CHAPTER 3

Results and Discussion

3.1 CHEMICAL CONSTITUENTS FROM THE LEAVES OF *PHYLLAGATHIS ROTUNDIFOLIA* (JACK) AND *PHYLLAGATHIS PRAETERMISSA* (WEBER)

The isolation of chemical constituents from the leaves of *Phyllagathis rotundifolia* and *P. praetermissa* are shown schematically in Figures 3.1.1 and 3.1.2, respectively. The isolated compounds were consisted of galloylated cyanogenic glucosides (**1**-**7**), gallotannins (**8**-**14**), ellagitannins (**15**-**18**), ellagic acid derivatives (**19**-**20**) and aromatic compounds (**21**-**22**) as tabulated in Table 3.1.1. Several compounds such as 3,6-di-*O*-galloyl-D-glucose (**9**), 1,2,3,6-tetra-*O*-galloyl-β-D-glucose (**13**), 6-*O*-galloyl-2,3-*O*-(*S*)-hexahydroxydiphenoyl-D-glucose (**15**), praecoxin B (**16**), 3'-*O*-methyl-3,4-methylenedioxyellagic acid 4'-*O*-β-D-glucopyranoside (**19**), 3,3',4-tri-*O*-methylellagic acid 4'-*O*-β-D-glucopyranoside (**20**) and gallic acid (**21**) were found in both species. Casuarinin (**18**) was found present only in the water soluble fraction of *P. praetermissa*. Previous phytochemical study on the leaves of *P. rotundifolia* (Ling *et al*, 2002) has reported seven prunasin based cyanogenic glucosides with galloyl esterification. These compounds have also been isolated from *P. rotundifolia* in current study. However, *P. praetermissa* collected from certain locality did not reveal their presence or only showed relatively low intensity of their presence. Hence these compounds can be used to differentiate *P. rotundifolia* from *P. praetermissa*.

The results also showed that both *P. rotundifolia* and *P. praetermissa* are rich in hydrolysable tannins (gallotannins and ellagitannins) as reported in the Cronquist's classification in 1968. Ellagic acid and its derivatives have been suggested as chemotaxonomic markers in Melastomataceae family (Lowry, 1968). 3'-*O*-methyl-3,4-methylenedioxyellagic acid 4'-*O*-β-D-glucopyranoside (**19**) and 3,3',4-tri-*O*-methylellagic acid 4'-*O*-β-D-glucopyranoside (**20**) were the two ellagic acid derivatives found in these two species that have potential to become systematic markers. The chemical structures of the compounds isolated from the leaves of *P. rotundifolia* and *P. praetermissa* are shown in Figure 3.1.3.

The compounds isolated from the leaves of both *Phyllagathis* species were characterised by ESI-MS^n that was operated in negative ionisation. The experimental *m/z* values of the molecular ions together with mass accuracy measurement and their MS^2 and MS^3 fragment ions are tabulated in Table 3.1.2. Fragment ions of galloylated cyanogenic glucosides were found to be similar for the isomeric compounds such as prunasin-digallate (**2**-**4**) and prunasin-trigallate (**5**-**6**). The trigalloyl-glucose analogues (compounds **10**-**12**) exhibited similar daughter ions in the MS^2 spectrum except that compound **12** showed a dominant fragment ion at *m/z* 483, which enables it to be differentiated from the other two analogues. Compounds (**15**-**18**) yielded a product ion at *m/z* 301, representing the fragment ion of ellagic acid and it was dominant in compound **15** and **16** (Table 3.1.2). The loss of *m/z* 44 $[M–H–COO]^-$ in gallic acid (**20**) resulted in a product ion at *m/z* 125 (Mämmelä *et al.*, 2000; Zywicki *et al.*, 2002; Soong and Barlow, 2005; Meyers *et al.*, 2006; He and Xia, 2007; Nuengchamnong and Ingkaninan, 2009). The isotopic and fragmentation patterns for galloylated cyanogenic glucosides (**1**-**7**), gallotannins (**8**-**14**), ellagitannins (**15**-**18**), ellagic acid derivatives (**19**-**20**) and aromatic acid (**21**-**22**) were described in the following section. The fragmentation scheme for each of the groups (galloylated cyanogenic glucosides,

gallotannins and ellagitannins) are summarized in Figures 3.1.1.15, 3.1.2.15 and 3.1.3.9, respectively.

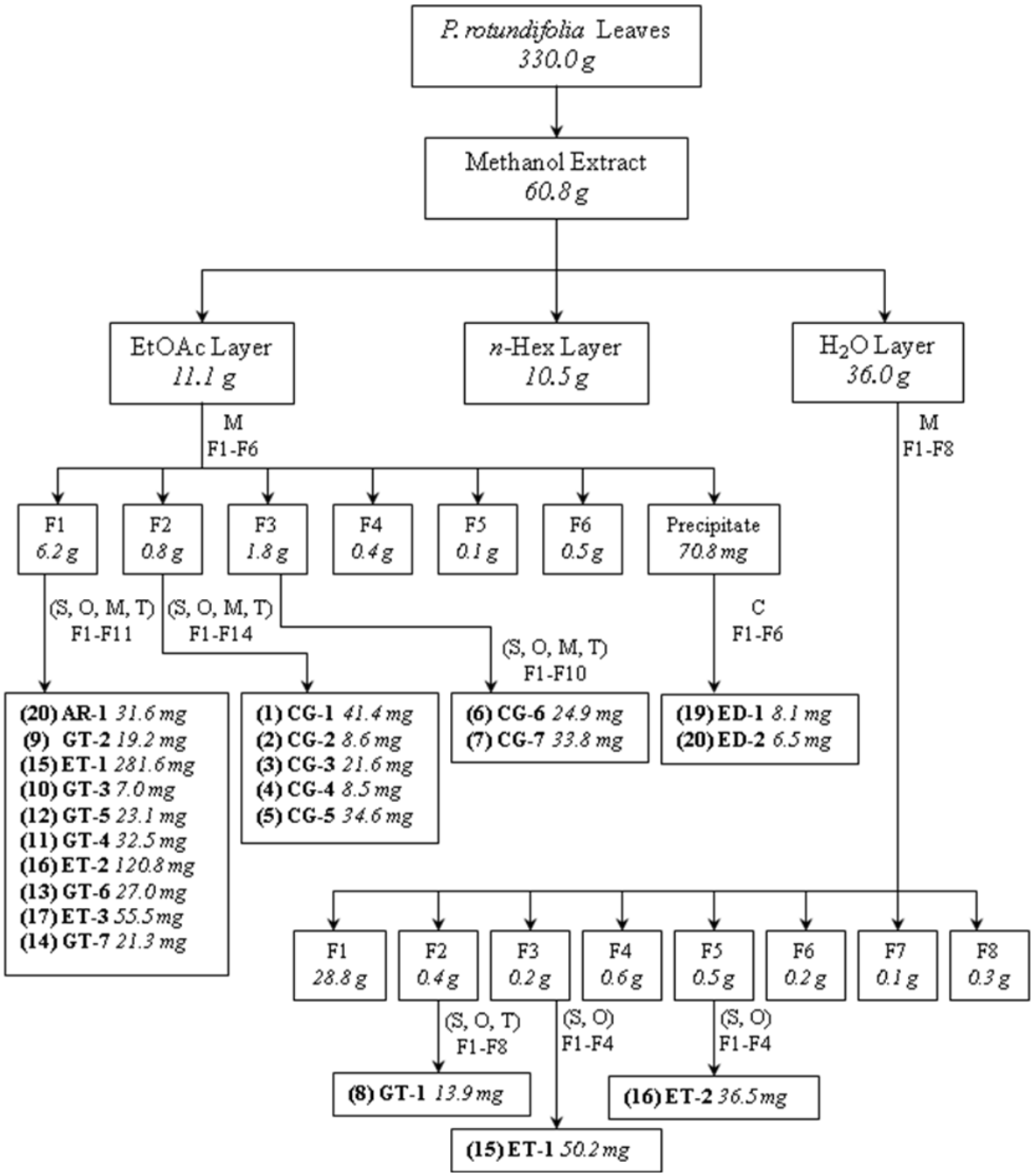


Figure 3.1.1: Fractionation and purification of constituents from the leaves of *Phyllagathis rotundifolia*. M–MCI gel CHP 20P; O–Chromatorex ODS; S–Sephadex LH-20; T–Toyopearl HW-40F; C–Semi-preparative reversed-phase Symmetry C-18 column; F–Fraction.

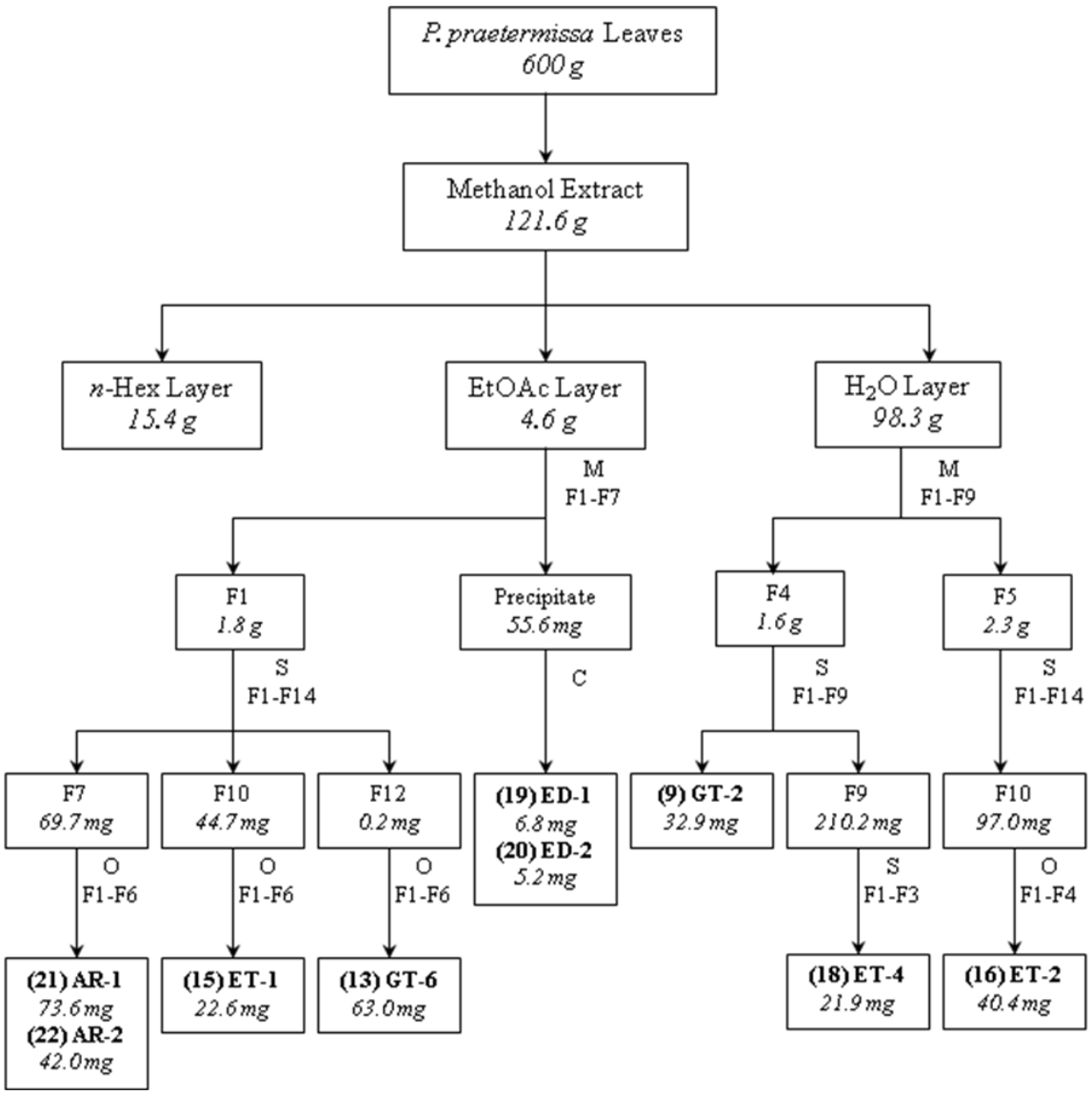
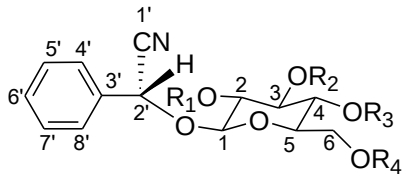
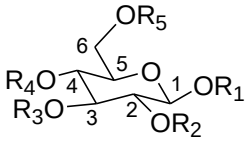


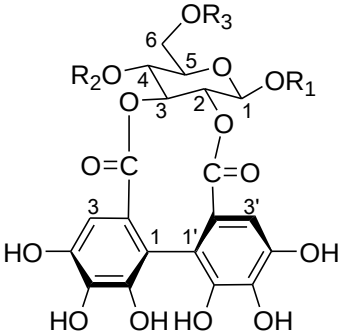
Figure 3.1.2: Fractionation and purification of constituents from the leaves of *Phyllagathis praetermissa*. M–MCI gel CHP 20P; O–Chromatorex ODS; S–Sephadex LH-20; C–Semi-preparative reversed-phase Symmetry C-18 column; F–Fraction.



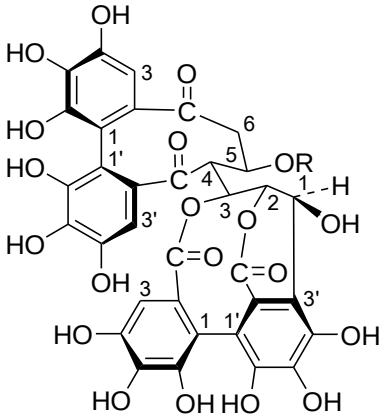
	R_1	R_2	R_3	R_4
1:	H	H	H	G
2:	G	H	H	G
3:	H	G	H	G
4:	H	H	G	G
5:	G	G	H	G
6:	H	G	G	G
7:	G	G	G	G



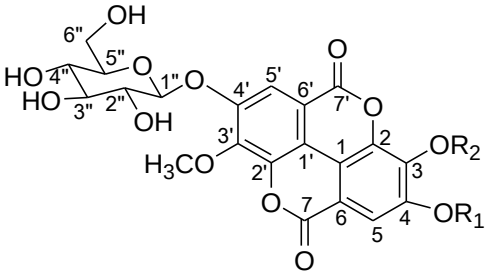
	R_1	R_2	R_3	R_4	R_5
8:	H	H	H	H	G
9:	H	H	G	H	G
10:	G	G	G	H	H
11:	G	H	H	G	G
12:	H	H	G	G	G
13:	G	G	G	H	G
14:	G	G	G	G	G



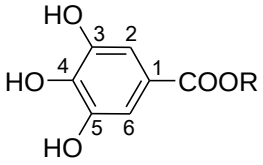
	R_1	R_2	R_3
15:	H	H	G
16:	H	G	G
17:	G	G	G



18



	R_1	R_2
19:	- CH_2 -	
20:	CH_3	CH_3



	R
21:	H
22:	CH_3

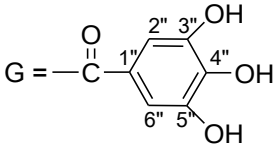


Figure 3.1.3: Compounds isolated from the leaves of *P. rotundifolia* and *P. praetermissa*.

Table 3.1.1: Summary of compounds isolated from the leaves of *P. rotundifolia* and *P. praetermissa*.

No.	Compound Name	Group	PR	PP	References
1	Prunasin 6'-*O*-gallate	CG-1	✓		Ling *et al.* 2002; Miller *et al.* 2006; Isaza *et al.* 2001
2	Prunasin 2',6'-di-*O*-gallate	CG-2	✓		Ling *et al.* 2002
3	Prunasin 3',6'-di-*O*-gallate	CG-3	✓		Ling *et al.* 2002
4	Prunasin 4',6'-di-*O*-gallate	CG-4	✓		Ling *et al.* 2002
5	Prunasin 2',3',6'-tri-*O*-gallate	CG-5	✓		Ling *et al.* 2002
6	Prunasin 3',4',6'-tri-*O*-gallate	CG-6	✓		Ling *et al.* 2002
7	Prunasin 2',3',4',6'-tetra-*O*-gallate	CG-7	✓		Ling *et al.* 2002
8	6-*O*-Galloyl-D-glucose	GT-1	✓		He *et al.* 2001; Nonaka and Nishioka 1983
9	3,6-Di-*O*-galloyl-D-glucose	GT-2	✓	✓	He *et al.* 2001
10	1,2,3-Tri-*O*-galloyl-β-D-glucose	GT-3	✓		Nawwar *et al.* 1994; Lee *et al.* 1990
11	1,4,6-Tri-*O*-galloyl-β-D-glucose	GT-4	✓		Nawwar *et al.* 1994; Lee *et al.* 1992; Nonaka *et al.* 1984
12	3,4,6-Tri-*O*-galloyl-D-glucose	GT-5	✓		He *et al.* 2001; Lee *et al.* 1991; Lee *et al.* 1989; Wilkins 1988
13	1,2,3,6-Tetra-*O*-galloyl-β-D-glucose	GT-6	✓	✓	Duan *et al.* 2004; Yoshida *et al.* 1991a; Saijo *et al.* 1990; Nishizawa *et al.* 1983
14	1,2,3,4,6-Penta-*O*-galloyl-β-D-glucose	GT-7	✓		Gao *et al.* 2007; Tanaka *et al.* 1985; Nishizawa *et al.* 1983; Saijo *et al.* 1990
15	6-*O*-Galloyl-2,3-*O*-(*S*)-hexahydroxydiphenoyl-D-glucose	ET-1	✓	✓	Yoshida *et al.* 1991a; Yoshida *et al.* 1991c
16	Praecoxin B	ET-2	✓	✓	Yoshida *et al.* 1991a; Hatano *et al.* 1991
17	Pterocarinin C	ET-3	✓		Yoshida *et al.* 1991a; Yoshida *et al.* 1995
18	Casuarinin	ET-4		✓	Nonaka *et al.* 1985; Okuda *et al.* 1983; Hatano *et al.* 1988; Yoshimura *et al.* 2008
19	3'-*O*-Methyl-3,4-methylenedioxyellagic acid 4'-*O*-β-D-glucopyranoside	ED-1	✓	✓	Khallouki *et al.* 2007; Li *et al.* 1999; Li *et al.* 2000
20	3,3',4-Tri-*O*-methylellagic acid 4'-*O*-β-D-glucopyranoside	ED-2	✓	✓	Li *et al.* 1999; Li *et al.* 2000
21	Gallic acid	AR-1	✓	✓	Yoshida *et al.* 1991c; Chanwitheesuk *et al.* 2007; Nawwar *et al.* 1982
22	Gallic acid methyl ester	AR-2		✓	Yoshida *et al.* 1991b; Khac *et al.* 1990
	Total:		**20**	**9**	

CG- cyanogenic glucosides; GT-gallotannins; ET-ellagitannins; ED-ellagic acid derivatives; AR-aromatic; PR-*P. rotundifolia*; PP-*P. praetermissa*

Table 3.1.2: Mass spectral data of compounds isolated from the leaves of *P. rotundifolia* and *P. praetermissa*.

No.	Compound	λ_{max} (nm)	Elemental composition	$[M–H]^-$	Mass accuracy (ppm)	MS^2	MS^3
1	Prunasin 6'-*O*-gallate	216, 275	$C_{21}H_{21}NO_{10}$	446.10913	-0.3	125(2), 169(22), 223(2), 241(1), 253(2), 270(1), 295(1), 313(100), 359(2), 419(5)	125(7), 151(3), 169(100), 193(2), 211(4), 223(15), 233(2), 241(8), 253(8), 295(9)
2	Prunasin 2',6'-di-*O*-gallate	216, 277	$C_{28}H_{25}NO_{14}$	598.11981	-0.7	169(2), 295(1), 297(1), 313(3), 405(2), 428(4), 446(100), 465(12), 554(0.2), 580(1)	125(1), 169(15), 223(1), 241(0.4), 253(2), 270(0.2), 295(1), 313(100), 359(0.3), 419(1)
3	Prunasin 3',6'-di-*O*-gallate	216, 276	$C_{28}H_{25}NO_{14}$	598.11981	-0.7	169(3), 295(1), 297(2), 313(6), 321(0.4), 405(2), 428(16), 446(100), 465(19), 554(0.4)	125(2), 169(20), 223(1), 241(0.4), 253(3), 270(0.3), 295(1), 313(100), 359(0.4), 419(1)
4	Prunasin 4',6'-di-*O*-gallate	216, 278	$C_{28}H_{25}NO_{14}$	598.11957	+1.1	169(5), 211(1), 295(3), 297(6), 313(11), 405(1), 428(47), 446(100), 465(56), 554(0.8)	125(2), 169(20), 223(1), 211(0.4), 241(0.4), 253(2), 295(1), 313(100), 359(1), 419(2)
5	Prunasin 2',3',6'-tri-*O*-gallate	216, 277	$C_{35}H_{29}NO_{18}$	750.13000	-1.6	279(1), 383(1), 428(1), 446(1), 449(6), 465(3), 553(3), 580(43), 598(100), 617(3)	169(2), 235(0.4), 295(1), 297(2), 313(6), 405(1), 428(21), 446(45), 465(0.3), 554(1)
6	Prunasin 3',4',6'-tri-*O*-gallate	216, 277	$C_{35}H_{29}NO_{18}$	750.13068	-0.7	383(1), 410(1), 428(3), 446(1), 449(2), 465(0.4), 553(1), 580(1), 598(100), 617(4)	169(2), 235(0.2), 295(1), 297(1), 313(4), 405(2), 428(14), 446(100), 465(16), 554(0.4)
7	Prunasin 2',3',4',6'-tetra-*O*-gallate	216, 278	$C_{42}H_{33}NO_{22}$	902.14069	-1.6	429(0.3), 447(0.4), 449(0.3), 564(1), 580(2), 598(3), 617(1), 732(20), 750(100), 769(1)	295(1), 383(1), 428(2), 447(2), 449(5), 465(7), 553(2), 580(38), 598(95), 617(100)
8	6-*O*-Galloyl-D-glucose	216, 274	$C_{13}H_{16}O_{10}$	331.06747	+1.2	117(0.2), 125(0.2), 151(0.2), 169(5), 193(2), 211(1), 241(1), 271(100), 295(0.2), 313(1)	125(0.3), 169(11), 193(0.6), 211(100), 223(0.2), 235(0.2), 253(1)
9	3,6-Di-*O*-galloyl-D-glucose	216, 277	$C_{20}H_{20}O_{14}$	483.07839	+0.7	169(2), 193(3), 211(2), 271(6), 313(4), 331(4), 405(0.6), 423(100), 439(1), 465(2)	169(4), 193(12), 211(26), 217(1), 235(2), 253(10), 271(88), 329(0.3), 378(0.4), 405(100)
10	1,2,3-Tri-*O*-galloyl-β-D-glucose	216, 278	$C_{27}H_{24}O_{18}$	635.08862	-0.6	169(2), 235(0.4), 271(0.4), 295(2), 313(3), 405(0.4), 421(6), 465(100), 483(23), 617(1)	125(4), 169(41), 193(4), 235(8), 295(31), 313(100), 321(3), 377(3), 421(82), 447(5)
11	1,4,6-Tri-*O*-galloyl-β-D-glucose	217, 273	$C_{27}H_{24}O_{18}$	635.08839	-0.9	169(5), 221(7), 271(5), 295(5), 313(13), 405(3), 423(32), 465(100), 483(82), 617(12)	125(2), 169(22), 193(3), 235(9), 253(1), 271(1), 295(13), 313(100), 405(2), 447(7)

12	3,4,6-Tri-*O*-galloyl-D-glucose	215, 276	$C_{27}H_{24}O_{18}$	635.08815	-1.3	169(9), 235(2), 271(4), 295(14), 313(9), 405(5), 423(30), 465(68), 483(100), 617(11)	125(1), 169(11), 193(8), 211(4), 241(3), 271(13), 313(16), 331(24), 423(100), 465(8)
13	1,2,3,6-Tetra-*O*-galloyl-β-D-glucose	217, 278	$C_{34}H_{28}O_{22}$	787.09864	-1.7	295(1), 403(2), 421(0.4), 429(1), 447(2), 465(3), 529(0.2), 573(4), 617(100), 635(31)	169(6), 235(5), 277(5), 295(10), 403(57), 421(12), 447(65), 465(80), 529(7), 573(100)
14	1,2,3,4,6-Penta-*O*-galloyl-β-D-glucose	217, 281	$C_{41}H_{32}O_{26}$	939.11006	+0.9	329(0.4), 439(0.4), 447(0.2), 515(0.2), 599(1), 601(0.2), 617(3), 725(1), 769(100), 787(8)	277(4), 403(5), 429(15), 447(20), 511(6), 555(7), 573(4), 599(47), 617(100), 725(25)
15	6-*O*-Galloyl-2,3-*O*-(*S*)-hexahydroxydiphenoyl-D-glucose	216, 263	$C_{27}H_{22}O_{18}$	633.07252	-1.3	229(1), 249(1), 257(1), 275(19) , 301(100), 319(0.4), 481(3) , 589(0.3), 615(1)	171(8), 185(5), 229(23), 257(100), 273(15), 284(3)
16	Praecoxin B	216, 275	$C_{34}H_{26}O_{22}$	785.08375	-0.7	275(15), 301(100), 319(3), 405(2), 423(32), 465(9), 483(75), 633(35), 741(1), 767(2)	171(30), 229(23), 257(100), 273(16)
17	Pterocarinin C	217, 277	$C_{41}H_{30}O_{26}$	937.09595	+0.7	275(1), 301(17), 313(1), 465(12), 483(7), 617(9), 635(100), 767(2), 785(19), 919(0.4)	169(1), 271(0.4), 295(3), 313(8), 405(1), 423(3), 465(100), 483(644), 591(0.3), 617(0.3)
18	Casuarinin	204, 272	$C_{41}H_{28}O_{26}$	935.07832	-1.4	301 (4), 571(10), 589(6), 615(9), 633(100), 659(23), 783(8), 855(7), 873(19), 891(7), 917(58)	275(24), 299(44), 301(2), 317(13), 329(27), 383(11), 419(11), 481(25), 553(15), 571(95), 589(48), 615(100)
19	3'-*O*-Methyl-3,4-methylenedioxyellagic acid 4'-*O*-β-D-glucopyranoside	250, 365	$C_{22}H_{18}O_{13}$	489.06738	-0.2	327[a], 312(100), 283(0.2),	171(4), 212(4), 240(89), 256(5), 284(100)
20	3,3',4-Tri-*O*-methylellagic acid 4'-*O*-β-D-glucopyranoside	247, 367	$C_{23}H_{22}O_{13}$	505.09872	-0.1	343[a], 328(100), 313(0.1)	171(0.2), 299(1), 313(100)
21	Gallic acid	216, 271	$C_7H_6O_5$	169.01428	+0.2	125(100)	-
22	Gallic acid methyl ester	216, 273	$C_8H_8O_5$	183.03003	+0.7	168(100)	-

[a] Represented the intense fragment ion due to the removal of glucose (*m/z* 162) and was selected for fragmentation in the MS^2, bold *m/z* values were the major predominant ion and underlined *m/z* values were the major MS^3 products from MS^2 major predominant ion.

3.1.1 Galloylated Cyanogenic Glucosides

(a) Prunasin 6'-*O*-gallate (1), CG-1

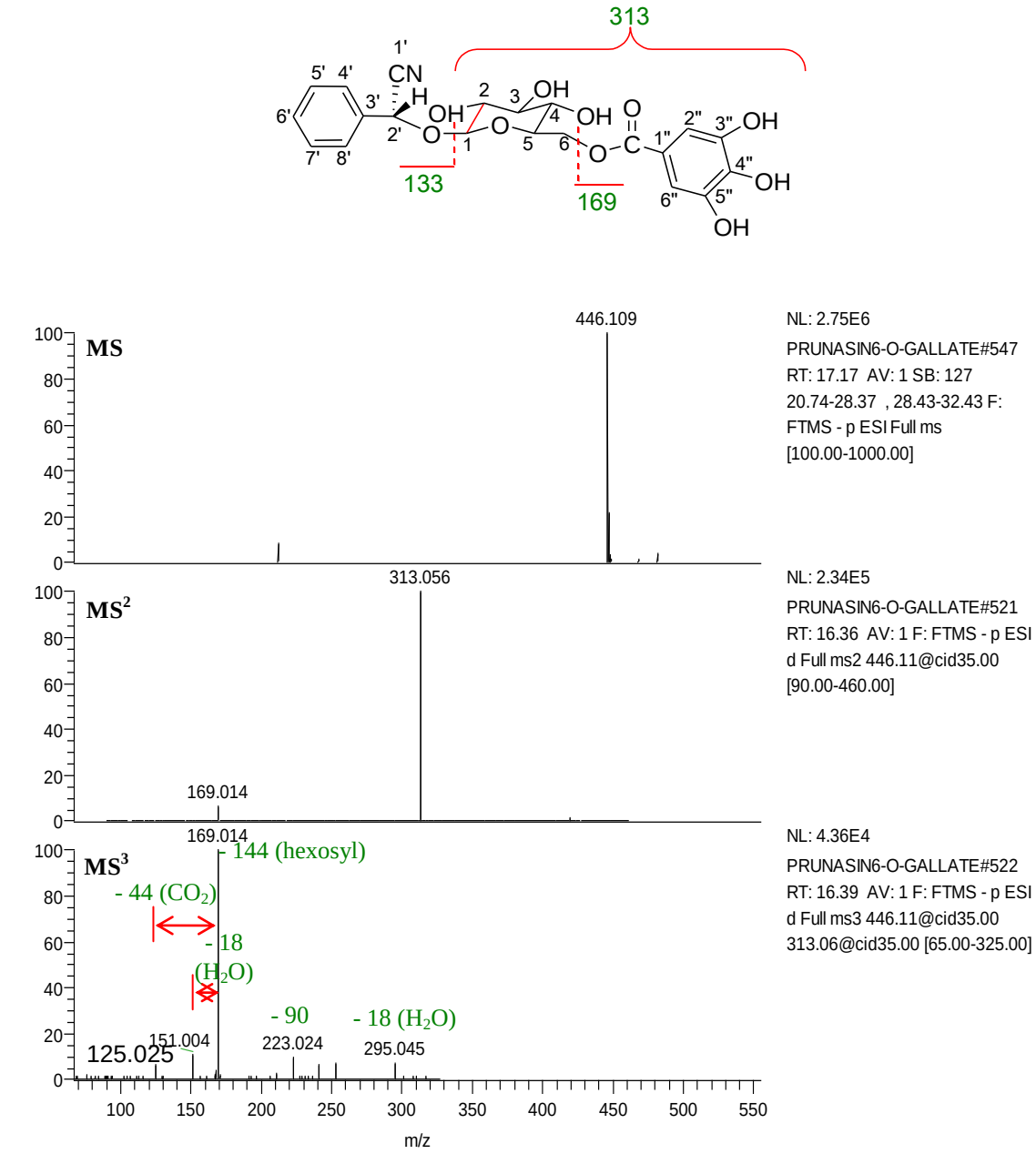


Figure 3.1.1.1: Fragmentation pattern for prunasin 6'-*O*-gallate (**1**).

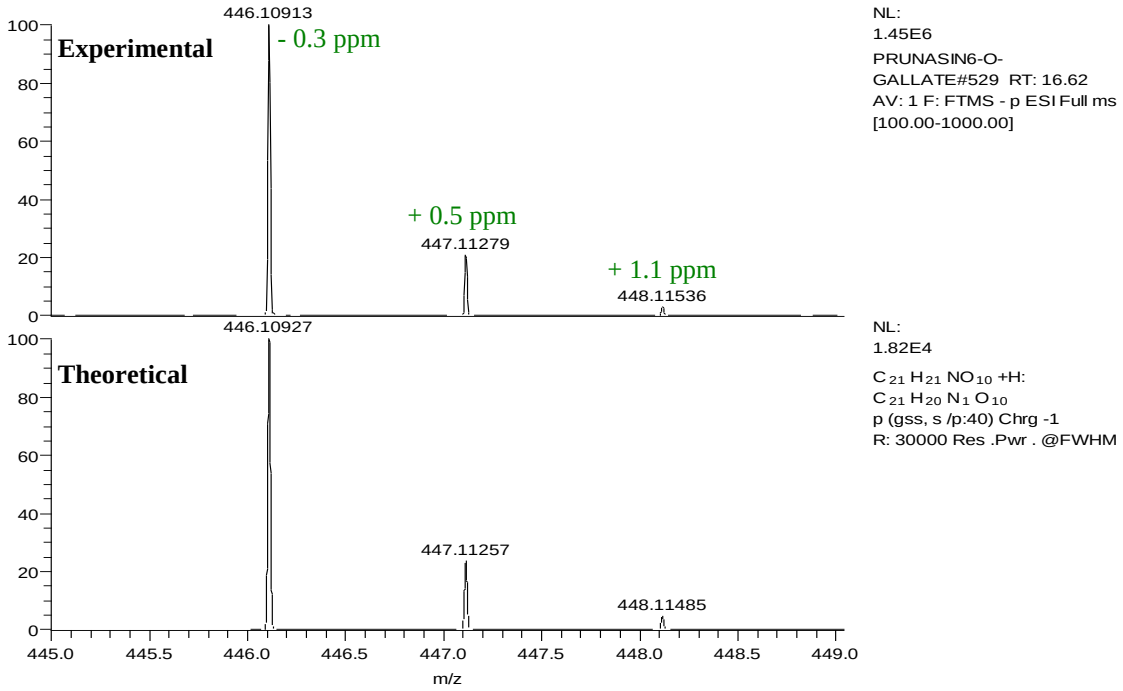
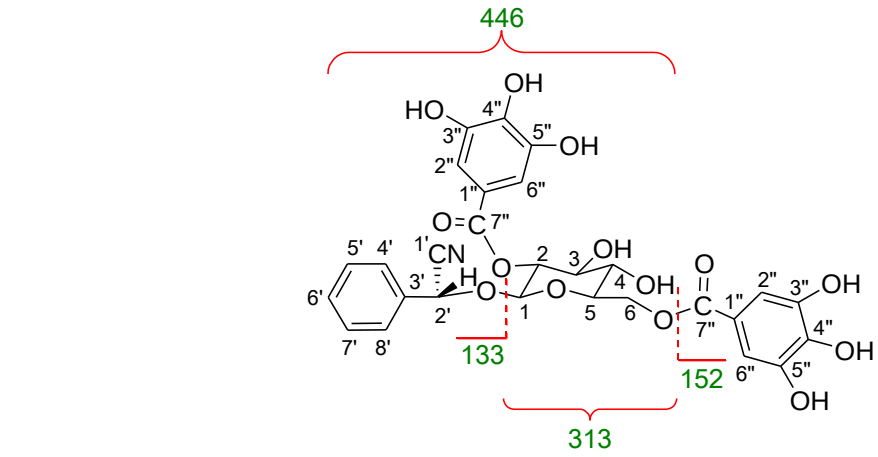


Figure 3.1.1.2: Full scan mass spectra (*m/z*) with isotopic pattern for prunasin 6'-*O*-gallate (**1**).

Prunasin 6'-*O*-gallate (**1**), CG-1, isolated as a white amorphous powder, with the molecular formula of $C_{21}H_{21}NO_{10}$ showed deprotonated $[M–H]^-$ ion at *m/z* 446. Compound CG-1 is consisted of a cyanohydrin, a galloyl group and a glucose moiety. The ion at *m/z* 446 was fragmented to a major product ion at *m/z* 313 in the MS^2 spectrum. This indicated the removal of the cyanohydrin group (Figure 3.1.1.1). In the MS^3 spectrum, the removal of a sugar moiety (*m/z* 144) produced a major ion at *m/z* 169, representing the deprotonated gallic acid. The occurrence of ion at *m/z* 151 was due to the dehydration from the ion at *m/z* 169. The calculated mass accuracy in ppm together with comparison between the experimental and theoretical isotopic pattern are shown in Figure 3.1.1.2. The ^{1}H-NMR showed a singlet oxymethine proton (δ 5.85) and multiplet signals arising from a monosubstituted benzene ring (δ 7.43-7.58) whereas ^{13}C-NMR exhibited six aromatic carbon signals (δ 128.6-134.5), a methine carbon signal (δ 118.8) and a quaternary carbon signal (δ 68.1) that were consistent with reported data for the aglycon structure (Tables

3.1.1.1 and 3.1.1.2, Appendix I). One galloyl group was esterified to C-6 position of the glucose. The deshielded H-6 in the glucose moiety and the upfield shift of C-5 showed that the galloyl unit was assigned at C-6 position. Thus, compound CG-1 was designated as prunasin 6'-*O*-gallate. (Isaza *et al.*, 2001; Ling *et al.*, 2002; Miller *et al.*, 2006).

(b) Prunasin 2',6'-di-*O*-gallate (2), CG-2



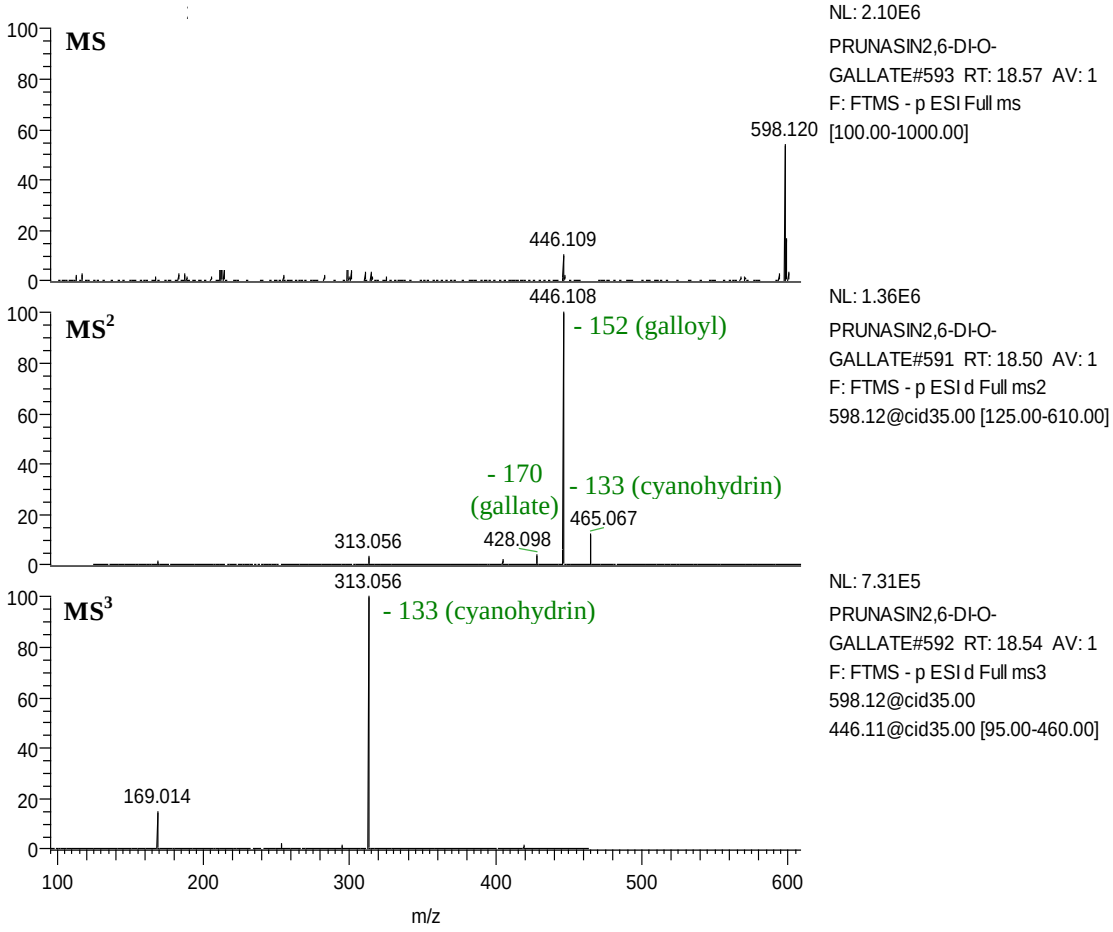


Figure 3.1.1.3: Fragmentation pattern for prunasin 2',6'-di-*O*-gallate (**2**).

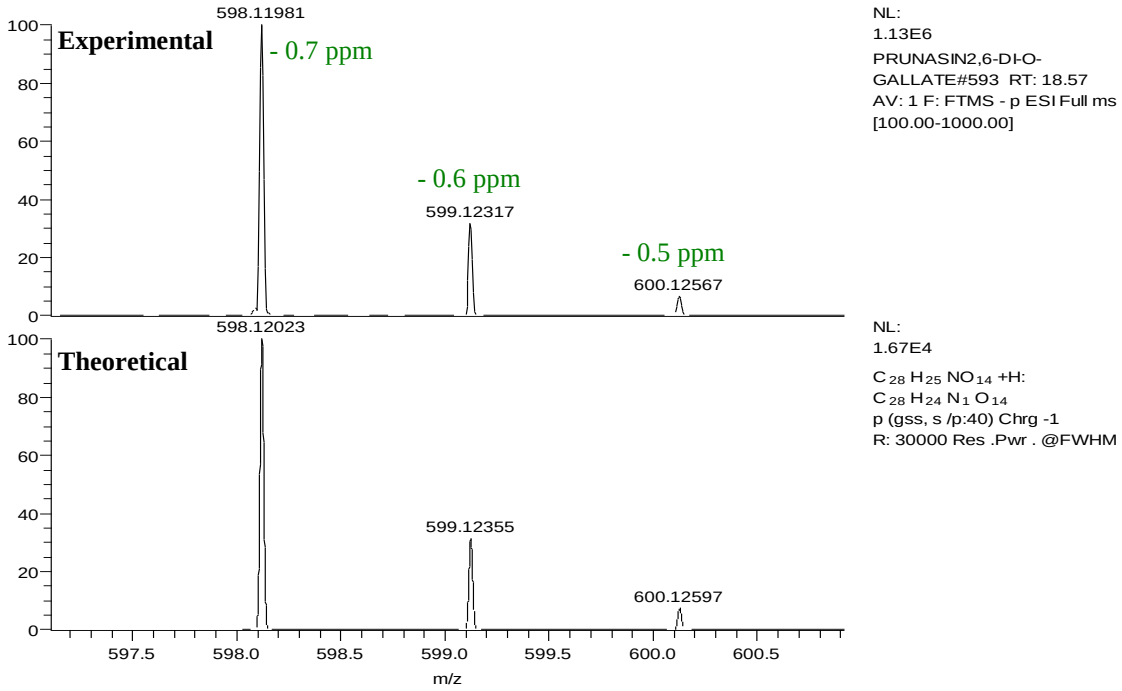


Figure 3.1.1.4: Full scan mass spectra (*m/z*) with isotopic pattern for prunasin 2',6'-di-*O*-gallate (**2**).

Prunasin 2',6'-di-*O*-gallate (**2**), CG-2, isolated as a tan amorphous powder, with the molecular formula of $C_{28}H_{25}NO_{14}$ which showed deprotonated $[M–H]^-$ ion at *m/z* 598. Compound CG-2 is consisted of a cyanohydrin, two galloyl groups and a glucose moiety. In the MS^2 spectrum, compound CG-2 was fragmented to a major product ion at *m/z* 446 which displayed the removal of a galloyl moiety (*m/z* 152). In addition, the mass loss of a cyanohydrin (*m/z* 133) and a gallate (*m/z* 170) were gave ion peaks at *m/z* 465 and 428, respectively (Figure 3.1.1.3). Further fragmentation on ion $[M–H]^-$ at *m/z* 446 led to cleavage of cyanohydrin (*m/z* 133) and produced a product ion at *m/z* 313 in the MS^3 spectrum. The ion at *m/z* 169 represented the deprotonated gallic acid and it corresponded to the mass loss of a sugar moiety (*m/z* 144). The fragmentation pattern of prunasin 2',6'-di-*O*-gallate (**2**) in the MS^3 spectrum was found similar to that of prunasin 6'-*O*-gallate (**1**) in the MS^2 spectrum. Figure 3.1.1.4 presents the calculated mass accuracy in ppm and the isotopic pattern for compound CG-2. The ^1H and ^{13}C-NMR spectroscopic data showed the

presence of an aglycon with an oxymethine proton (δ 5.90) and five aromatic protons from benzene ring (δ 7.17-7.46) together with six aromatic carbons (δ 128.4-134.2), a methine carbon at δ 118.7 and a quaternary carbon at δ 68.4 (Tables 3.1.1.1 and 3.1.1.2, Appendix I). The upfield shift of C-1, C-3 and C-5 and downfield shift of H-2 and H-6 in the glucose unit of CG-2 indicated the presence of two galloyl groups at position C-2 and C-6, respectively. Thus, compound CG-2 was assigned as prunasin 2',6'-di-*O*-gallate (Ling *et al.*, 2002).

(c) Prunasin 3',6'-di-*O*-gallate (3), CG-3

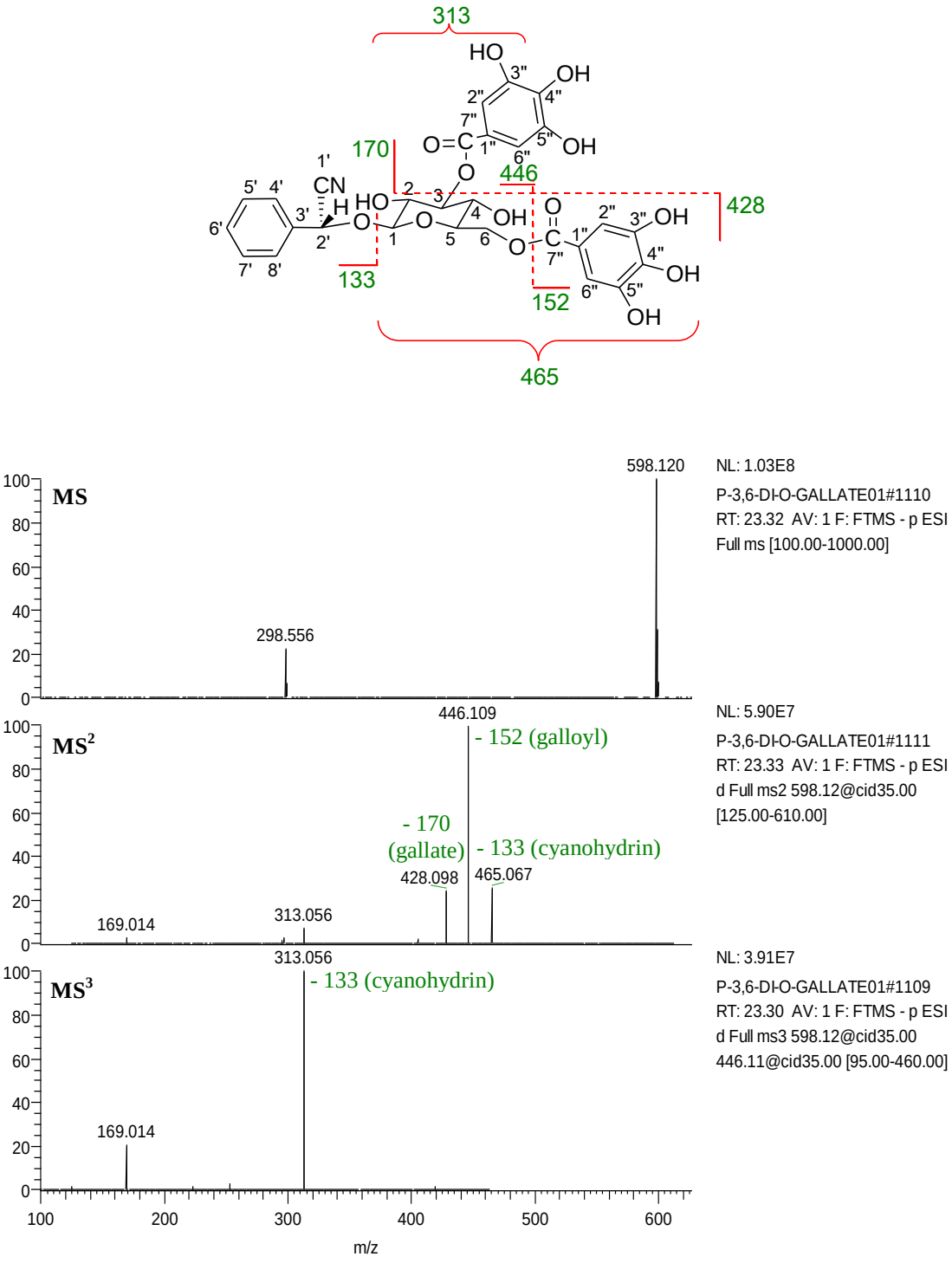


Figure 3.1.1.5: Fragmentation pattern for prunasin 3',6'-di-*O*-gallate (**3**).

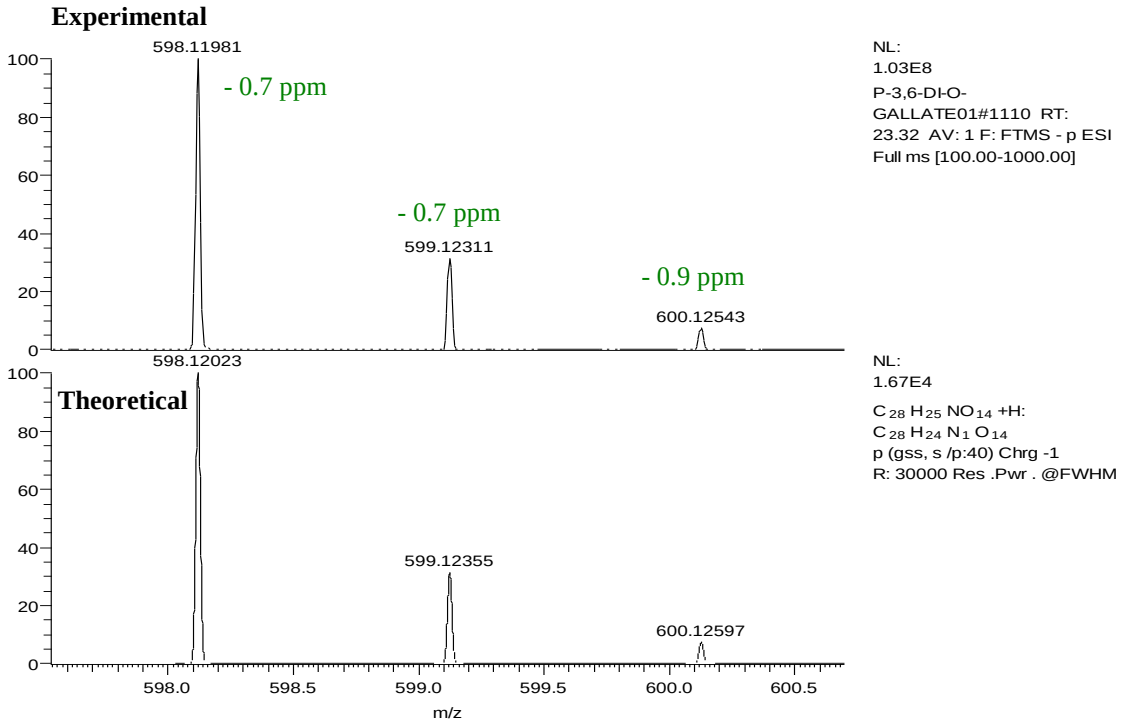


Figure 3.1.1.6: Full scan mass spectra (*m/z*) with isotopic pattern for prunasin 3',6'-di-*O*-gallate (**3**).

Prunasin 3',6'-di-*O*-gallate (**3**), CG-3, a tan amorphous powder, with the molecular formula of $C_{28}H_{25}NO_{14}$ showed deprotonated molecular $[M–H]^-$ ion at *m/z* 598. Compound CG-3 is consisted of a cyanohydrin, two galloyl moieties and a glucose moiety. Parent ion at *m/z* 598 was fragmented to a major ion at *m/z* 446 in MS^2 spectrum that was resulted from the mass loss of a galloyl moiety (*m/z* 152). The other product ions such as *m/z* 465 and 428 showed the removal of a cyanohydrin (*m/z* 133) and gallate (*m/z* 170), respectively in MS^2 spectrum (Figure 3.1.1.6). Collision induced dissociation was further applied to the fragment ion at *m/z* 446 and the resulting MS^3 spectrum showed two major product ions at *m/z* 313 and 169. The ion at *m/z* 313 and 169 were related to the mass loss of a cyanohydrin (*m/z* 133) and sugar (*m/z* 144), respectively. In Figures 3.1.1.3 and 3.1.1.5, similar fragmentation characteristic is found among prunasin 2',6'-di-*O*-gallate (**2**) and prunasin 3',6'-di-*O*-gallate (**3**). The isotopic pattern and calculated mass accuracy for prunasin 3',6'-di-*O*-gallate (**3**) are shown in Figure 3.1.1.6. The ^1H and ^{13}C-NMR signals of the aglycon

(Appendix I) consisting of a singlet oxymethine proton (δ 5.92), five aromatic protons from benzene ring (δ 7.46-7.61) and six aromatic carbon signals (δ 128.5-134.4), a methine carbon (δ 118.8) and a quaternary carbon (δ 68.3) are shown in Tables 3.1.1.1 and 3.1.1.2. Compound CG-3 has galloyl ester linkages at C-3 and C-6 due to the downfield chemical shift of H-3 and H-6, as well as upfield shift of C-2, C-4 and C-5 in the glucose moiety. Therefore, compound CG-3 was characterised as prunasin 3',6'-di-*O*-gallate (Ling *et al.*, 2002).

(d) Prunasin 4',6'-di-*O*-gallate (4), CG-4

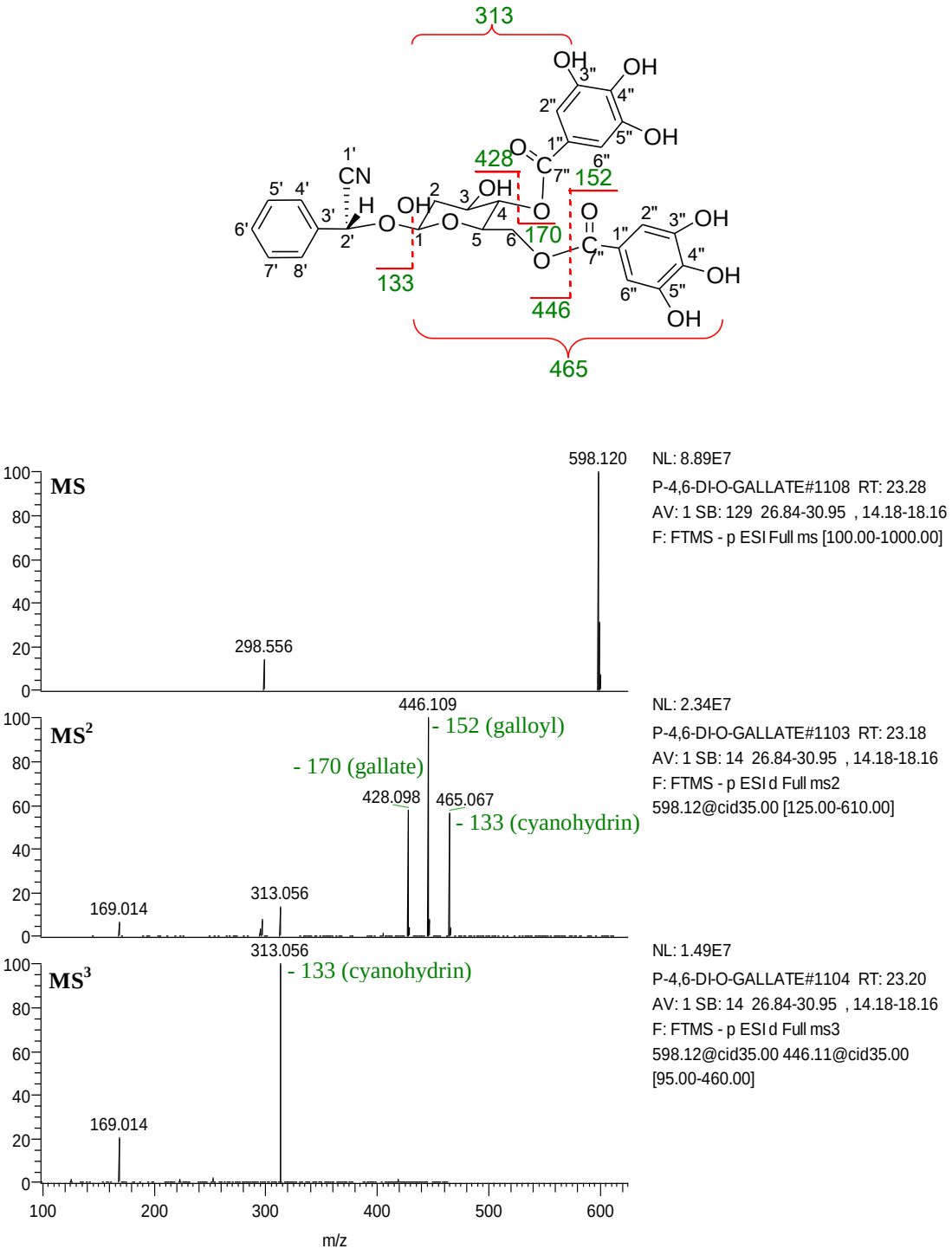


Figure 3.1.1.7: Fragmentation pattern for prunasin 4',6'-di-*O*-gallate (**4**).

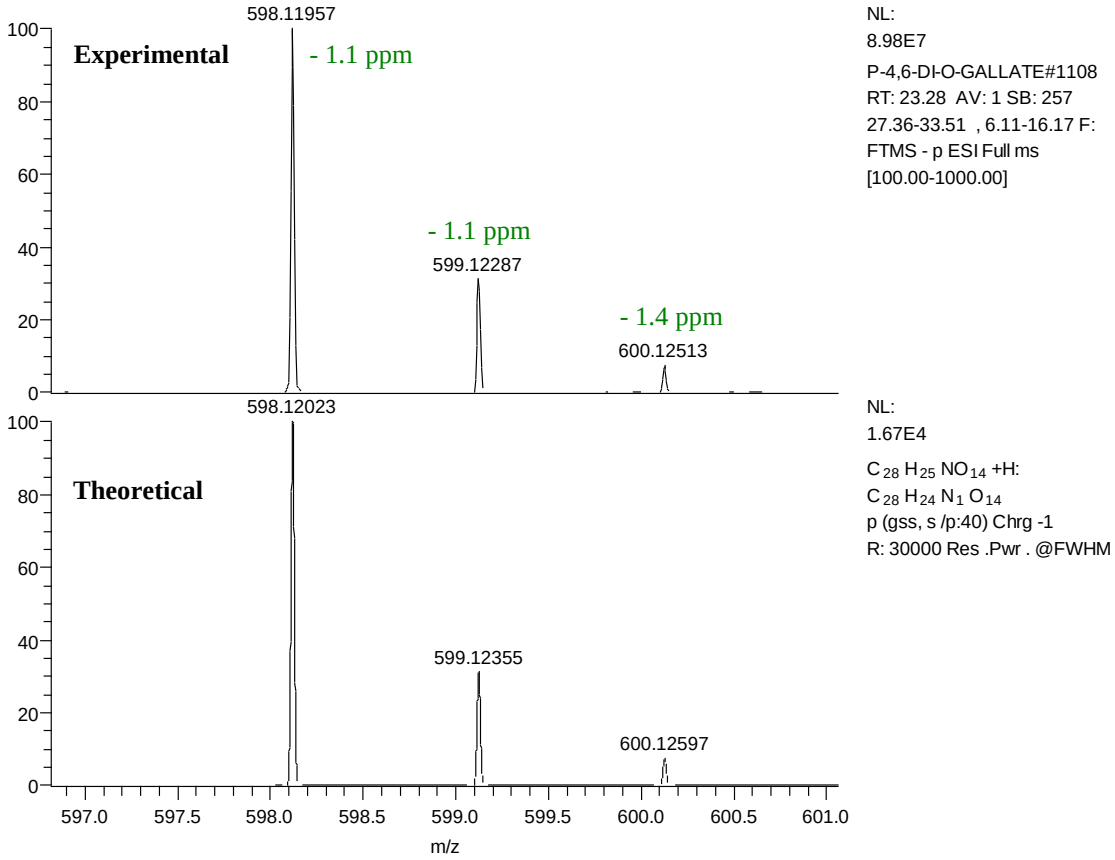
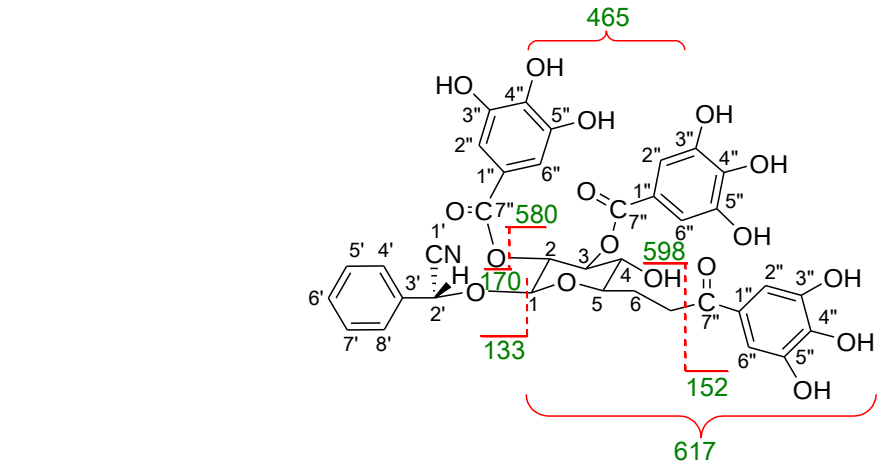


Figure 3.1.1.8: Full scan mass spectra (*m/z*) with isotopic pattern for prunasin 4',6'-di-*O*-gallate (**4**).

Prunasin 4',6'-di-*O*-gallate (**4**), CG-4, a tan amorphous powder, with the molecular formula of $C_{21}H_{21}NO_{10}$ showed deprotonated molecular $[M–H]^-$ ion at *m/z* 598. Compound CG-4 consisted of a cyanohydrin, two galloyl groups and a glucose moiety. It has the same elemental composition with compound **2** and **3** which showed the parent ion at *m/z* 598. The fragment ions at *m/z* 446, 465 and 428 resulting from *m/z* 598 in MS^2 spectrum (Figure 3.1.1.7) represented the mass loss of a galloyl moiety (*m/z* 152), cyanohydrin (*m/z* 133) and gallate (*m/z* 170). Fragmentation of major ion at *m/z* 446 in MS^3 spectrum gave ions at *m/z* 313 and 169 which were attributed by the mass loss of cyanohydrin (*m/z* 133) and sugar (*m/z* 144), respectively. Compounds **2**, **3** and **4** were found to exhibit similar fragmentation due to their isomeric nature. Figure 3.1.1.8 shows

the isotopic pattern of compound CG-4 together with the calculated mass accuracy. The ^{1}H and ^{13}C-NMR spectral data (Appendix I) for compound CG-4 are presented in Tables 3.1.1.1 and 3.1.1.2. The proton and carbon signals of the aglycon included a singlet oxymethine proton (δ 5.89), five aromatic protons from benzene ring (δ 7.45-7.60), six aromatic carbons (δ 128.6-134.4), a methine carbon (δ 118.8) and a quaternary carbon (δ 68.2). Both compounds CG-4 and CG-3 displayed similar chemical shifts for the aglycon. The downfield shift of H-4 and H-6, and upfield shift of C-2, C-3 and C-5 suggested the galloyl linkages at C-4 and C-6 of the glucose unit. Thus, compound CG-4 was identified as prunasin 4',6'-di-*O*-gallate (Ling *et al.*, 2002).

(e) Prunasin 2',3',6'-tri-*O*-gallate (5), CG-5



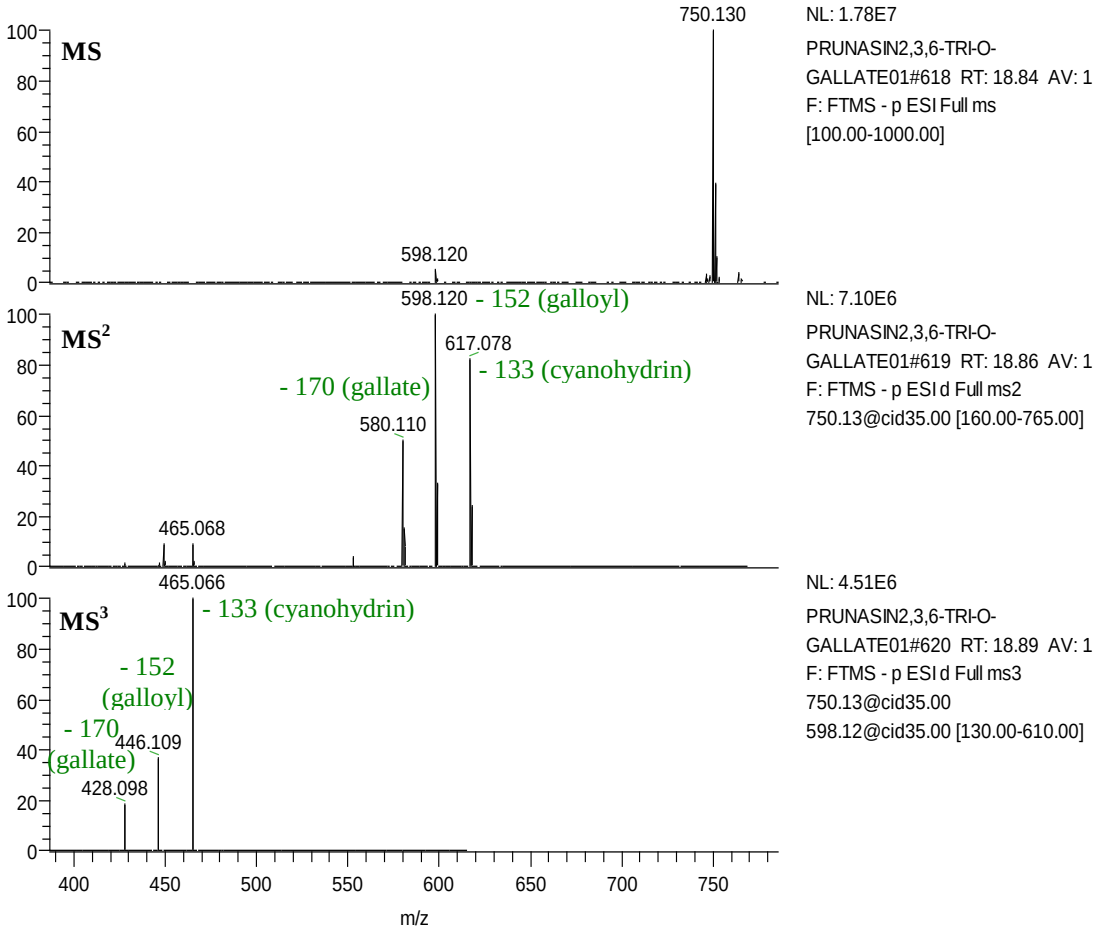


Figure 3.1.1.9: Fragmentation pattern for prunasin 2',3',6'-tri-*O*-gallate (**5**).

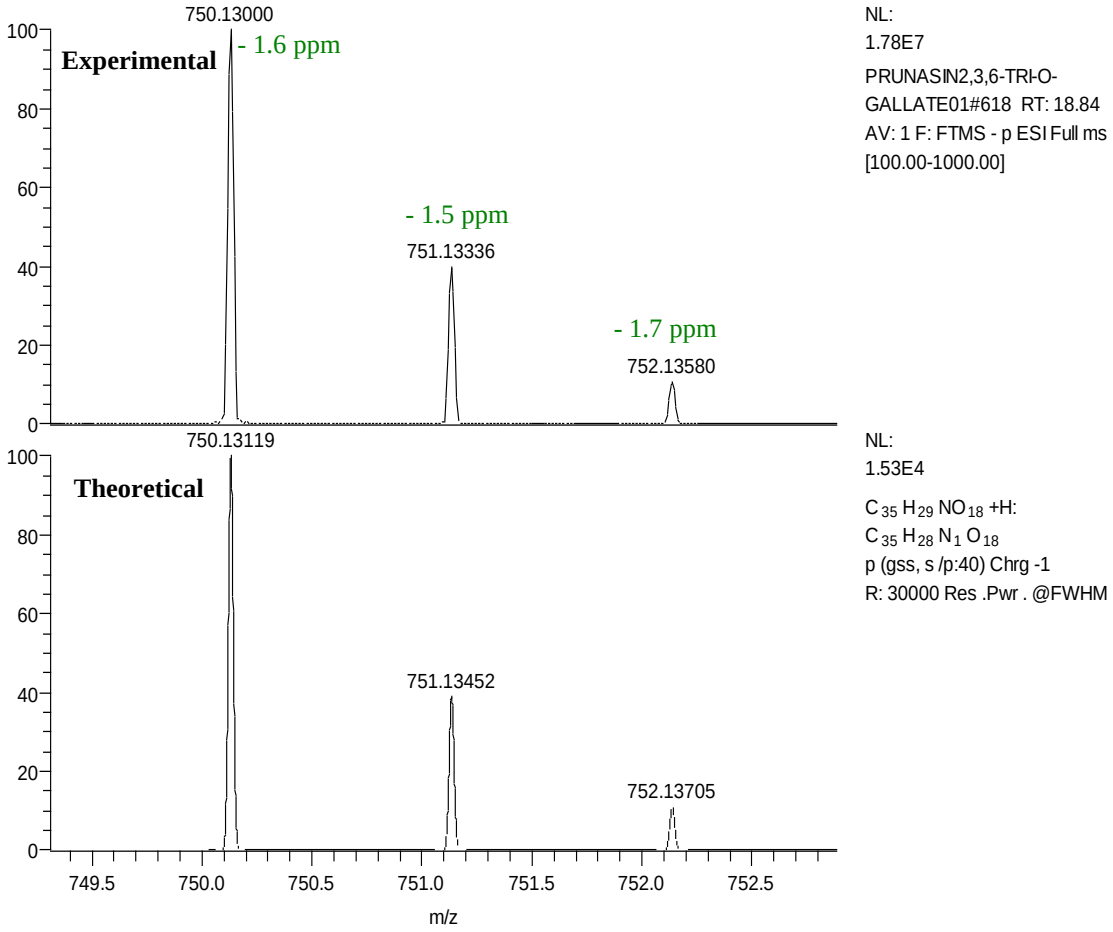
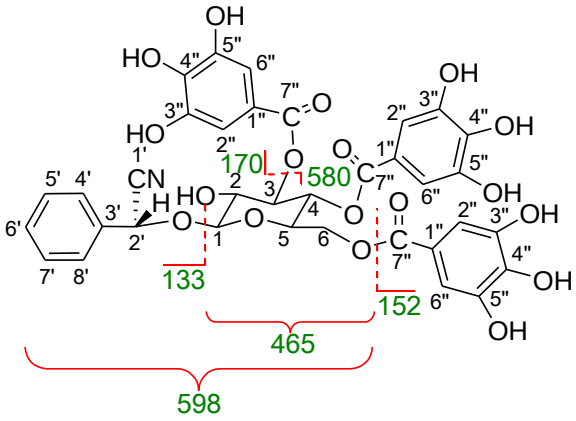


Figure 3.1.1.10: Full scan mass spectra (*m/z*) with isotopic pattern for prunasin 2',3',6'-tri-*O*-gallate (**5**).

Prunasin 2',3',6'-tri-*O*-gallate (**5**), CG-5, a tan amorphous powder, with the molecular formula of $C_{35}H_{29}NO_{18}$ showed molecular $[M–H]^-$ ion at *m/z* 750. Compound CG-5 is consisted of one cyanohydrin, three galloyl groups and a glucose moiety. The ion at *m/z* 598, 580 and 617 from parent ion at *m/z* 750 presented the removal of a galloyl moiety (*m/z* 152), a gallate (*m/z* 170) and a cyanohydrin (*m/z* 133), respectively (Figure 3.1.1.9). Further fragmentation of major ion at *m/z* 598 gave ion at *m/z* 465, 446 and 428 in the MS^3 spectrum. These *m/z* values were corresponded to the mass loss of a cyanohydrin (*m/z* 133), galloyl moiety (*m/z* 152) and gallate (*m/z* 170). The MS^3 mass spectrum of prunasin 2',3',6'-tri-*O*-gallate (**5**) was also found similar to that of prunasin-digallate (**2-4**).

The isotopic pattern and the mass accuracy measurement of compound CG-5 are presented in Figure 3.1.1.10. The proton signals at δ 5.95 and δ 7.30-7.46 represented the aglycon oxymethine and the other five aromatic protons from benzene ring, respectively (Appendix I). The ^{13}C-NMR spectrum exhibited six aromatic carbon signals (δ 128.0-134.2), a methine carbon (δ 118.4) and a quaternary carbon (δ 68.7). The NMR spectral data of compound CG-5 is tabulated in Tables 3.1.1.1 and 3.1.1.2. Based on the downfield chemical shift of H-2, H-3 and H-6 and the upfield shift of C-4, the galloyl groups were determined to be esterified at C-2, C-3 and C-6 of the glucose moiety. As a result, compound CG-5 was identified as prunasin 2',3',6'-tri-*O*-gallate (Ling *et al.*, 2002).

(f) Prunasin 3',4',6'-tri-*O*-gallate (6), CG-6



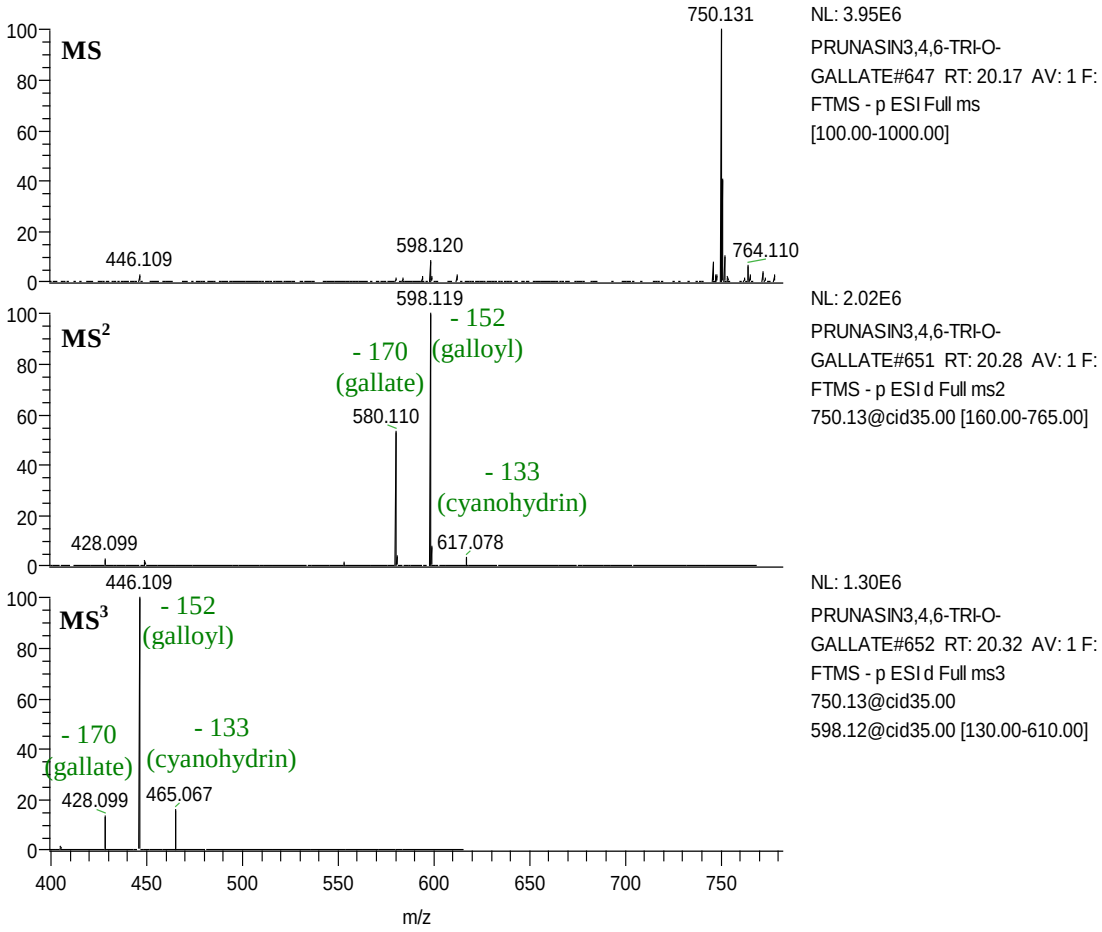


Figure 3.1.1.11: Fragmentation pattern for prunasin 3',4',6'-tri-*O*-gallate (**6**).

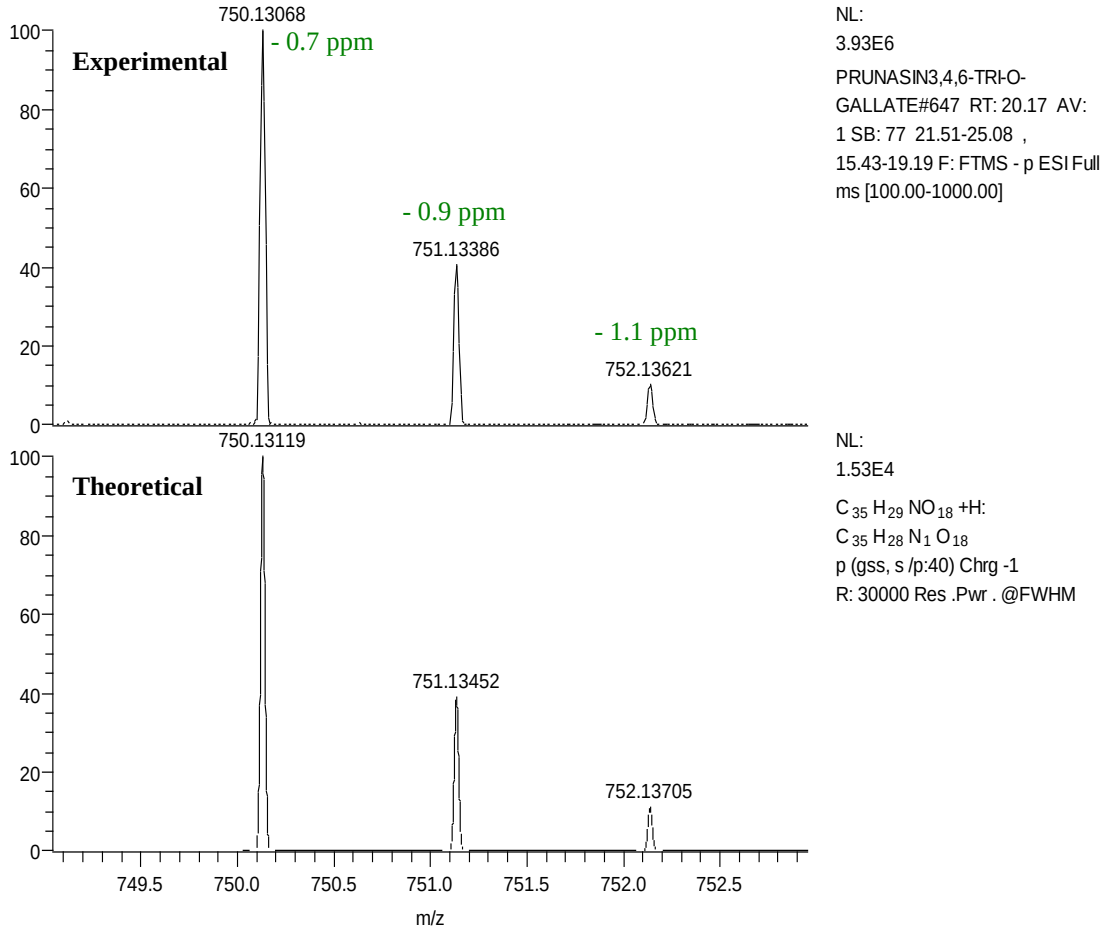
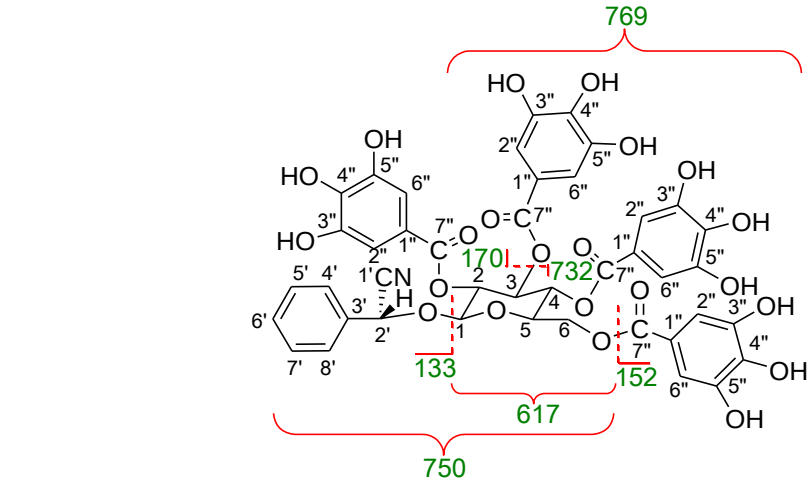


Figure 3.1.1.12: Full scan mass spectra (*m/z*) with isotopic pattern for prunasin 3',4',6'-tri-*O*-gallate (**6**).

Prunasin 3',4',6'-tri-*O*-gallate (**6**), CG-6, a tan amorphous powder, with the molecular formula of $C_{35}H_{29}NO_{18}$ showed deprotonated molecular $[M–H]^-$ ion at *m/z* 750. Compound CG-6 is consisted of a cyanohydrin, three galloyl groups and a glucose moiety. Similar to compound **5**, the parent ion at *m/z* 750 was fragmented to major ions at *m/z* 598, 580 and 617 which indicated the mass loss of a galloyl moiety (*m/z* 152), gallate (*m/z* 170) and cyanohydrin (*m/z* 133), respectively (Figure 3.1.1.11). As compared to the MS^3 spectrum of compound **5**, the product ion *m/z* 446 in MS^3 spectrum was found to be more intense in compound **6**. Hence, compound **6** was more susceptible to cleavage of the galloyl group (*m/z* 152) instead of cyanohydrin group (*m/z* 133). Basically, the fragmentation

characteristic of compounds **5** and **6** were similar since they were isomers and the isotopic pattern of compound CG-6 is presented in Figure 3.1.1.12. The ^{1}H and ^{13}C-NMR spectral data for compound CG-6 are summarized in Tables 3.1.1.1 and 3.1.1.2. Comparatively to compound CG-5, compound CG-6 has downfield proton signals for the agylcon that included a singlet oxymethine proton signal at δ 5.97 and five aromatic protons from benzene ring at δ 7.46-7.62 (Appendix I). Similarly, the aglycon indicated six aromatic carbon signals at δ 128.6-134.3, a methine carbon at δ 118.7 and a quaternary carbon at δ 68.4. The galloyl esterifications at C-3, C-4 and C-6 were determined by the downfield signals at H-3, H-4 and H-6 as well as the upfield shift at C-2. Based on the ^{1}H and ^{13}C-NMR spectral, compound CG-6 was determined as prunasin 3',4',6'-tri-*O*-gallate (Ling *et al.*, 2002).

(g) Prunasin 2',3',4',6'-tetra-*O*-gallate (7), CG-7



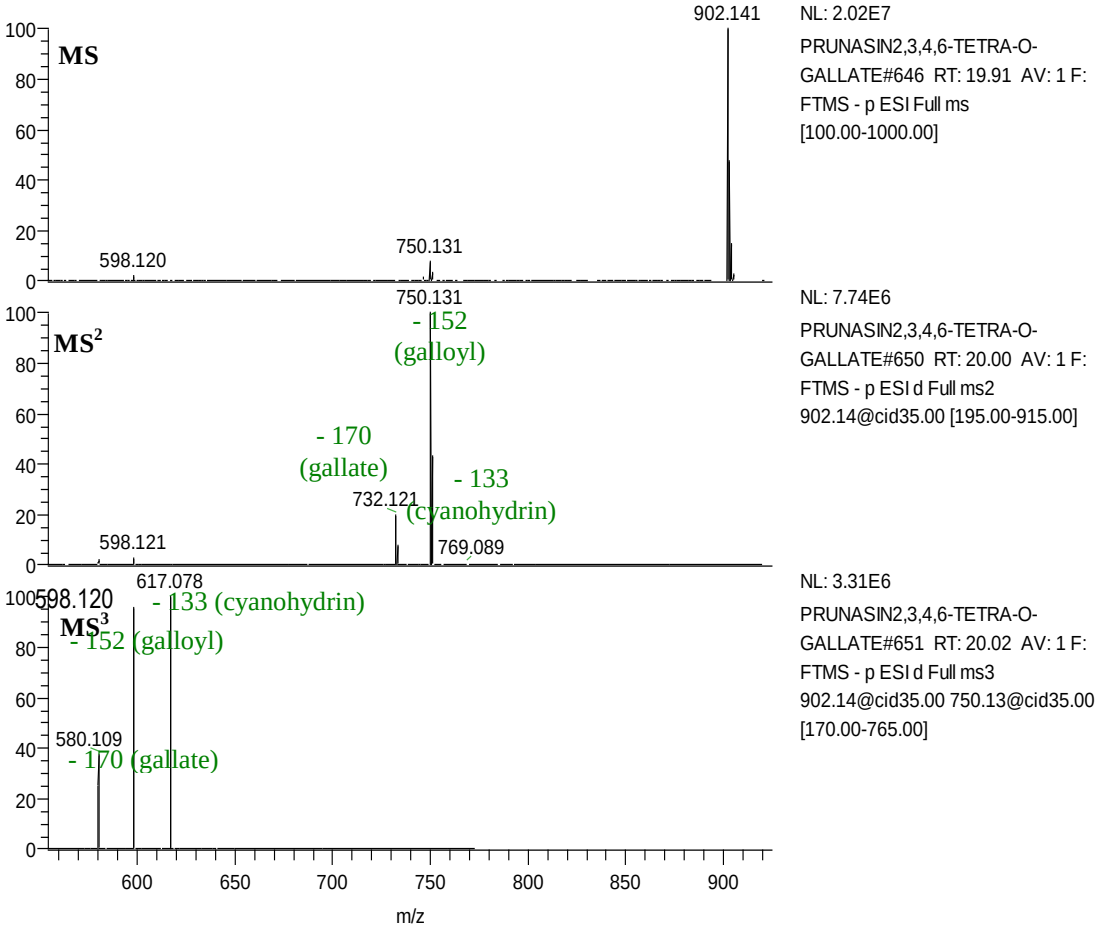


Figure 3.1.1.13: Fragmentation pattern for prunasin 2',3',4',6'-tetra-*O*-gallate (**7**).

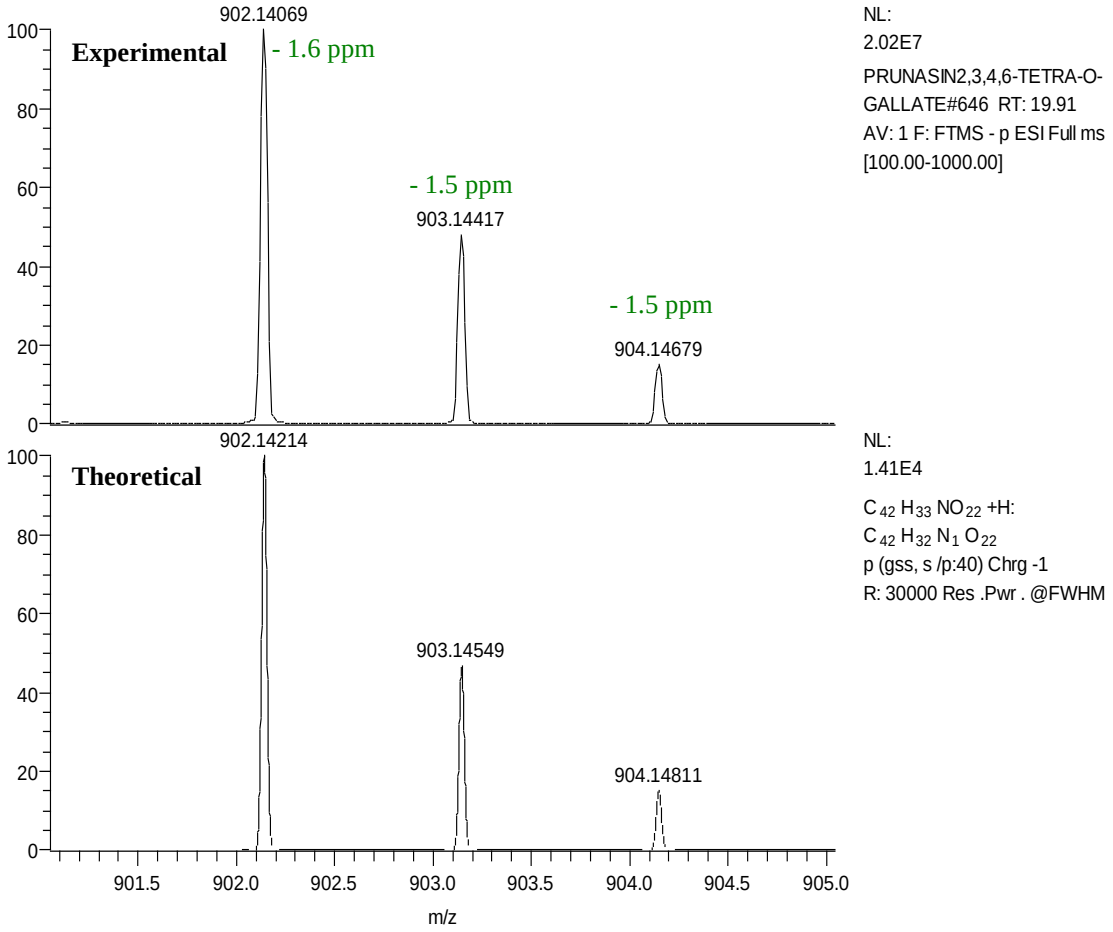


Figure 3.1.1.14: Full scan mass spectra (*m/z*) with isotopic pattern for prunasin 2',3',4',6'-tetra-*O*-gallate (**7**).

Prunasin 2',3',4',6'-tetra-*O*-gallate (**7**), CG-7, a tan amorphous powder, with the molecular formula of $C_{42}H_{33}NO_{22}$ showed deprotonated molecular $[M–H]^-$ ion at *m/z* 902. Compound CG-7 is consisted of a cyanohydrin, four galloyl groups and a glucose moiety. A major product ion, *m/z* 750 resulted from the fragmentation of ion at *m/z* 902 indicated the removal of a galloyl group (*m/z* 152). The other ions signals at *m/z* 732 and 769 exhibited the mass loss of a gallate (*m/z* 170) and cyanohydrin (*m/z* 133), respectively (Figure 3.1.1.13). In the MS^3 spectrum, the fragmentation of major ions at *m/z* 750 gave ion at *m/z* 617, 598 and 580. These ions were corresponded to the mass loss of a cyanohydrin (*m/z* 133), galloyl moiety (*m/z* 152) and gallate (*m/z* 170), respectively. The comparison of

isotopic pattern between experimental and theoretical mass spectra was confirmed the elemental composition of compound CG-7 (Figure 3.1.1.14). The ^{1}H and ^{13}C-NMR data for compound CG-7 are displayed in Tables 3.1.1.1 and 3.1.1.2. Five aromatic proton signals from benzene ring (δ 7.30-7.45), a singlet oxymethine proton (δ 6.00) as well as signals for six aromatic carbons (δ 128.1-134.1), a methine carbon (δ 118.3) and a quaternary carbon (δ 68.8) supported the presence of the aglycon (Appendix I). Accordingly, four galloyl esterifications corresponded to the downfield shift of H-2, H-3, H-4 and H-6 which demonstrated esterifcation at C-2, C-3, C-4 and C-6 of the glucose moiety. Hence, compound CG-7 was deduced as prunasin 2',3',4',6'-tetra-*O*-gallate (Ling *et al.*, 2002).

(h) Fragmentation Scheme of Galloylated Cyanogenic Glucosides

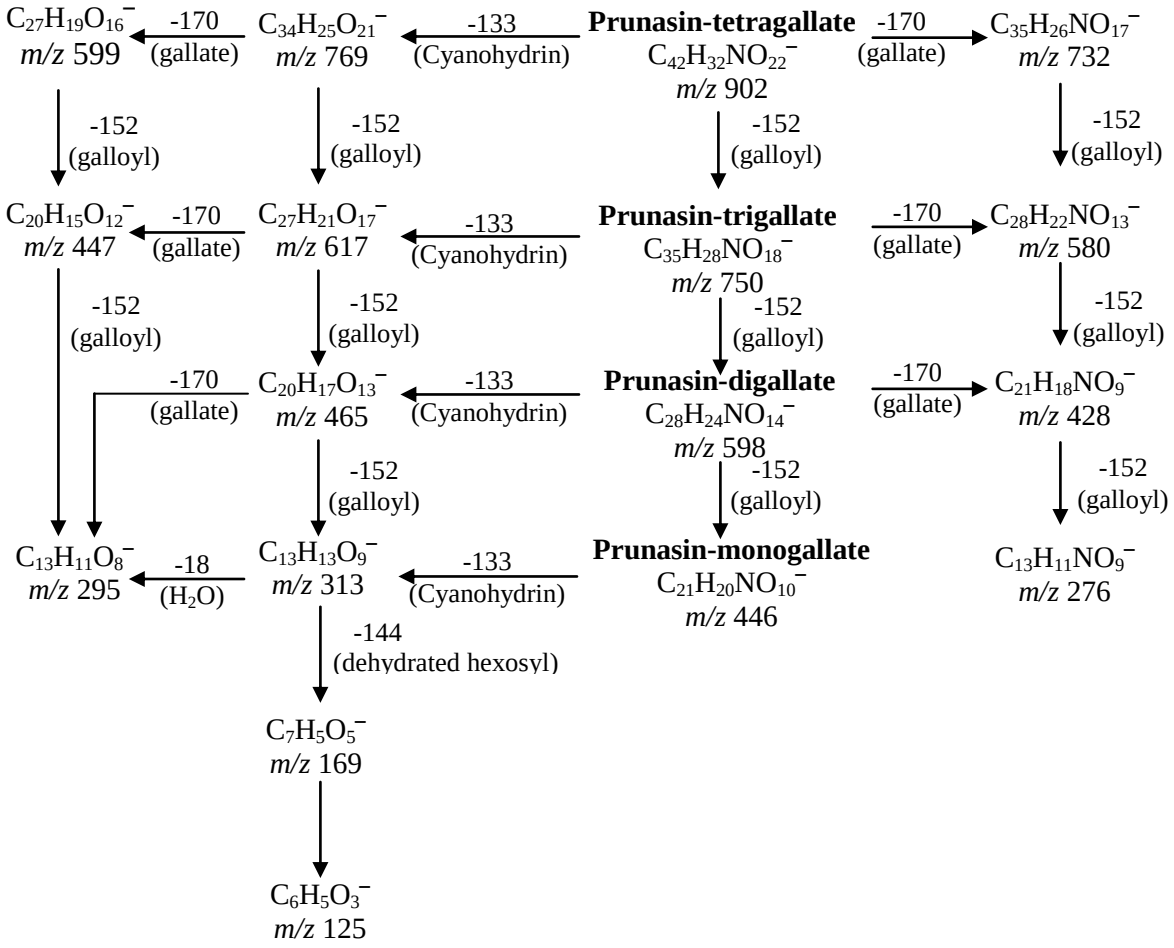


Figure 3.1.1.15: Fragmentation scheme for galloylated cyanogenic glucosides (**1-7**) observed by data dependent MS^n with dynamic exclusion in negative mode.

Figure 3.1.1.15 summarizes the fragmentation scheme for the galloylated cyanogenic glucosides (**1-7**) based on ESI-MS^n spectra (Figures 3.1.1.1-3.1.1.14). A previously reported MS^n analysis in positive mode for cyanogenic glucosides gave $[M+Na]^+$ and $[M+NH_4]^+$, while the collision-induced dissociation (CID) showed fragmentation of HCN and aglycone from the sodium adduct parent ion and subsequently derived $[M+Na–HCN]^+$ and [dehydrated glucose + Na] $^+$ (Franks *et al.*, 2005). So far, no fragmentation

study has been reported for galloylated cyanogenic glucosides. In this study, the galloylated cyanogenic glucosides [prunasin 6'-*O*-gallate (**1**), prunasin 2',6'-di-*O*-gallate (**2**), prunasin 3',6'-di-*O*-gallate (**3**), prunasin 4',6'-di-*O*-gallate (**4**), prunasin 2',3',6'-tri-*O*-gallate (**5**), prunasin 3',4',6'-tri-*O*-gallate (**6**) and prunasin 2',3',4',6'-tetra-*O*-gallate (**7**)] were fragmented using negative ionisation that generated deprotonated parent ions $[M–H]^-$. Compounds **1**, **2**-**4**, **5**-**6** and **7** were referred to as prunasin-monogallate, prunasin-digallate, prunasin-trigallate and prunasin-tetragallate, respectively (Figure 3.1.1.15). The ESI-MS spectrum of prunasin-tetragallate showed $[M–H]^-$ ion at *m/z* 902, while the fragment ions at $[M–H–133]^-$, $[M–H– 52]^-$ and $[M–H–170]^-$ indicated the mass loss of cyanohydrin, galloyl moiety and gallate, respectively (Figure 3.1.1.15). Based on the ESI-MS^n spectra of galloylated cyanogenic glucosides (Figures 3.1.1.1, 3.1.1.3, 3.1.1.5, 3.1.1.7, 3.1.1.9, 3.1.1.11 and 3.1.1.13), similar fragmentation patterns were also observed in prunasin-trigallate (*m/z* 750) and prunasin-digallate (*m/z* 598) but not in prunasin-monogallate (*m/z* 446). The major fragment ion of prunasin-monogallate at *m/z* 313 $[M–H–133]^-$ underwent a neutral mass loss of *m/z* 144 (dehydrated hexosyl) indicating scission of a hexosyl unit to yield deprotonated gallic acid (Figure 3.1.1.1 and Table 3.1.2). The deprotonated gallic acid at *m/z* 169 further underwent decarboxylation with mass loss of *m/z* 44 to form a trihydroxyphenol moiety at *m/z* 125. Product ions at *m/z* 769, 617, 465 and 313 were also found in the fragmentation scheme of gallotannins (Figures 3.1.1.15 and 3.1.2.15). These product ions of galloylated cyanogenic glucosides are believed to demonstrate a similar fragmentation pathway as presented in Figure 3.1.2.15.

Table 3.1.1.1: ^{1}H-NMR Spectral data for compounds CG-1 to CG-7 (**1-7**) in Acetone-d_6.

	δ_H						
	CG-1 (1)	**CG-2 (2)**	**CG-3 (3)**	**CG-4 (4)**	**CG-5 (5)**	**CG-6 (6)**	**CG-7 (7)**
Aglycon							
H-2'	5.85 s	5.90 s	5.92 s	5.89 s	5.95 s	5.97 s	6.00 s
H-4', -8'	7.55-7.58 m	7.19-7.46 m	7.58-7.61 m	7.57-7.60 m	7.30-7.46 m	7.59-7.62 m	7.30-7.45 m
H-5', -7'	7.43-7.45 m	7.19-7.46 m	7.46-7.48 m	7.45-7.47 m	7.30-7.46 m	7.46-7.48 m	7.30-7.45 m
H-6'	7.43-7.45 m	7.19-7.46 m	7.46-7.48 m	7.45-7.47 m	7.30-7.46 m	7.46-7.48 m	7.30-7.45 m
Glucose							
H-1	4.41 d (6.9)	4.61 d (7.8)	4.62 d (7.8)	4.52 d (7.8)	5.02 d (7.8)	4.77 d (8.1)	5.18 d (8.1)
H-2	3.66-3.31 m	5.13 dd (7.8, 9.0)	3.68 dd (7.8, 9.3)	3.53 dd (7.8, 8.7)	5.31 dd (7.8, 9.6)	3.84 dd (8.1, 9.0)	5.46 dd (8.1, 9.6)
H-3	3.31-3.66 m	3.74-3.86 m	5.17 dd (8.7, 9.3)	3.76 dd (8.7, 9.3)	5.47 dd (8.7, 9.6)	5.52 dd (9.0, 9.6)	5.80 t (9.6)
H-4	3.31-3.66 m	3.74-3.86 m	3.84 dd (8.7, 9.9)	5.15 dd (9.3, 9.9)	4.05 dd (8.7, 9.3)	5.40 dd (9.6, 9.9)	5.59 t (9.6)
H-5	3.31-3.66 m	3.74-3.86 m	3.77 m	3.92 m	3.97 m	4.16 m	4.35 m
H-6a	4.59 dd (1.8, 12.0)	4.45 br d (12.0)	4.50 dd (4.8, 12.0)	4.21 dd (6.0, 12.3)	4.60 dd (3.9, 12.0)	4.34 dd (5.4, 12.0)	4.43 dd (5.4, 12.0)
H-6b	4.38 dd (5.4, 12.0)	4.52 br d (12.0)	4.61 dd (2.1, 12.0)	4.49 dd (2.1, 12.3)	4.64 br d (12.0)	4.50 dd (1.5, 12.0)	4.55 br d (12.0)
Galloyl							
H-2", -6"	7.20 s	7.11, 7.13 s	7.14, 7.22 s	7.11, 7.20 s	7.01, 7.04, 7.24 s	7.00, 7.03, 7.22 s	6.94, 7.01, 7.02, 7.24 s

Table 3.1.1.2: ^{13}C-NMR Spectral data for compounds CG-1 to CG-7 (**1**-**7**) in Acetone-d_6.

	δ_C						
	CG-1 (1)	**CG-2 (2)**	**CG-3 (3)**	**CG-4 (4)**	**CG-5 (5)**	**CG-6 (6)**	**CG-7 (7)**
Aglycon							
C-1'	118.8	118.7	118.8	118.8	118.4	118.7	118.3
C-2'	68.1	68.4	68.3	68.2	68.7	68.4	68.8
C-3'	134.5	134.2	134.4	134.4	134.2	134.3	134.1
C-4', -8'	128.6	128.4	128.5	128.6	128.0	128.6	128.1
C-5', -7'	129.8	129.8	129.8	129.9	129.8	129.9	129.9
C-6'	130.6	130.6	130.6	130.7	130.6	130.7	130.7
Glucose							
C-1	101.8	101.8	102.0	101.7	100.3	100.3	100.0
C-2	74.5	75.2	74.4	74.7[a]	72.2	72.8[c]	72.1
C-3	77.6	74.3	73.4	73.4[a]	75.4[b]	73.3[c]	73.2[d]
C-4	71.0	70.8	69.4	71.8	69.2	69.6	69.4
C-5	75.4	78.3	75.3	75.3[a]	75.8[b]	75.5[c]	73.4[d]
C-6	64.2	64.1	64.0	63.6	63.7	63.2	62.9
Galloyl							
C-1"	121.7	121.1, 121.9	121.6, 121.9	121.3, 121.6	120.8, 121.2, 121.6	120.5, 121.2, 121.5	120.3, 120.4, 120.6, 121.4
C-2", -6"	110.0	109.8, 109.9	109.9, 110.0	110.0, 110.1	109.9, 110.0, 110.1	109.9, 110.0, 110.1	109.9, 110.0, 110.1 × 2
C-3", -5"	146.0	145.8, 145.9	145.8, 146.0	145.8, 145.9	145.7, 145.8, 145.9	145.7, 145.8, 145.9	145.6, 145.7, 145.8, 145.9
C-4"	138.8	138.9, 139.1	138.6, 138.8	138.8, 138.9	138.6, 138.7, 138.9	138.7, 138.8, 139.1	138.8, 138.9, 139.0, 139.1
C-7"	166.7	166.5, 166.8	166.4, 166.7	166.0, 166.5	165.5, 166.1, 166.6	165.8, 166.1, 166.4	165.4, 165.6, 165.8, 166.4

a-c values with the same superscript may be interchanged.

3.1.2 Gallotannins

(a) 6-O-Galloyl-D-glucose (8), GT-1

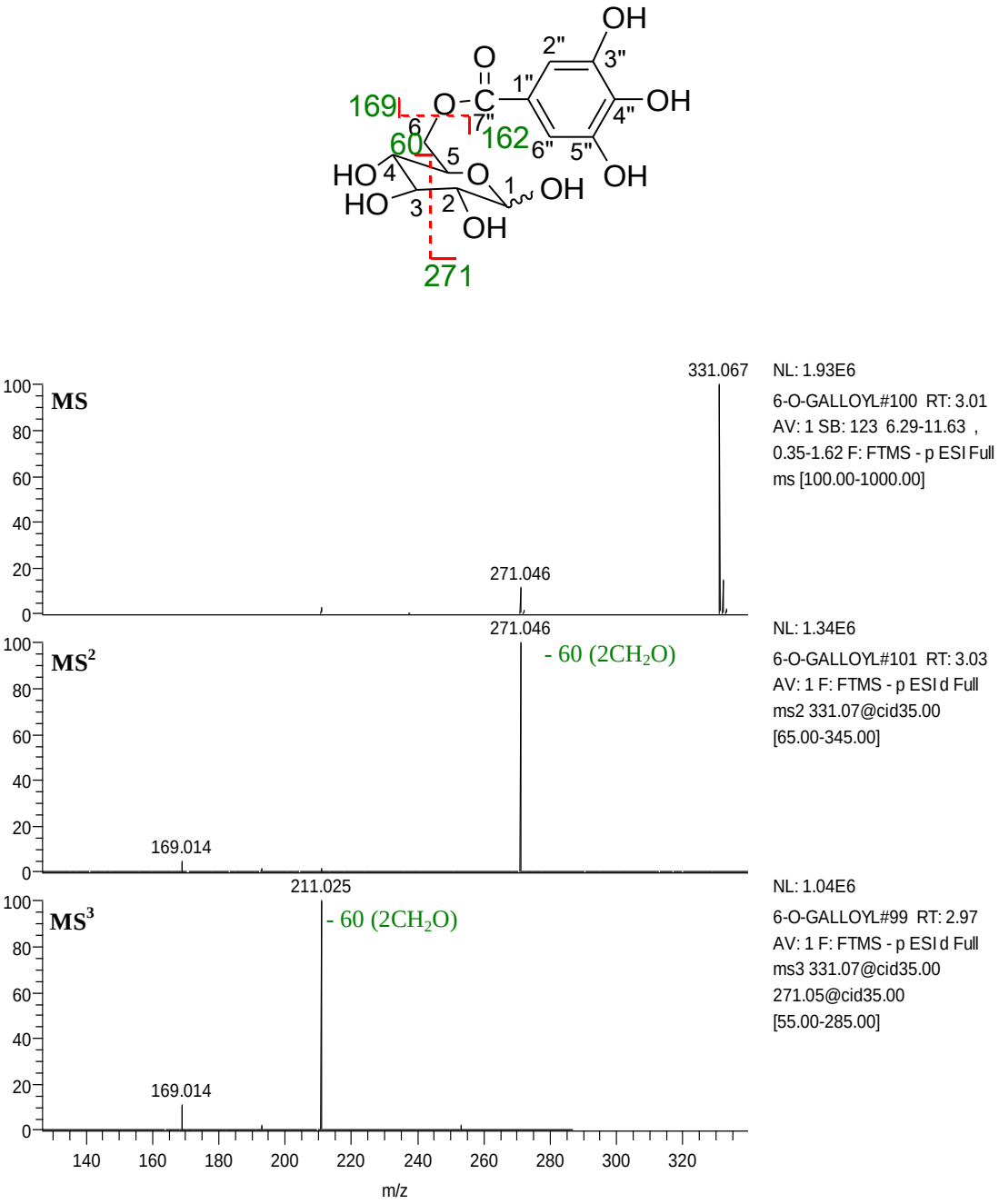


Figure 3.1.2.1: Fragmentation pattern for 6-*O*-galloyl-D-glucose (**8**).

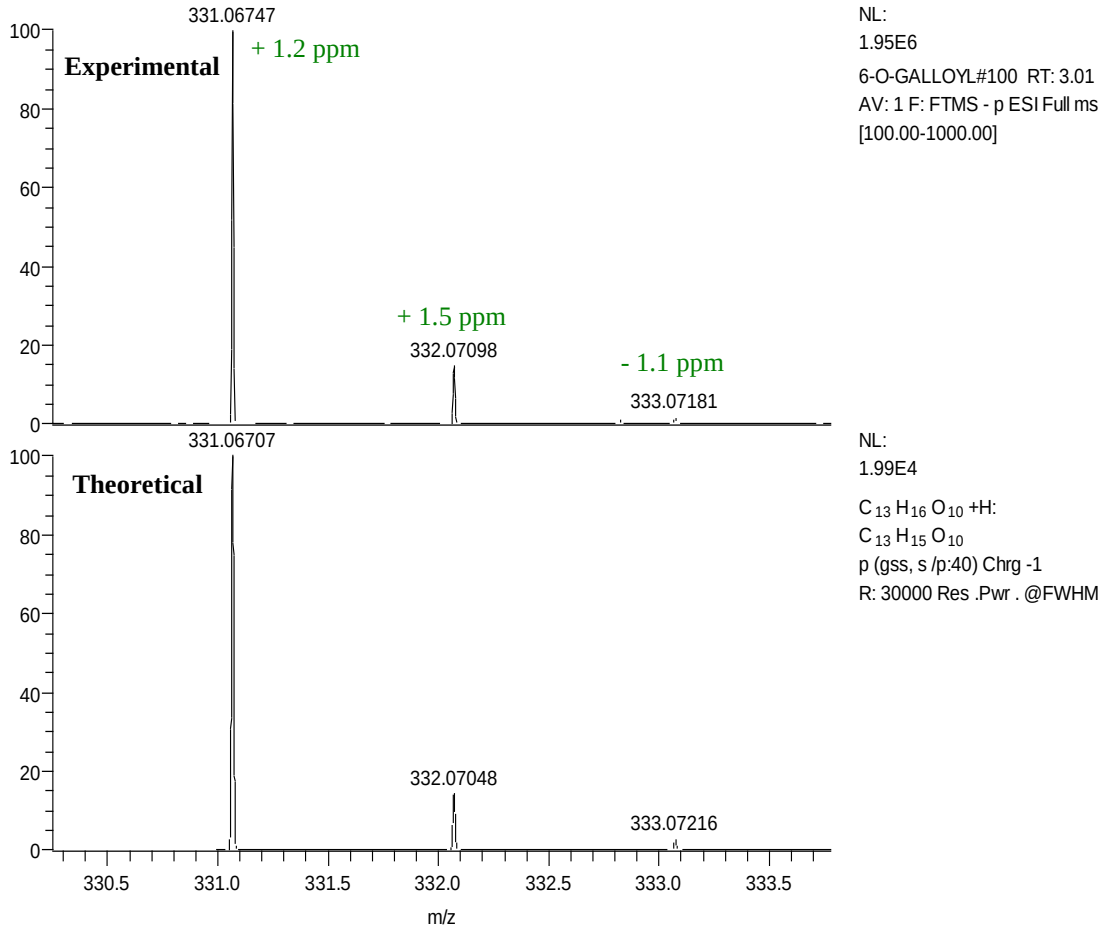


Figure 3.1.2.2: Full scan mass spectra (*m/z*) with isotopic pattern for 6-*O*-galloyl-D-glucose (**8**).

6-*O*-Galloyl-D-glucose (**8**), GT-1, an amorphous powder, with the molecular formula of $C_{13}H_{16}O_{10}$ showed deprotonated molecular $[M–H]^-$ ion at *m/z* 331. Compound GT-1 has one galloyl moiety esterified to a glucose unit. The ion at *m/z* 331 was fragmented to *m/z* 271 after the loss of two formaldehyde (CH_2O) groups which represented part of the glucose fragmentation (Taylor *et al.*, 2005). Further fragmentation of ion at *m/z* 271 yielded ions at *m/z* 211 and 169. The ion at *m/z* 211 indicated the mass loss of two formaldehyde (CH_2O) groups whereas fragment ion at *m/z* 169 exhibited deprotonated gallic acid (Figure 3.1.2.1). The elemental composition of compound GT-1 was further characterised by comparing the isotopic pattern with theoretical mass spectra

(Figure 3.1.2.2). The ^{1}H and ^{13}C-NMR showed α and β-anomeric proton signal for glucose moiety at δ 5.10, d, J = 3.6 Hz and at δ 4.54, d, J = 7.8 Hz, and α, β-anomeric carbon signals at δ 93.4 and δ 97.9 (Table 3.1.2.1). Proton signals in the range of δ 3.16-4.35 represented the H-2, H-3 and H-4 of the glucose moiety. A singlet signal at δ 7.10 corresponding to 2H was attributed by H-2'' and H-6'' of the galloyl group which suggested the presence of one galloyl moiety (Appendix I). Therefore, compound GT-1 was characterised as 6-*O*-galloyl-D-glucose (Nonaka *et al.*, 1983; He *et al.*, 2001).

(b) 3,6-Di-*O*-galloyl-D-glucose (9), GT-2

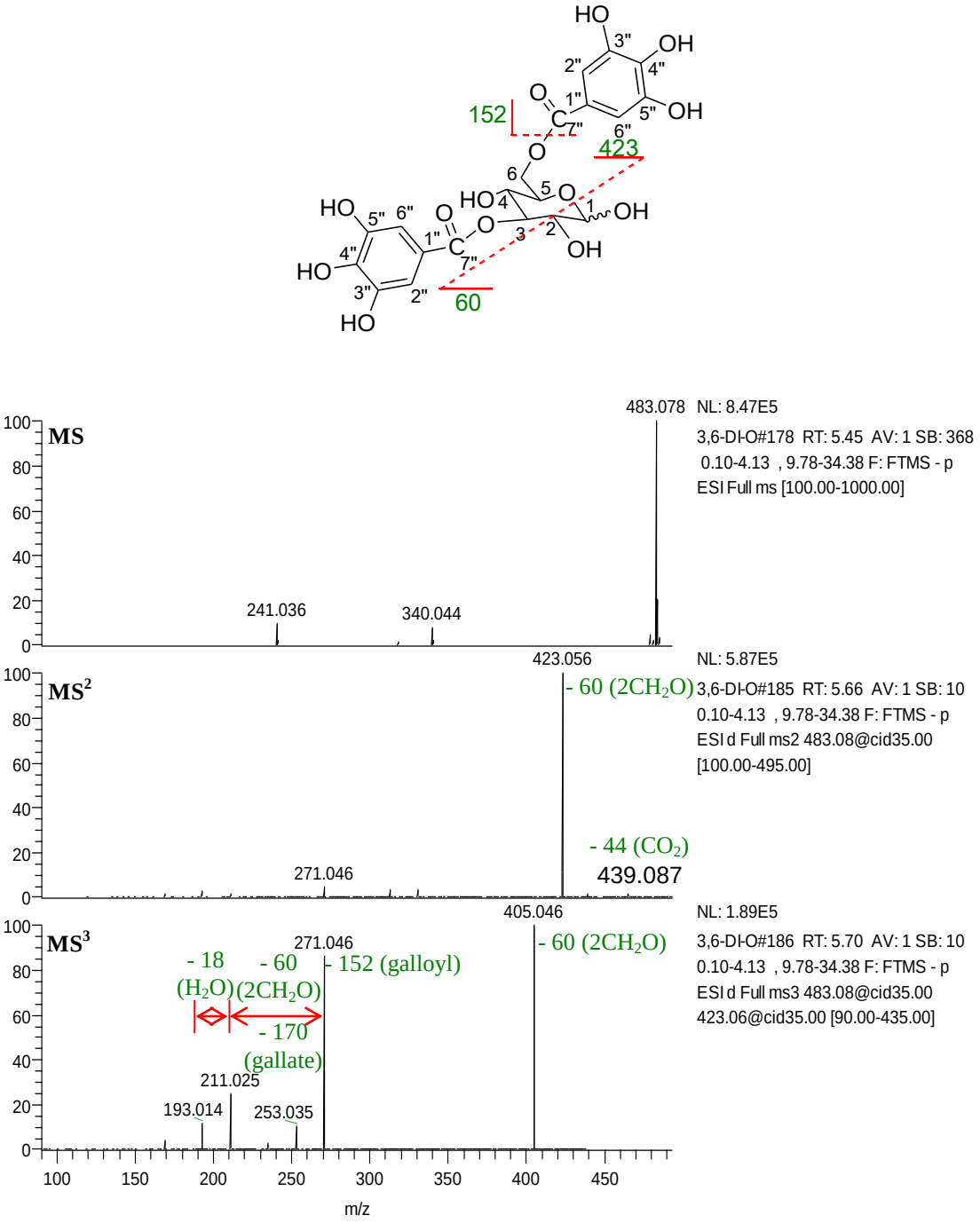


Figure 3.1.2.3: Fragmentation pattern for 3,6-di-*O*-galloyl-D-glucose (**9**).

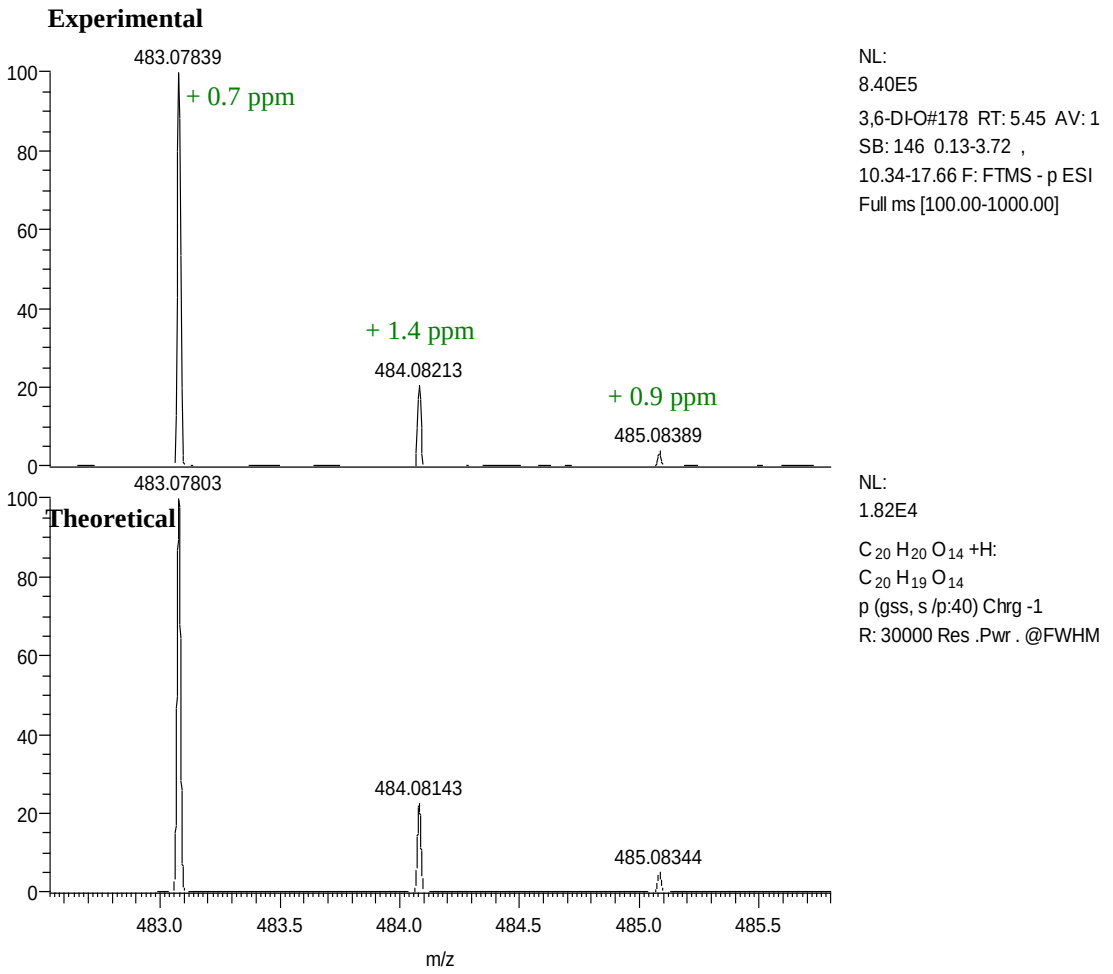


Figure 3.1.2.4: Full scan mass spectra (*m/z*) with isotopic pattern for 3,6-di-*O*-galloyl-D-glucose (**9**).

3,6-Di-*O*-galloyl-D-glucose (**9**), GT-2, a light brown amorphous powder, with the molecular formula of $C_{20}H_{20}O_{14}$ showed deprotonated molecular $[M–H]^-$ ion at *m/z* 483. Compound GT-2 consists of two galloyl groups and one glucose unit. Similar to compound **8**, ion at *m/z* 483 displayed the mass loss of two formaldehyde (CH_2O) groups (Taylor *et al.*, 2005) and yielded a ion at *m/z* 423. The MS^2 spectrum also exhibited the decarboxylation of parent ion at *m/z* 483 and formed *m/z* 439 (Figure 3.1.2.3). Two major fragment ions at *m/z* 405 and 271 in MS^3 spectrum indicated the removal of a water and galloyl moiety (*m/z* 152), respectively. The product ion at *m/z* 211 was suggested to result from the two formaldehyde (CH_2O) groups from *m/z* 271. The ion *m/z* 169 was also

observed in the MS^3 spectrum which corresponded to deprotonated gallic acid. The isotopic pattern of compound GT-2 is shown in Figure 3.1.2.4. The ^{1}H-NMR displayed two doublet anomeric proton signals at δ 5.21, J = 3.6 Hz and δ 4.67, J = 7.5 Hz and two singlets at δ 7.12 and δ 7.13 assignable to two galloyl groups (Table 3.1.2.1). The two anomeric carbon signals at δ 93.5 and δ 98.2 indicated that compound GT-2 was present in a mixture of α and β-anomers (Table 3.1.2.2, Appendix I). The positions of the galloyl groups were determined to be at C-3 and C-6 upon the downfield shift of H-3 and H-6. Thus, the structure of compound GT-2 was determined to be 3,6-di-*O*-galloyl-D-glucose (He *et al.*, 2001).

(c) 1,2,3-Tri-*O*-galloyl-β-D-glucose (10), GT-3

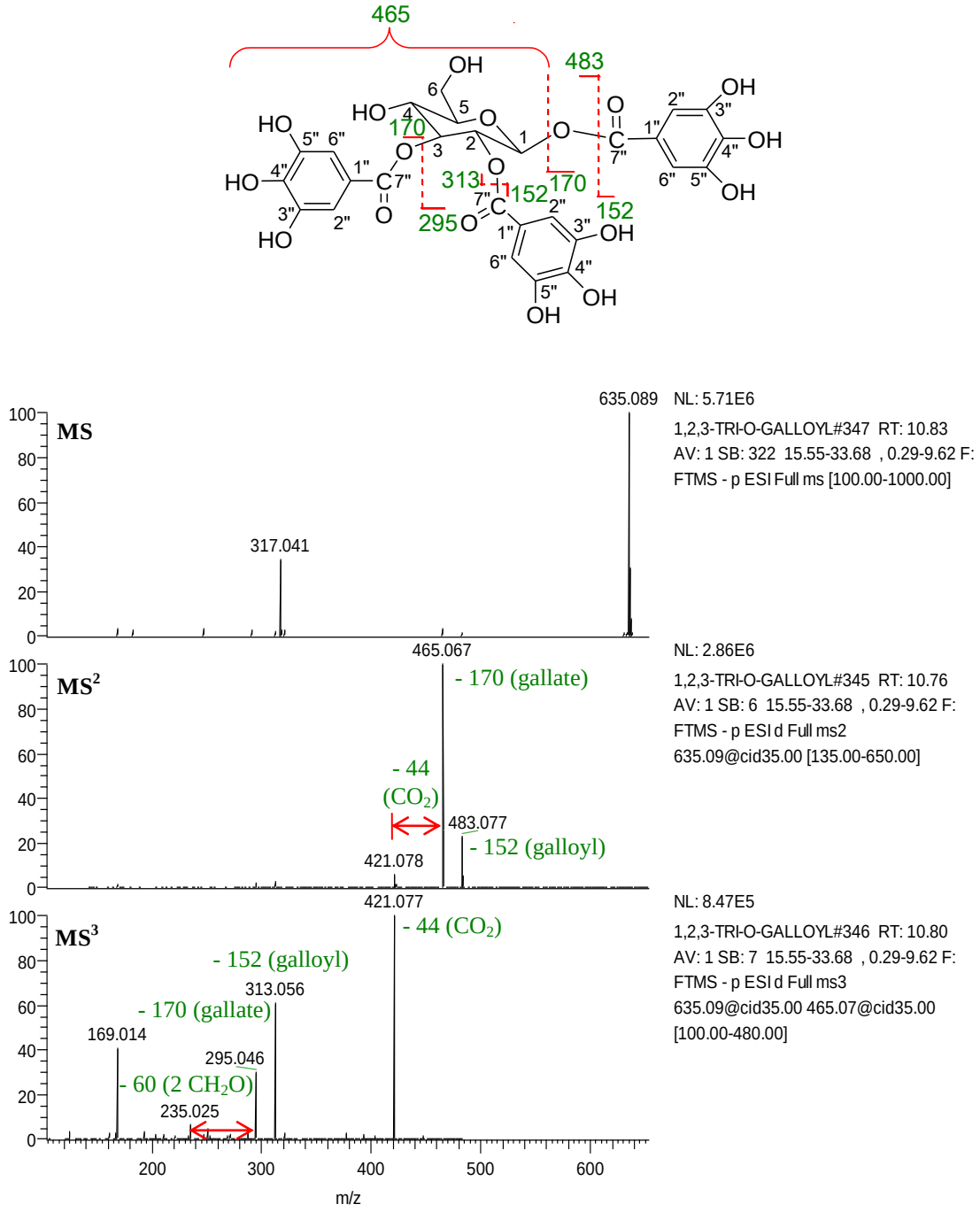


Figure 3.1.2.5: Fragmentation pattern for 1,2,3-tri-*O*-galloyl-β-D-glucose (**10**).

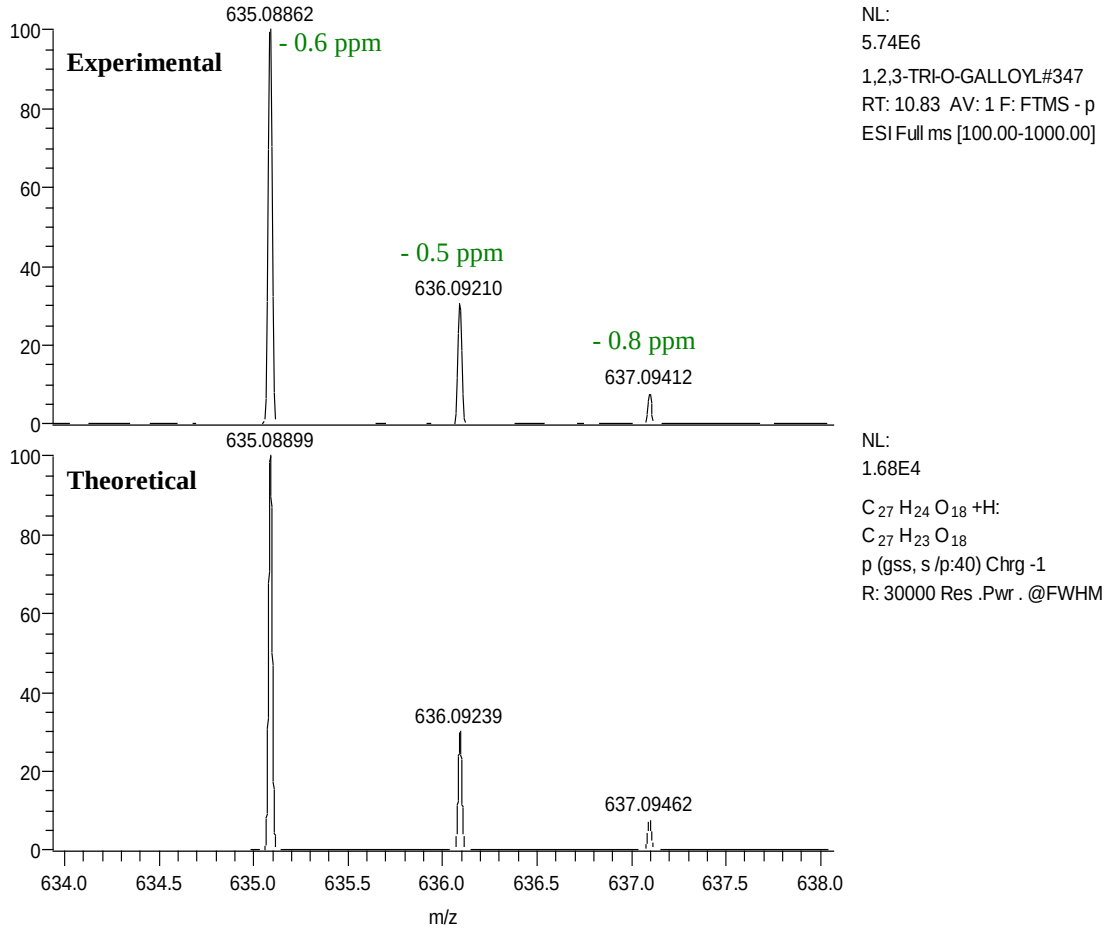


Figure 3.1.2.6: Full scan mass spectra (*m/z*) with isotopic pattern for 1,2,3-tri-*O*-galloyl-β-D-glucose (**10**).

1,2,3-Tri-*O*-galloyl-D-glucose (**10**), GT-3, a light brown amorphous powder, with the molecular formula of $C_{27}H_{24}O_{18}$ showed deprotonated molecular $[M–H]^-$ ion at *m/z* 635. Compound GT-3 contains three galloyl groups and one glucose unit. This trigalloyl-glucose has deprotonated ion at *m/z* 635 and the fragment ions showed the removal of a gallate (*m/z* 170) and galloyl moiety (*m/z* 152) at *m/z* 465 and 483, respectively (Figure 3.1.2.5). Decarboxylation of ion at *m/z* 465 yielded *m/z* 421. The mass loss of a gallate (*m/z* 170) and galloyl group (*m/z* 152) also occurred in MS^3 spectrum which corresponded to *m/z* 313 and 295, respectively. Two formaldehyde (CH_2O) groups were suggested to be removed from ion at *m/z* 295 and formed fragment ion at *m/z* 235. The ion at *m/z* 169 was

contributed by deprotonated gallic acid. Compound GT-3 was further characterised by comparing the isotopic pattern (Figure 3.1.2.6). The three singlet signals in ^{1}H-NMR at δ 6.97, δ 7.06 and δ 7.08 indicated the presence of three galloyl groups (Appendix I). The downfield shift of H-1, H-2 and H-3 indicated the esterification of galloyl groups at C-1, C-2 and C-3 position of glucose moiety (Tables 3.1.2.1 and 3.1.2.2; Appendix I). The stereochemistry of the anomeric proton was shown to be in β-configuration due to the large coupling constant (J = 8.4 Hz) at H-1. Therefore, compound GT-3 was deduced as 1,2,3-tri-*O*-galloyl-β-D-glucose (Lee *et al.*, 1990; Nawwar *et al.*, 1994).

(d) 1,4,6-Tri-*O*-galloyl-β-D-glucose (11), GT-4

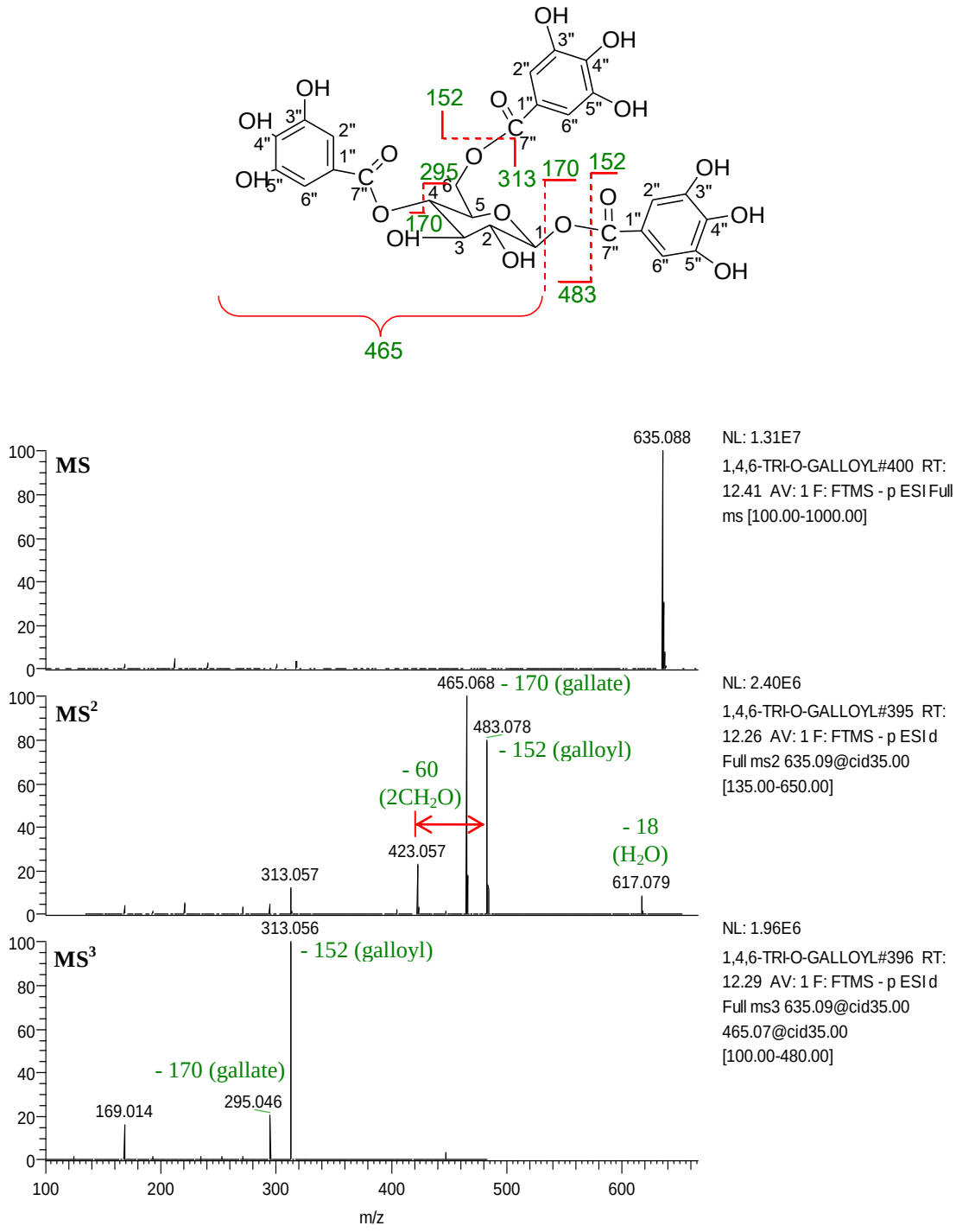


Figure 3.1.2.7: Fragmentation pattern for 1,4,6-tri-*O*-galloyl-β-D-glucose (**11**).

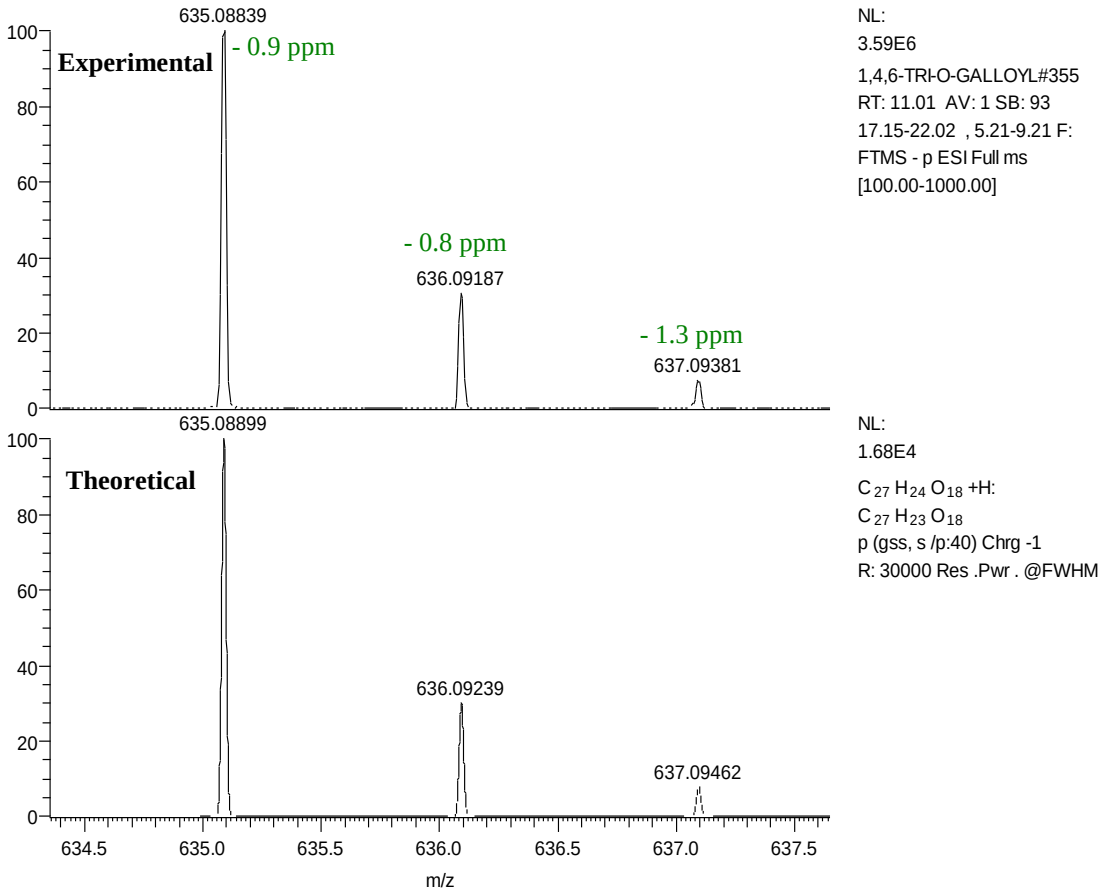


Figure 3.1.2.8: Full scan mass spectra (*m/z*) with isotopic pattern for 1,4,6-tri-*O*-galloyl-β-D-glucose (**11**).

1,4,6-Tri-*O*-galloyl-D-glucose (**11**), GT-4, a light brown amorphous powder, with the molecular formula of $C_{27}H_{24}O_{18}$ showed deprotonated molecular $[M–H]^-$ ion at *m/z* 635. Compound GT-4 contains three galloyl groups and one glucose unit as shown by its fragmentation pattern in the removal of a gallate (*m/z* 170) and galloyl moiety (*m/z* 152) at *m/z* 465 and 483, respectively (Figure 3.1.2.7). The fragment ion at *m/z* 423 was due to the mass loss of two formaldehyde (CH_2O) groups from part of the glucose moiety (Taylor *et al.*, 2005). Both compounds **10** and **11** gave the product ions at *m/z* 313, 295 and 169 which corresponded to the removal of a galloyl group (*m/z* 152), gallate (*m/z* 170) and contribution of deprotonated gallic acid, respectively. However, the major ion at *m/z* 421 in MS^3 spectrum of compound **10** did not exist in compound **11**. The mass accuracy and

isotopic pattern of compound GT-4 are shown in Figure 3.1.2.8. The NMR data of compound **11** showed a downfield doublet (δ 5.84, *J* = 8.5 Hz) and triplet (δ 5.26, *J* = 9.6 Hz) proton signals corresponding to the anomeric H-1 and H-4, respectively (Table 3.1.2.1). In addition to an upfield triplet (δ 3.72, *J* = 8.57 Hz) and downfield shift of H-6, three galloyl groups were found to be attached to C-1, C-4 and C-6 position. The β-configuration was observed by a large coupling constant (*J* = 8.5 Hz) at H-1. Therefore, the data for compound GT-4 data was consistent as reported for 1,4,6-tri-*O*-galloyl-β-D-glucose (Nonaka *et al.*, 1984; Lee *et al.*, 1992; Nawwar *et al.*, 1994).

(e) 3,4,6-Tri-*O*-galloyl-D-glucose (12), GT-5

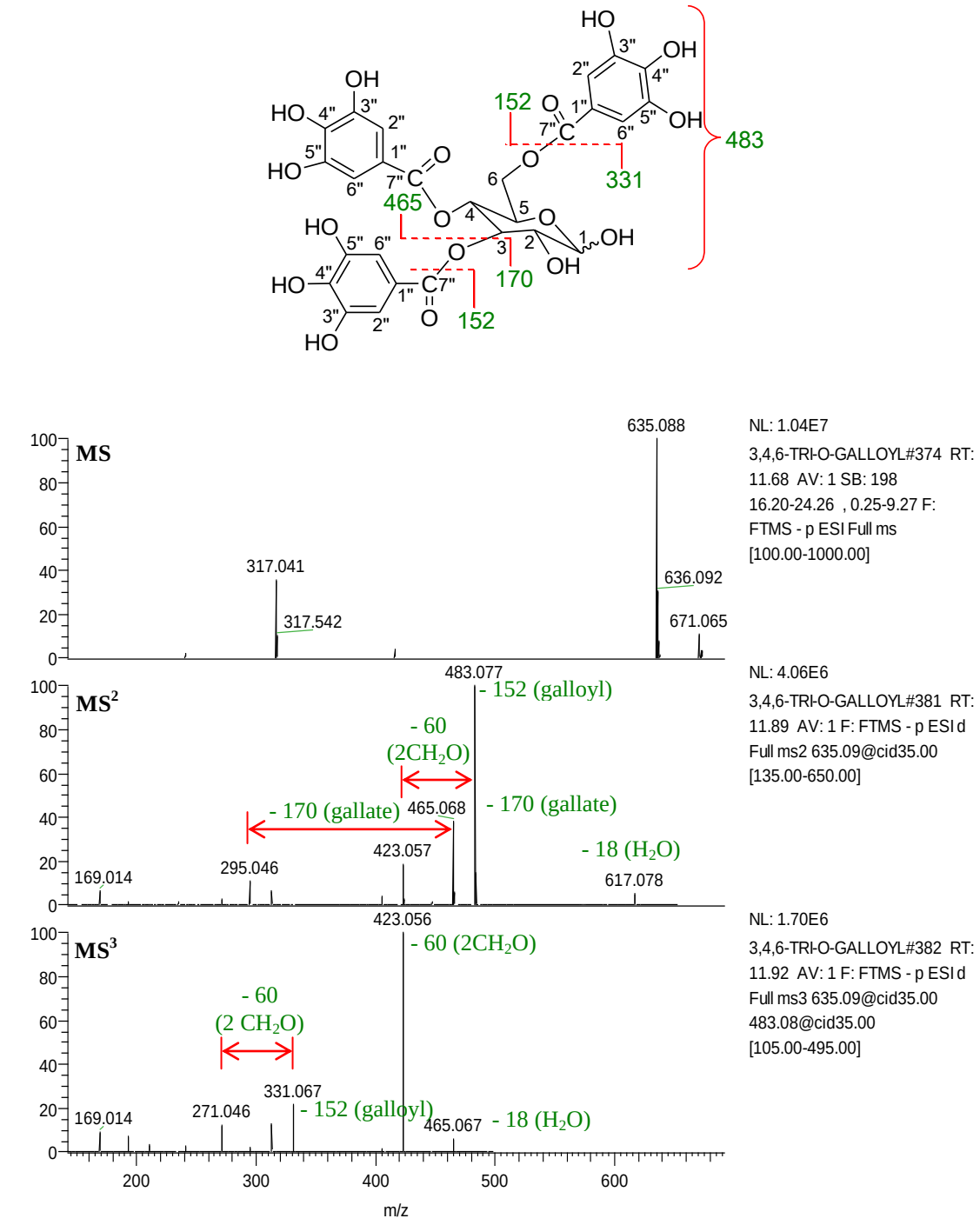


Figure 3.1.2.9: Fragmentation pattern for 3,4,6-tri-*O*-galloyl-D-glucose (**12**).

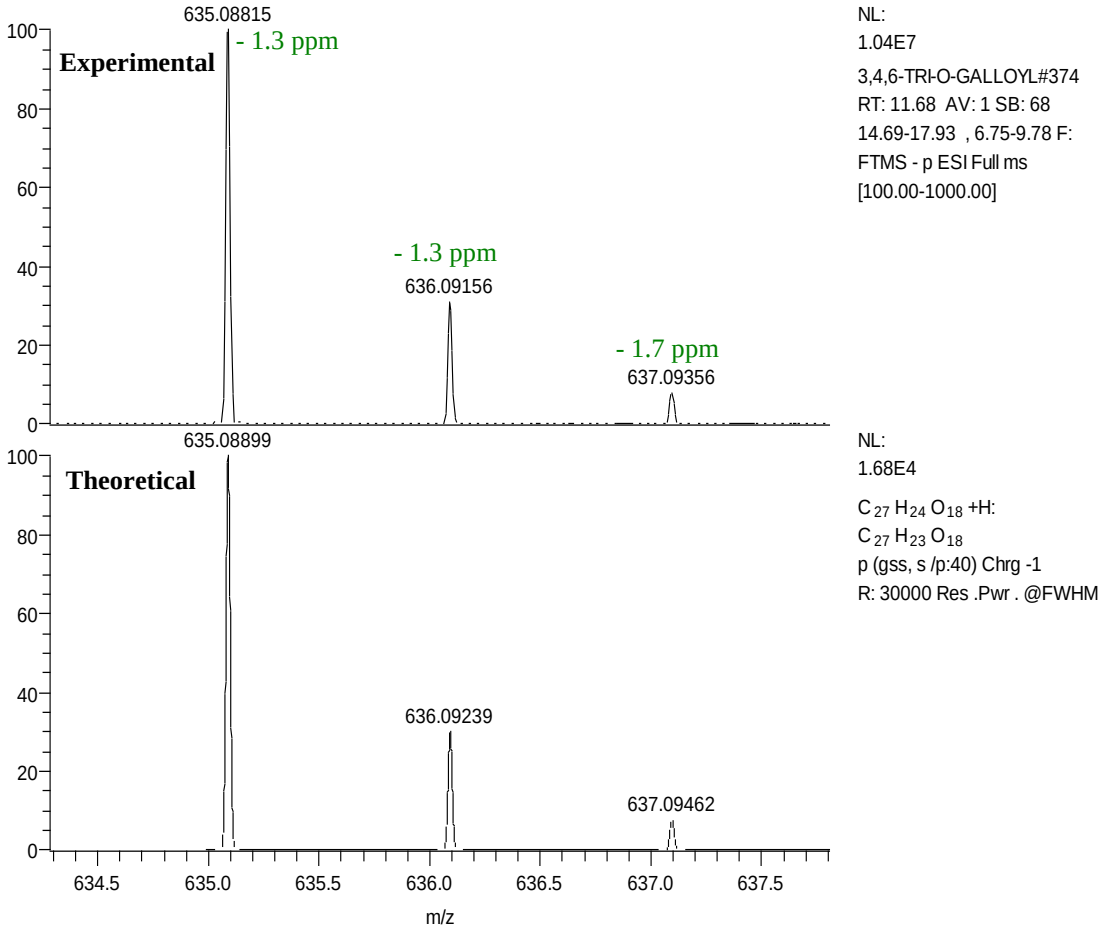
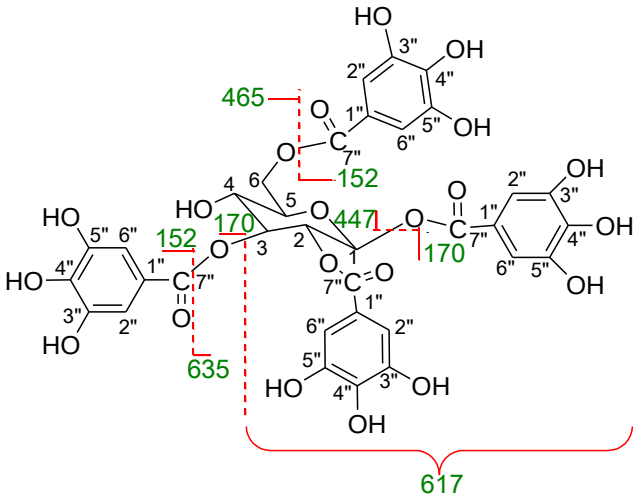


Figure 3.1.2.10: Full scan mass spectra (*m/z*) with isotopic pattern for 3,4,6-tri-*O*-galloyl-D-glucose (**12**).

3,4,6-Tri-*O*-galloyl-D-glucose (**12**), GT-5, an off-white amorphous powder, with the molecular formula of $C_{27}H_{24}O_{18}$ showed deprotonated molecular $[M–H]^-$ ion at *m/z* 635. Compound GT-5 contains three galloyl groups and one glucose unit which showed a major fragment ion at *m/z* 483 instead of *m/z* 465 indicating that it was more susceptible to the mass loss of a galloyl group (*m/z* 152) than a gallate (*m/z* 170). The occurrence of *m/z* 423 in MS^3 mass spectrum was due to the removal of two formaldehyde (CH_2O) groups which was part of the glucose fragmentation pattern (Taylor *et al.*, 2005). The presence of fragment ion at *m/z* 331 indicated the mass loss of a galloyl moiety (*m/z* 152) whereas the ion at *m/z* 169 showed the presence of deprotonated gallic acid (Figure 3.1.2.9). The

calculated mass accuracy and the isotopic pattern of compound **12** are shown in Figure 3.1.2.10. The presence of galloyl groups and sugar moiety was indicated by the observation of six aromatic singlets at δ 7.00 - 7.15 and aliphatic signals at δ 3.62 – 5.71 which were integrated for seven protons (Appendix I). The carbon signals at δ 93.6 and δ 98.3 were corresponded to α and β-form of the anomeric carbon in the glucose moiety (Table 3.1.2.1, Appendix I). Thus, compound GT-5 was identified as 3,4,6-tri-*O*-galloyl-D-glucose (Wilkins, 1988; Lee *et al.*, 1989; Lee *et al.*, 1991; He *et al.*, 2001).

(f) 1,2,3,6-Tetra-*O*-galloyl-β-D-glucose (13), GT-6



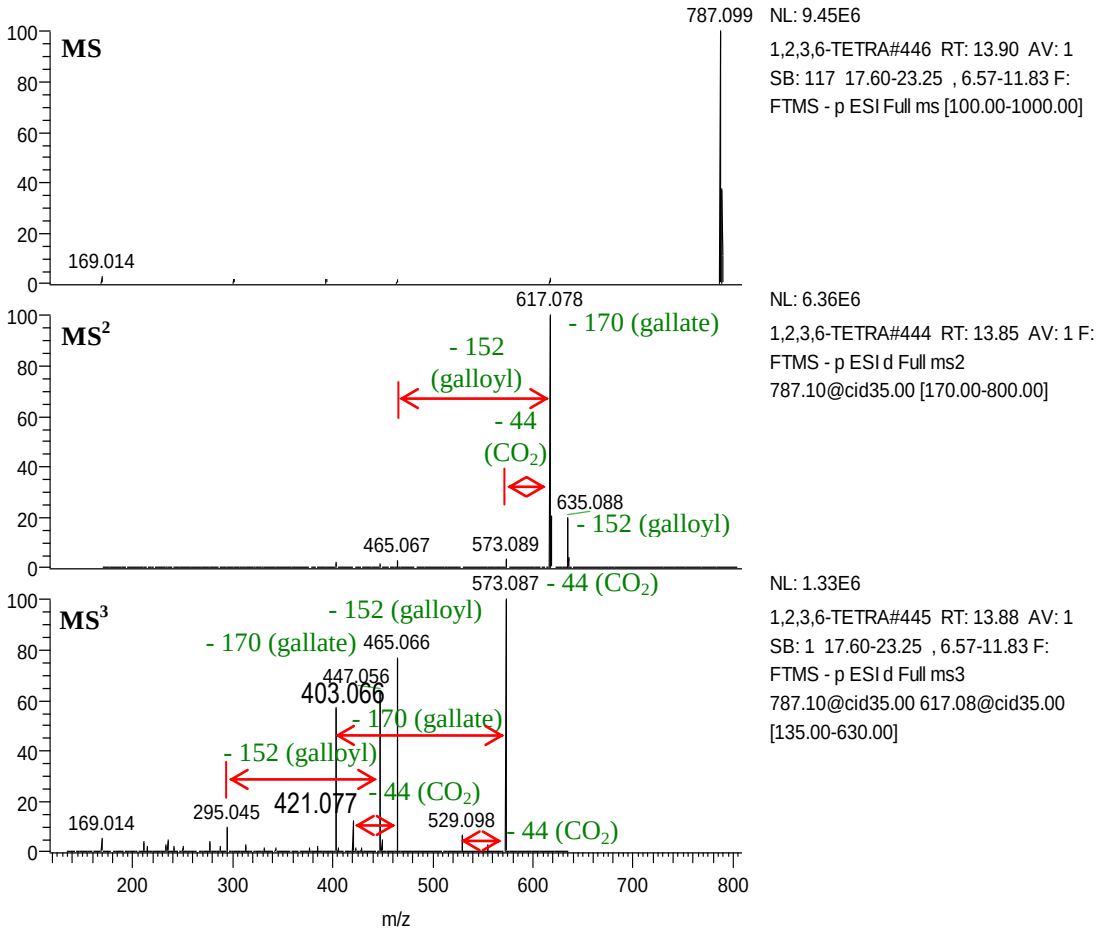


Figure 3.1.2.11: Fragmentation pattern for 1,2,3,6-tetra-*O*-galloyl-β-D-glucose (**13**).

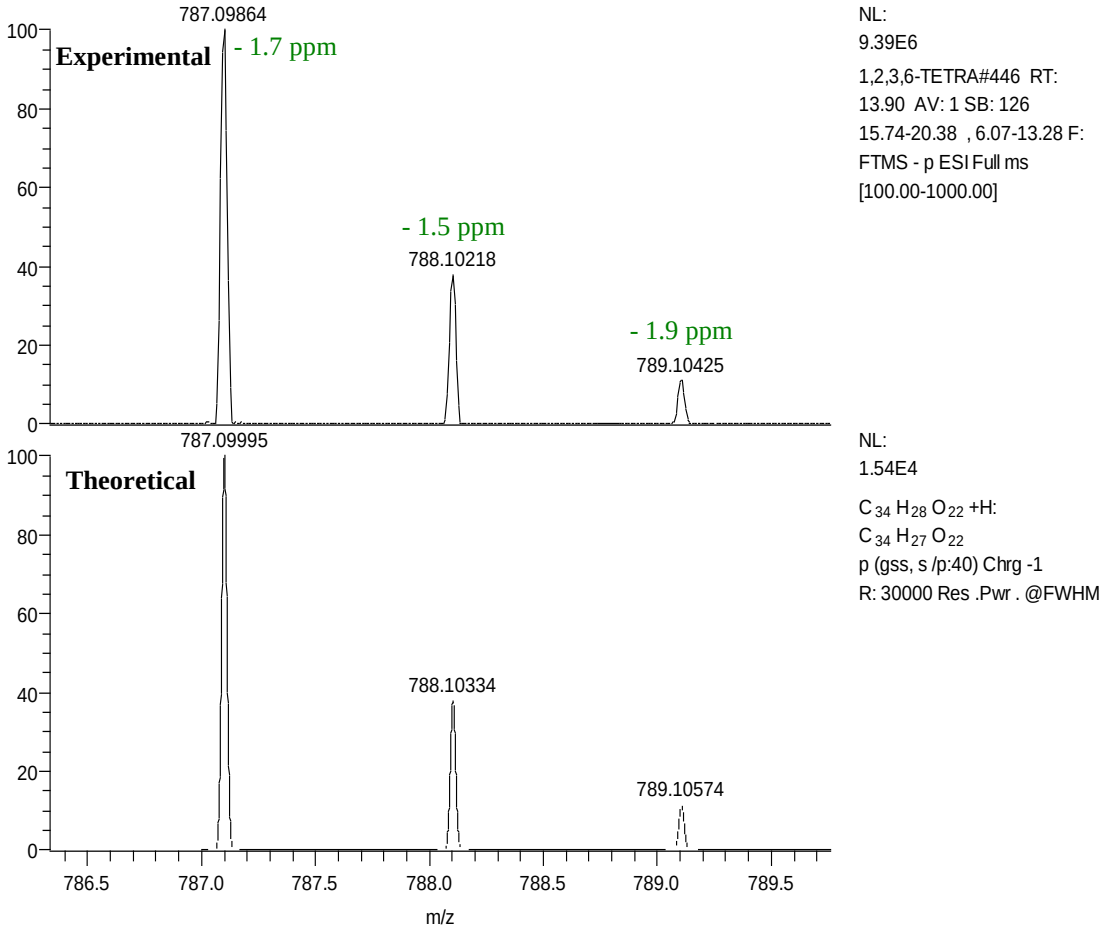
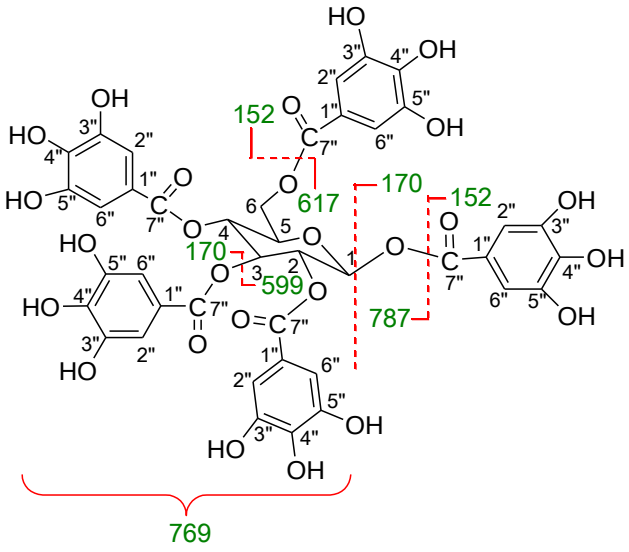


Figure 3.1.2.12: Full scan mass spectra (*m/z*) with isotopic pattern for 1,2,3,6-tetra-*O*-galloyl-β-D-glucose (**13**).

1,2,3,6-Tetra-*O*-galloyl-D-glucose (**13**), GT-6, a pale brown amorphous powder, with the molecular formula of $C_{34}H_{28}O_{22}$ showed deprotonated molecular $[M–H]^-$ ion at *m/z* 787. Compound GT-6 consists of four galloyl groups and one glucose unit. The fragment ion at *m/z* 617 and 635 were the major ions in MS^2 spectrum which represented the removal of a gallate (*m/z* 170) and galloyl moiety (*m/z* 152), respectively. Decarboxylation occurred in MS^3 spectrum which yielded ions at *m/z* 573, 529 and 421. The mass loss of a galloyl moiety (*m/z* 152) and gallate (*m/z* 170) were also indicated by the presence of fragment ions at *m/z* 465, 447, 403 and 295 (Figure 3.1.2.11). Deprotonated gallic acid was also found at *m/z* 169 in the MS^3 mass spectrum. The calculated mass

accuracy and the comparison of isotopic pattern for compound GT-6 are presented in Figure 3.1.2.12. The ^{1}H-NMR data showed four singlet of 2H each at δ 6.98, δ 7.00, δ 7.06 and δ 7.07 which indicated the presence of four galloyl groups (Table 3.1.2.1, Appendix I). The location of the galloyl groups was confirmed by considerably downfield shift of H-1, H-2, H-3 and H-6 which indicated the galloyl esterification at C-1, C-2, C-3 and C-6 positions (Table 3.1.2.2). Large coupling constant (J = 8.4 Hz) at H-1 was observed and the stereochemistry at the anomeric center was deduced to be β-configuration. Thus, compound GT-6 was characterised as 1,2,3,6-tetra-*O*-galloyl-β-D-glucose (Nishizawa *et al.*, 1983; Saijo *et al.*, 1990; Yoshida *et al.*, 1991a; Duan *et al.*, 2004).

(g) 1,2,3,4,6-Penta-*O*-galloyl-β-D-glucose (14), GT-7



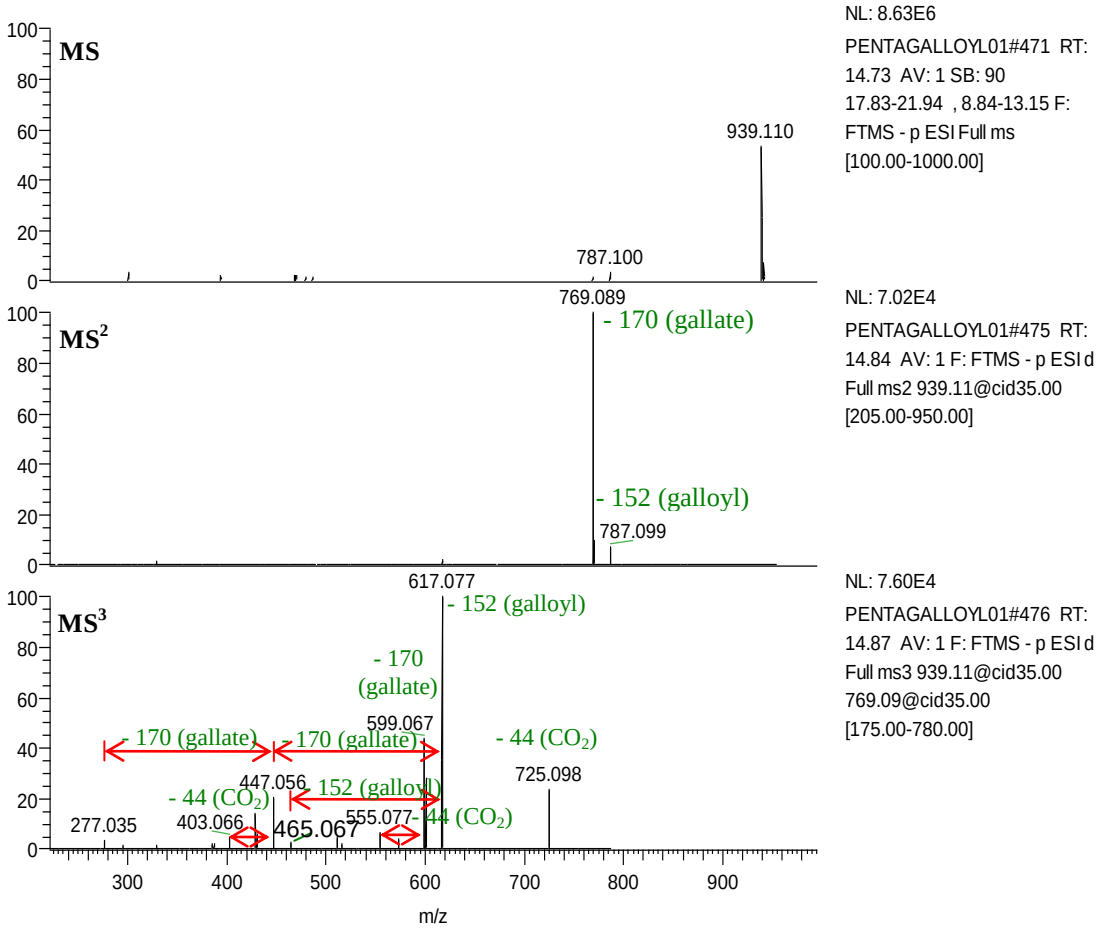


Figure 3.1.2.13: Fragmentation pattern for 1,2,3,4,6-penta-*O*-galloyl-β-D-glucose (**14**).

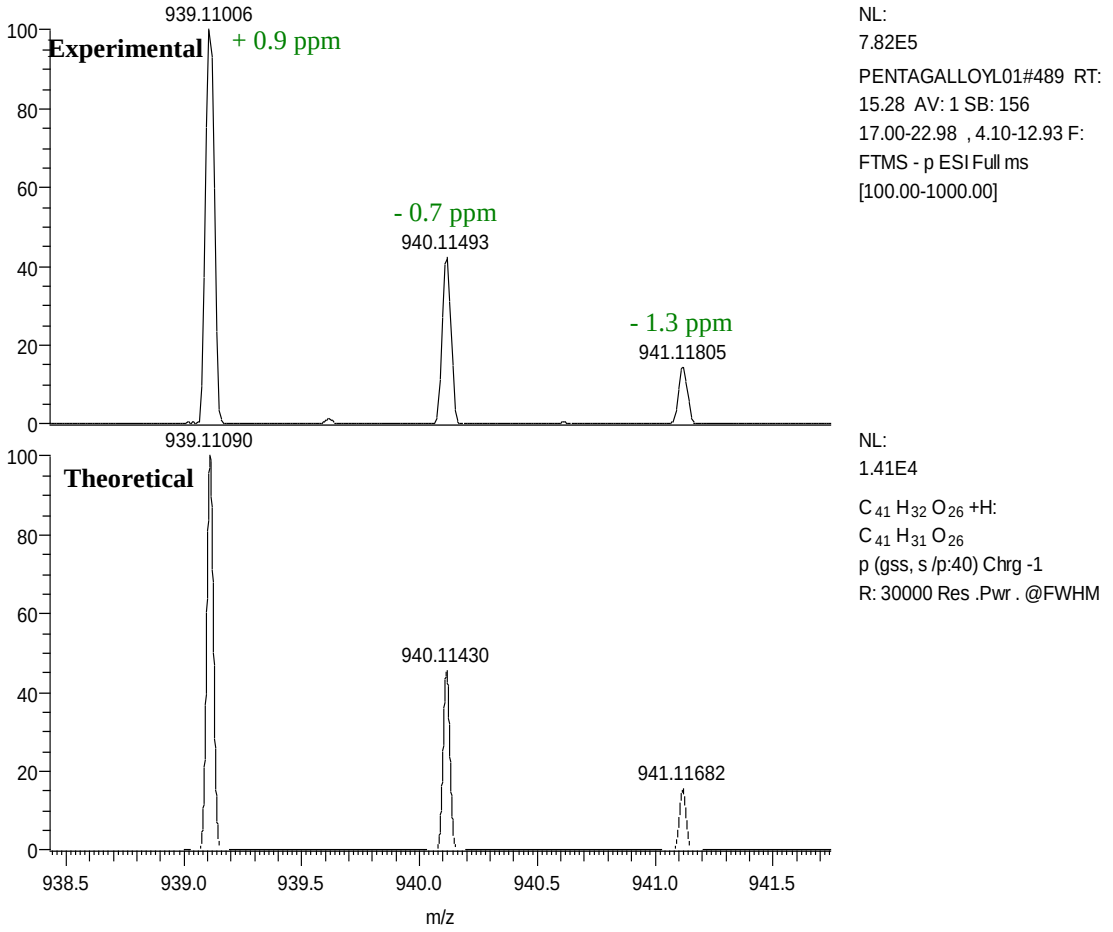


Figure 3.1.2.14: Full scan mass spectra (*m/z*) with isotopic pattern for 1,2,3,4,6-penta-*O*-galloyl-β-D-glucose (**14**).

1,2,3,4,6-Penta-*O*-galloyl-D-glucose (**14**), GT-7, an off white amorphous powder has the molecular formula of $C_{41}H_{32}O_{26}$ and showed deprotonated molecular $[M–H]^-$ ion at *m/z* 939. Compound GT-7 has five galloyl groups esterified to a glucose unit. The two major fragment ions at *m/z* 769 and 787 in MS^2 spectrum were corresponded to the removal of a gallate (*m/z* 170) and galloyl (*m/z* 152) group, respectively (Figure 3.1.2.13). Further in the MS^3 spectrum, the product ions at *m/z* 617, 599, 465, 447 and 277 were due to the removal of gallate (*m/z* 170) and galloyl (*m/z* 152) unit. Besides, the fragment ions at *m/z* 725, 555 and 403 showed the mass loss of CO_2 group. The mass accuracy and isotopic pattern for compound GT-7 are shown in Figure 3.1.2.14. Tables 3.1.2.1 and 3.1.2.2 show

the ^{1}H- and ^{13}C-NMR data of compound GT-7. Five singlet aromatic proton singlets (δ 6.97, δ 7.01, δ 7.06, δ 7.11 and δ 7.18) corresponded to the protons in the galloyl moieties (Appendix I). The downfield shift of H-1, H-2, H-3, H-4 and H-6 of the glucose moiety corresponded to galloyl linkages with glucose at C-1, C-2, C-3, C-4 and C-6. The β-configuration of the glucose was confirmed by the large coupling constant (J = 8.4 Hz) at H-1. Comparatively, compound **14** has downfield chemical shifts for H-1 to H-6 of the glucose moiety as compared to compound **8**-**18** due to esterification of five galloyl groups. Hence, compound GT-7 was assigned as 1,2,3,4,6-penta-*O*-galloyl-β-D-glucose (Nishizawa *et al.*, 1983; Tanaka *et al.*, 1985; Saijo *et al.*, 1990; Gao *et al.*, 2007).

(h) Fragmentation Schemes of Gallotannins

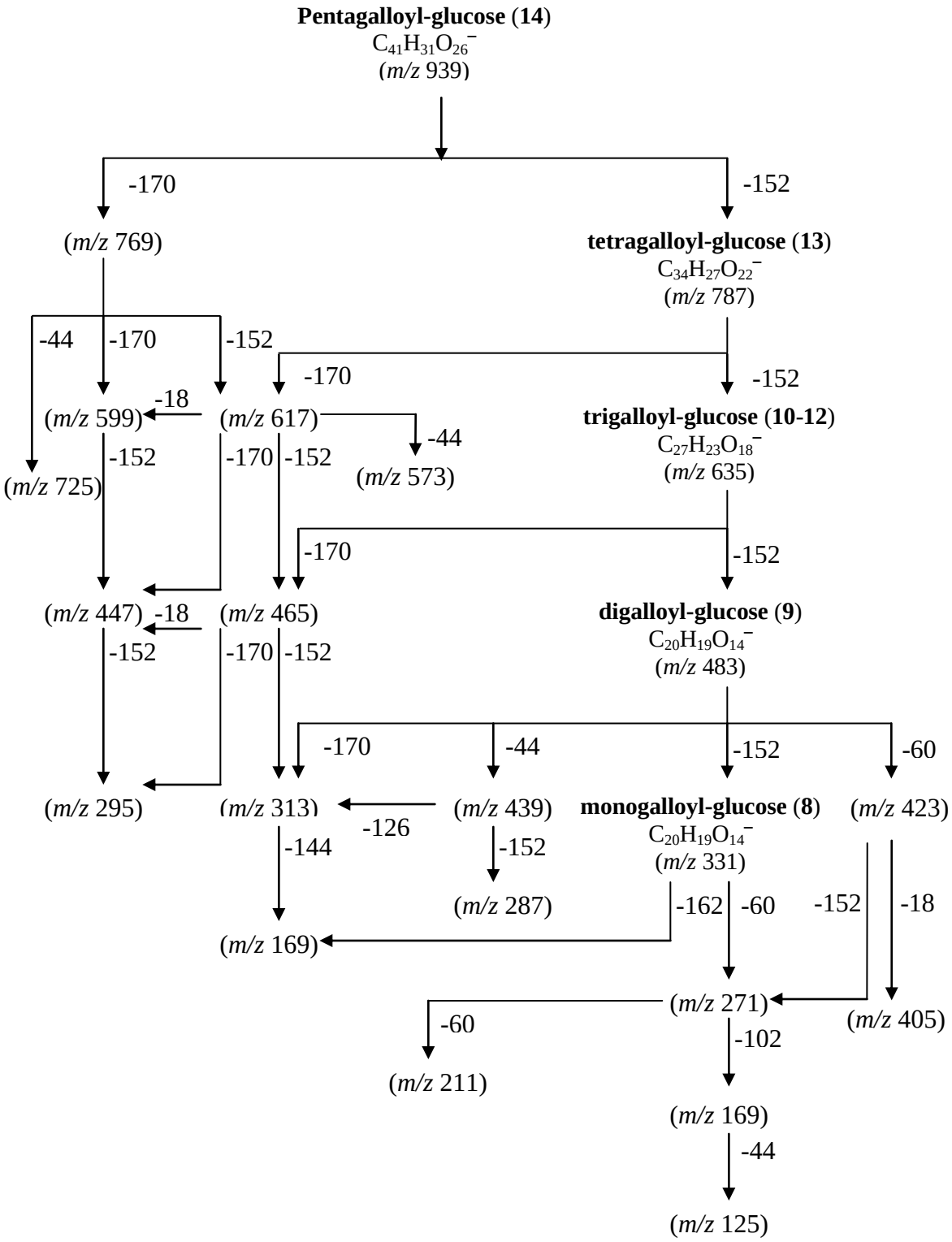


Figure 3.1.2.15: Fragmentation pathway of gallotannins (**8**-**14**) observed by data dependent MS^n with dynamic exclusion in negative mode. (The mass loss of 170, 152, 144, 60, 44 and 18 represented the removal of gallate, galloyl, hexosyl-H_2O, $2CH_2O$, CO_2 and H_2O, respectively.

Gallotannins or galloyl esters of glucose are known to show elimination of multiple galloyl [M–H–152] $^{-}$ and gallate [M–H–170] $^{-}$ moieties in MS^n analysis (Salminen *et al.*, 1999; Soong and Barlow, 2005; Meyers *et al.*, 2006). In this study, the schematic fragmentation of 1,2,3,4,6-penta-*O*-galloyl-β-D-glucose (**14**, *m/z* 939) showed the mass loss of several galloyl (*m/z* 152) and gallate (*m/z* 170) groups (Figure 3.1.2.13 and Table 3.1.2). The fragmentation pathways of the gallotannins are shown in Figure 3.1.2.15, where the monogalloyl-glucose is represented by 6-*O*-galloyl-D-glucose (**8**), digalloyl-glucose by 3,6-di-*O*-galloyl-D-glucose (**9**), trigalloyl-glucose by 1,2,3-tri-*O*-galloyl-β-D-glucose (**10**), 1,4,6-tri-*O*-galloyl-β-D-glucose (**11**) and 3,4,6-tri-*O*-galloyl-D-glucose (**12**), tetragalloyl-glucose by 1,2,3,6-tetra-*O*-galloyl-β-D-glucose (**13**) and pentagalloyl-glucose by 1,2,3,4,6-penta-*O*-galloyl-β-D-glucose (**14**). The pentagalloyl-glucose (*m/z* 939) showed consistent mass loss of galloyl moieties to tetragalloyl-glucose (*m/z* 787), trigalloyl-glucose (*m/z* 635), digalloyl-glucose (*m/z* 483) and monogalloyl-glucose (*m/z* 331). Additionally, the neutral mass loss of gallic acid [M–H–170] $^{-}$ moieties in pentagalloyl-glucose (*m/z* 939), tetragalloyl-glucose (*m/z* 787), trigalloyl-glucose (*m/z* 635), digalloyl-glucose (*m/z* 483) were also illustrated. Fragmentation of digalloyl-glucose (**9**) at *m/z* 483 contributed to two major product ions at *m/z* 423 and 439. The product ion at *m/z* 423 was suggested to be derived from the removal of two formaldehyde moieties [M–H–60] $^{-}$ (Taylor *et al.*, 2005) from the glucose core, while that at *m/z* 439 was formed as a result of decarboxylation [M–H–44] $^{-}$. The fragment ion at *m/z* 423 further underwent independent losses of [M–H–152] $^{-}$ (*m/z* 271) and [M–H–18] $^{-}$ (*m/z* 405) (Figure 3.1.2.15 and Table 3.1.2). The monogalloyl-glucose showed elimination of glucosyl moiety and formed the deprotonated gallic acid at *m/z* 169 (Figure 3.1.2.15). At the same time, fragment ions at *m/z* 271 [M–H–60] $^{-}$ and *m/z* 211 [M–H–60–60] $^{-}$ represented the major product ions in MS^2 and MS^3, respectively (Figure 3.1.2.15 and Table 3.1.2). The same observation had also been described previously

for monogalloyl-glucose and digalloyl-glucose (Mämmelä *et al.*, 2000; Soong and Barlow, 2005; Zywicki *et al.*, 2002). The fragment ion at *m/z* 211 was suggested to be due to the removal of two formaldehyde (CH_2O) groups in the glucose moiety. Further, the ion at *m/z* 271 was fragmented to produce a signal at *m/z* 169, indicating the loss of a glucosyl unit (He and Xia, 2007; Kool *et al.*, 2010, Mullen *et al.*, 2003; Simirgiotis and Schmeda-Hirschmann, 2010). This ion then underwent removal of CO_2 to generate an ion at *m/z* 125.

Table 3.1.2.1: ^{1}H-NMR Spectral data for compounds GT-1 to GT-7 (**8**-**14**) in Acetone-d_6.

	δ_H								
	GT-1 (8)	**GT-2 (9)**		**GT-3 (10)**	**GT-4 (11)**	**GT-5 (12)**		**GT-6 (13)**	**GT-7 (14)**
	(α, β-form)	**(α -form)**	**(β-form)**			**(α-form)**	**(β-form)**		
Glucose									
H-1	5.10 d (3.6) 4.54 d (7.8)	5.21 d (3.6)	4.67 d (7.5)	6.08 d (8.4)	5.84 d (8.5)	5.32 d (4.5);	4.86 d (7.8)	6.15 d (8.4)	6.31 d (8.4)
H-2	3.16-4.35 m	3.53 dd (3.6, 9.3)	3.33 dd (7.5, 9.6)	5.38 dd (8.4, 9.9)	3.72 t (8.7)	3.82 d (3.0);	3.62 d (9.0)	5.45 dd (8.1, 9.9)	5.63 dd (8.4, 9.6)
H-3	3.16-4.35 m	5.13 t (9.3)	5.12 t (9.6)	5.59 t (9.9)	3.98 t (9.3)	5.71 d (9.3);	5.50 t (9.6)	5.68 dd (8.4, 9.9)	6.01 t (9.6)
H-4	3.16-4.35 m	4.04 t (9.3)	3.79 t (9.3)	3.75-4.02 m	5.26 t (9.6)	5.29-5.40 m	5.29-5.40 m	4.10 dd (8.4, 9.6)	5.65 t (9.6)
H-5	3.16-4.35 m	4.17 m	3.91 m	3.75-4.02 m	4.13 m	4.09-4.55 m	4.09-4.55 m	4.15 m	4.57 m
H-6a	3.44-4.49 m	4.13 br d (12.0)	4.11 br d (12.0)	3.75-4.02 m	4.45 br d (12.9)	4.09-4.55 m	4.09-4.55 m	4.63 dd (1.8, 12.0)	4.36 dd (5.1, 12.9)
H-6b	3.44-4.49 m	4.41 dd (1.8, 12.0)	4.42 dd (2.1, 12.0)	3.75-4.02 m	4.17 dd (5.1, 12.9)	4.09-4.55 m	4.09-4.55 m	4.57 dd (5.2, 12.0)	4.53 dd (2.1, 12.9)
Galloyl									
H-2", -6"	7.10	7.12, 7.13		6.97, 7.06, 7.08	7.12, 7.15, 7.17	7.00, 7.02, 7.03, 7.11, 7.14, 7.15		6.98, 7.00, 7.06, 7.07	6.97, 7.01, 7.06, 7.11, 7.18 s

Table 3.1.2.2: ^{13}C-NMR Spectral data for compounds GT-1 to GT-7 (**8-14**) in Acetone-d_6.

	δ_C									
	GT-1 (8)		**GT-2 (9)**		**GT-3 (10)**	**GT-4 (11)**	**GT-5 (12)**		**GT-6 (13)**	**GT-7 (14)**
	(α-form)	**(β-form)**	**(α-form)**	**(β-form)**			**(α-form)**	**(β-form)**		
Glucose										
C-1	93.4	97.9	93.5	98.2	93.4	95.4	93.6	98.3	93.5	93.2
C-2	71.3	74.8	73.8	76.4	71.8	74.0	72.8	73.8	71.7	71.7
C-3	74.5	77.5	72.5	72.5	76.2	75.4	74.0	75.9	75.8	73.2
C-4	70.4	71.1	72.2	75.4	69.1	71.5	68.5	72.0	69.1	69.2
C-5	73.4	75.8	73.0	73.0	78.6	73.9	70.0	70.1	75.9	73.9
C-6	64.5	64.6	68.5	63.8	61.7	63.4	63.4	63.4	63.9	62.7
Galloyl										
C-1"	121.4		121.6, 121.8		121.4, 120.8, 120.2	121.5, 121.4, 120.7	120.8, 121.5, 121.7		120.6, 120.7, 121.2, 121.4	119.8, 120.3, 120.4, 120.5, 121.3,
C-2", C-6"	109.7		109.9, 110.1		110.2, 110.0, 109.9	110.3, 110.2, 109.9	110.0, 110.0, 110.1		109.9, 110.0, 110.2, 110.5	109.7, 109.9, 110.0, 110.1, 110.2
C-3", C-5"	146.0		145.8, 145.9		145.9, 145.8, 145.7	146.0, 145.9, 145.8	145.8, 145.9, 146.0		145.7, 145.8, 145.9, 146.0	145.7 × 2, 145.8, 145.9, 146.0
C-4"	138.8		138.6, 138.8		139.4, 138.9, 138.7	139.3, 138.8, 138.7	138.7, 138.8, 139.1		138.7, 139.0, 139.5, 139.8	138.8, 139.0, 139.1, 139.2, 139.6
C-7"	167.0		166.1, 166.5		166.2, 165.8, 165.0	166.4, 166.0, 165.4	165.9, 166.5, 166.6		165.7, 165.8, 166.1, 166.2	164.9, 165.6, 165.7, 165.9, 166.4

3.1.3 Ellagitannins

(a) 6-*O*-Galloyl-2,3-*O*-(*S*)-hexahydroxydiphenoyl-D-glucose (15), ET-1

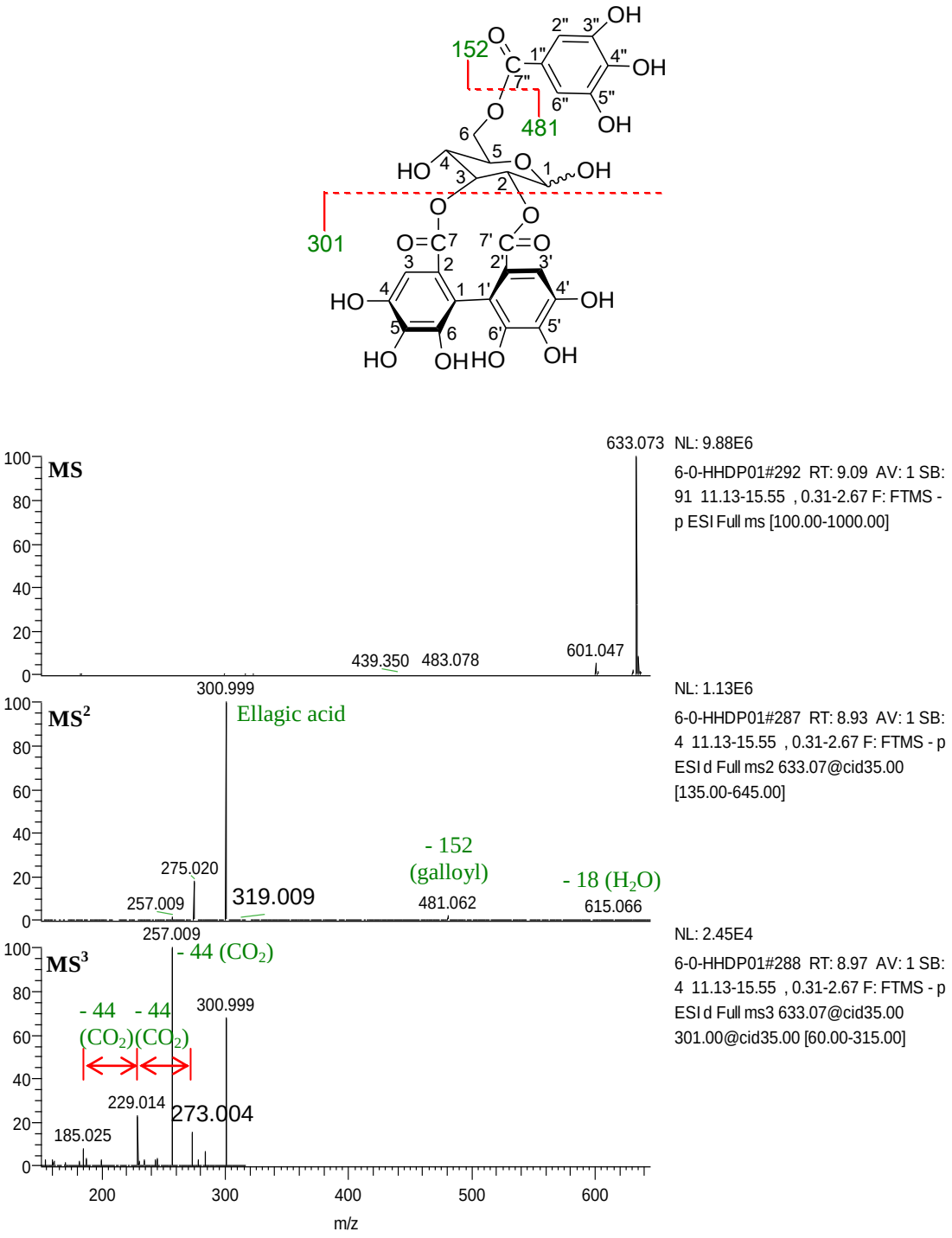


Figure 3.1.3.1: Fragmentation pattern for 6-*O*-galloyl-2,3-*O*-(*S*)-hexahydroxydiphenoyl-D-glucose (**15**).

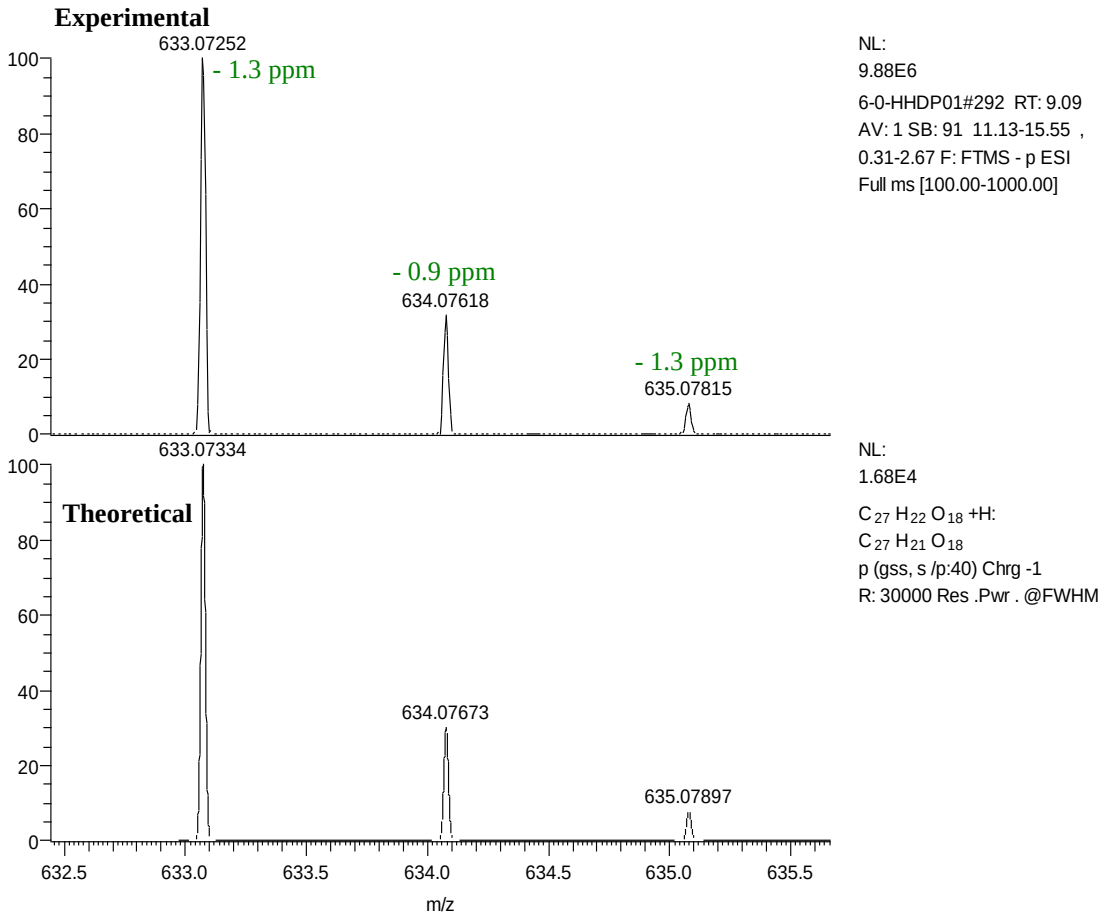
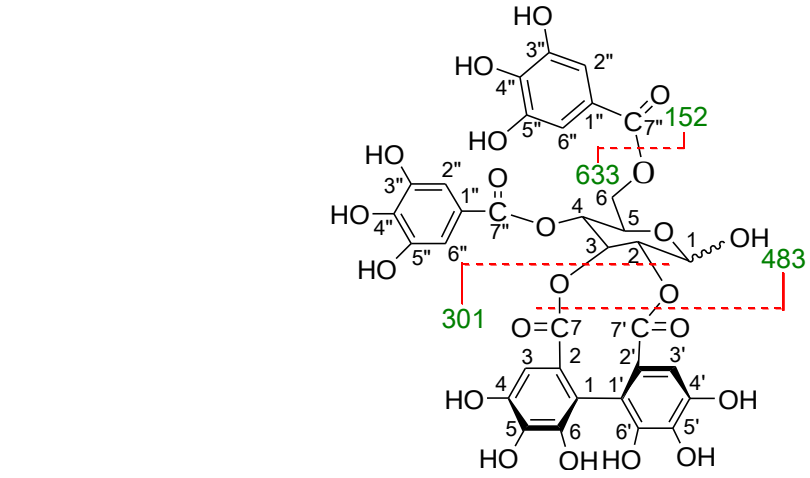


Figure 3.1.3.2: Full scan mass spectra (*m/z*) with isotopic pattern for 6-*O*-galloyl-2,3-*O*-(*S*)-hexahydroxydiphenoyl-D-glucose (**15**).

6-*O*-Galloyl-2,3-*O*-(*S*)-hexahydroxydiphenoyl-D-glucose (**15**), ET-1, an off-white amorphous powder, with the molecular formula of $C_{27}H_{22}O_{18}$ showed deprotonated molecular $[M–H]^-$ ion at *m/z* 633. Compound ET-1 has a galloyl and a HHDP group that were esterified to a glucose moiety. The ions at *m/z* 319 and 275 were resulted from the lactonization and decarboxylation of HHDP before formation of ellagic acid (*m/z* 301) (Figures 3.1.3.1 and 3.1.3.9). The relatively low abundance of fragment ion at *m/z* 481 indicated that compound **15** was less susceptible to the removal of galloyl group (*m/z* 152). Further fragmentation of ellagic acid (*m/z* 152) in MS^3 produced fragment ions at *m/z* 257, 229 and 185 which indicated the mass loss of CO_2 group (*m/z* 44). A fragment ion at *m/z* 273 was also observed in the MS^3 which indicated the removal of CO group (*m/z* 28). The

mass accuracy was calculated and isotopic pattern of 6-*O*-galloyl-2,3-*O*-(*S*)-hexahydroxydiphenoyl-D-glucose (**15**) is presented in Figure 3.1.3.2. Most of the proton and carbon signals occurred in duplicate were due to the presence of a mixture of α and β-anomers (Tables 3.1.3.1 and 3.1.3.2). The two doublet signals at δ 5.40, J = 3.6 Hz and δ 5.04, J = 8.1 Hz represented the α and β-anomeric proton, respectively (Appendix I). The signal at δ 7.15 was attributed to H-2" and H-6" of the galloyl group. Two singlet signals at δ 6.58 and δ 6.68 indicated the presence of a HHDP group. A remarkable downfield shift of H-6 showed galloyl esterification at C-6 position of the glucose unit. Compound ET-1 was deduced to be in (*S*)-configuration since 2,3-*O*-(*S*)-HHDP was more thermodynamically stable than 2,3-*O*-(*R*)-HHDP (Immel and Khanbabaee, 2000). Morever, the NMR spectroscopic data (Yoshida *et al.*, 1991a; Yoshida *et al.*, 1991c) were corresponded to those reported in the literatures. Thus, compound ET-1 was assigned as 6-*O*-galloyl-2,3-*O*-(*S*)-hexahydroxydiphenoyl-D-glucose.

(b) Praecoxin B (16), ET-2



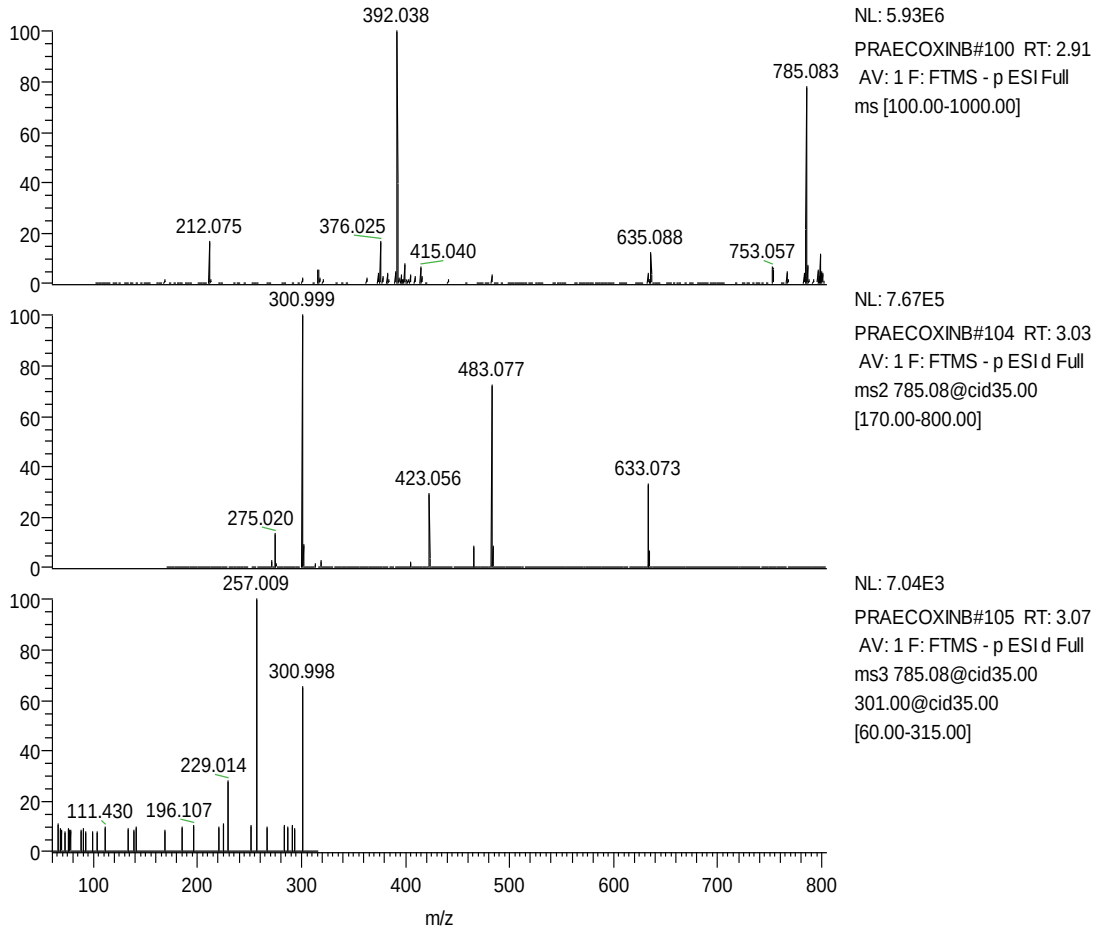


Figure 3.1.3.3: Fragmentation pattern for praecoxin B (**16**).

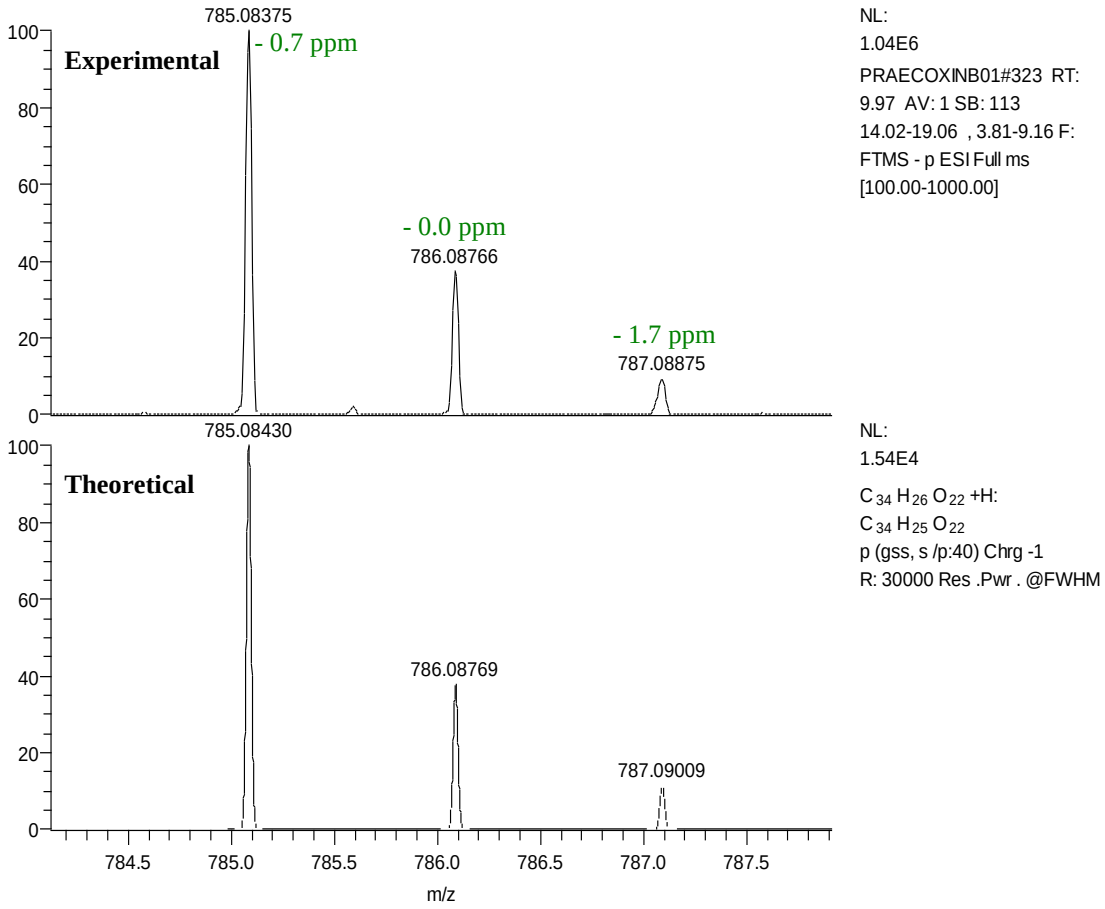
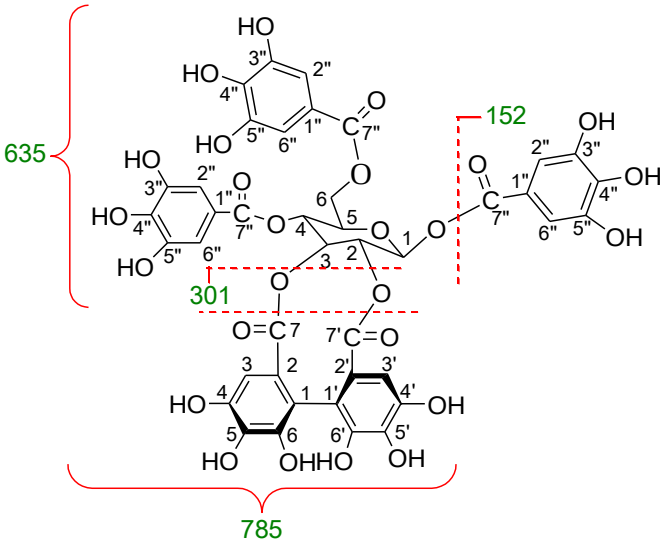


Figure 3.1.3.4: Full scan mass spectra (*m/z*) with isotopic pattern for praecoxin B (**16**).

Praecoxin B (**16**), ET-2, a light brown powder, with the molecular formula of $C_{34}H_{26}O_{22}$ showed deprotonated molecular $[M–H]^-$ ion at *m/z* 785. Compound ET-2 contains two galloyl and one HHDP group that are attached to a glucose unit. It showed fragment ions at *m/z* 301 (ellagic acid), 483 (digalloyl-glucose) and 633 (*m/z* 785-152, loss of galloyl). Fragment ion at *m/z* 423 which was also observed in digalloyl-glucose (**9**) showed the fragmentation of glucose (loss of 2 CH_2O). Further, the ions at *m/z* 319 and 275 were due to the lactonization and decarboxylation of HHDP group (Figures 3.1.3.3 and 3.1.3.9). The relatively high abundance of fragment ion at *m/z* 301 in MS^3 indicated the tendency of HHDP to form ellagic acid. In addition, the ions at *m/z* 257 and 229 indicated the mass loss of CO_2 group (*m/z* 44) while that of *m/z* 273 likely showed the removal of CO

group (*m/z* 28). The calculated mass accuracy and the isotopic pattern of ET-2 are displayed in Figure 3.1.3.4. The ^{1}H- and ^{13}C-NMR spectral data as shown in Tables 3.1.3.1 and 3.1.3.2 indicated the presence of a mixture of α and β-anomers. This was supported by the observation of two anomeric proton doublets at δ 5.51, J = 3.6 Hz and δ 5.19, J = 8.4 Hz as well as two anomeric carbon signals at δ 91.3 and δ 94.9 (Appendix I). The occurrence of two galloyl groups and one HHDP moiety was due to four proton signals at δ 7.11-7.16 and two proton signals at δ 6.39 and δ 6.61, respectively. As compared to compound ET-1, compound ET-2 has relatively downfield shift of H-4 in the glucose unit which indicated that the galloyl esterification occurred not only at C-6 but also at C-4 (Table 3.1.3.1 and 3.1.3.2). As mentioned earlier, the HHDP moiety was more stable in the (*S*)-configuration (Immel and Khanbabaee, 2000). Thus, compound ET-2 was characterised as praecoxin B (Hatano *et al.*, 1991; Yoshida *et al.*, 1991a).

(c) Pterocarinin (17) C, ET-3



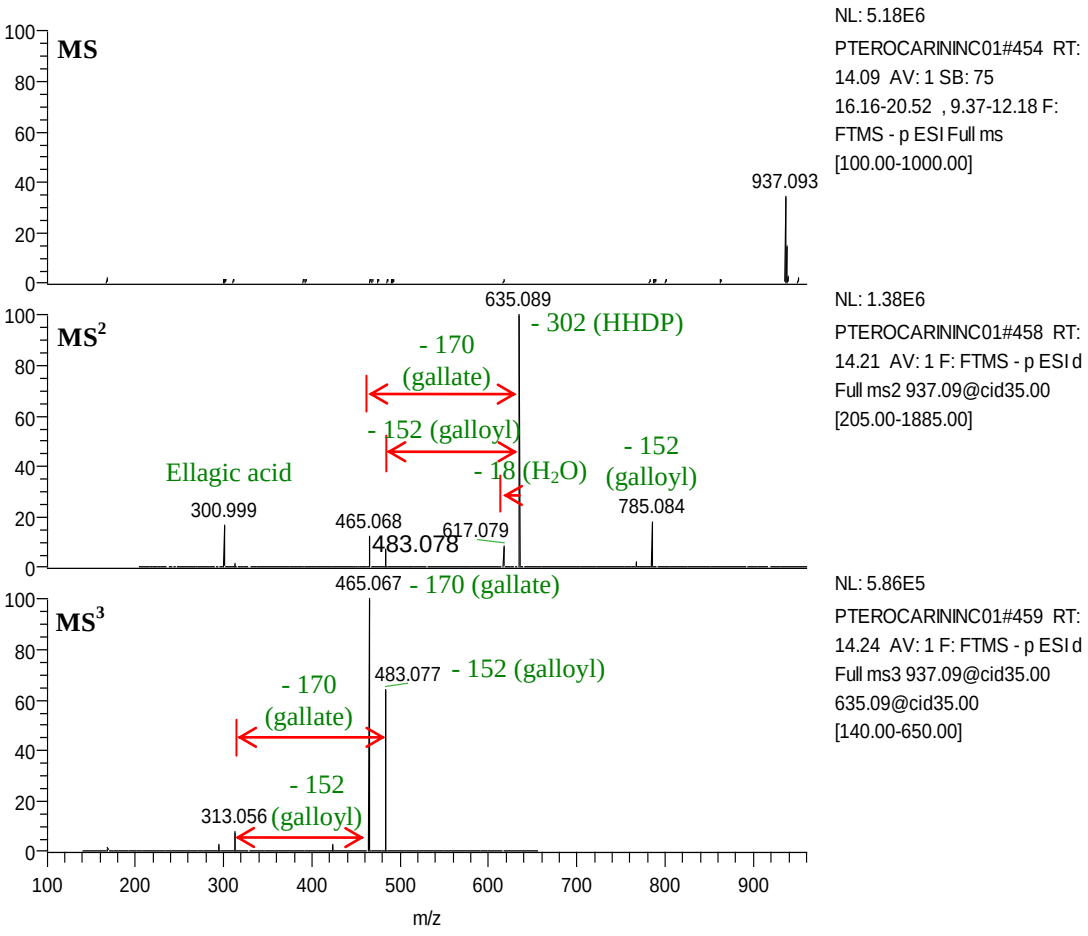


Figure 3.1.3.5: Fragmentation pattern for pterocarinin C (**17**).

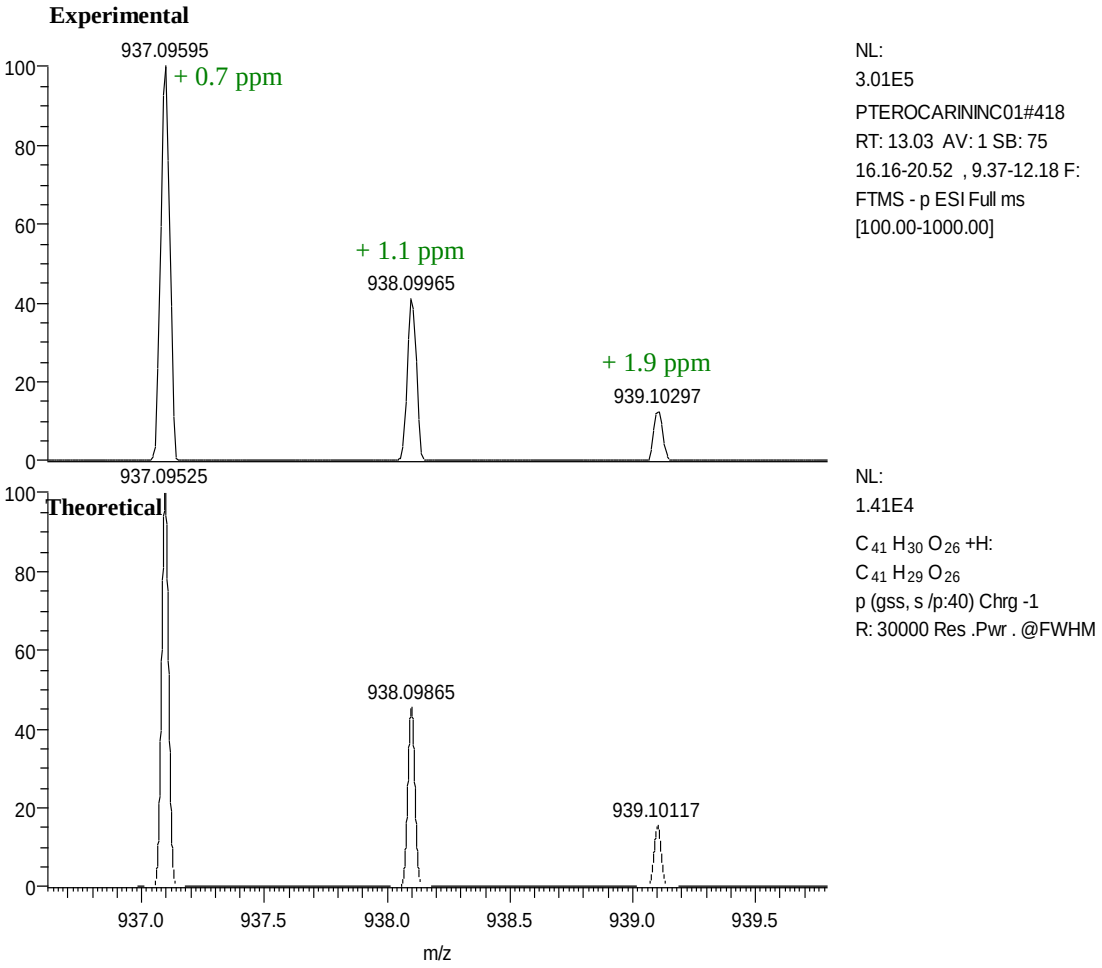
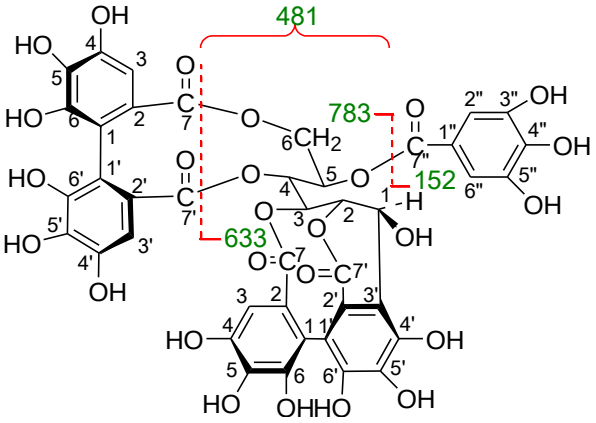


Figure 3.1.3.6: Full scan mass spectra (*m/z*) with isotopic pattern for pterocarinin C (**17**).

Pterocarinin C (**17**), ET-3, a tan amorphous powder, with the molecular formula of $C_{41}H_{30}O_{26}$ showed deprotonated molecular $[M–H]^-$ ion at *m/z* 937. Compound ET-3 has three galloyl groups and a HHDP moiety which are linked to a glucose unit. The parent ion gave the major fragment ion at *m/z* 635 (mass loss of HHDP) which corresponded to trigalloyl-glucose (Figure 3.1.3.5). This fragmentation pattern was different with compounds **15** and **16** which showed dominant fragment ion at *m/z* 301 (ellagic acid). Nevertheless, the product ions of MS^3 were similar to that of trigalloyl-glucose which gave ions at *m/z* 483, 465 and 313 (Figures 3.1.2.5, 3.1.2.7, 3.1.2.9) indicating the removal of gallate (*m/z* 170) and, or galloyl (*m/z* 152) groups. Figure 3.1.3.6 presents the calculated mass accuracy and isotopic pattern for compound ET-3. The anomeric proton doublet at δ

6.35, J = 8.4 Hz was assigned to the β-form of glucose (Table 3.1.3.1). The ^{1}H-NMR data of compound ET-3 showed the presence of three galloyl groups (δ 7.08, δ 7.13, δ 7.16, each 2H) and one HHDP moiety (δ 6.43, δ 6.45 each 1H, singlet) (Appendix I). The downfield shift of H-1, H-4 and H-6 in glucose moiety revealed that the galloyl groups were esterified at C-1, C-4 and C-6 position (Table 3.1.3.2). The (*S*)-configuration of HHDP was taken for compound ET-3 based on the explanation by Immel and Khanbabaee, 2000 and the comparison of NMR spectroscopic data with literatures (Yoshida *et al.*, 1991a; Yoshida *et al.*, 1995). Thus, compoumd ET-3 was established as pterocarinin C.

(d) Casuarinin (18), ET-4



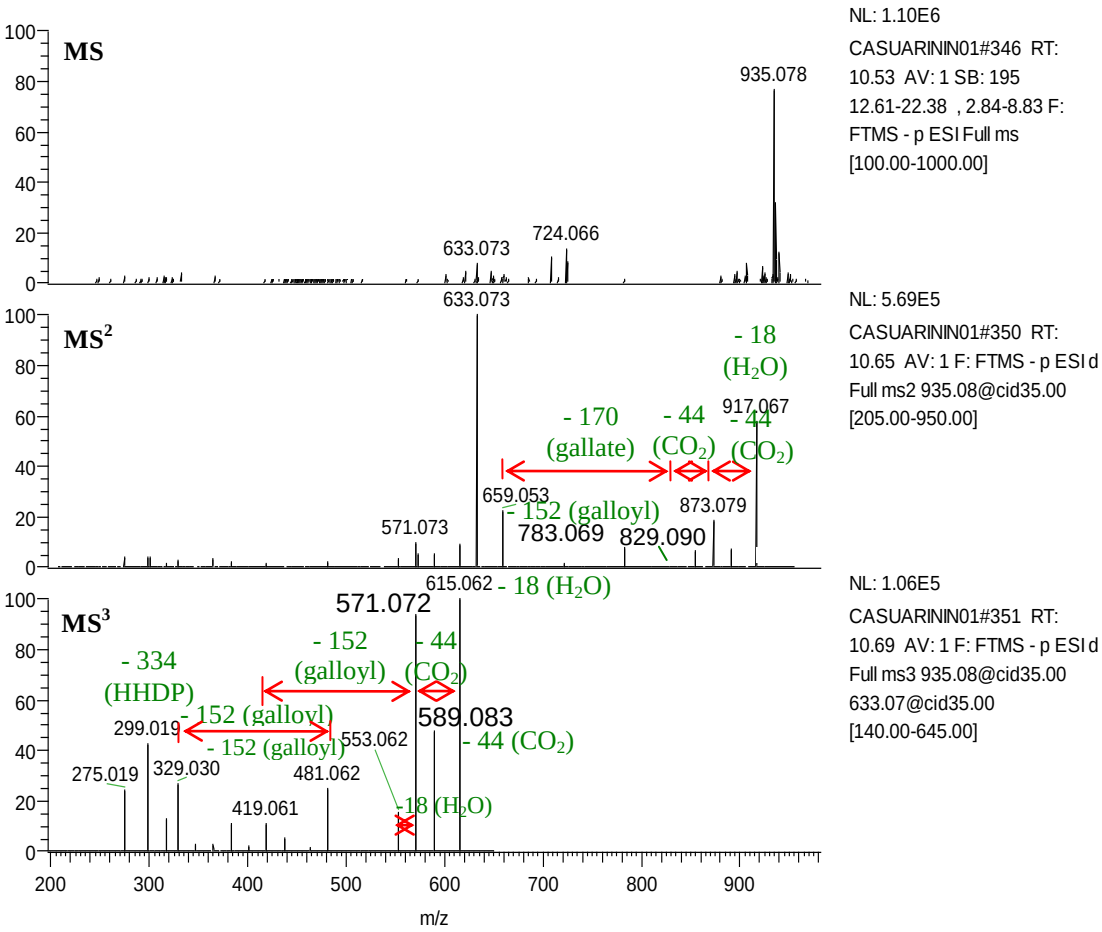


Figure 3.1.3.7: Fragmentation pattern for casuarinin (**18**).

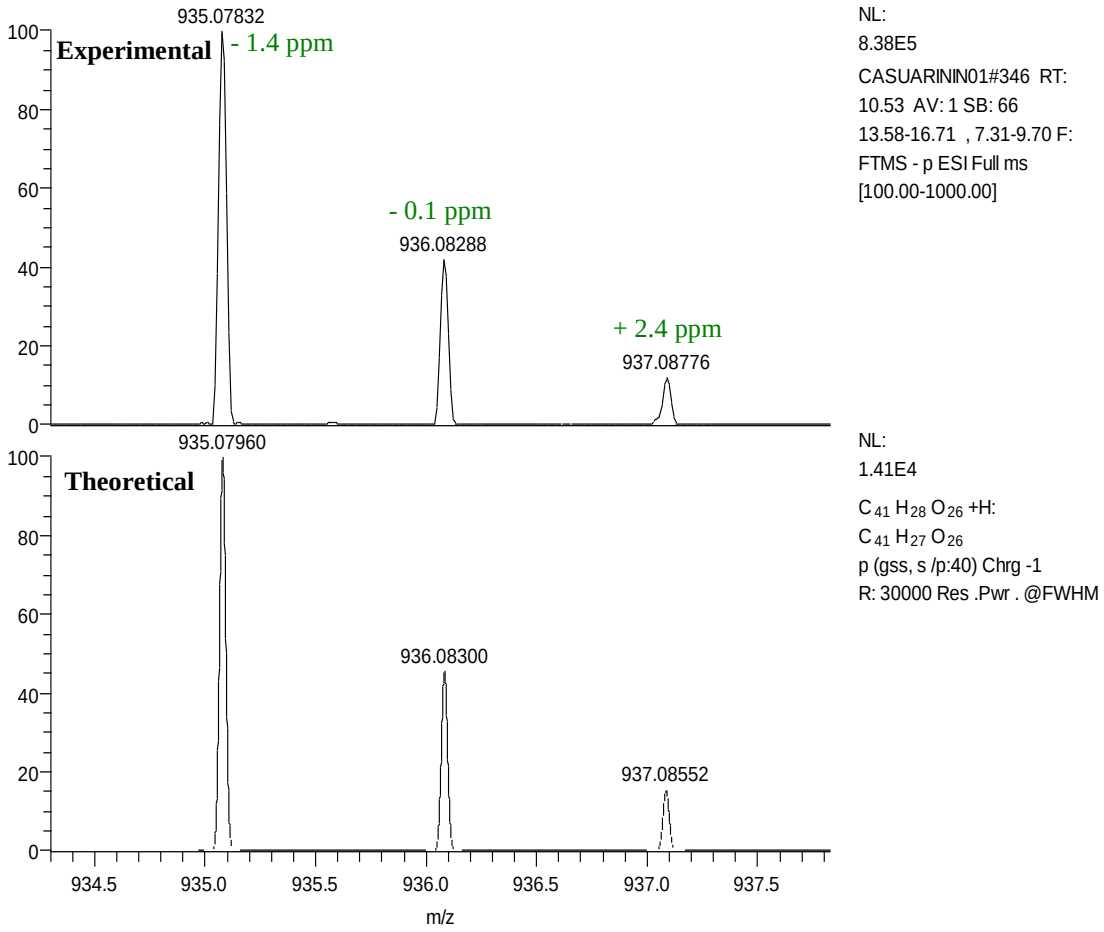


Figure 3.1.3.8: Full scan mass spectra (*m/z*) with isotopic pattern for casuarinin (**18**).

Casuarinin (**18**), ET-4, a pale yellow amorphous powder, with the molecular formula of $C_{41}H_{28}O_{26}$ showed deprotonated molecular $[M–H]^-$ ion at *m/z* 935. Compound ET-4 was different with compounds **15-17** in that it has a galloyl and two HHDP moieties attached to an open ring glucose core. The parent ion gave ions at *m/z* 633, 917, 873 and 783 which showed the mass loss of HHDP, H_2O, CO_2 and galloyl moiety, respectively (Figure 3.1.3.7). In addition, an open ring of monogalloyl-glucose fragment was also shown in MS^3 at *m/z* 299. The mass loss of H_2O (*m/z* 18) and CO_2 (*m/z* 44) yielded the product ions at *m/z* 615, 598, 571 and 553. The MS^3 spectrum also showed the removal of gallate (*m/z* 170) and galloyl (*m/z* 152) group. The mass accuracy for compound ET-4 was calculated and the isotopic pattern was compared with the theoretical mass spectra as

shown in Figure 3.1.3.8. The ^{1}H and ^{13}C-NMR data of compound ET-4 are tabulated in Tables 3.1.3.1 and 3.1.3.2, respectively. Three singlets at δ 6.52, δ 6.56 and δ 6.77 were attributed to the protons in HHDP group, while the signal at δ 7.12 was attributed to the protons in galloyl group (Appendix I). The glucose proton signals of glucose were observed in the range of δ 4.07-5.65. Correspondingly, six glucose carbons were observed in the region of δ 64.4-76.9 with the absence of a hemiacetal carbon which was expected in the range of δ 90.0-100.0. The NMR data also included five ester carbon signals at δ 165.3, δ 166.1, δ 168.7, δ 169.1 and δ 169.6 which verified the occurrence of the open-chain glucose in compound ET-4. The downfield shift of carbon signal of HHDP-3' at δ 117.2 was due to the carbon linkage with C-1 of glucose. Based on the data, compound ET-4 was deduced as casuarinin (Okuda *et al.*, 1983; Nonaka *et al.*, 1985; Hatano *et al.*, 1988; Yoshimura *et al.*, 2008).

(d) Fragmentation Schemes of Ellagitannins

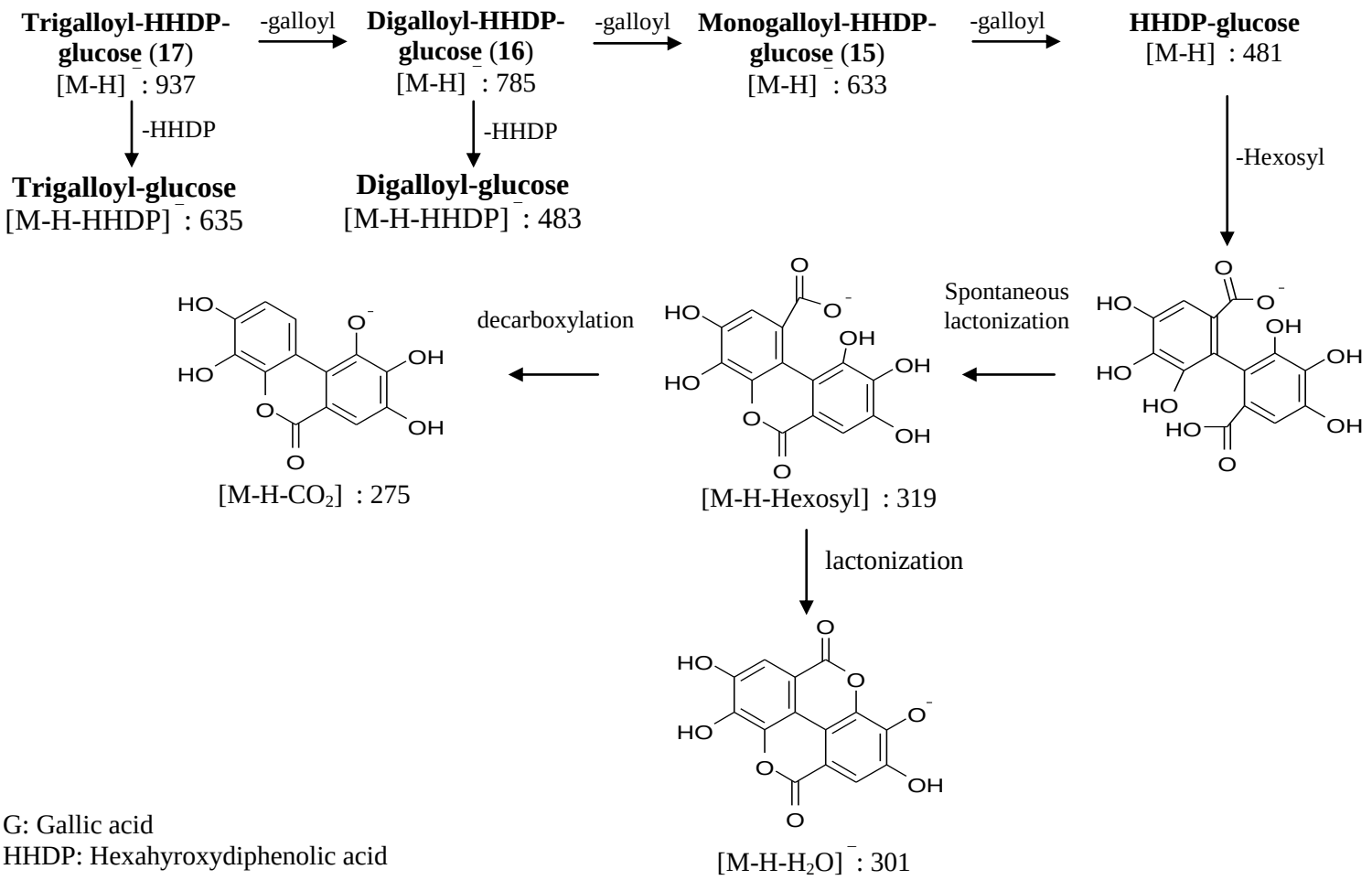


Figure 3.1.3.9: Fragmentation pattern of ellagitannins (**15-17**) in negative mode.

The ellagitannins (**15-17**) demonstrate esterification of the HHDP group at the C-2 and C-3 of the glucose moiety (Figure 3.1.3.9). 6-*O*-galloyl-2,3-*O*-(*S*)-hexahydroxydiphenoyl-D-glucose (**15**), praecoxin B (**16)** and pterocarinin C (**17**) were grouped as monogalloyl-HHDP-glucose, digalloyl-HHDP-glucose and trigalloyl-HHDP-glucose, respectively. The ESI-MS depicted the subsequent mass loss of galloyl moiety (m/z 152), from trigalloyl-HHDP-glucose (*m/z* 937) to yield digalloyl-HHDP-glucose (*m/z* 785), then monogalloyl-HHDP-glucose (*m/z* 633) and finally HHDP-glucose (*m/z* 481) as shown in Figure 3.1.3.9 and Table 3.1.2. The ellagitannins were also fragmented to become gallotannins by losing the mass of 302 (HHDP) and consequently resulted in fragment ions at *m/z* 635 (trigalloyl-glucose) and *m/z* 483 (digalloyl-glucose). These daughter ions are suggested to undergo similar fragmentation as the gallotannins (Figure 3.1.2.15). The monogalloyl-HHDP-glucose did not show the mass of monogalloyl-glucose and instead two dominant fragment ions at *m/z* 301 and 275 were observed. The ion at *m/z* 301 (loss of glucose and galloyl groups) was the result of spontaneous dilactonization of HHDP moiety. The partial lactonization of HHDP unit yielded a product ion at *m/z* 319 and subsequent decarboxylation formed an ion at *m/z* 275 $[M–H–44]^-$. As shown in Table 3.1.2, the MS^3 of *m/z* 301 in compounds **15** and **16** gave prominent ions at *m/z* 273, 257 and 229, similar to deprotonated ellagic acid (Mämmelä *et al.*, 2000; Mullen *et al.*, 2003; Soong and Barlow, 2005; He and Xia, 2007, Simirgiotis and Schmeda-Hirschmann, 2010). These ions were suggested to be corresponded to $[M–H–CO]^-$, $[M–H–CO_2]^-$ and $[M–H–CO_2–CO]^-$, respectively. Casuarinin (**18**) has two HHDP groups attached to an open ring glucose structure which was also applicable to the fragmentation pathway of ellagitannins (Figure 3.1.3.9). The MS^2 showed high relative abundance of ions at *m/z* 633 (Table 3.1.2) and the subsequent fragmentation showed the similarity to 6-*O*-galloyl-2,3-*O*-(*S*)-hexahydroxydiphenoyl-D-glucose (**15**).

Table 3.1.3.1: ^{1}H-NMR Spectral data for compounds ET-1 to ET-4 (**15-18**) in Acetone-d_6.

	δ_H					
	ET-1 (15)		**ET-2 (16)**		**ET-3 (17)**	**ET-4 (18)**
	(α-form)	**(β-form)**	**(α-form)**	**(β-form)**		
Glucose						
H-1	5.40 d (3.6)	5.04 d (8.1)	5.51 d (3.6)	5.19 d (8.4)	6.35 d (8.4)	5.65 d (4.8)
H-2	4.94 dd (3.6, 9.6)	4.75 dd (8.1, 9.6)	5.09 dd (3.6, 9.6)	4.90 dd (8.4, 9.6)	5.21 t (8.4)	4.68 dd (1.8, 4.8)
H-3	5.41 dd (9.6, 9.9)	5.07 t (9.6)	5.64 t (9.6)	5.37 t (9.6)	5.40 dd (8.4, 9.3)	5.36 m
H-4	3.92 t (9.9)	3.91 t (9.6)	5.52 t (10.0)	5.46 t (9.6)	5.68 t (9.3)	5.36 m
H-5	4.24 m	3.82 m	4.54 ddd (2.0, 4.0, 10.0)	4.22 ddd (2.0, 5.0, 9.5)	4.53 dd (4.2, 9.3)	5.36 m
H-6a	4.57 dd (1.8, 12.0)	4.53 dd (1.5, 12.0)	4.52 dd (2.0, 12.0)	4.52 dd (2.0, 12.0)	4.56 d (12.9)	4.81 dd (3.0, 13.0)
H-6b	4.48 dd (4.5, 12.0)	4.45 dd (4.5, 12.0)	4.29 dd (4.0, 12.0)	4.28 dd (5.0, 12.0)	4.34 dd (4.2, 12.9)	4.07 d (13.0)
Galloyl						
H-2", -6"	7.15	7.15	7.16 × 2	7.11, 7.14	7.08, 7.13, 7.16	7.12
HHDP						
H-3, -3'	6.58, 6.68	6.58, 6.68	6.39, 6.61	6.39, 6.61	6.43, 6.45	6.52, 6.56, 6.77

HHDP-Hexahydroxyldiphenoyl

Table 3.1.3.2: ^{13}C-NMR Spectral data for compounds ET-1 to ET-4 (**15-18**) in Acetone-d_6.

	δ_C					
	ET-1 (15)		**ET-2 (16)**		**ET-3 (17)**	**ET-4 (18)**
	(α-form)	**(β-form)**	**(α-form)**	**(β-form)**		
Glucose						
C-1	91.4	94.9	91.3	94.9	92.0	66.9
C-2	75.3	77.7	75.1	77.7	75.4	76.9
C-3	78.0	80.3	75.3	77.6	77.4	69.8
C-4	68.2	68.6	68.7	68.5	67.9	74.0
C-5	70.6	75.1	68.4	73.1	74.2	71.1
C-6	63.9	63.9	63.1	63.2	62.7	64.4
Galloyl						
C-1"	121.5, 121.6		120.5, 120.6, 121.3, 121.4		119.9, 120.6, 121.3	124.7
C-2", -6"	109.7, 109.8		109.6, 109.9		110.0, 110.1, 110.2	109.9 x2
C-3", -5"	145.8, 145.9		145.8, 145.9		145.9, 146.0, 146.1	145.2 x2
C-4"	138.6, 138.7		138.7, 139.2		138.8, 139.3, 139.7	139.4
HHDP						
C-1, C-1'	114.3, 114.4		114.2, 114.3, 114.5, 114.6		114.4, 114.6	115.1, 116.0, 116.1, 116.3
C-2, C-2'	126.9, 127.4		126.5, 126.6, 126.7		126.1, 126.3	119.8, 120.5, 126.7, 127.2
C-3, C-3'	107.2, 107.5, 107.6		107.1, 107.2, 107.6		107.1, 107.2	105.3, 107.0, 108.5, 117.2[a]
C-4, C-4'	144.9, 145.0		144.9, 145.0		145.0, 145.7	145.7, 145.9 x2, 146.0
C-5, C-5'	136.0, 135.9		135.9, 136.0, 136.1, 136.2		136.2, 136.3	134.9, 135.9, 136.8, 138.7
C-6, C-6'	144.1, 144.3		144.2, 144.3		144.5, 145.0	143.4, 143.6, 144.4, 144.5,
Ester carbonyl	166.6, 166.7, 169.1, 169.5		165.7, 166.5, 169.0, 169.1, 168.8, 168.9		164.9, 165.6, 166.4, 168.6, 169.0	165.3, 166.1, 168.7, 169.1, 169.6

HHDP-Hexahydroxyldiphenoyl; [a] Signal due to the carbon linked to C-1 of glucose.

3.1.4 Ellagic Acid Derivatives

(a) 3'-*O*-Methyl-3,4-methylenedioxyellagic acid 4'-*O*-β-D-glucopyranoside (19), ED-1

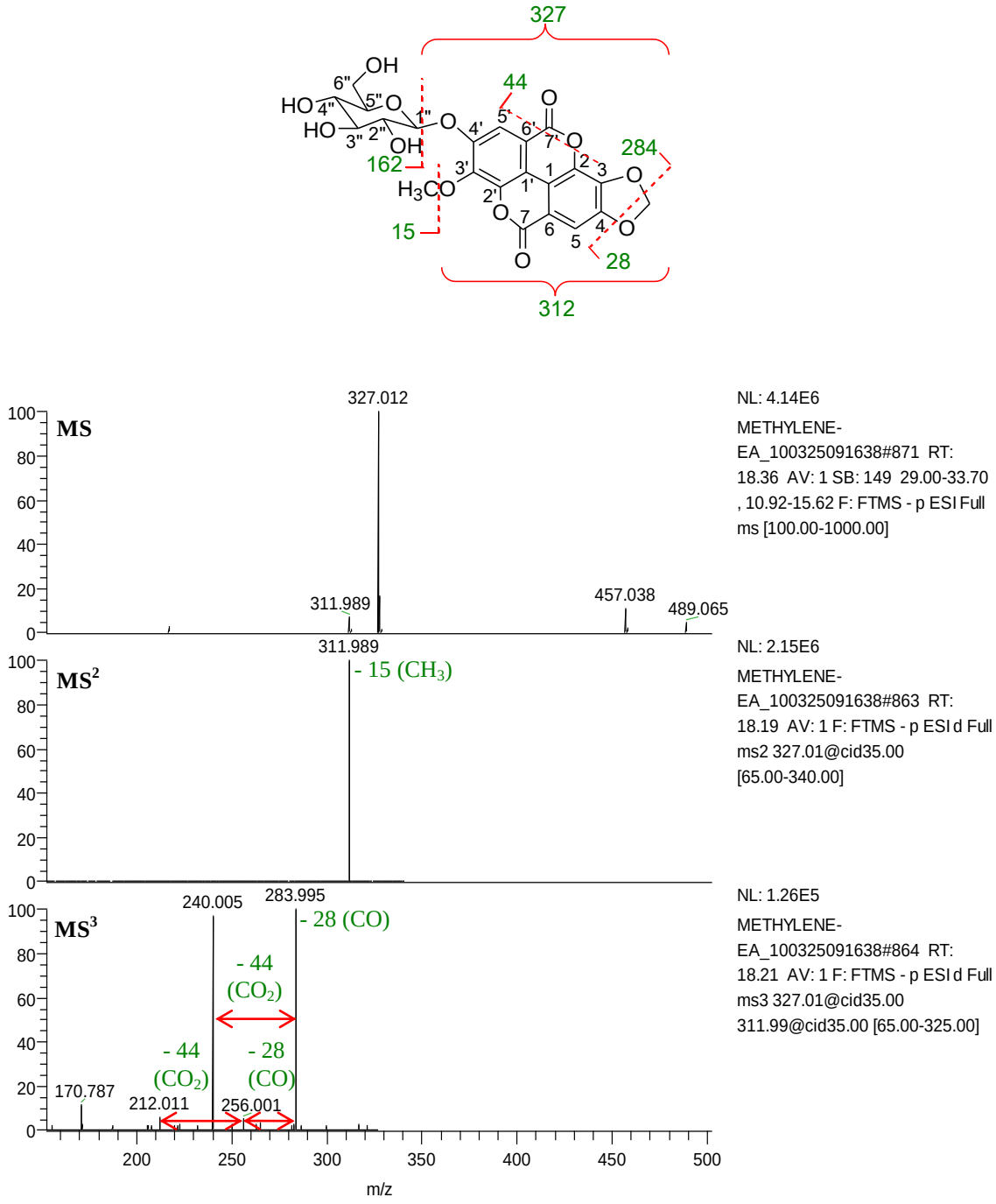


Figure 3.1.4.1: Fragmentation pattern for 3'-*O*-methyl-3,4-methylenedioxyellagic acid 4'-*O*-β-D-glucopyranoside (**19**).

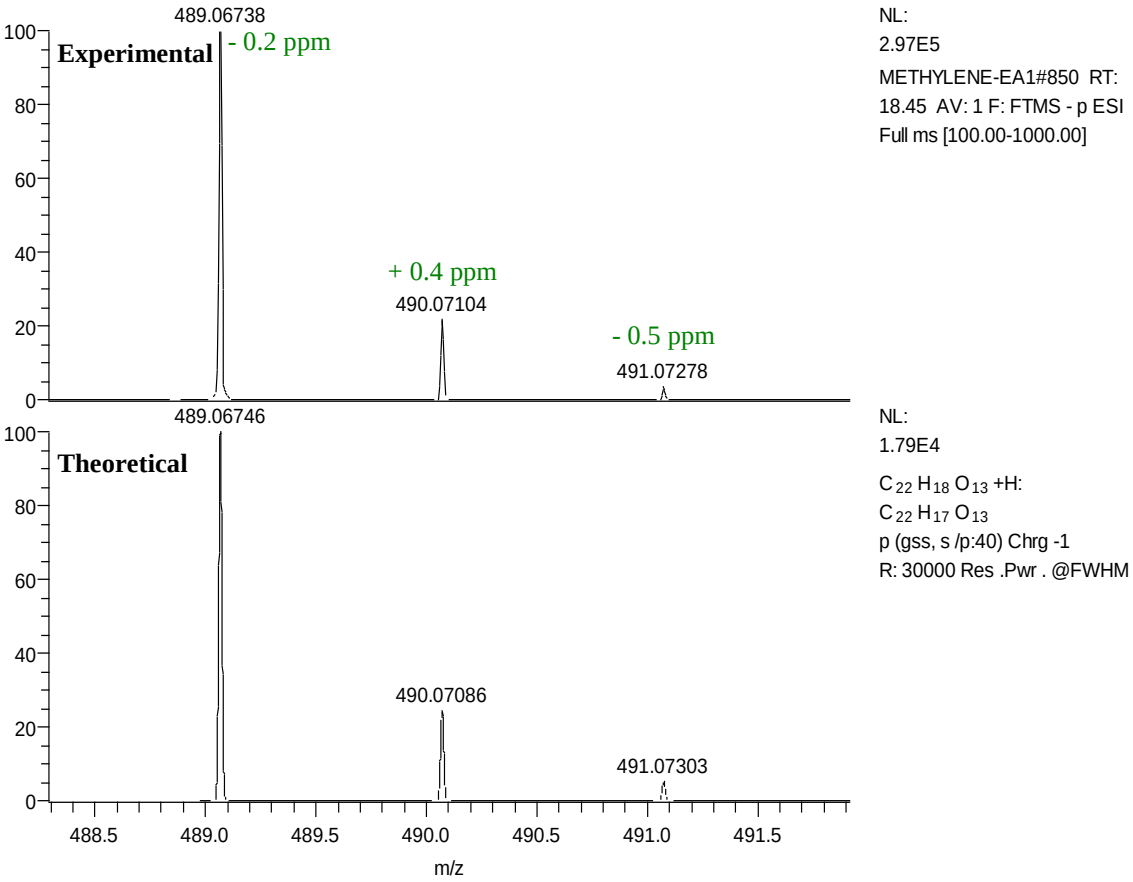


Figure 3.1.4.2: Full scan mass spectra (*m/z*) with isotopic pattern for 3'-*O*-methyl-3,4-methylenedioxyellagic acid 4'-*O*-β-D-glucopyranoside (**19**).

3'-*O*-Methyl-3,4-methylenedioxyellagic acid 4'-*O*-β-D-glucopyranoside (**19),** ED-1, obtanined as granular crystals, with the molecular formula of $C_{22}H_{18}O_{13}$ showed deprotonated molecular $[M–H]^-$ ion at *m/z* 489. Compound ED-1 has a methyl and a methylene group in the ellagic acid skeleton. It was dominated by an intense ion peak at *m/z* 327 due to the removal of a glucose moiety $[M–H–162]^-$. Upon fragmentation, the major ion peak yielded a ion at *m/z* 312, indicating the loss of a CH_3 group $[M–H–15]^-$. Further fragmentation of ion at *m/z* 312 showed the cleavage of the methylene group which contributed to ion at *m/z* 284. Subsequent decarboxylation of ion at *m/z* 284 gave signal at *m/z* 240 with the mass loss of 44. The mass accuracy in ppm and isotopic pattern are presented in Figure 3.1.4.2. In the ^{1}H-NMR spectrum, a methoxy proton singlet (δ 4.09) and a methylene proton singlet δ 6.39 between C-3 and C-4 were observed (Appendix I). The anomeric carbon at δ 101.5 was particularly shifted

downfield (Table 3.1.4.1). Consequently, compound ED-1 was determined as 3'-*O*-methyl-3,4-methylenedioxyellagic acid 4'-*O*-β-D-glucopyranoside (Li *et al.*, 1999; Li *et al.*, 2000; Khallouki *et al.*, 2007).

(b) 3,3',4-Tri-*O*-methylellagic acid 4'-*O*-β-D-glucopyranoside (20), ED-2

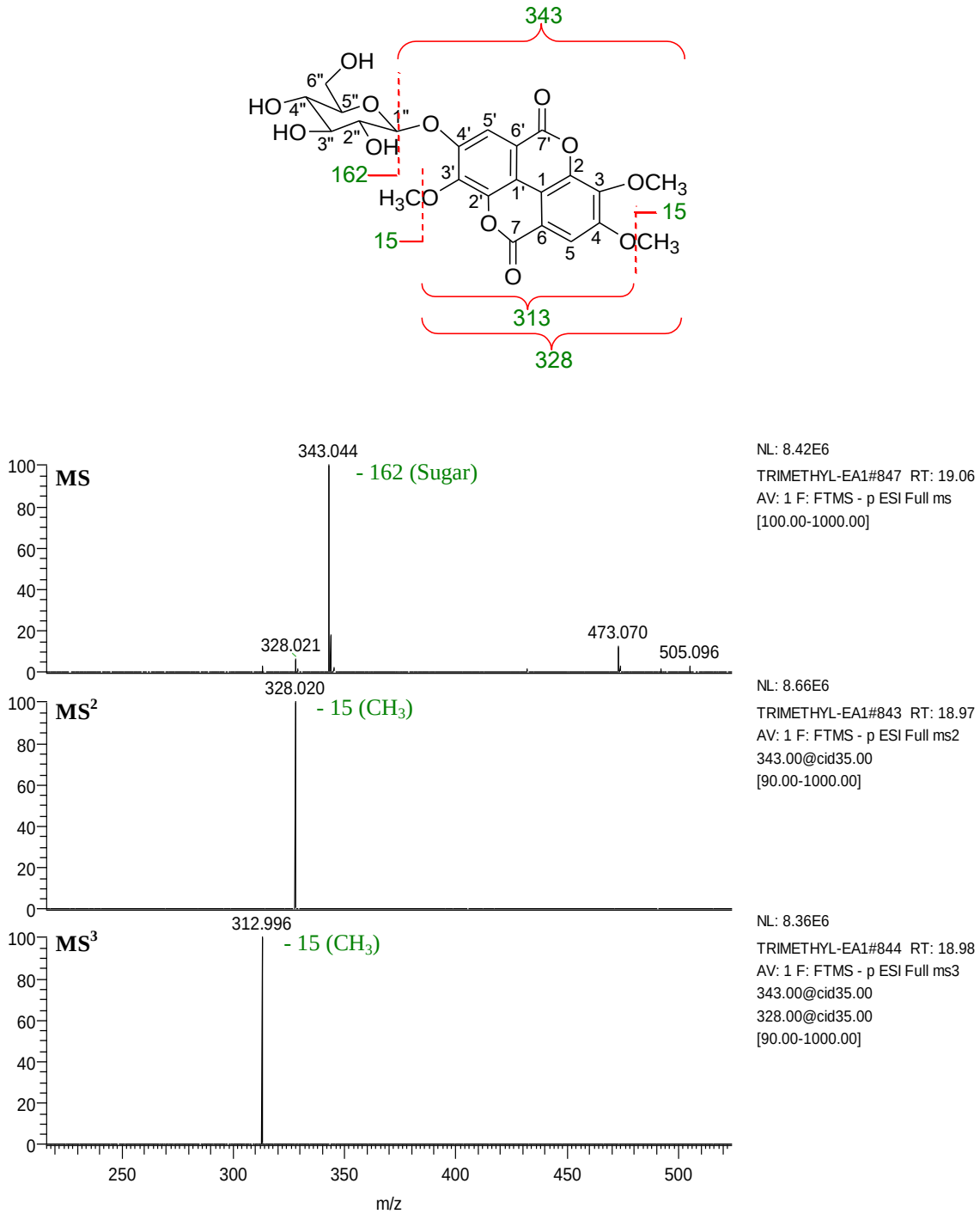


Figure 3.1.4.3: Fragmentation pattern for 3,3',4-tri-*O*-methylellagic acid 4'-*O*-β-D-glucopyranoside (**20**).

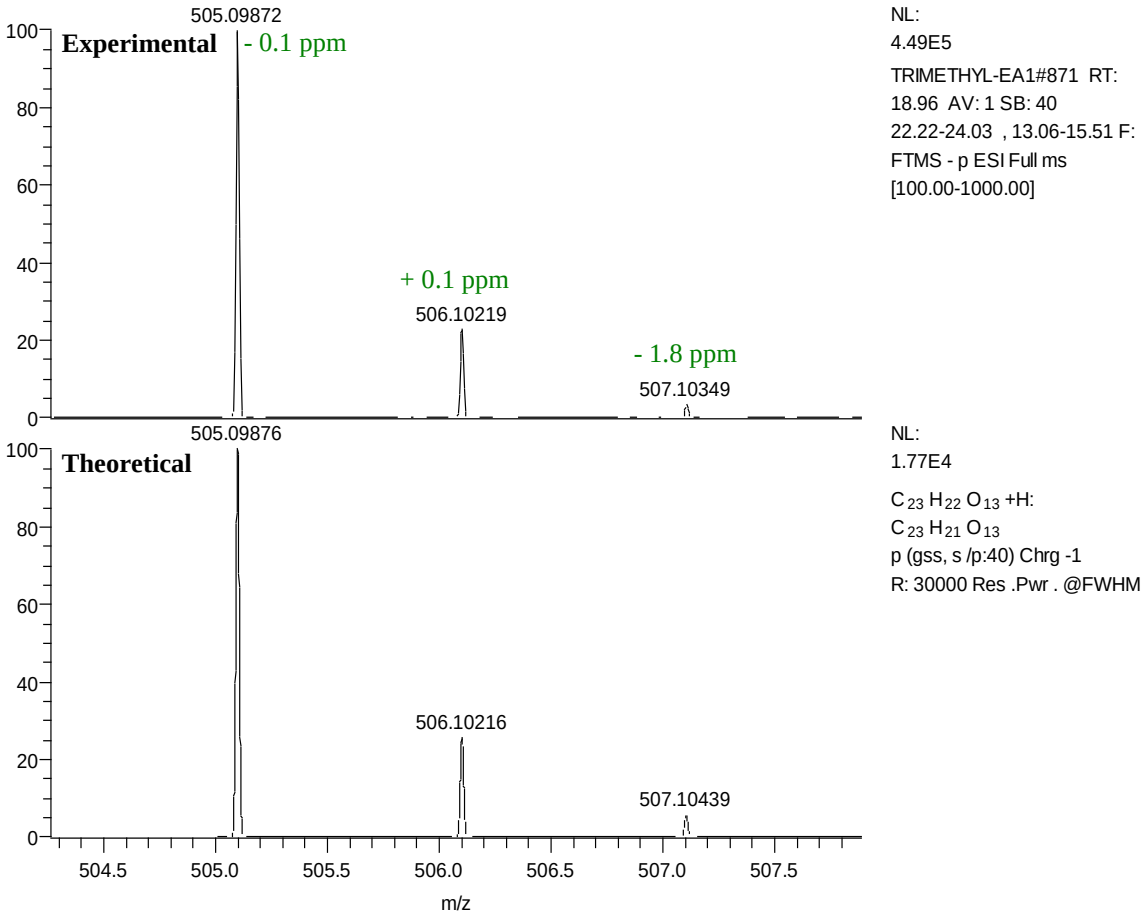


Figure 3.1.4.4: Full scan mass spectra (*m/z*) with isotopic pattern for 3,3',4-tri-*O*-methylellagic acid 4'-*O*-β-D-glucopyranoside (**20**).

3,3',4-Tri-*O*-methylellagic acid 4'-*O*-β-D-glucopyranoside (**20**), ED-2, obtained as granular crystals, with the molecular formula of $C_{23}H_{22}O_{13}$ showed deprotonated molecular ion $[M–H]^-$ at *m/z* 505. Compound ED-2 has three methyl groups on ellagic acid which is attached to a glucose moiety. The deprotonated parent ion at *m/z* 505 and its MS/MS are presented in Figure 3.1.4.3. It showed the dominant ion peak at *m/z* 343 due to the removal of a glucose moiety $[M–H–162]^-$ and this ion was further fragmented to ion at *m/z* 328 and 313 as a result of the mass loss of a CO_2 group (*m/z* 44). The isotopic pattern and the mass accuracy of compound ED-2 are presented in Figure 3.1.4.4. Three methoxy groups (3-OMe, 4-OMe and 3''-OMe) were assigned at δ 4.04, δ 4.00 and δ 4.09, respectively in the ^{1}H-NMR (Table 3.1.4.1). The coupling constant (J = 6.9 Hz) for the anomeric proton (H-1'') at δ 5.15 indicated the β-form of

sugar moiety (Appendix I). The glucosyl linkage at position C-4' was supported by downfield shift of H-5'. Thus, compound ED-2 was characterised as 3,3',4-tri-*O*-methylellagic acid 4'-*O*-β-D-glucopyranoside (Li *et al.*, 1999; Li *et al.*, 2000).

Table 3.1.4.1: ^{1}H-NMR and ^{13}C-NMR Spectral data for compounds ED-1 (**19**) to ED-2 (**20)** in DMSO-d_6.

	δ_H		δ_C	
	ED-1 (19)	**ED-2 (20)**	**ED-1 (19)**	**ED-2 (20)**
Aglycon				
1	-	-	115.7	113.1
2	-	-	131.6	141.4
3	-	-	138.6	141.0
4	-	-	150.8	154.5
5	7.53 s	7.63 s	104.2	107.8
6	-	-	112.0	112.8
7	-	-	158.4	158.6
1'	-	-	113.8	113.8
2'	-	-	141.5	141.4
3'	-	-	142.0	141.9
4'	-	-	152.2	152.1
5'	7.83 s	7.82 s	112.7	112.5
6'	-	-	112.5	112.2
7'	-	-	157.7	158.4
-CH_2-	6.39 s	-	104.7	-
3-OMe	-	4.04 s	-	61.5
4-OMe	-	4.00 s	-	60.0
3'-OMe	4.09 s	4.09 s	61.9	61.9
Glucose				
1"	5.15 d (6.9)	5.16 d (6.9)	101.5	101.5
2"	3.41	3.41	73.5	73.5
3"	3.39	3.39	76.7	76.7
4"	3.22	3.25	69.7	69.7
5"	3.45	3.45	77.5	77.5
6a"	3.69 br d (11.4)	3.69 br d (10.2)	60.7	60.7
6b"	3.51 br d (9.0)	3.51 br d (11.7)	-	-

3.1.5 Aromatic Compounds

(a) Gallic acid (21), AR-1

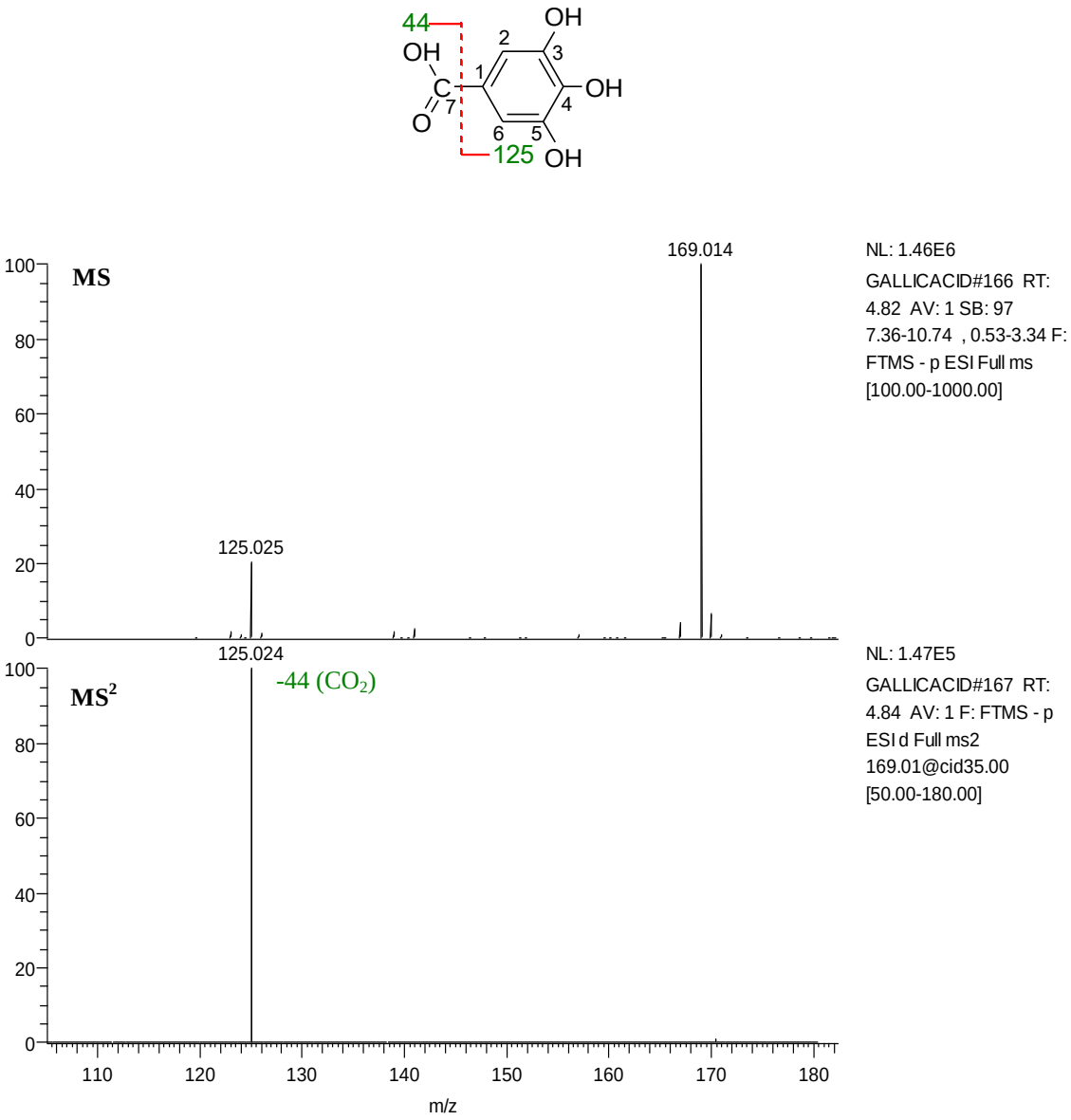


Figure 3.1.5.1: Fragmentation pattern for gallic acid (**21**).

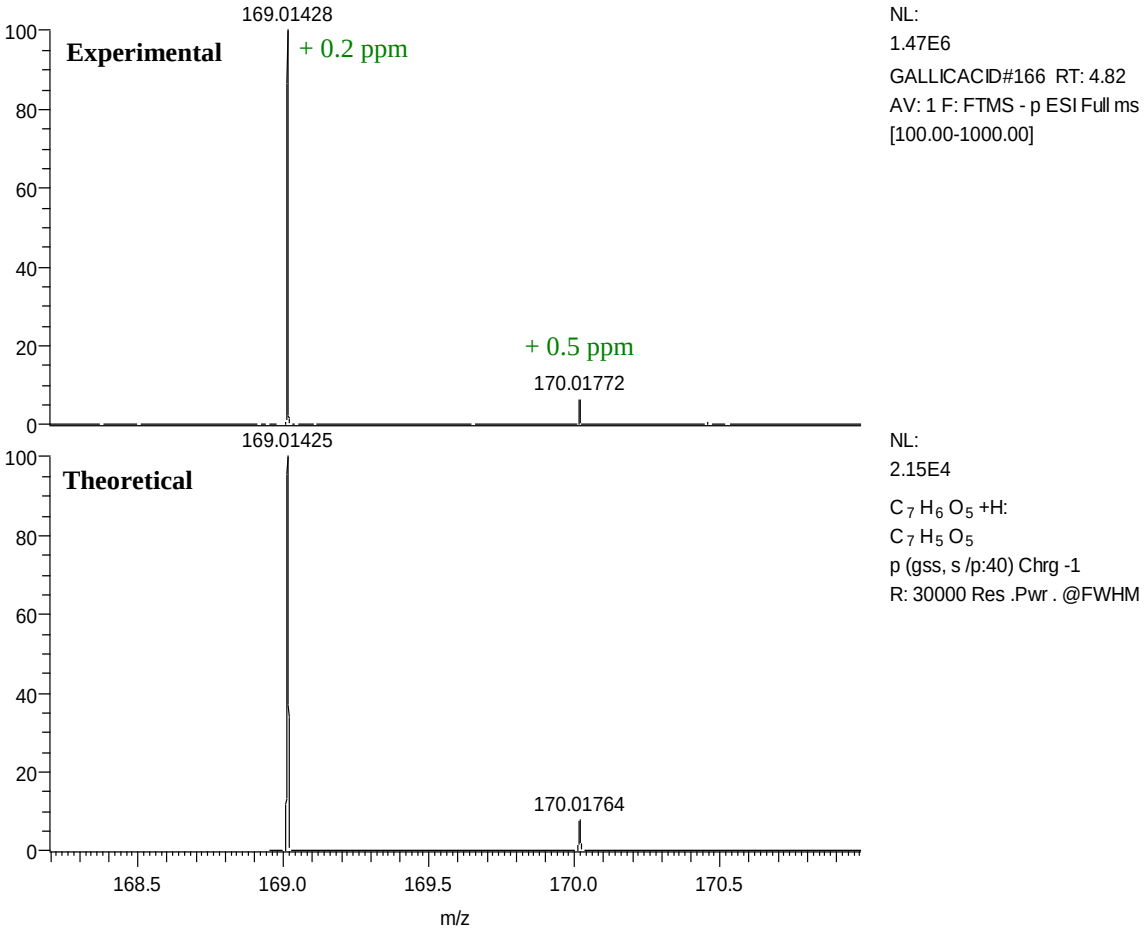


Figure 3.1.5.2: Full scan mass spectra (*m/z*) with isotopic pattern for gallic acid (**21**).

Gallic acid (**21**), AR-1, colorless needles, with the molecular formula of $C_7H_6O_5$ showed deprotonated $[M–H]^-$ ion at *m/z* 169. Fragmentation of ion at *m/z* 125 $[M–H–COO]^-$ (Figure 3.1.5.1) indicated a mass loss of *m/z* 44 (CO_2) (Mämmelä *et al.*, 2000; Zywicki *et al.*, 2002; Soong and Barlow, 2005; Meyers *et al.*, 2006; He and Xia, 2007; Nuengchamnong and Ingkaninan, 2009). The isotopic pattern and the calculated mass accuracy for compound AR-1 are shown in Figure 3.1.5.2. The ^{1}H and ^{13}C-NMR spectral data are shown in Table 3.1.5.1 and Appendix I. A proton singlet at δ 7.14 corresponded to the H-2 and H-6 of the galloyl moiety. Thus, compound AR-1 was determined as gallic acid (Nawwar *et al.*, 1982; Yoshida *et al.*, 1991c; Chanwitheesuk *et al.*, 2007).

(b) Gallic acid methyl ester (22), AR-2

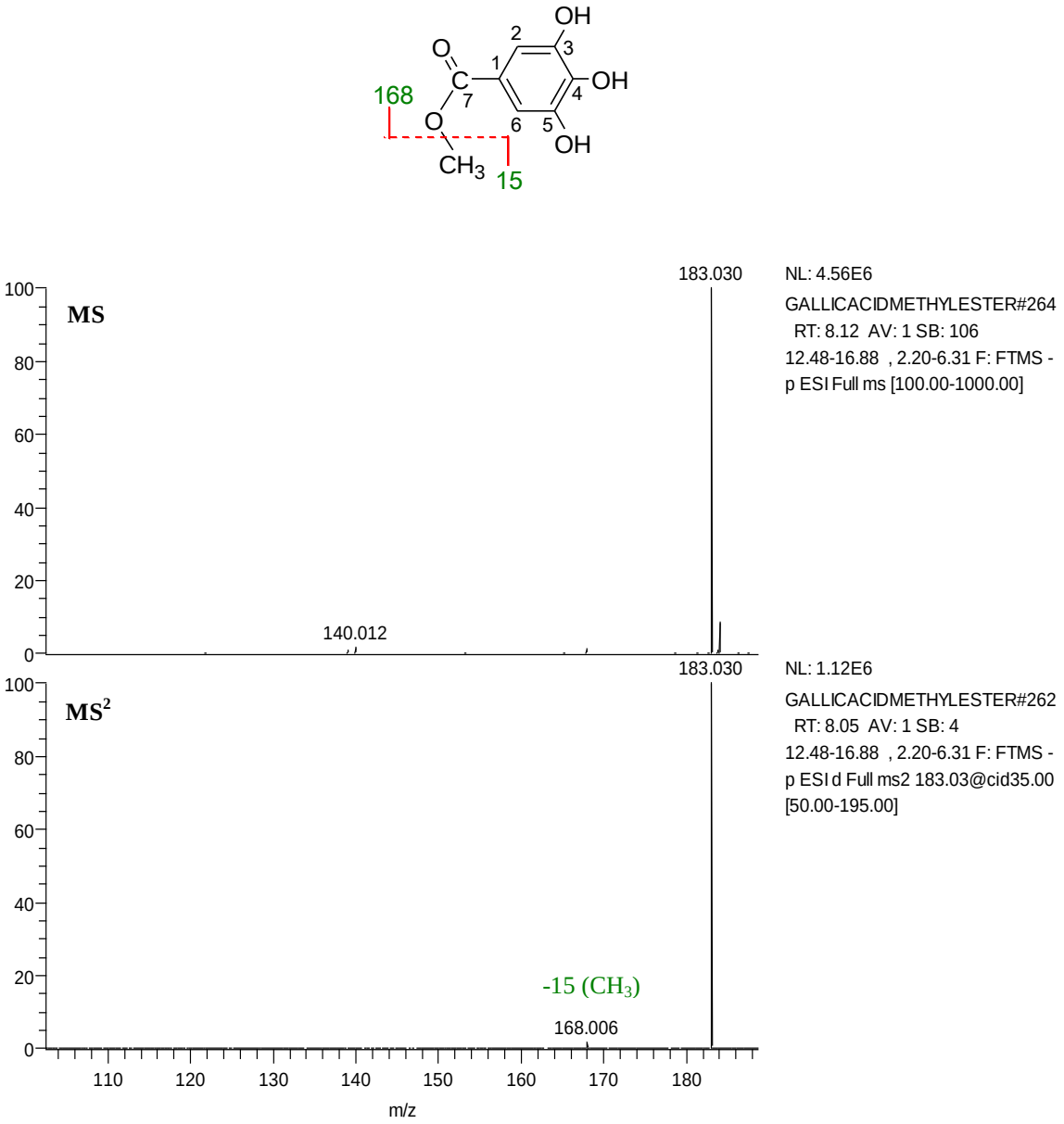


Figure 3.1.5.3: Fragmentation pattern for gallic acid methyl ester (**22**).

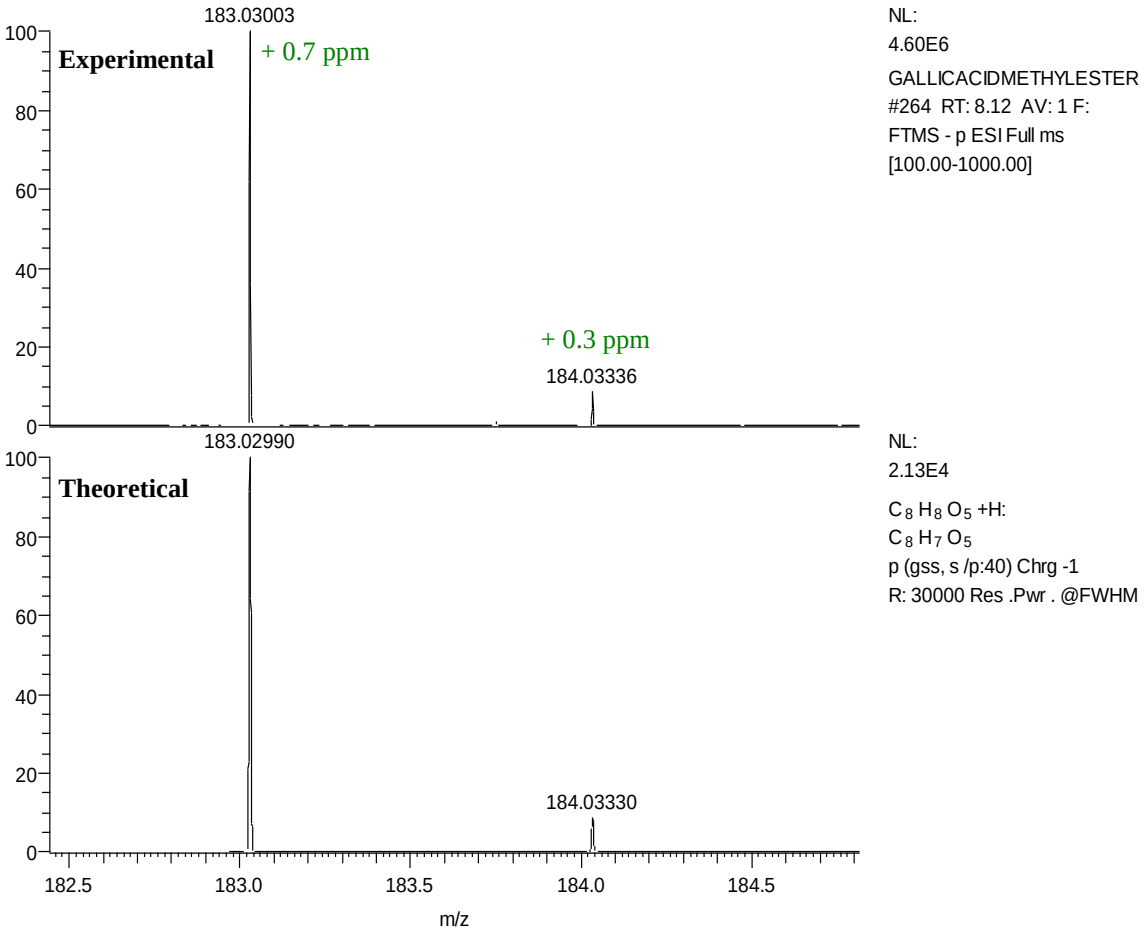


Figure 3.1.5.4: Full scan mass spectra (*m/z*) with isotopic pattern for gallic acid methyl ester (**22**).

Gallic acid methyl ester (**22**), AR-2, a white amorphous powder, with the molecular formula of $C_8H_8O_5$ showed deprotonated molecular ion $[M–H]^-$ at *m/z* 183. The fragment ion at *m/z* 168 was resulted from the mass loss of *m/z* 44 $[M–H–COO]^-$ (Figure 3.1.5.3). Figure 3.1.5.4 shows the characteristic isotopic pattern and the calculated mass accuracy of compound AR-2. This compound was formed by methanolysis of galloyl esters during extraction and separation process (Mueller-Harvey, 2001). In the ^{1}H-NMR spectrum, a 2H signal at δ 7.10 indicated the H-2 and H-6 of the galloyl while the signal at δ 3.75 showed the presence of a methoxy group (Table 3.1.5.1, Appendix I). Consequently, compound AR-2 was identified as gallic acid methyl ester (Khac *et al.*, 1990; Yoshida *et al.*, 1991b).

Table 3.1.5.1: ^{1}H-NMR and ^{13}C-NMR Spectral data for AR-1 to AR-2 (**21**-**22**) in Acetone-d_6.

	δ_H		δ_C	
	AR-1 (21)	**AR-2 (22)**	**AR-1 (21)**	**AR-2 (22)**
Galloyl				
1	-	-	121.8	121.2
2, 6	7.14	7.10	109.8	109.8
3, 5	-	-	145.9	145.9
4	-	-	138.6	138.8
COOH	-	-	168.4	167.6
OMe	-	3.75	-	51.9