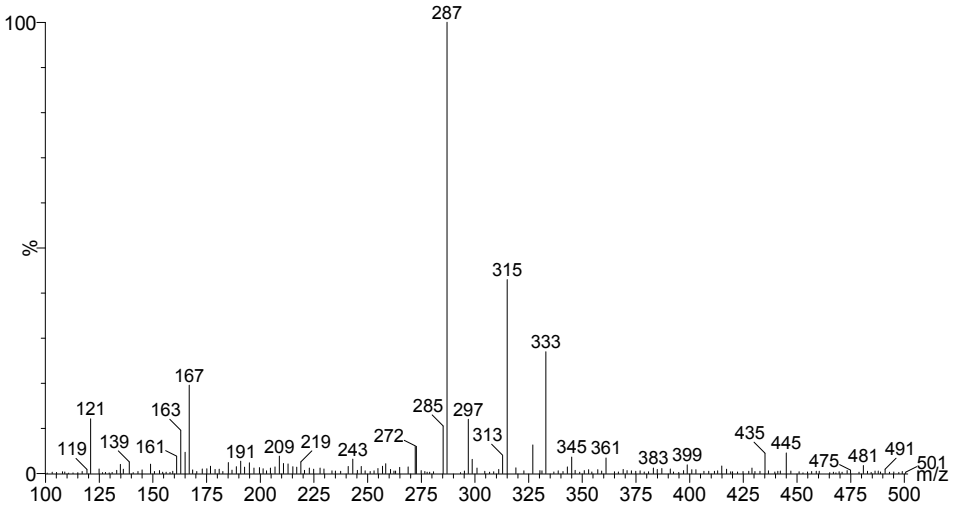
Supplementary Information

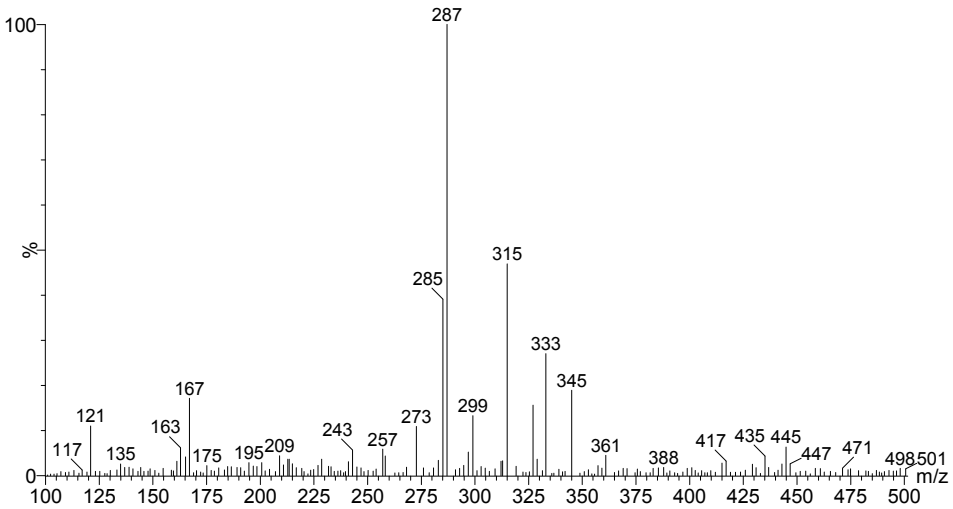
Supplementary figure 1

High cone voltage fragmentation of vitexin and isovitexin products from incubating UDP-glucose and [2',3',5',6'-$^2H_4$]2-hydroxynaringenin with *D. incanum* soluble proteins. The fragmentation pattern of [2',3',5',6'-$^2H_4$]vitexin (**8**) and [2',3',5',6'-$^2H_4$]isovitexin (**9**) shows a molecular ion of $[M+4-H]^-$ 435 due to incorporation of four deuterium atoms, and ions at $[M+4-90-H]^-$ 345 and $[M+4-120]^-$ 315 proving the hexose is *C*-linked.

[2',3',5',6'-$^2H_4$]vitexin (**8**)



[2',3',5',6'-$^2H_4$]isovitexin (**9**)

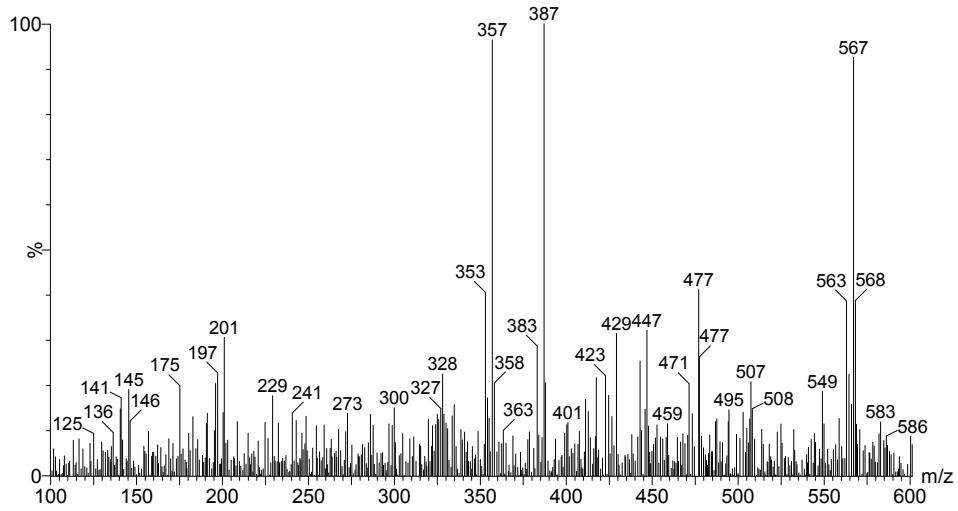


Supplementary figure 2

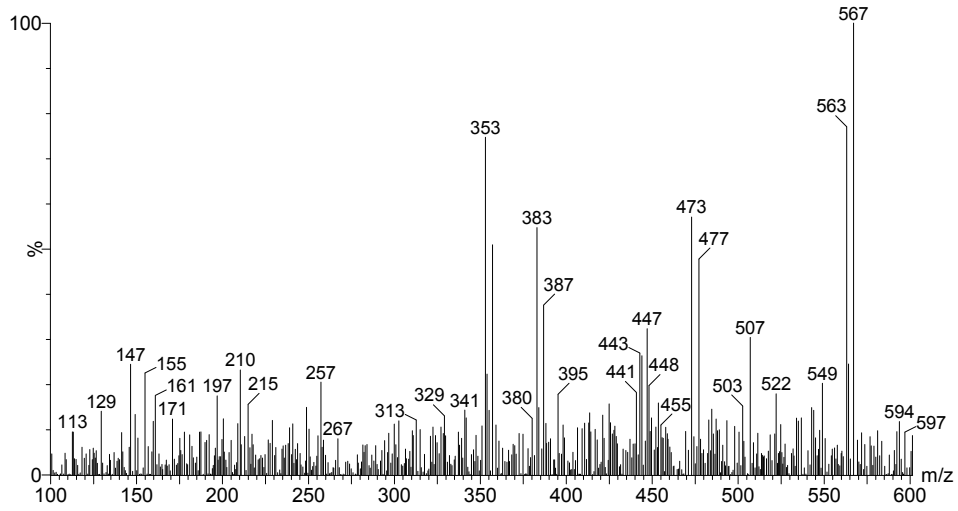
High cone voltage fragmentation of [2',3',5',6'-$^2H_4$]ischaftoside (**4**) and [2',3',5',6'-$^2H_4$]soschaftoside (**5**) products from incubating UDP-arabinose and [2',3',5',6'-$^2H_4$]6-*C*-glucosyl-2-hydroxynaringenin (**10**) with *D. incanum* soluble proteins. The fragmentation pattern shows molecular ions of $[M+4-H]^-$ 567 and ions at $[M+4-60-H]^-$ 507, $[M+4-90-H]^-$ 477, $[M+4-120-H]^-$ 447, $[M+4-180-H]^-$ 387 and [M+4-

210]$^-$ 357 proving both sugars are *C*-linked. The analogous fragmentation patterns relating to [M-H]$^-$ 563 are unlabelled, co-eluting natural products.

[2',3',5',6'-$^2H_4$]isoschaftoside (**4**)

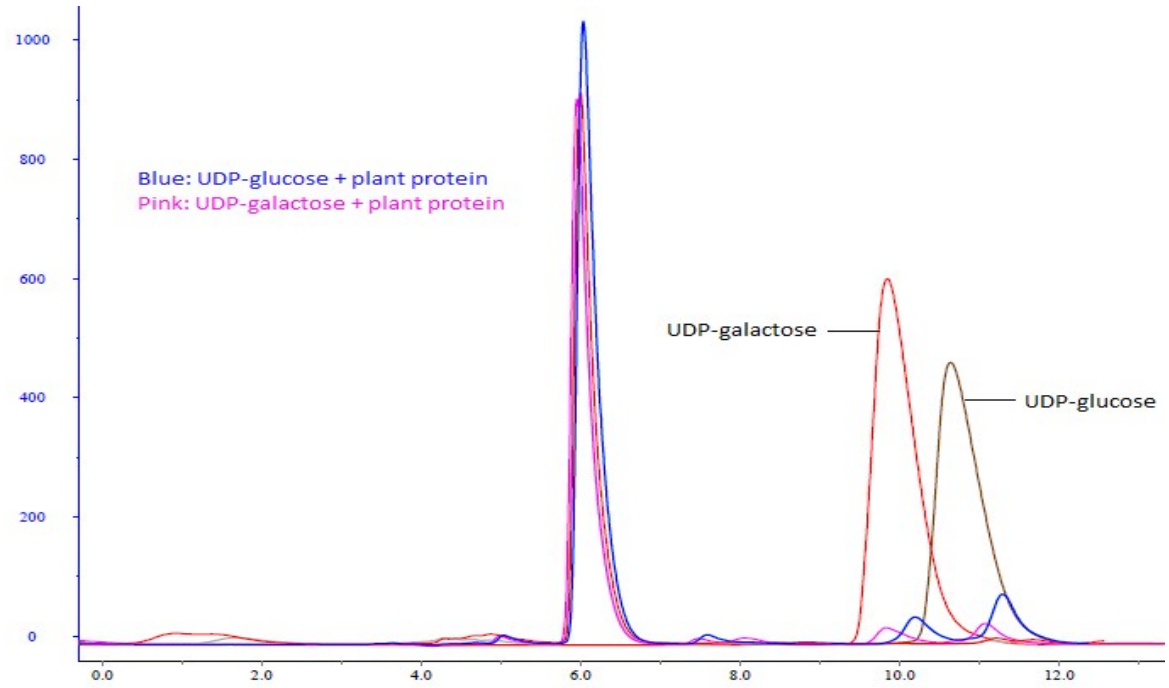


[2',3',5',6'-$^2H_4$]schaftoside (**5**)



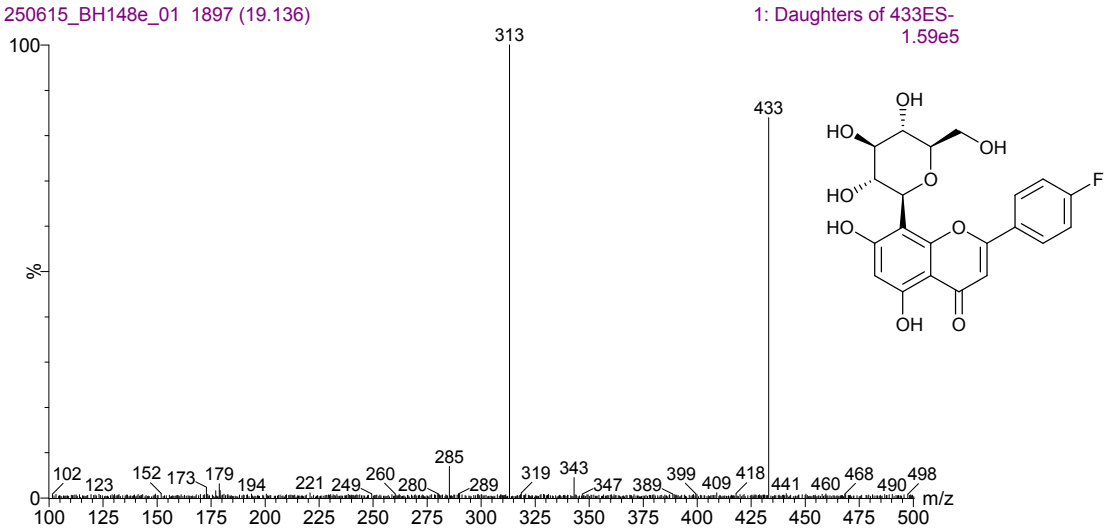
Supplementary figure 3

HPLC traces demonstrating conversion of UDP-glucose to UDP-galactose (blue) and UDP-galactose to UDP-glucose (pink) by *D. incanum* soluble protein fraction.

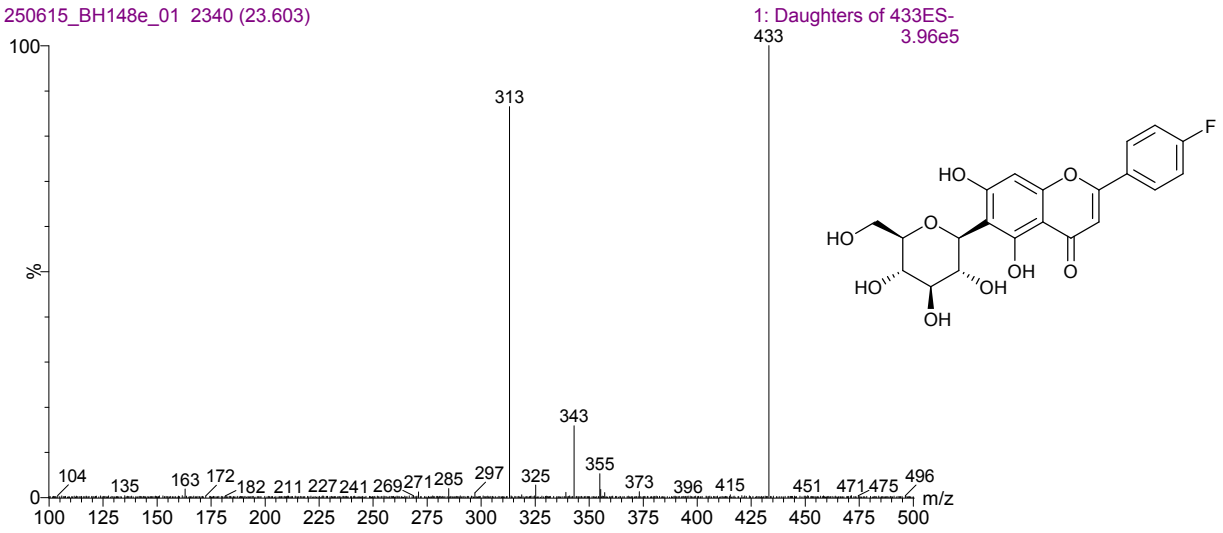


Supplementary figure 4

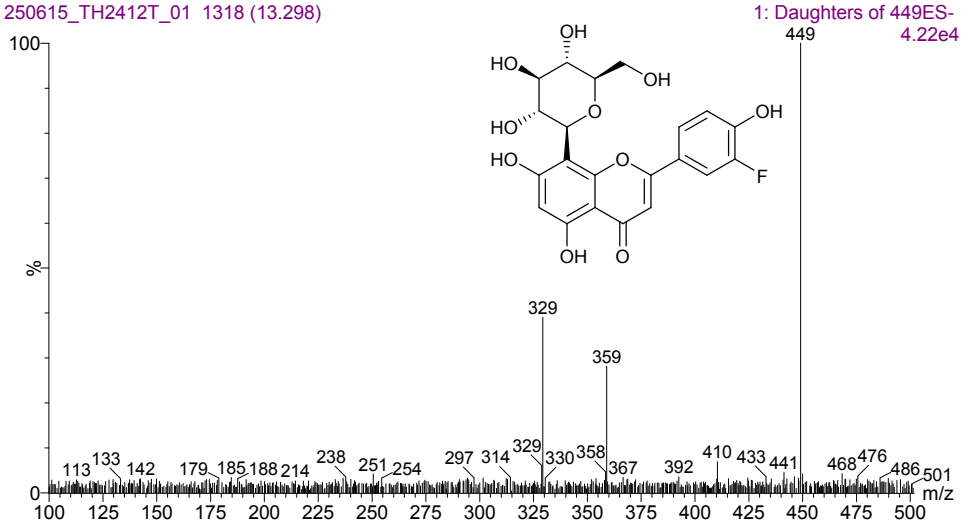
High cone voltage ESIMS fragmentation of **11a** $[M-H]^-$ 433, demonstrating *C*- linked hexose through fragment ions of $[M-H-90]^-$ 343 and $[M-H-120]^-$ 313.



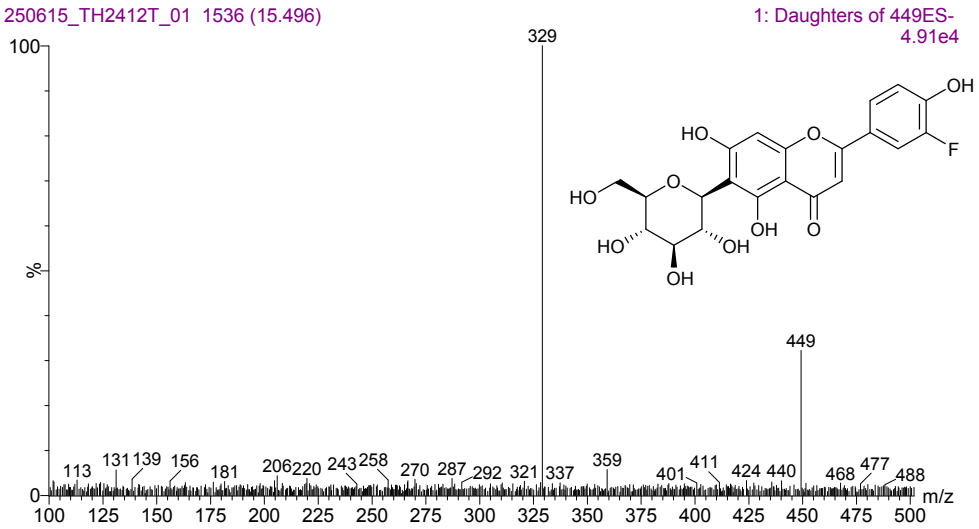
High cone voltage ESIMS fragmentation of **11b** $[M-H]^-$ 433, demonstrating *C*- linked hexose through fragment ions of $[M-H-90]^-$ 343 and $[M-H-120]^-$ 313.



High cone voltage ESIMS fragmentation of **12a** $[M-H]^-$ 449, demonstrating *C*- linked hexose through fragment ions of $[M-H-90]^-$ 359 and $[M-H-120]^-$ 329.

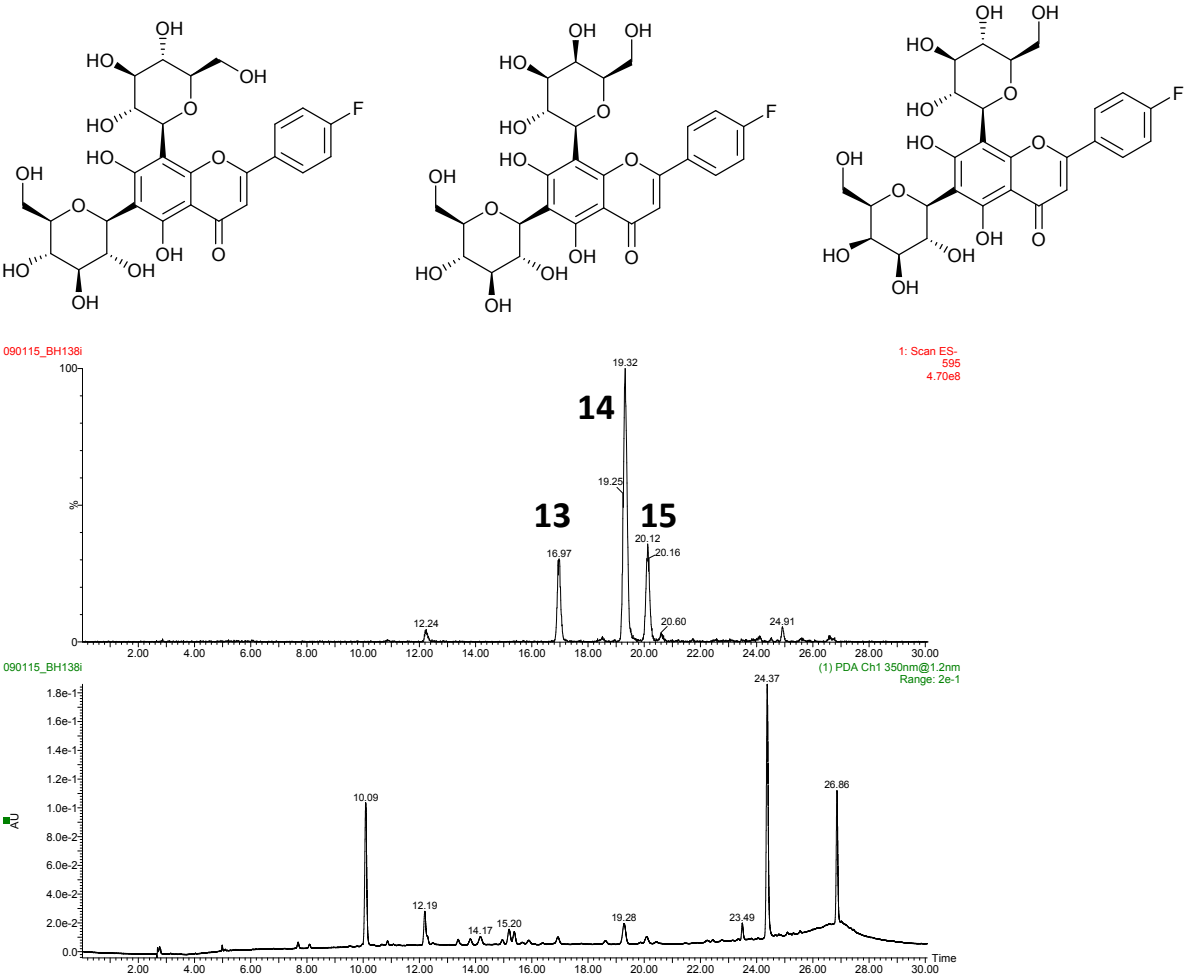


High cone voltage ESIMS fragmentation of **12b** $[M-H]^-$ 449, demonstrating *C*- linked hexose through fragment ions of $[M-H-90]^-$ 359 and $[M-H-120]^-$ 329.

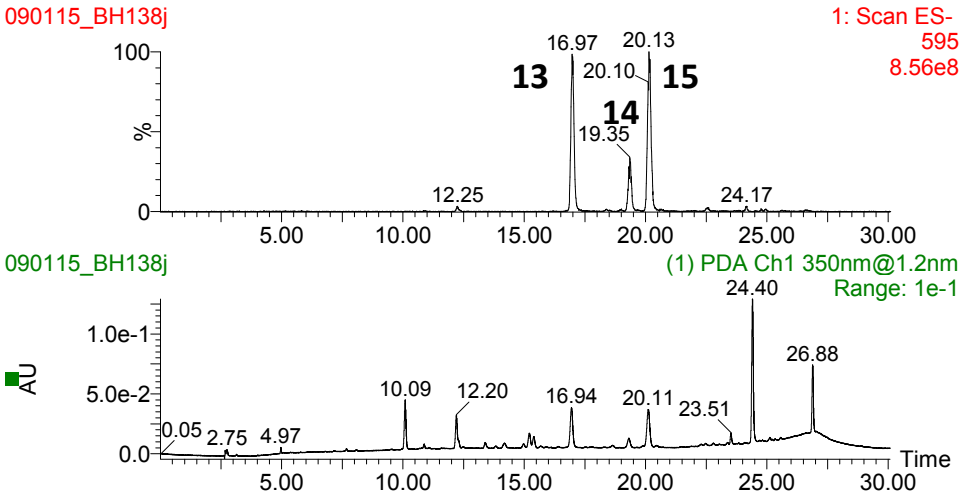


Supplementary figure 5A

LCMS PDA trace (lower) and ion trace (higher) showing 3 novel compounds (**13**, **14** and **15**) with molecular weight $[M-H]^-$ 595 derived from UDP-glucose, **11a** and *D. incanum* root protein.

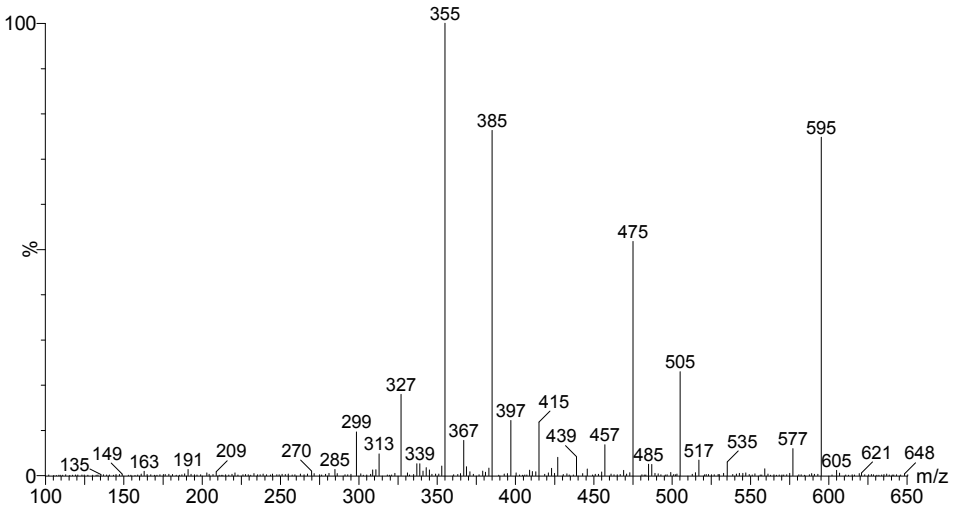


LCMS PDA trace (lower) and ion trace (higher) showing 3 novel compounds (**13**, **14** and **15**) with molecular weight $[M-H]^-$ 595 derived from UDP-galactose, **11a** and *D. incanum* root protein.

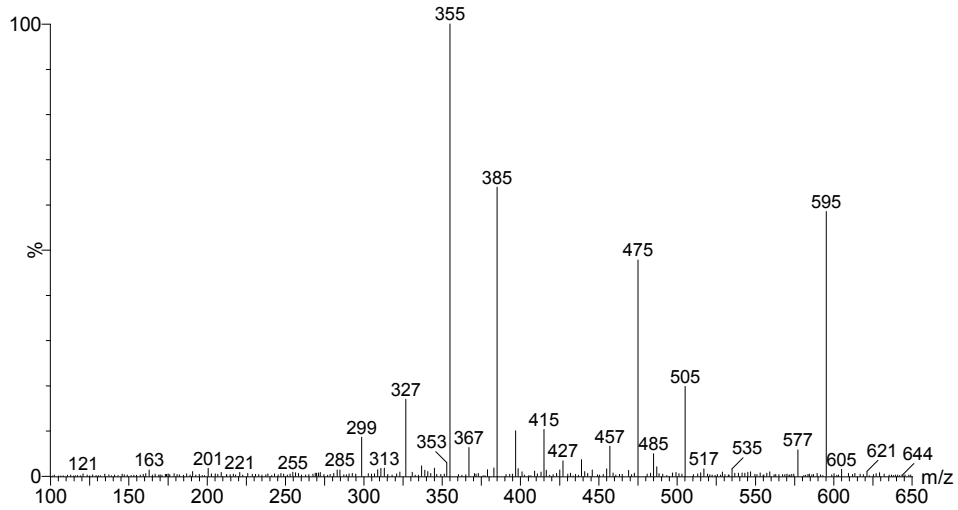


High cone voltage ESIMS mass spectrum of the three novel metabolites **13**, **14**, and **15**, from the ion trace at 595 amus, demonstrating they are di-*C*-linked hexose moieties through fragment ions at $[M-H-90]^-$ 505, $[M-H-120]^-$ 475, $[M-H-180]^-$ 415, $[M-H-210]^-$ 385 and $[M-H-240]^-$ 355.

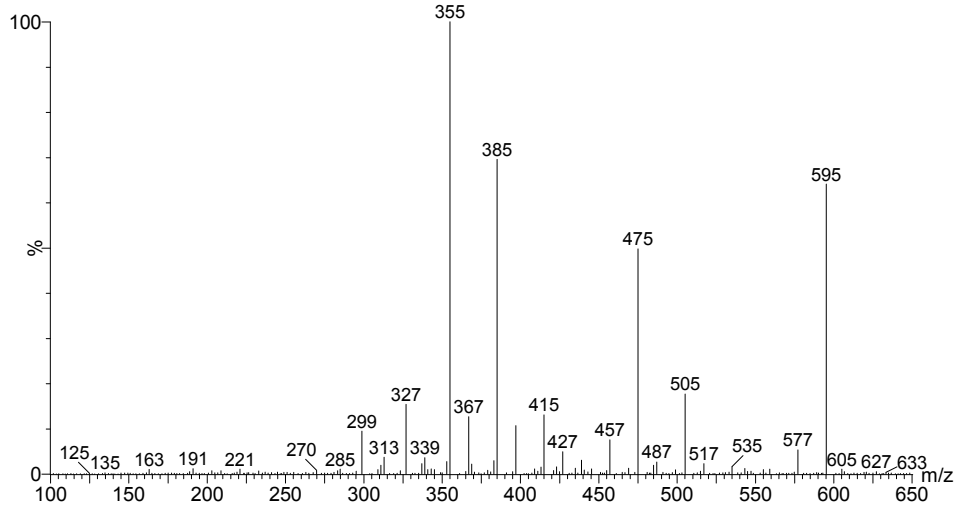
Metabolite **13**



Metabolite **14**

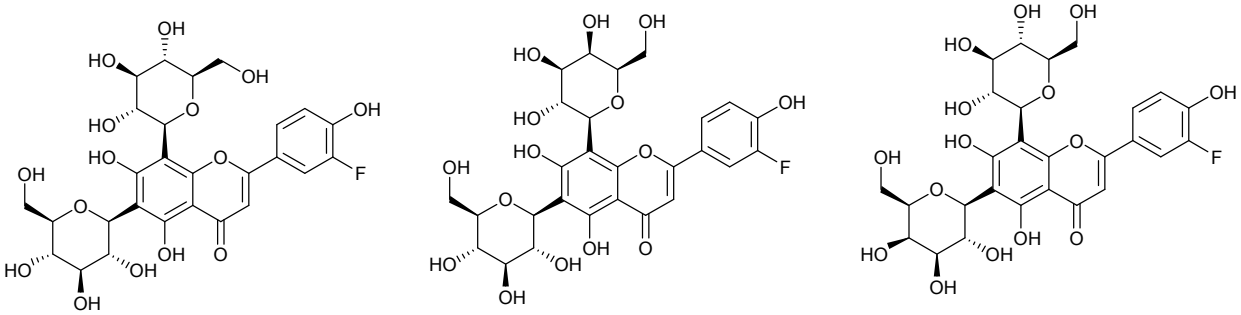


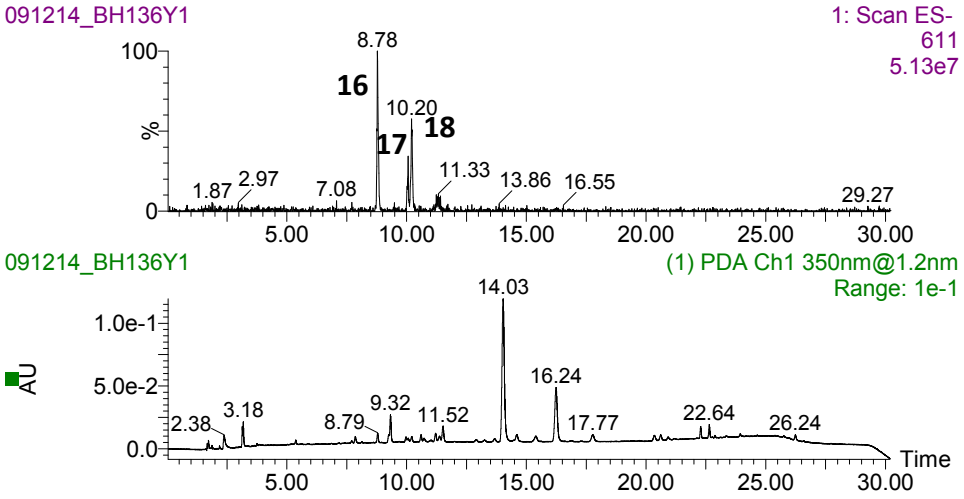
Metabolite **15**



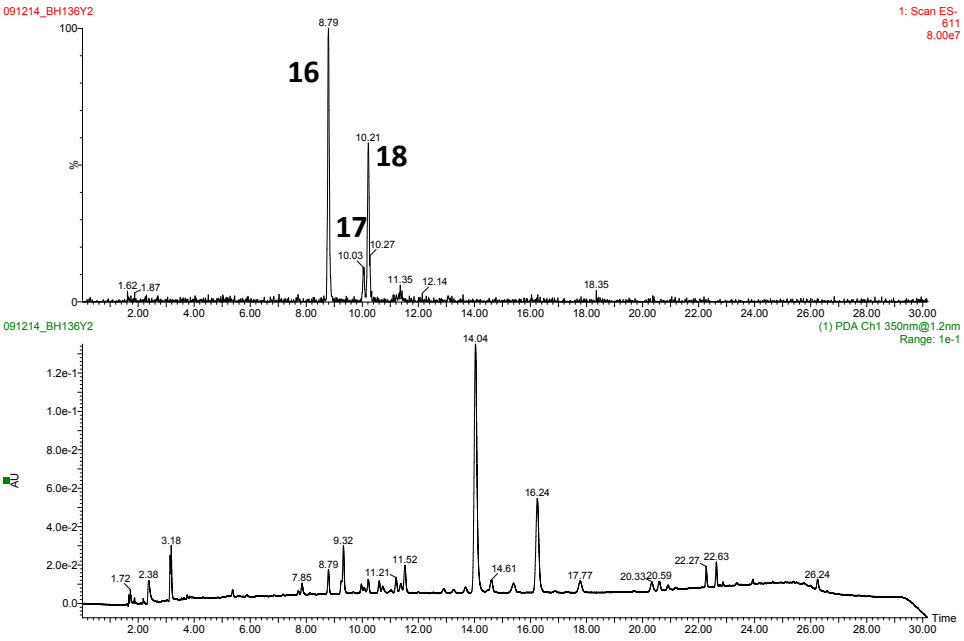
Supplementary figure 5B

LCMS PDA trace (lower) and ion trace (higher) showing 3 novel compounds (**16**, **17** and **18**) with molecular weight $[M-H]^-$ 611 derived from UDP-glucose, **12a** and *D. incanum* leaf protein.



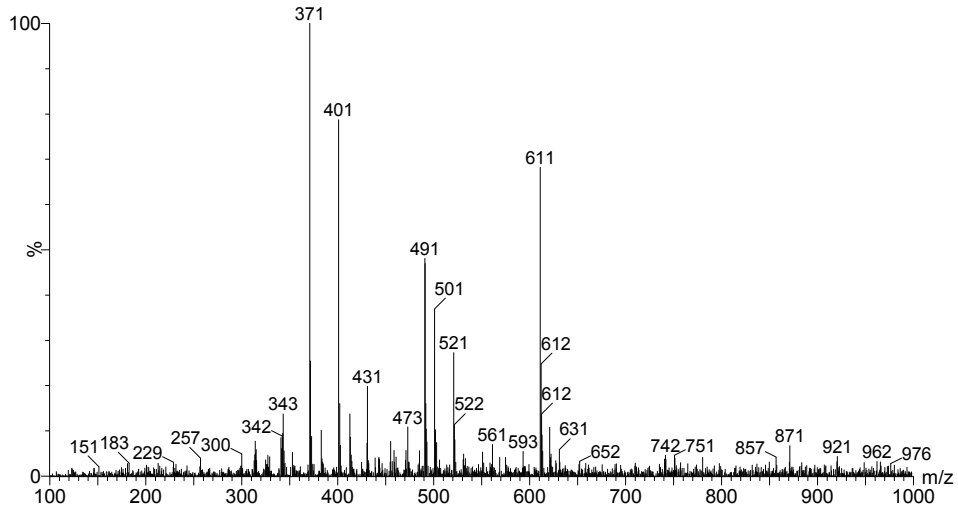


LCMS PDA trace (lower) and ion trace (higher) showing 3 novel compounds (**16**, **17** and **18**) with molecular weight $[M-H]^-$ 611 derived from UDP-galactose, **12a** and *D. incanum* leaf protein.

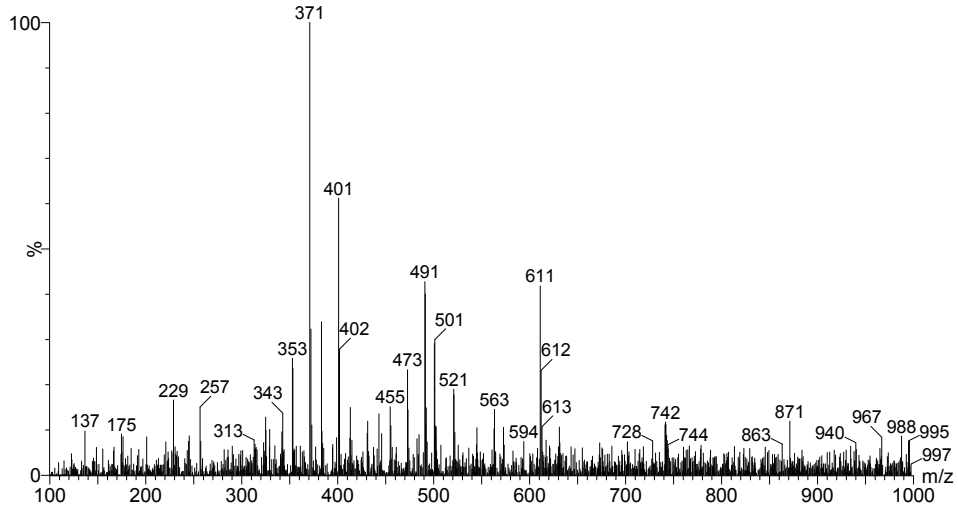


High cone voltage ESIMS mass spectrum of the three novel metabolites **16**, **17**, and **18**, from the ion trace at 611 amus, demonstrating they are di-*C*-linked hexose moieties through fragment ions at $[M-H-90]^-$ 521, $[M-H-120]^-$ 491, $[M-H-180]^-$ 431, $[M-H-210]^-$ 401 and $[M-H-240]^-$ 371.

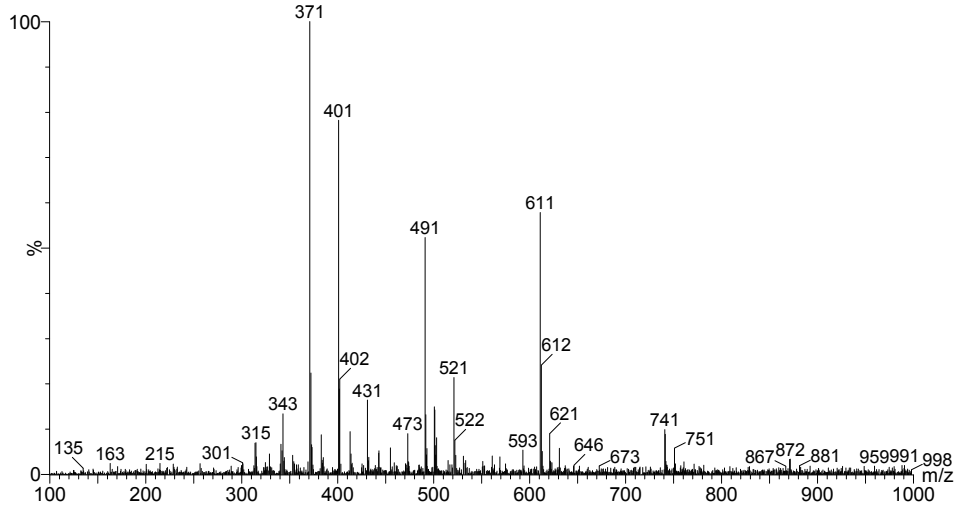
Metabolite **16**



Metabolite **17**

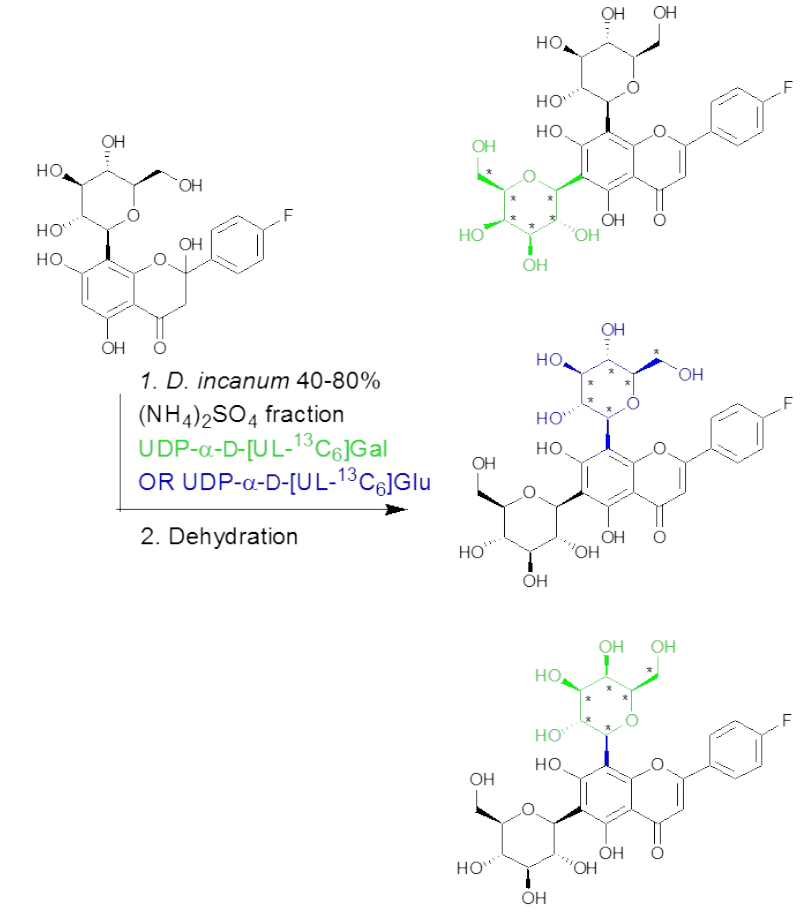


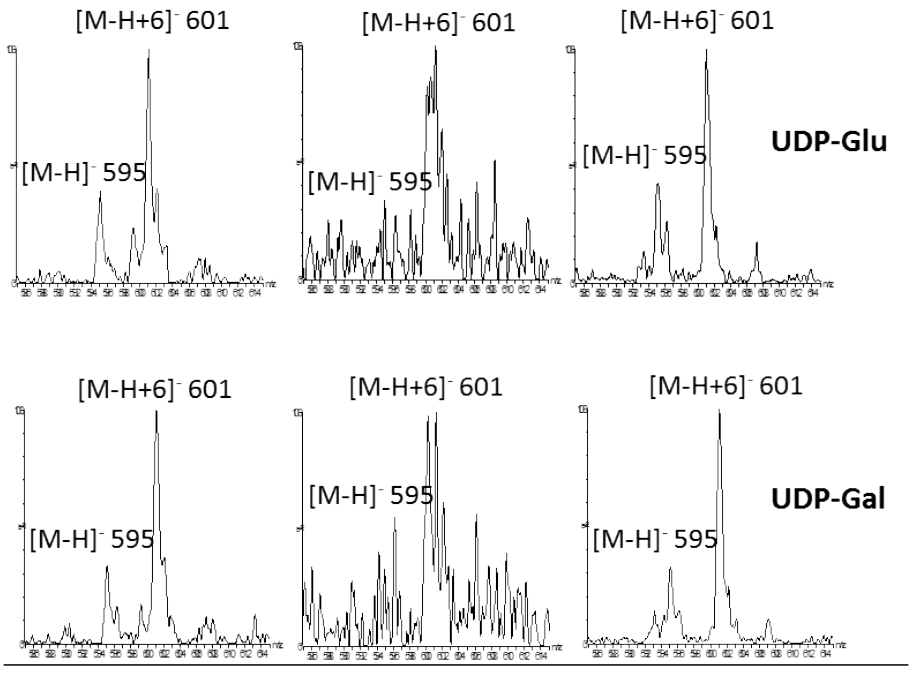
Metabolite **18**



Supplementary figure 6A

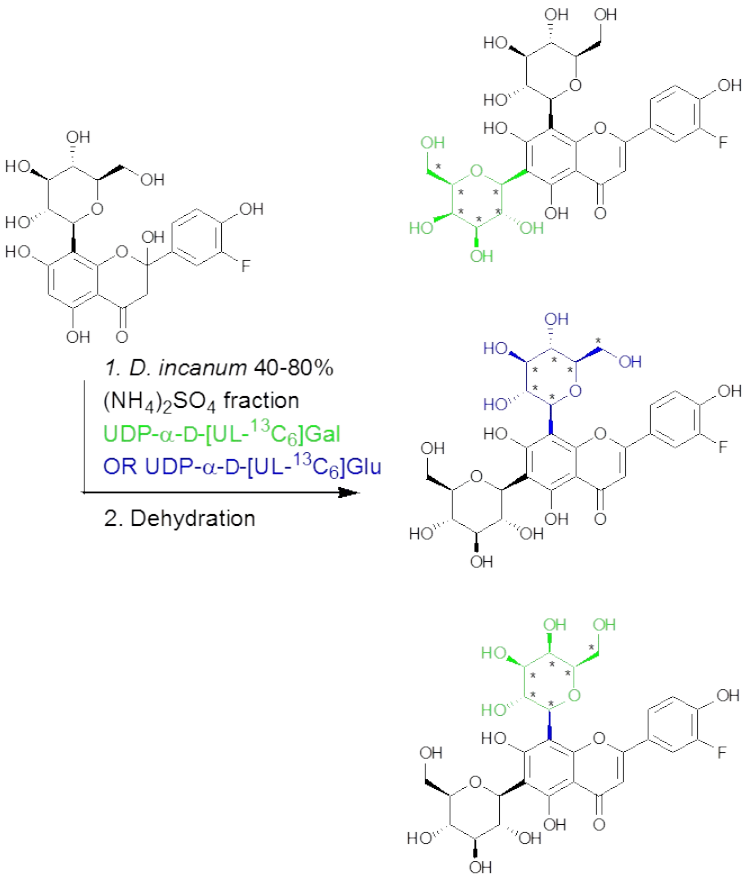
ESIMS spectra showing incorporation of UDP-α-D-[UL-$^{13}C_6$]glucose or UDP-α-D-[UL-$^{13}C_6$]galactose into metabolites **13**, **14** and **15** by the soluble leaf protein of *D. incanum*.

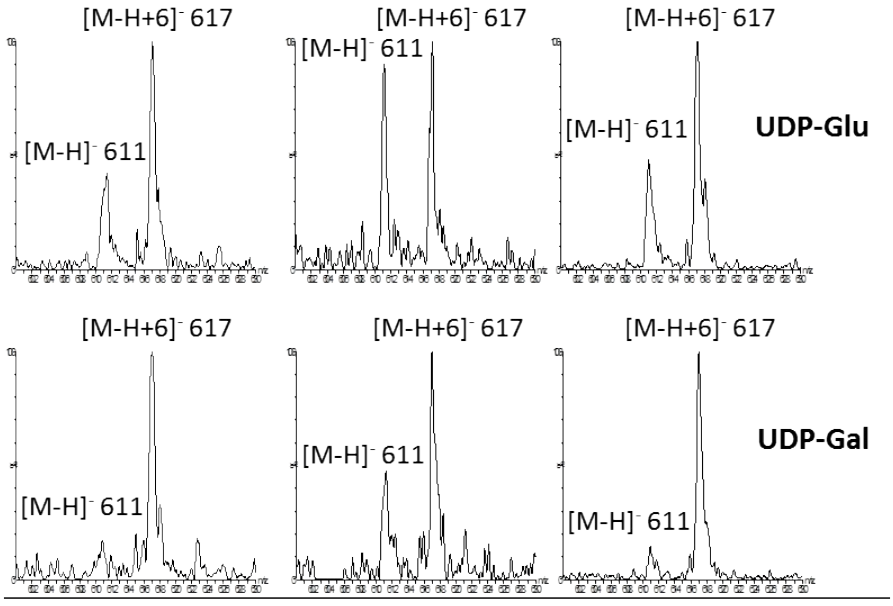




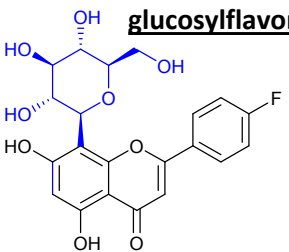
Supplementary figure 6B

ESIMS spectra showing incorporation of UDP-α-D-[UL-$^{13}C_6$]glucose or UDP-α-D-[UL-$^{13}C_6$]galactose into metabolites **16**, **17** and **18** by the soluble root protein of *D. incanum*.

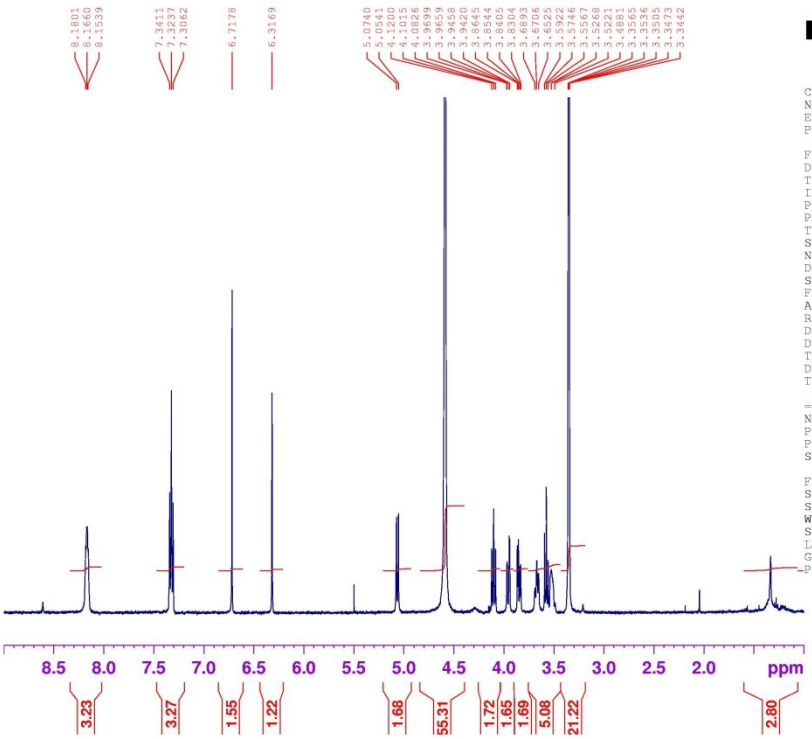




NMR Spectra of novel fluorinated *C*-glucosylflavones

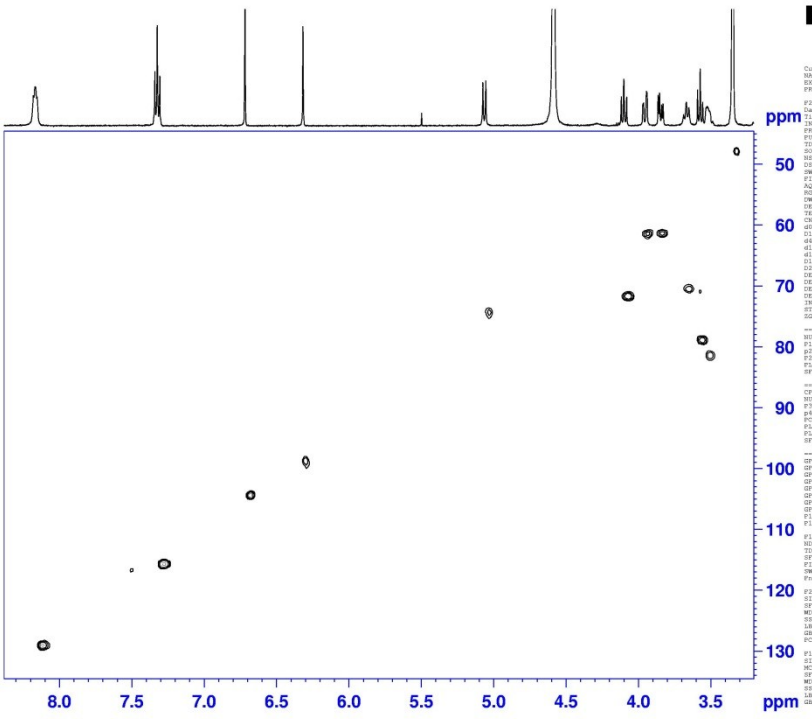


4’-fluoro-8-*C*-glucosylchrysin **11b**

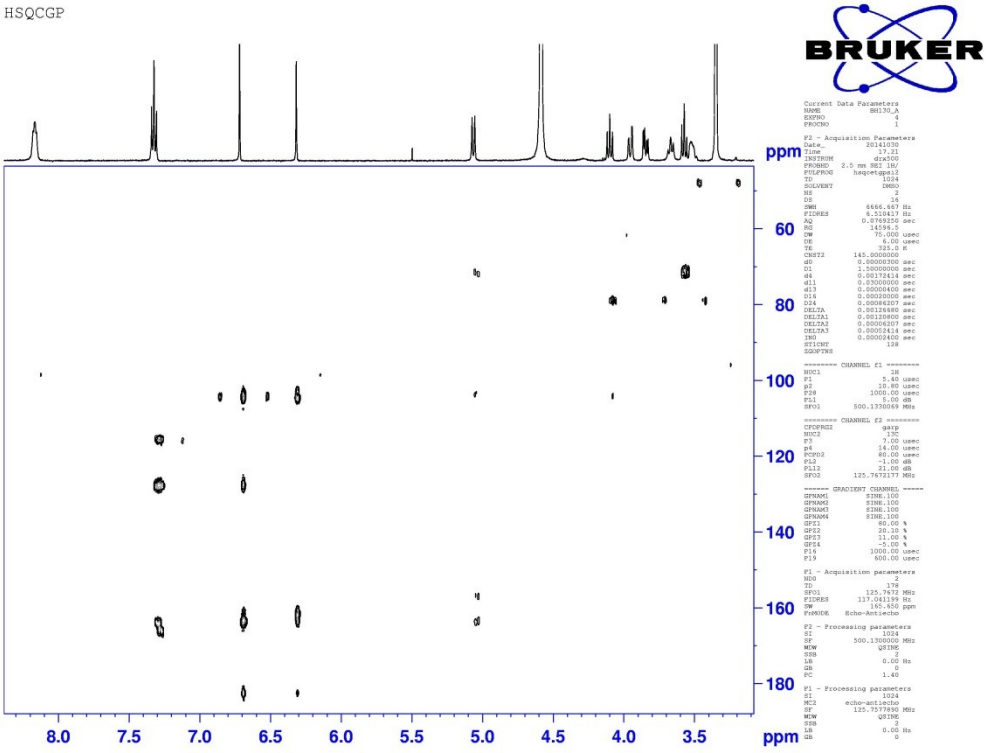


```
Current Data Parameters
NAME                  BH130_A
EXPNO                      10
PROCNO                      1

F2 - Acquisition Parameters
Date_               20141105
Time                   17.26
INSTRUM                drx500
PROBHD    5 mm BBO BB-1H
PULPROG                 zg30
TD                     65536
SOLVENT                 DMSO
NS                        16
DS                         2
SWH                10330.578 Hz
FIDRES             0.157632 Hz
AQ                 3.1720407 sec
RG                     574.7
DW                  48.400 usec
DE                   6.00 usec
TE                   325.0 K
D1                1.00000000 sec
TD0                        1

======== CHANNEL f1 ========
NUC1                      1H
P1                    9.88 usec
PL1                  -2.00 dB
SFO1           500.1330885 MHz

F2 - Processing parameters
SI                     32768
SF                500.1295850 MHz
WDW                       EM
SSB                        0
LB                    0.30 Hz
GB                         0
PC                      1.00
```

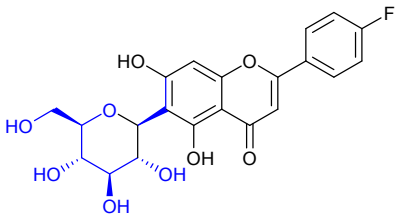
HSQCGP

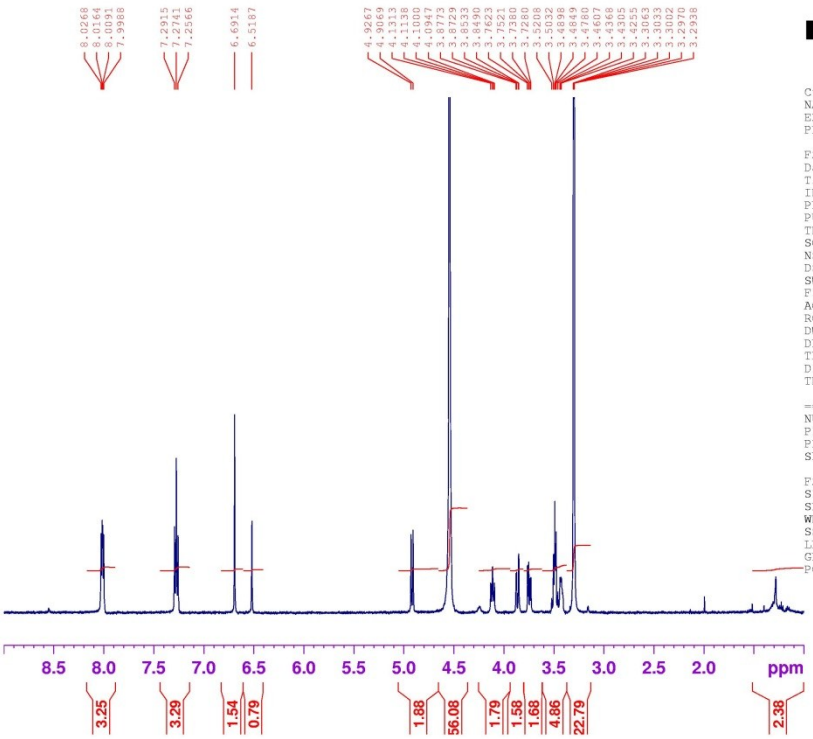


```
Current Data Parameters
NAME                  BH130_A
EXPNO                       4
PROCNO                      1

F2 - Acquisition Parameters
Date_               20141030
Time                   17.21
INSTRUM                drx500
PROBHD   2.5 mm BBI 1H/
PULPROG          hsqcetgp.2
TD                       1024
SOLVENT                 DMSO
NS                         2
DS                        16
SWH                5666.667 Hz
FIDRES             5.510417 Hz
AQ                 0.0769250 sec
RG                    14596.5
DW                   75.000 usec
DE                    6.00 usec
TE                   325.0 K
CNST2              145.0000000
d0                 0.00000300 sec
D1                 1.50000000 sec
d4                 0.00172414 sec
d11                0.05000000 sec
d13                0.00000400 sec
D16                0.00020000 sec
D24                0.00086207 sec
DELTA              0.00129885 sec
DELTA1             0.00129800 sec
DELTA2             0.00086207 sec
DELTA7             0.00052414 sec
INO                0.00000400 sec
ST1CNT                  128

======== CHANNEL f1 ========
NUC1                      1H
P1                    5.40 usec
p2                   10.80 usec
P28                1000.00 usec
PL1                   5.00 dB
SFO1           500.1330069 MHz

======== CHANNEL f2 ========
CPDPRG2                 garp
NUC2                     13C
P3                    7.00 usec
p4                   14.00 usec
PCPD2                80.00 usec
PL2                  -1.00 dB
PL12                 21.00 dB
SFO2           125.7671277 MHz

====== GRADIENT CHANNEL =====
GPNAM1               SINE.100
GPNAM6               SINE.100
GPNAM7               SINE.100
GPNAM8               SINE.100
GPZ1                 80.00 %
GPZ2                 20.10 %
GPZ7                 11.00 %
GPZ14                -5.00 %
P16                1000.00 usec
P19                 600.00 usec

F1 - Acquisition parameters
NDO                        2
TD                       178
SFO1           125.7672 MHz
FIDRES         117.041199 Hz
SW             165.650 ppm
FnMODE     Echo-Antiecho

F2 - Processing parameters
SI                       1024
SF             500.1300000 MHz
WDW                    QSINE
SSB                        2
LB                    0.00 Hz
GB                         0
PC                      1.40

F1 - Processing parameters
SI                       1024
MC2          echo-antiecho
SF             125.7577830 MHz
WDW                    QSINE
SSB                        2
LB                    0.00 Hz
GB                         0
```



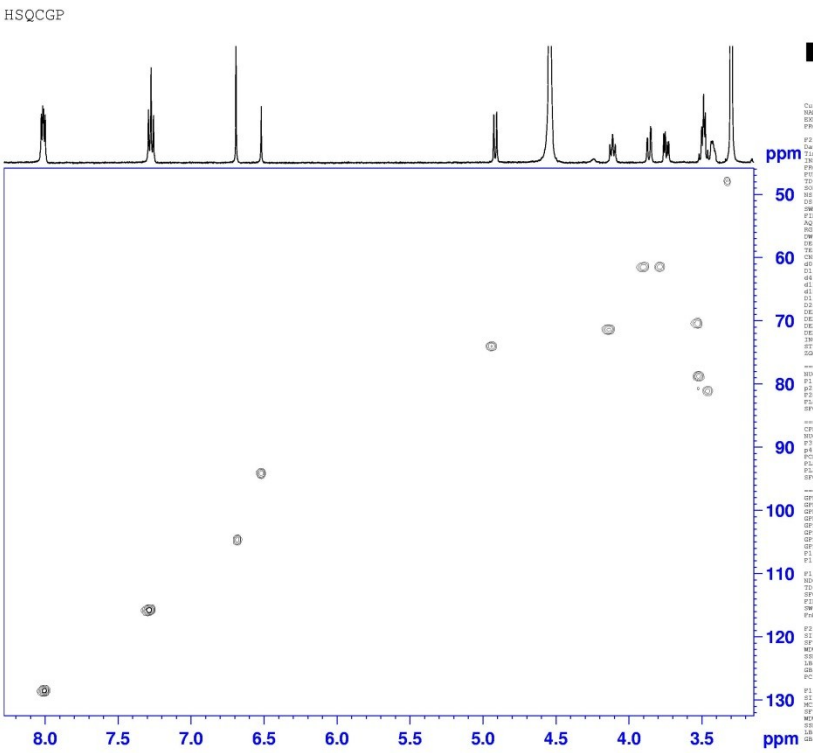
4’-fluoro-6-*C*-glucosylchrysin **11c**



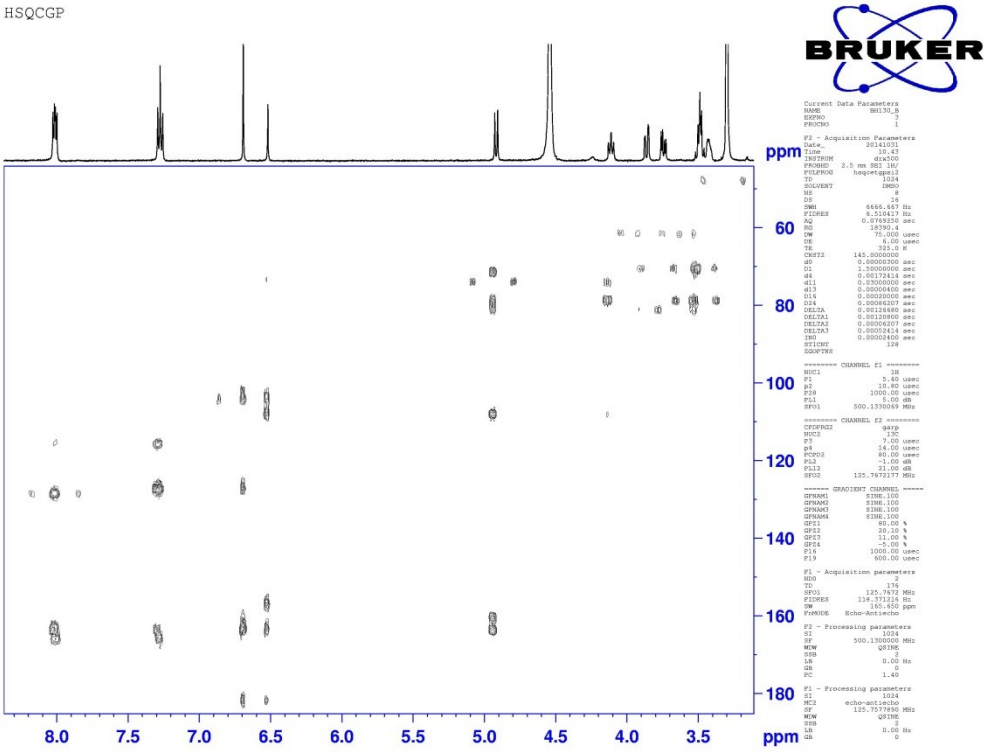



```
Current Data Parameters
NAME                  BH130_B
EXPNO                      10
PROCNO                      1

F2 - Acquisition Parameters
Date_               20141106
Time                   16.17
INSTRUM                drx500
PROBHD    5 mm BBO BB-1H
PULPROG                 zg30
TD                     65536
SOLVENT                 DMSO
NS                        16
DS                         2
SWH                10330.578 Hz
FIDRES             0.157632 Hz
AQ                 3.1720407 sec
RG                     724.1
DW                  48.400 usec
DE                   6.00 usec
TE                   325.0 K
D1                1.00000000 sec
TD0                        1

========== CHANNEL f1 =========
NUC1                       1H
P1                      9.88 usec
PL1                    -2.00 dB
SFO1           500.1330885 MHz

F2 - Processing parameters
SI                         32768
SF                500.1296102 MHz
WDW                           EM
SSB                            0
LB                          0.30 Hz
GB                             0
PC                          1.00
```

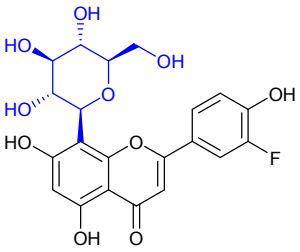
HSQCGP



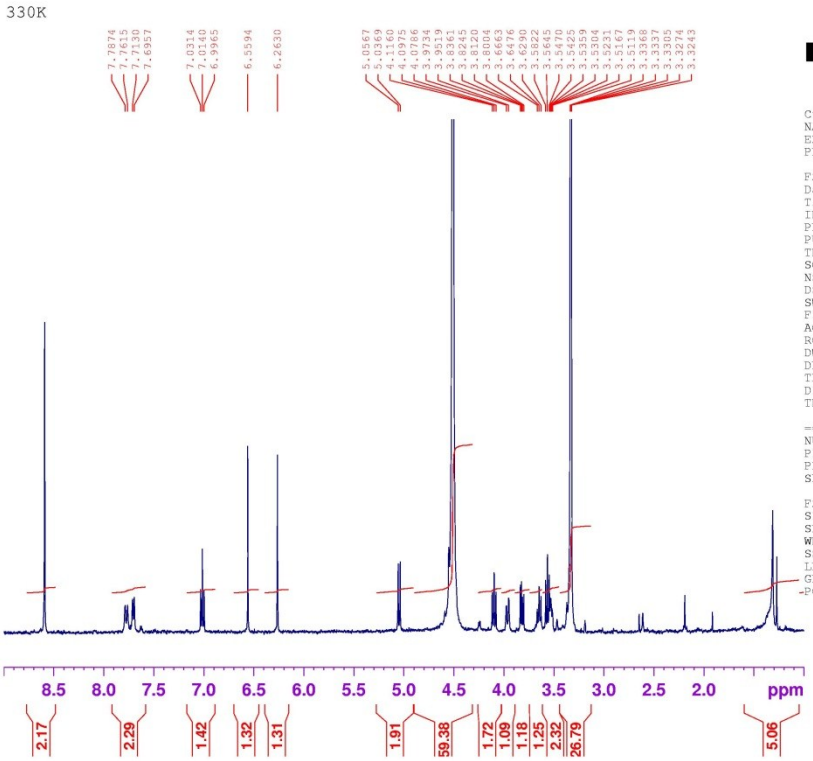

```
Current Data Parameters
NAME                  BH130_B
EXPNO                       3
PROCNO                      1

F2 - Acquisition Parameters
Date_               20141031
Time                   10.43
INSTRUM                drx500
PROBHD   2.5 mm BBI 1H/
PULPROG          hsqcetgpsi2
TD                       1024
SOLVENT                 DMSO
NS                         8
DS                        16
SWH                6666.667 Hz
FIDRES             6.510417 Hz
AQ                 0.0769250 sec
RG                    18390.4
DW                   75.000 usec
DE                    6.00 usec
TE                   325.0 K
CNST2              145.0000000
d0                 0.00000300 sec
D1                 1.50000000 sec
d4                 0.00172414 sec
d11                0.03000000 sec
d13                0.00000400 sec
D16                0.00020000 sec
D24                0.00068207 sec
DELTA              0.00126680 sec
DELTA1             0.00120800 sec
DELTA2             0.00006207 sec
DELTA3             0.00052414 sec
ING                0.00002400 sec
ST1CNT                  128
SSGPTW8

========== CHANNEL f1 =========
NUC1                       1H
P1                      9.40 usec
P2                     19.80 usec
P34                 1000.00 usec
PL1                    -0.65 dB
SFO1           500.1330069 MHz

========== CHANNEL f2 =========
CPDPRG2                 garp
NUC2                     13C
P3                      7.00 usec
p4                     14.00 usec
PCPD2                  80.00 usec
PL2                    -1.00 dB
PL12                   21.00 dB
SFO2           125.7672177 MHz

====== GRADIENT CHANNEL =====
GPNAM1              SINE.100
GPNAM2              SINE.100
GPNAM3              SINE.100
GPNAM4              SINE.100
GPZ1                   80.00 %
GPZ2                   20.10 %
GPZ3                   11.00 %
GPZ4                   -5.00 %
P16                 1000.00 usec
P19                  600.00 usec

F1 - Acquisition parameters
NDO                        2
TD                       176
SFO1           125.7672 MHz
FIDRES         138.371216 Hz
SW             165.650 ppm
FnMODE     Echo-Antiecho

F2 - Processing parameters
SI                       1024
SF            500.1300000 MHz
WDW                    QSINE
SSB                        2
LB                      0.00 Hz
GB                         0
PC                      1.40

F1 - Processing parameters
SI                       1024
MC2        echo-antiecho
SF            125.7577890 MHz
WDW                    QSINE
SSB                        2
LB                      0.00 Hz
GB                         0
```



3’-fluorovitexin **12b**

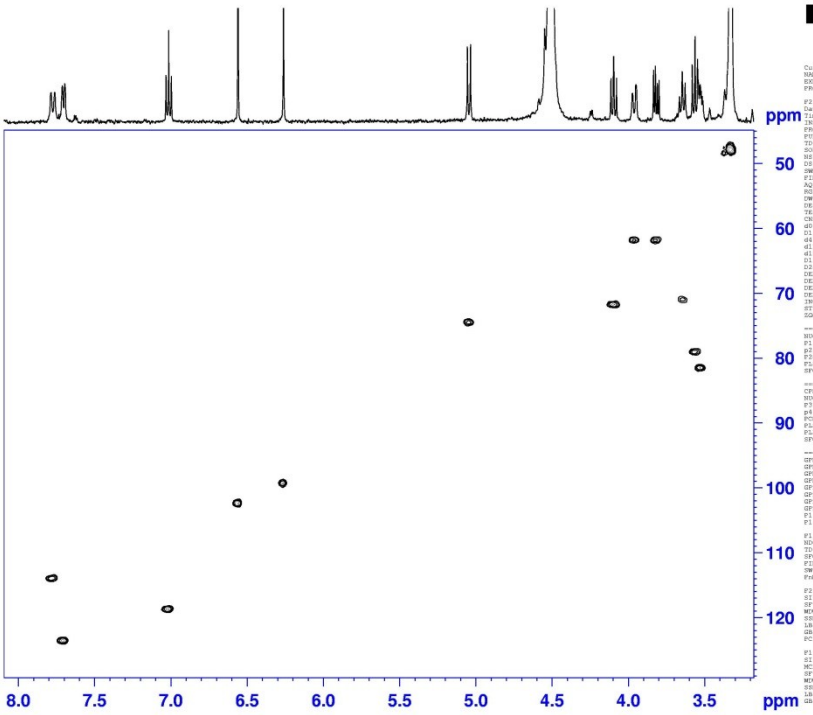


330K

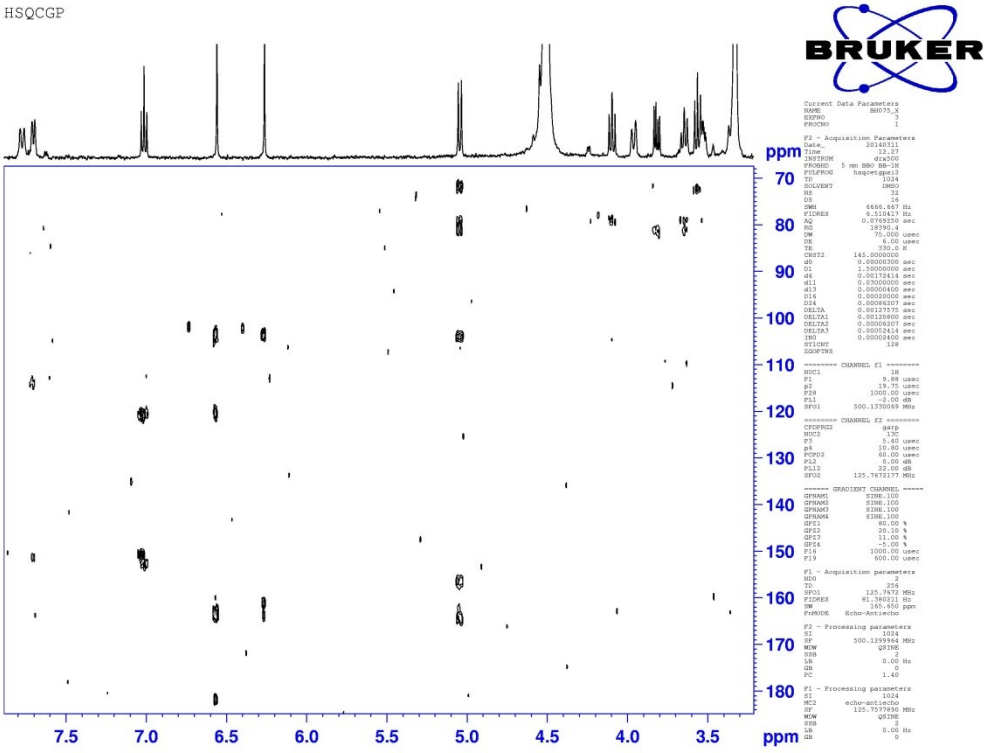



```
Current Data Parameters
NAME                  BH075_X
EXPNO                       1
PROCNO                      1

F2 - Acquisition Parameters
Date_              20140311
Time                  12.13
INSTRUM               drx500
PROBHD    5 mm BBO BB-1H
PULPROG                zg30
TD                    65536
SOLVENT                MeOD
NS                       64
DS                        2
SWH               10330.578 Hz
FIDRES             0.157632 Hz
AQ                3.1720407 sec
RG                     2048
DW                  48.400 usec
DE                   6.00 usec
TE                  330.0 K
D1                1.00000000 sec
TD0                       1

========= CHANNEL f1 =========
NUC1                      1H
P1                      5.40 usec
PL1                     5.00 dB
SFO1           500.1330885 MHz

F2 - Processing parameters
SI                        32768
SF             500.1300000 MHz
WDW                         EM
SSB                          0
LB                      1.00 Hz
GB                           0
PC                      1.00
```

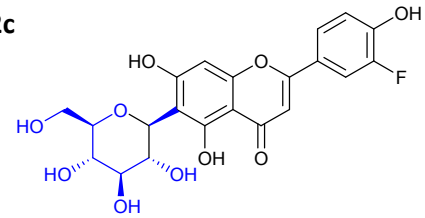
HSQCGP

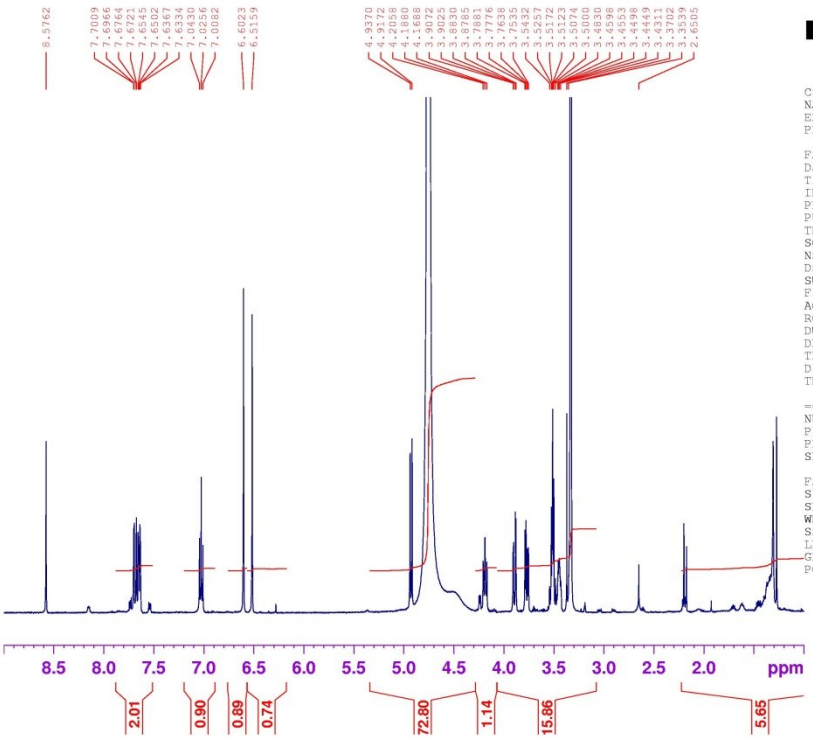



```
Current Data Parameters
NAME                  BH075_X
EXPNO                       3
PROCNO                      1

F2 - Acquisition Parameters
Date_              20140311
Time                  12.27
INSTRUM               drx500
PROBHD    5 mm BBO BB-1H
PULPROG          hsqcetgpsi2
TD                      1024
SOLVENT                 DMSO
NS                        32
DS                        16
SWH               6666.667 Hz
FIDRES            6.510417 Hz
AQ                0.0769250 sec
RG                   18390.4
DW                  75.000 usec
DE                   6.00 usec
TE                  330.0 K
CNST2             145.0000000
d0                0.00000300 sec
D1                1.50000000 sec
d4                0.00172414 sec
d11               0.03000000 sec
d13               0.00000400 sec
D16               0.00020000 sec
D24               0.00068207 sec
DELTA             0.00127575 sec
DELTA1            0.00128805 sec
DELTA2            0.00006207 sec
DELTA3            0.00052414 sec
ING               0.00002400 sec
STSIGMA                    128
SSGPTW8

========= CHANNEL f1 =========
NUC1                      1H
P1                      9.88 usec
p2                     19.76 usec
P34                  1000.00 usec
PL1                     0.00 dB
SFO1           500.1330869 MHz

========= CHANNEL f2 =========
CPDPRG2                 garp
NUC2                     13C
P3                      5.40 usec
p4                     10.80 usec
PCPD2                  60.00 usec
PL2                     0.00 dB
PL12                   22.00 dB
SFO2           125.7672177 MHz

====== GRADIENT CHANNEL =====
GPNAM1               SINE.100
GPNAM2               SINE.100
GPNAM3               SINE.100
GPNAM4               SINE.100
GPZ1                   80.00 %
GPZ2                   20.10 %
GPZ3                   11.00 %
GPZ4                   -5.00 %
P16                  1000.00 usec
P19                   600.00 usec

F1 - Acquisition parameters
ND0                         2
TD                        256
SFO1           125.7672 MHz
FIDRES           81.380211 Hz
SW               165.650 ppm
FnMODE     Echo-Antiecho

F2 - Processing parameters
SI                       1024
SF             500.1299964 MHz
WDW                    QSINE
SSB                        2
LB                      0.00 Hz
GB                         0
PC                      1.40

F1 - Processing parameters
SI                       1024
MC2          echo-antiecho
SF             125.7577890 MHz
WDW                    QSINE
SSB                        2
LB                      0.00 Hz
GB                         0
```



3’-fluoroisovitexin **12c**

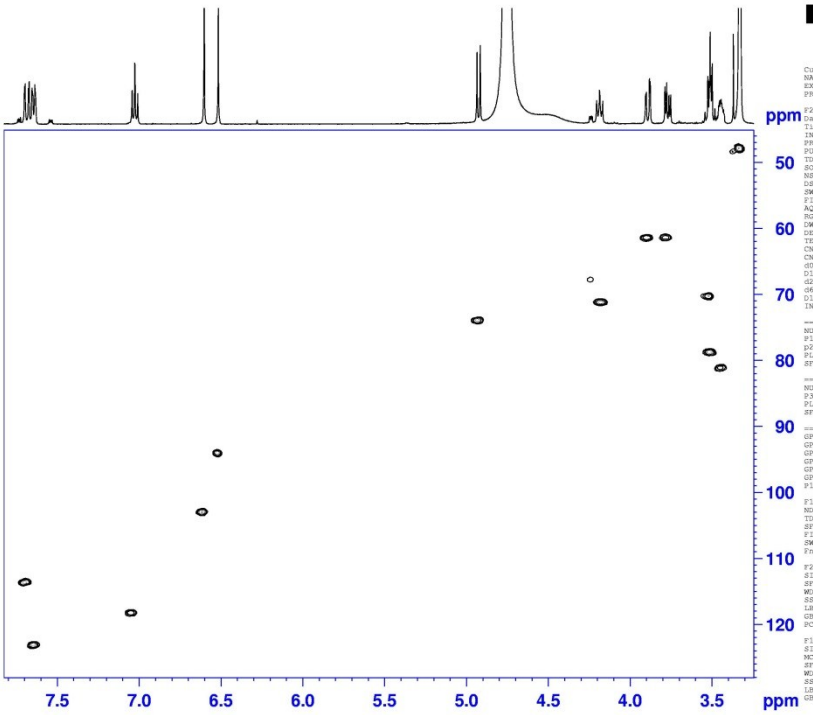




BRUKER

```
Current Data Parameters
NAME                  BH075_Y
EXPNO                       5
PROCNO                      1

F2 - Acquisition Parameters
Date_              20140211
Time                  16.50
INSTRUM               drx500
PROBHD    5 mm BBO BB-1H
PULPROG                zg30
TD                    65536
SOLVENT               CDCl3
NS                       64
DS                        2
SWH              10330.578 Hz
FIDRES               0.157632 Hz
AQ                3.1720407 sec
RG                      256
DW                   48.400 usec
DE                    6.00 usec
TE                   310.0 K
D1               1.00000000 sec
TD0                       1

======== CHANNEL f1 ========
NUC1                      1H
P1                     9.88 usec
PL1                   -2.00 dB
SFO1           500.1330885 MHz

F2 - Processing parameters
SI                    32768
SF             500.1300000 MHz
WDW                      EM
SSB                       0
LB                     0.30 Hz
GB                        0
PC                     1.00
```

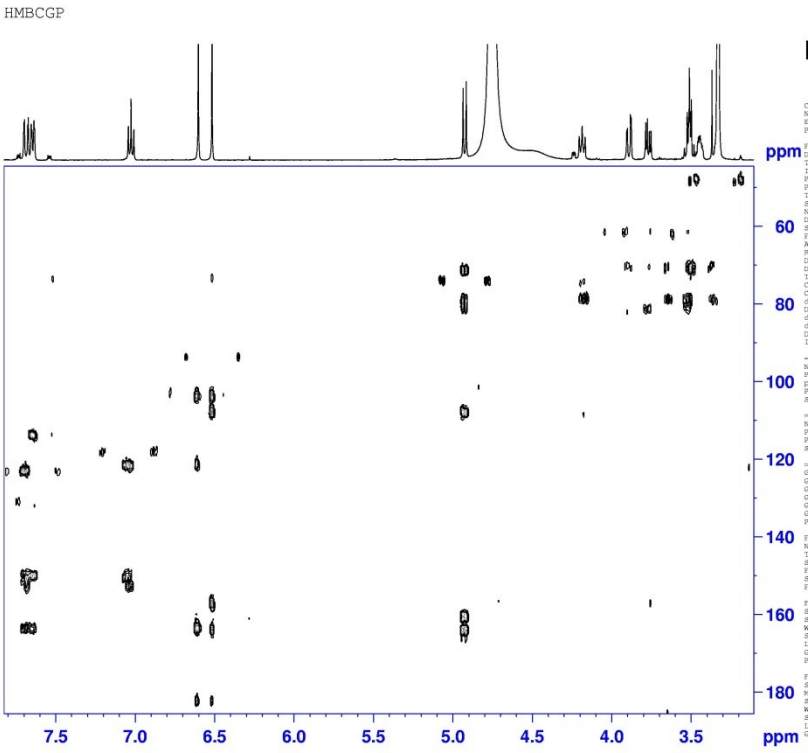
HMBCGP



BRUKER

```
Current Data Parameters
NAME                  BH075_Y
EXPNO                       2
PROCNO                      1

F2 - Acquisition Parameters
Date_              20140210
Time                  17.48
INSTRUM               drx500
PROBHD   2.5 mm BBI 1H/
PULPROG    hmbcgplpndqf
TD                    4096
SOLVENT               DMSO
NS                     256
DS                      16
SWH              6510.417 Hz
FIDRES               1.589457 Hz
AQ                0.3146996 sec
RG                   26008
DW                   76.800 usec
DE                    6.00 usec
TE                   310.0 K
CNST2           145.0000000
CNST13           10.0000000
d0                0.00000300 sec
D1                1.50000000 sec
d2                0.00344828 sec
d6                0.05000000 sec
D16               0.00020000 sec
IN0               0.00001796 sec

======== CHANNEL f1 ========
NUC1                      1H
P1                     5.40 usec
p2                    10.80 usec
PL1                    5.00 dB
SFO1           500.1331508 MHz

======== CHANNEL f2 ========
NUC2                     13C
P3                     7.00 usec
PL2                   -1.00 dB
SFO2           125.7703443 MHz

====== GRADIENT CHANNEL =====
GPNAM1              SINE.100
GPNAM2              SINE.100
GPNAM3              SINE.100
GPZ1                  50.00 %
GPZ2                  30.00 %
GPZ3                  40.15 %
P16                 1000.00 usec

F1 - Acquisition parameters
ND0                       2
TD                      128
SFO1           125.7703 MHz
FIDRES         232.774673 Hz
SW                 222.095 ppm
FnMODE                   QF

F2 - Processing parameters
SI                     2048
SF             500.1300000 MHz
WDW                    SINE
SSB                       0
LB                     0.00 Hz
GB                        0
PC                     1.00

F1 - Processing parameters
SI                     1024
MC2                      QF
SF             125.7577890 MHz
WDW                    SINE
SSB                       0
LB                     0.00 Hz
GB                        0
```

HMBCGP



BRUKER

```
Current Data Parameters
NAME                  BH075_Y
EXPNO                       2
PROCNO                      1

F2 - Acquisition Parameters
Date_              20140210
Time                    17.48
INSTRUM                 drx500
PROBHD    2.5 mm BBI 1H/
PULPROG      hmbcgplpndqf
TD                        4096
SOLVENT                   DMSO
NS                         256
DS                          16
SWH                   6510.417 Hz
FIDRES                1.589457 Hz
AQ                    0.3146996 sec
RG                        26008
DW                       76.800 usec
DE                        6.00 usec
TE                       310.0 K
CNST2               145.0000000
CNST13               10.0000000
d0                    0.00000300 sec
D1                    1.50000000 sec
d2                    0.00344828 sec
d6                    0.05000000 sec
D16                   0.00020000 sec
IN0                   0.00001790 sec

======== CHANNEL f1 ========
NUC1                        1H
P1                         5.40 usec
p2                        10.80 usec
PL1                        5.00 dB
SFO1               500.1331508 MHz

======== CHANNEL f2 ========
NUC2                       13C
P3                         7.00 usec
PL2                       -1.00 dB
SFO2               125.7703443 MHz

====== GRADIENT CHANNEL =====
GPNAM1               SINE.100
GPNAM2               SINE.100
GPNAM3               SINE.100
GPZ1                      50.00 %
GPZ2                      30.00 %
GPZ3                      40.10 %
P16                     1000.00 usec

F1 - Acquisition parameters
ND0                         2
TD                        128
SFO1               125.7703 MHz
FIDRES             232.774673 Hz
SW                    222.095 ppm
FnMODE                     QF

F2 - Processing parameters
SI                        2048
SF                 500.1300000 MHz
WDW                       SINE
SSB                          0
LB                        0.00 Hz
GB                           0
PC                        1.00

F1 - Processing parameters
SI                        1024
MC2                        QF
SF                 125.7577890 MHz
WDW                       SINE
SSB                          0
LB                        0.00 Hz
GB                           0
```