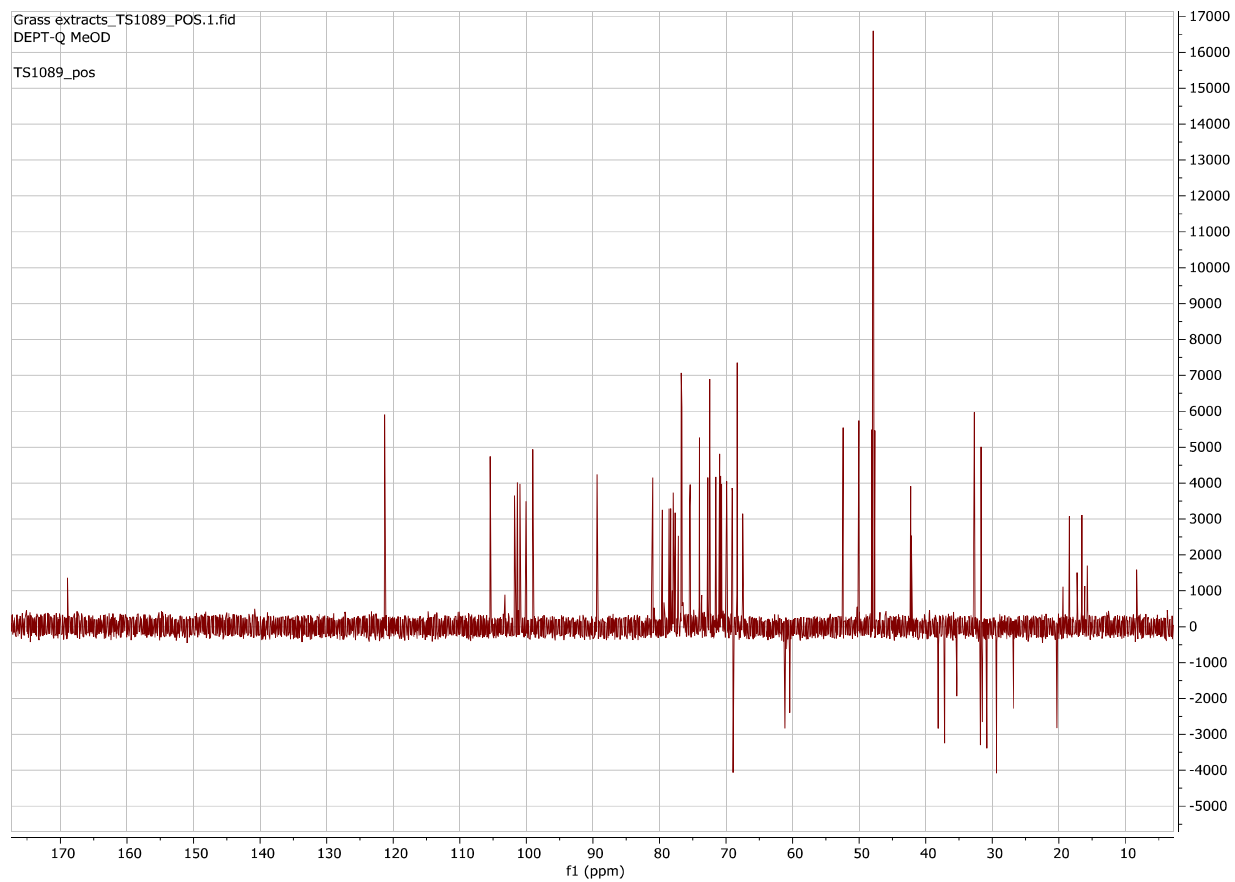


Figure S61. DEPT-Q (^{13}C) NMR spectrum for the intact molecule of the switchgrass saponin SS1089.

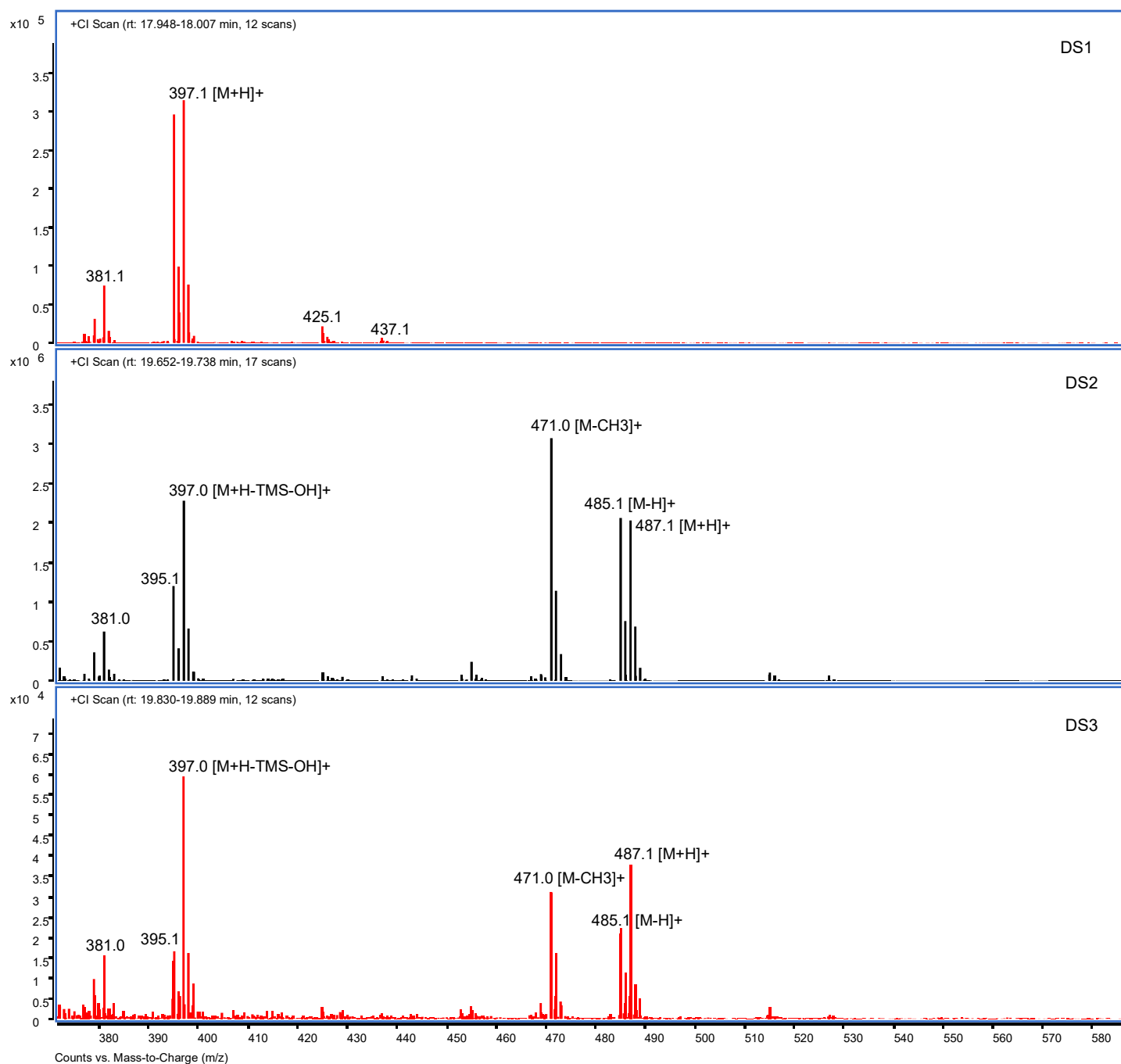


Figure S62. The (CI) GC-MS spectral patterns of the sapogenin peaks detected in the standard and switchgrass samples. The key precursor ion for each TMS derivatized sapogenin peak were documented in Table 2.

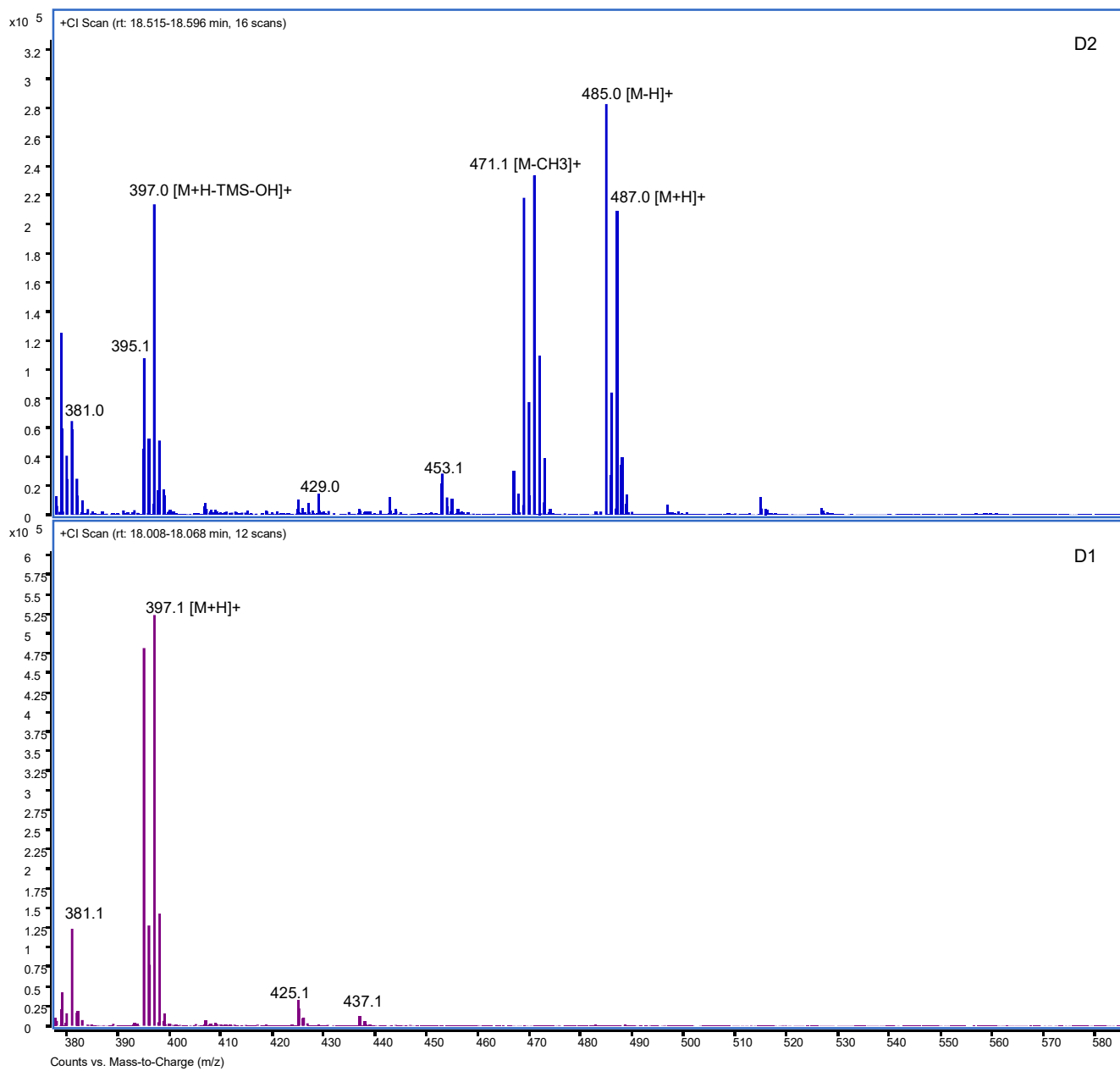


Figure S62 (continued).

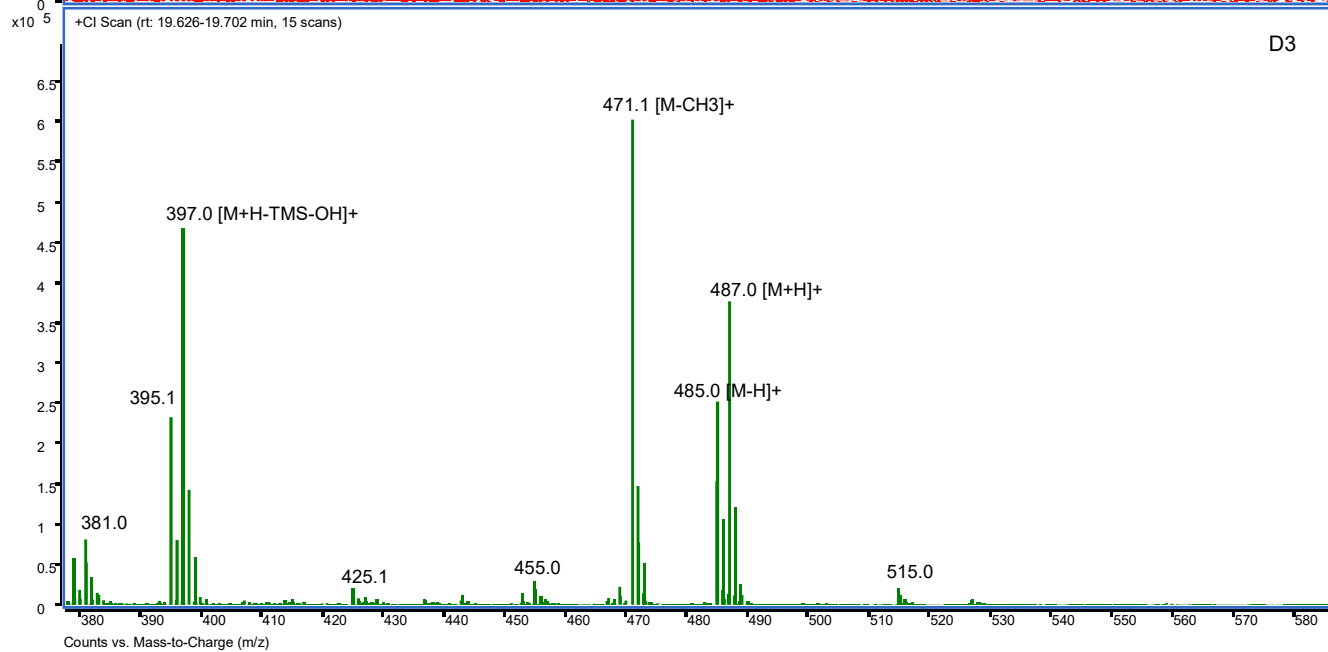
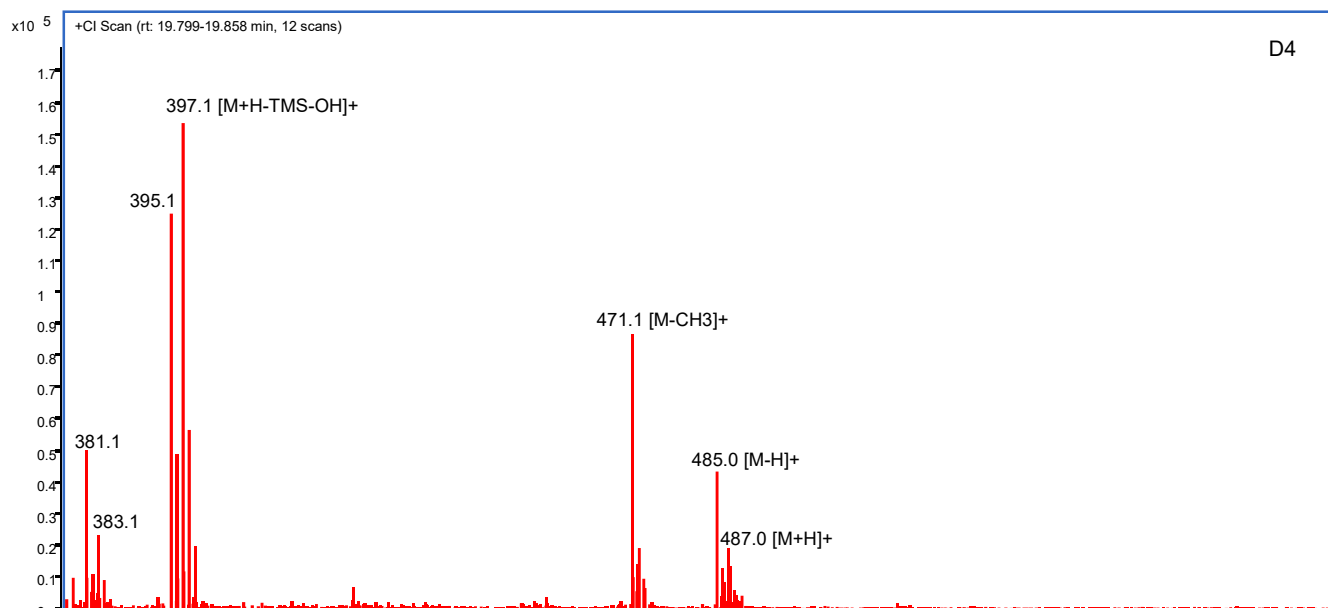


Figure S62 (continued).

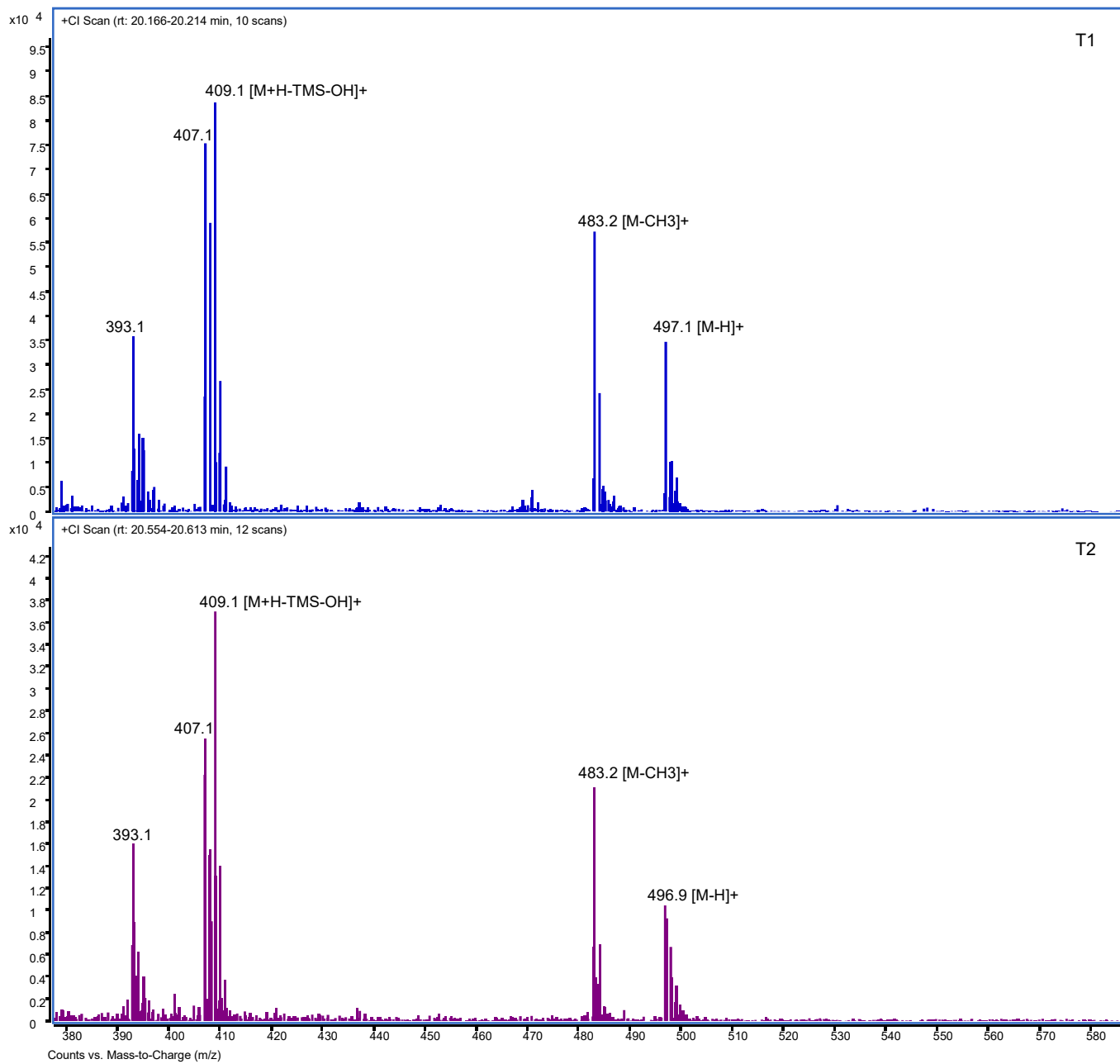


Figure S62 (continued).

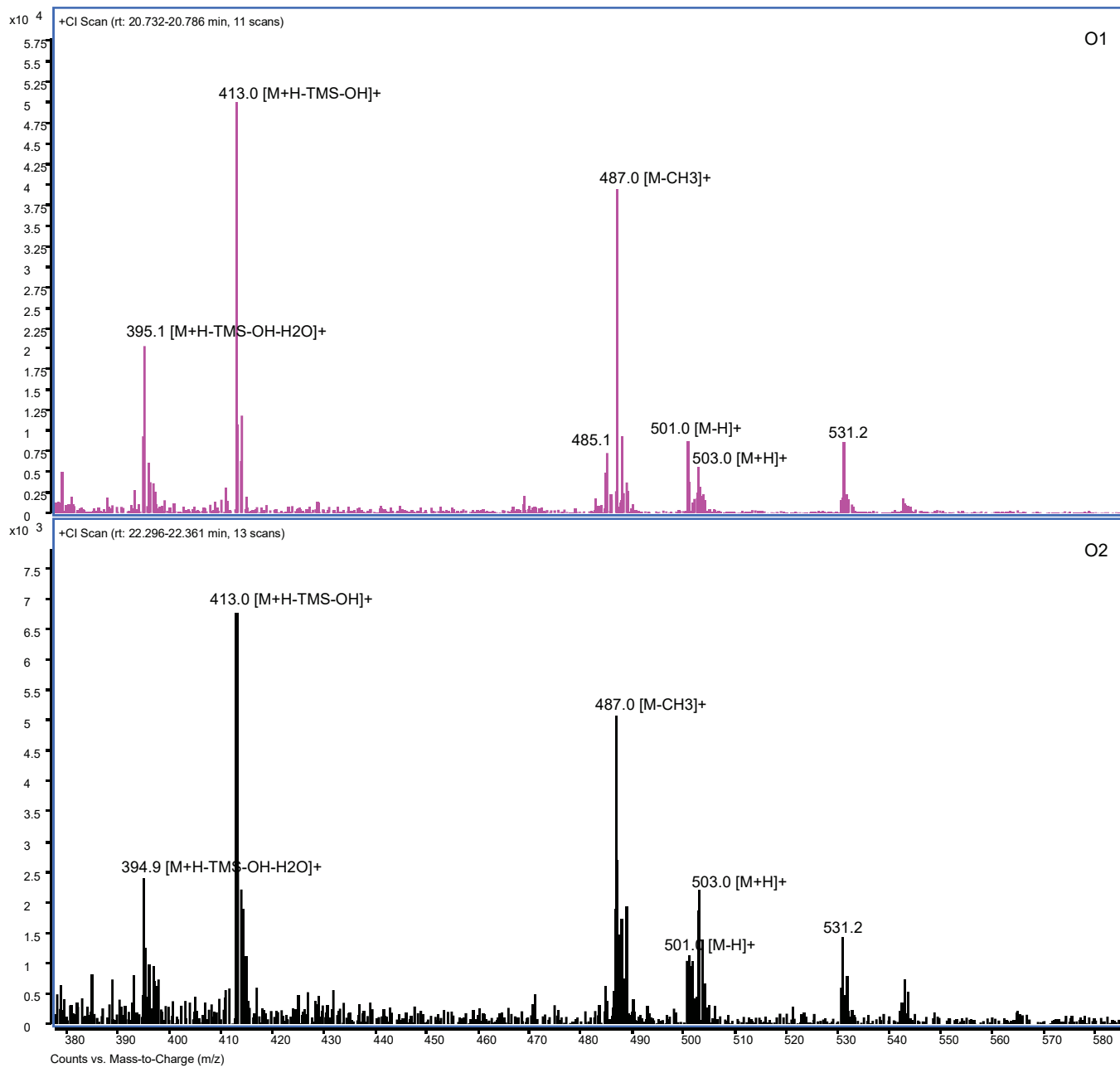


Figure S62 (continued).