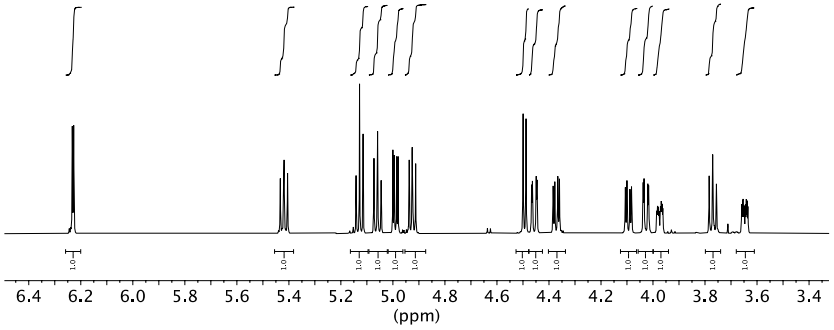
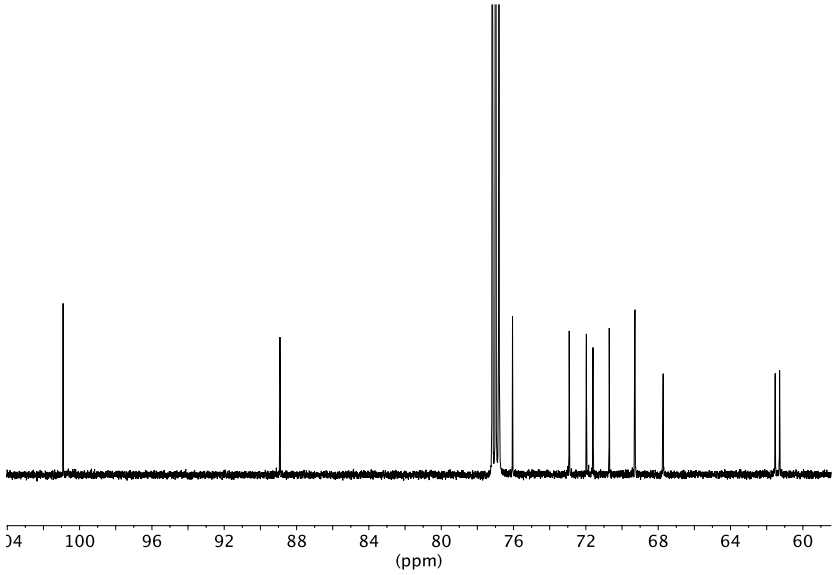
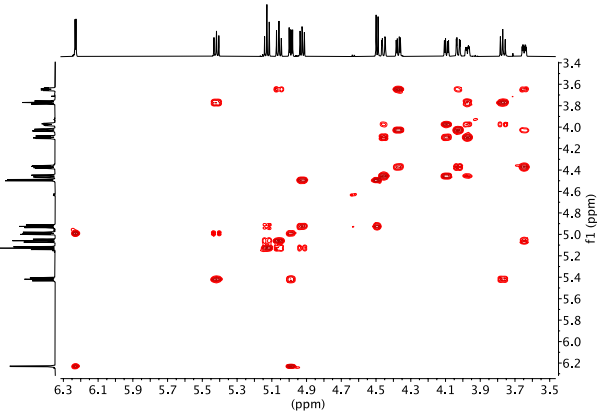
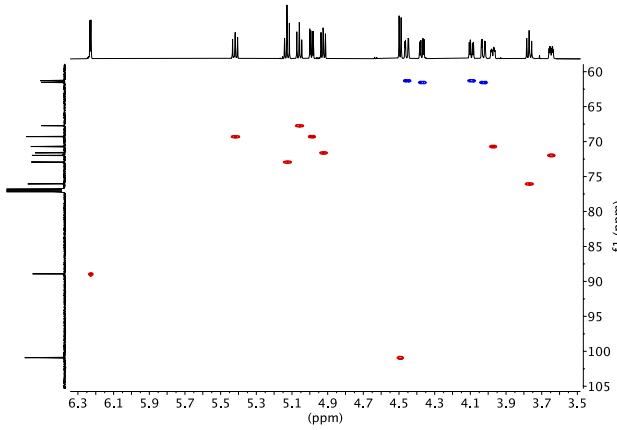
NMR Data Analysis By and For Synthetic Chemists: Is there a better way?

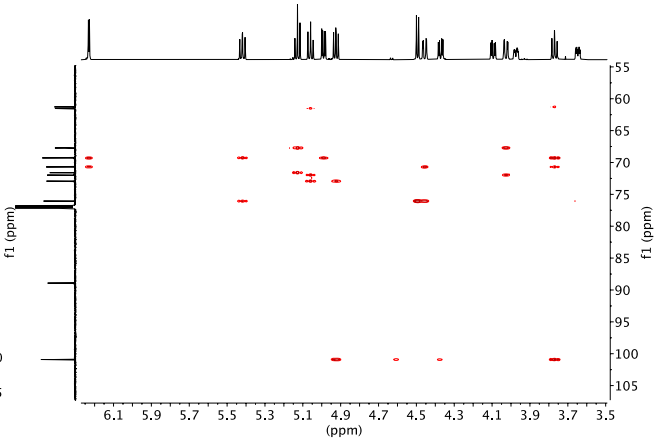
Eric Hughes and Alan Kenwright


6.4 6.2 6.0 5.8 5.6 5.4 5.2 5.0 4.8 4.6 4.4 4.2 4.0 3.8 3.6 3.4
(ppm)


100 96 92 88 84 80 76 72 68 64 60
(ppm)

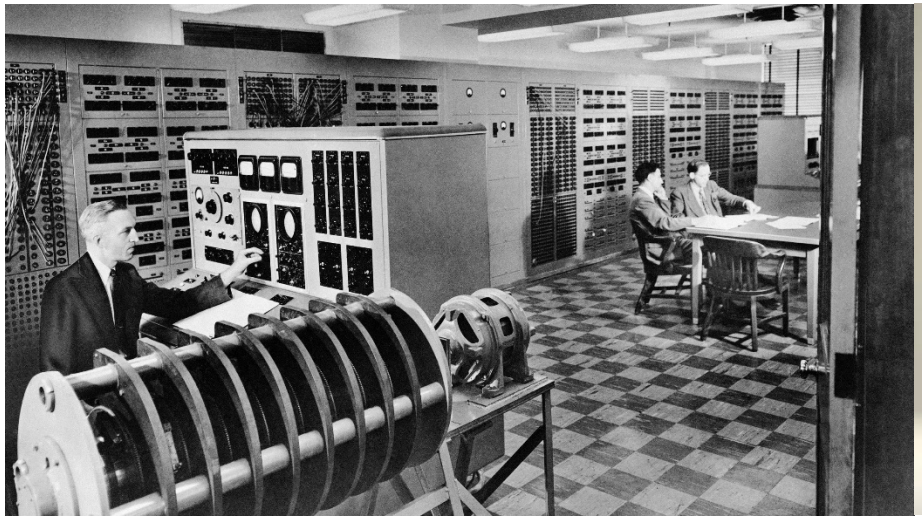

6.3 6.1 5.9 5.7 5.5 5.3 5.1 4.9 4.7 4.5 4.3 4.1 3.9 3.7 3.5
(ppm)
f1 (ppm)


6.3 6.1 5.9 5.7 5.5 5.3 5.1 4.9 4.7 4.5 4.3 4.1 3.9 3.7 3.5
(ppm)
f1 (ppm)


6.1 5.9 5.7 5.5 5.3 5.1 4.9 4.7 4.5 4.3 4.1 3.9 3.7 3.5
(ppm)
f1 (ppm)

A Mismatch?

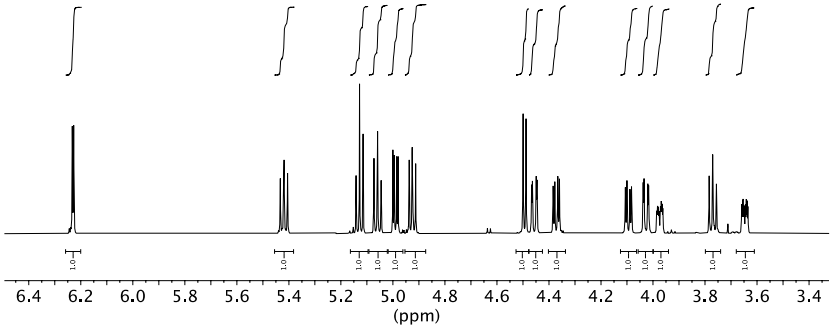
People in NMR Labs

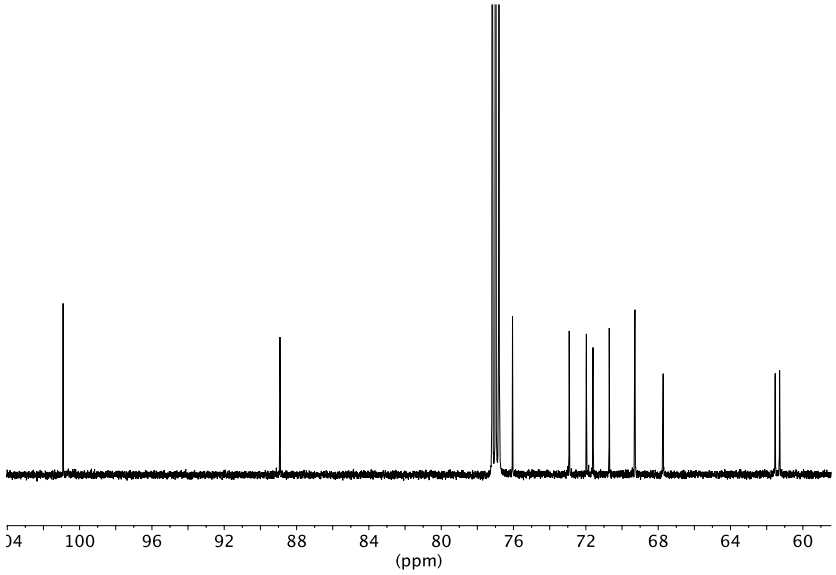


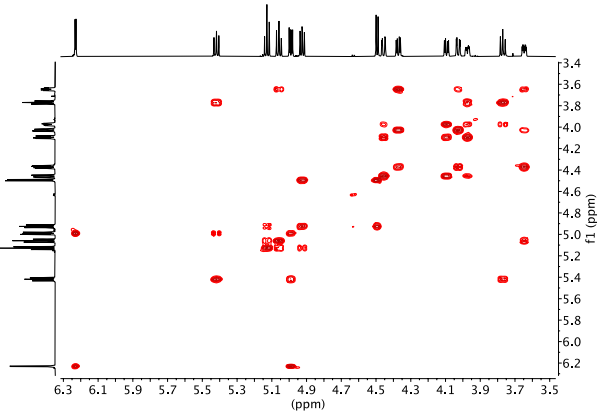
Know a lot about NMR theory so should be well placed to interpret spectra (in theory). Think in terms of spectral features.

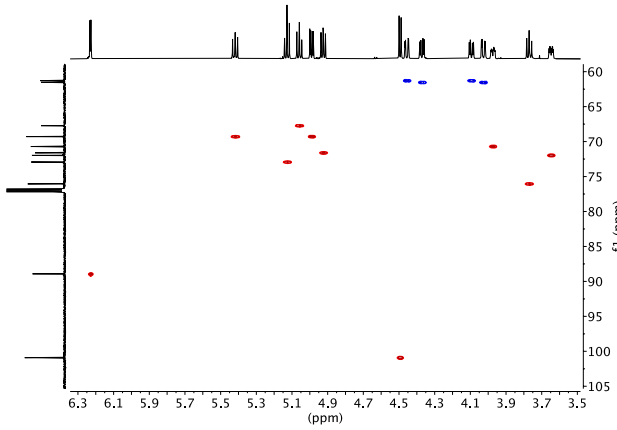
Synthetic Chemists

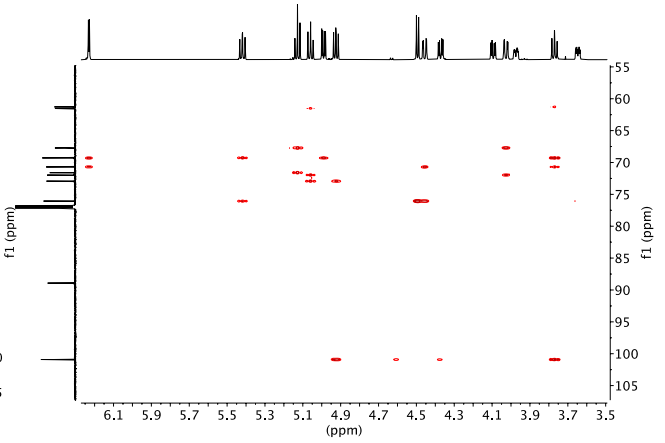
Do a lot of spectral assignment
Don't necessarily know a lot about NMR theory. Think in terms of structural fragments.


6.4 6.2 6.0 5.8 5.6 5.4 5.2 5.0 4.8 4.6 4.4 4.2 4.0 3.8 3.6 3.4
(ppm)


100 96 92 88 84 80 76 72 68 64 60
(ppm)


6.3 6.1 5.9 5.7 5.5 5.3 5.1 4.9 4.7 4.5 4.3 4.1 3.9 3.7 3.5
(ppm)
f1 (ppm)

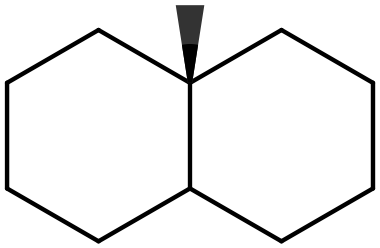

6.3 6.1 5.9 5.7 5.5 5.3 5.1 4.9 4.7 4.5 4.3 4.1 3.9 3.7 3.5
(ppm)
f1 (ppm)

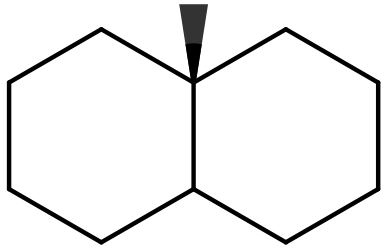

6.1 5.9 5.7 5.5 5.3 5.1 4.9 4.7 4.5 4.3 4.1 3.9 3.7 3.5
(ppm)
f1 (ppm)

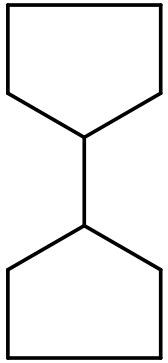
Constitution Verification

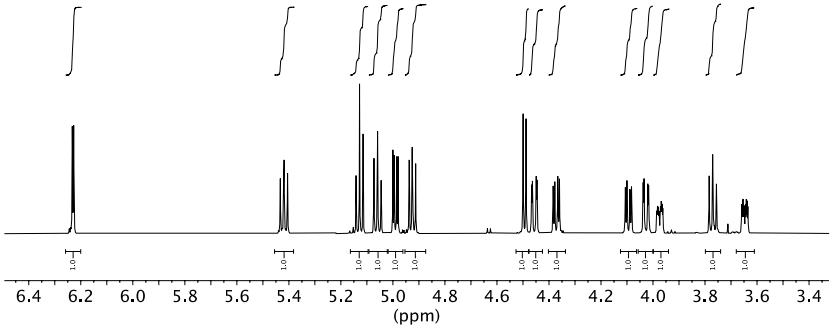
"Constitution" – pure connectivity of atoms within a molecule – no information on spatial arrangement e.g. cis/trans, E/Z, chirality, etc. The "flat" molecule.

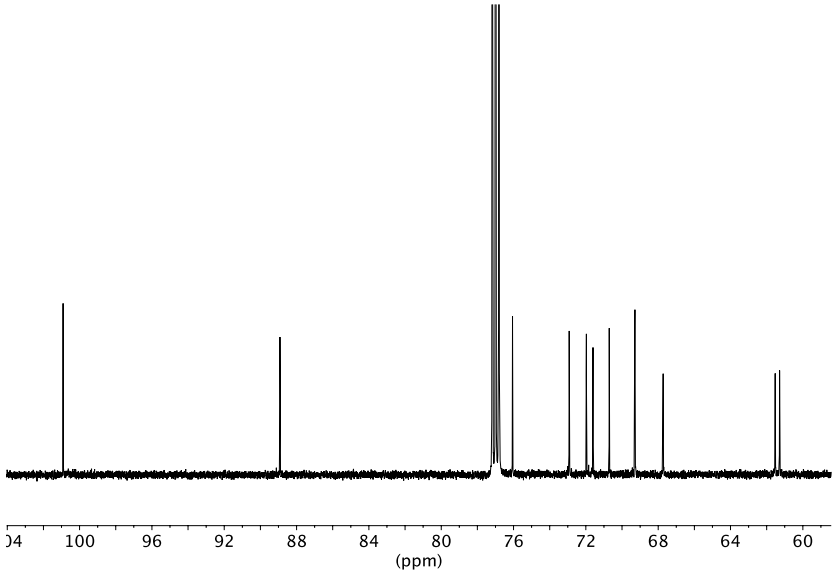
What are we verifying? That the spectral dataset is consistent with the postulated structure and that an assignment of the spectra can be made on that basis

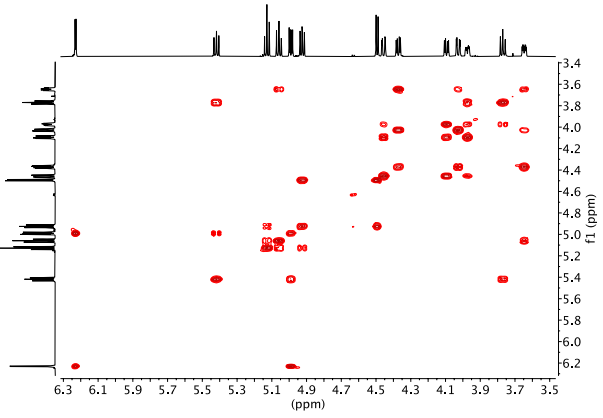


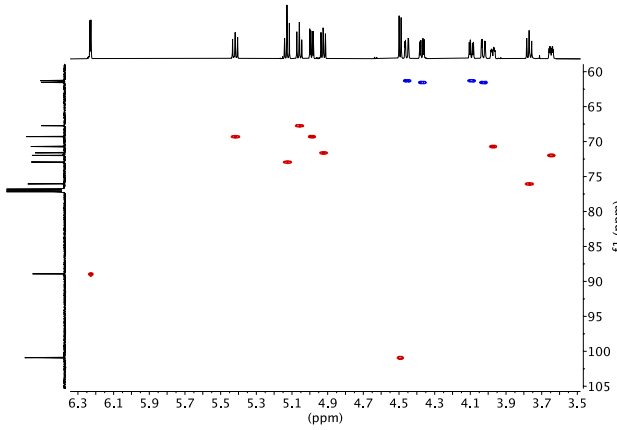


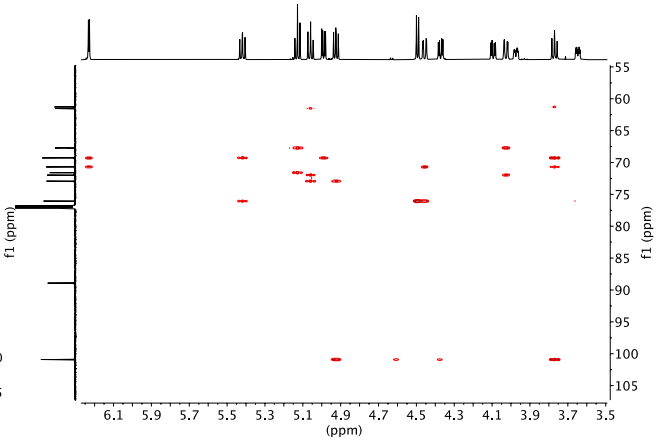


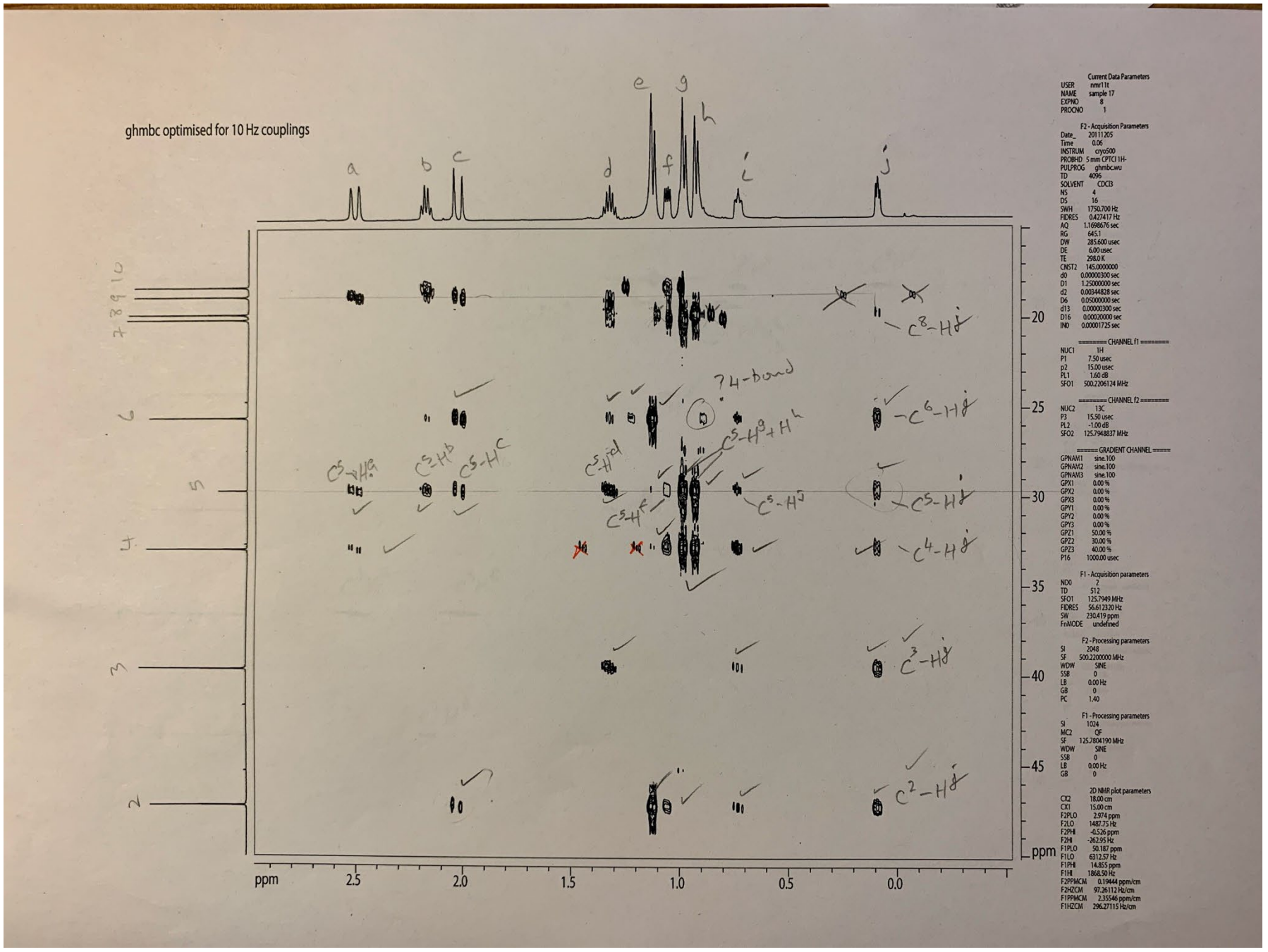

6.4 6.2 6.0 5.8 5.6 5.4 5.2 5.0 4.8 4.6 4.4 4.2 4.0 3.8 3.6 3.4
(ppm)

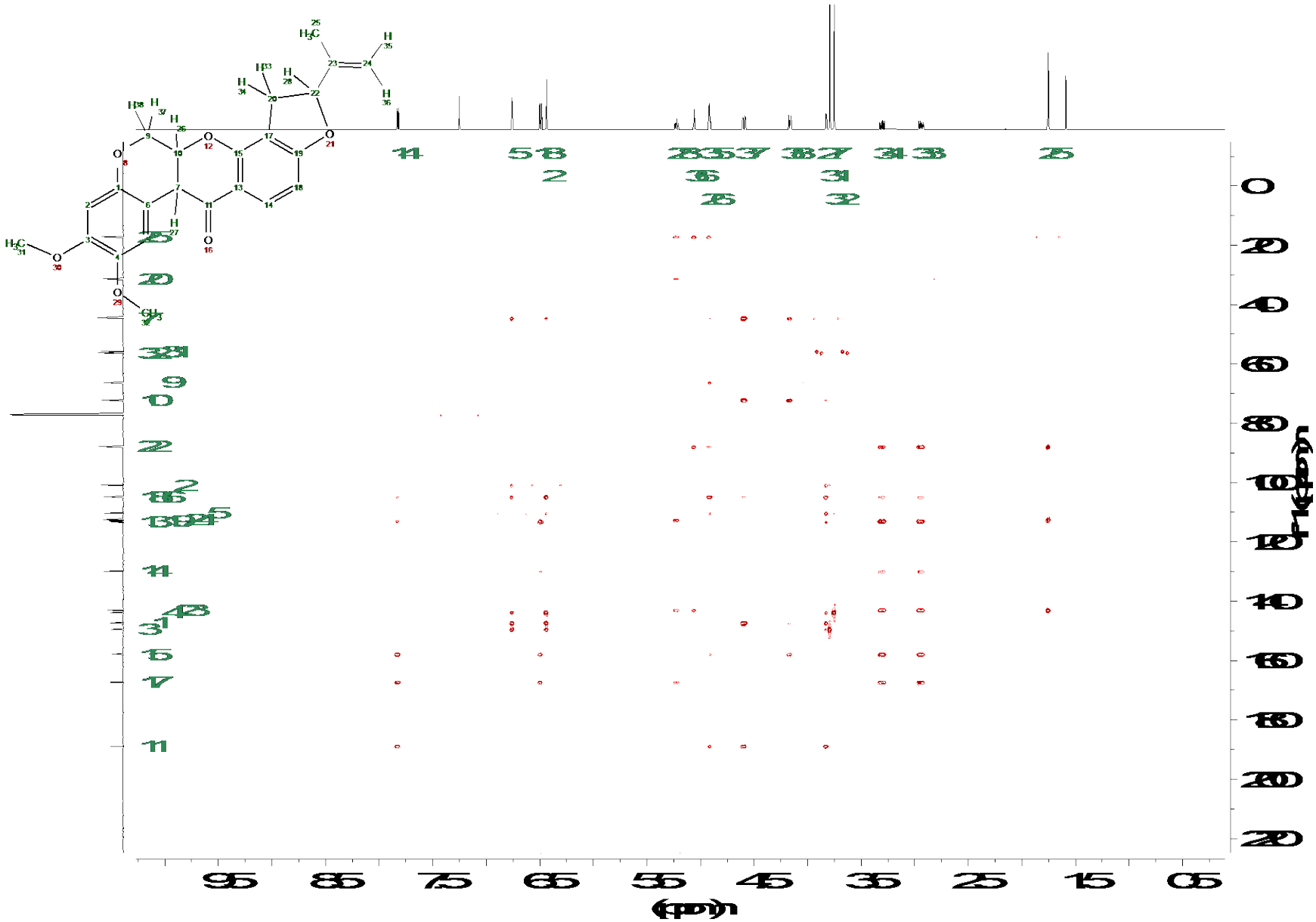

100 96 92 88 84 80 76 72 68 64 60
(ppm)


6.3 6.1 5.9 5.7 5.5 5.3 5.1 4.9 4.7 4.5 4.3 4.1 3.9 3.7 3.5
(ppm)
f1 (ppm)


6.3 6.1 5.9 5.7 5.5 5.3 5.1 4.9 4.7 4.5 4.3 4.1 3.9 3.7 3.5
(ppm)
f1 (ppm)

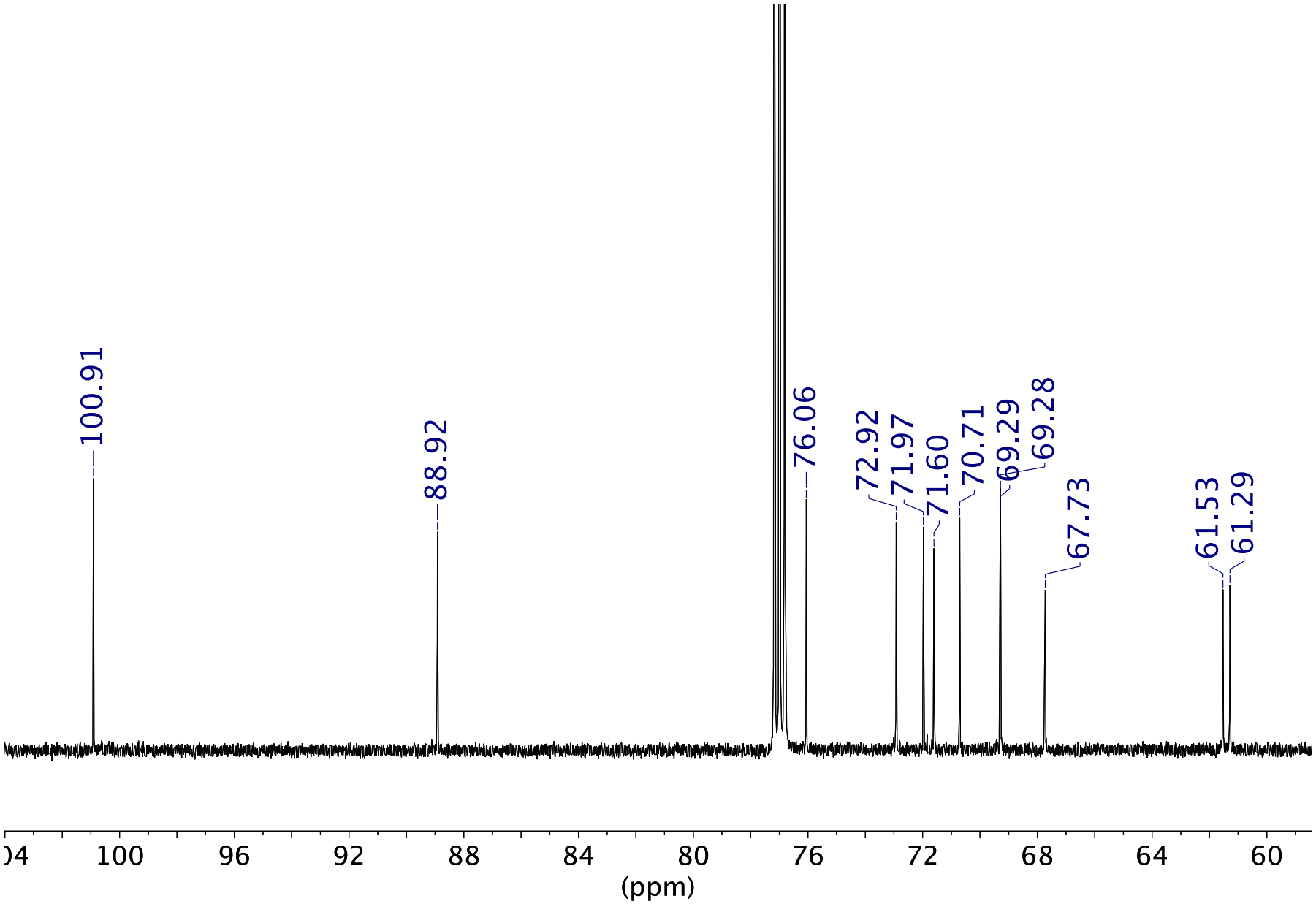

6.1 5.9 5.7 5.5 5.3 5.1 4.9 4.7 4.5 4.3 4.1 3.9 3.7 3.5
(ppm)
f1 (ppm)

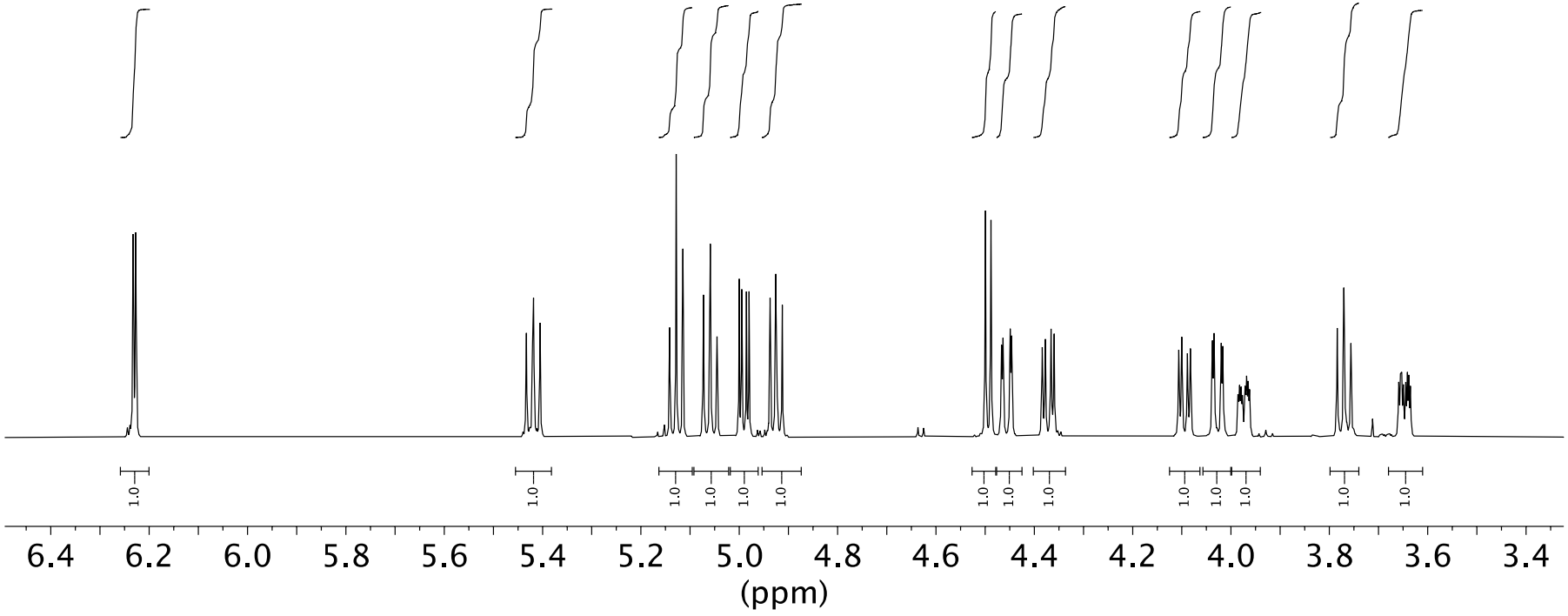

ghmbc optimised for 10 Hz couplings
a
b
c
d
e
f
g
h
i
j
ppm
2.5
2.0
1.5
1.0
0.5
0.0
20
25
30
35
40
45
ppm
74-bond
Current Data Parameters
USER nmr11t
NAME sample 17
EXPNO 8
PROCNO 1
F2 - Acquisition Parameters
Date_ 20111205
Time 0.06
INSTRUM cryo500
PROBHD 5 mm CPTCI 1H-
PULPROG ghmbc.wu
TD 4096
SOLVENT CDCl3
NS 4
DS 16
SWH 1750.700 Hz
FIDRES 0.427417 Hz
AQ 1.1698676 sec
RG 645.1
DW 285.600 usec
DE 6.00 usec
TE 298.0 K
CNST2 145.0000000
d0 0.00000300 sec
D1 1.25000000 sec
d2 0.00344828 sec
D6 0.05000000 sec
d13 0.00000300 sec
D16 0.00020000 sec
IN0 0.00001725 sec
======== CHANNEL f1 ========
NUC1 1H
P1 7.50 usec
p2 15.00 usec
PL1 1.60 dB
SFO1 500.2206124 MHz
======== CHANNEL f2 ========
NUC2 13C
P3 15.50 usec
PL2 -1.00 dB
SFO2 125.7948837 MHz
====== GRADIENT CHANNEL =====
GPNAM1 sine.100
GPNAM2 sine.100
GPNAM3 sine.100
GPX1 0.00 %
GPX2 0.00 %
GPX3 0.00 %
GPY1 0.00 %
GPY2 0.00 %
GPY3 0.00 %
GPZ1 50.00 %
GPZ2 30.00 %
GPZ3 40.00 %
P16 1000.00 usec
F1 - Acquisition parameters
ND0 2
TD 512
SFO1 125.7949 MHz
FIDRES 56.612320 Hz
SW 230.419 ppm
FnMODE undefined
F2 - Processing parameters
SI 2048
SF 500.2200000 MHz
WDW SINE
SSB 0
LB 0.00 Hz
GB 0
PC 1.40
F1 - Processing parameters
SI 1024
MC2 QF
SF 125.7804190 MHz
WDW SINE
SSB 0
LB 0.00 Hz
GB 0
2D NMR plot parameters
CX2 18.00 cm
CX1 15.00 cm
F2PLO 2.974 ppm
F2LO 1487.75 Hz
F2PHI -0.526 ppm
F2HI -262.95 Hz
F1PLO 50.187 ppm
F1LO 6312.57 Hz
F1PHI 14.855 ppm
F1HI 1868.50 Hz
F2PPMCM 0.19444 ppm/cm
F2HZCM 97.26112 Hz/cm
F1PPMCM 2.35546 ppm/cm
F1HZCM 296.27115 Hz/cm

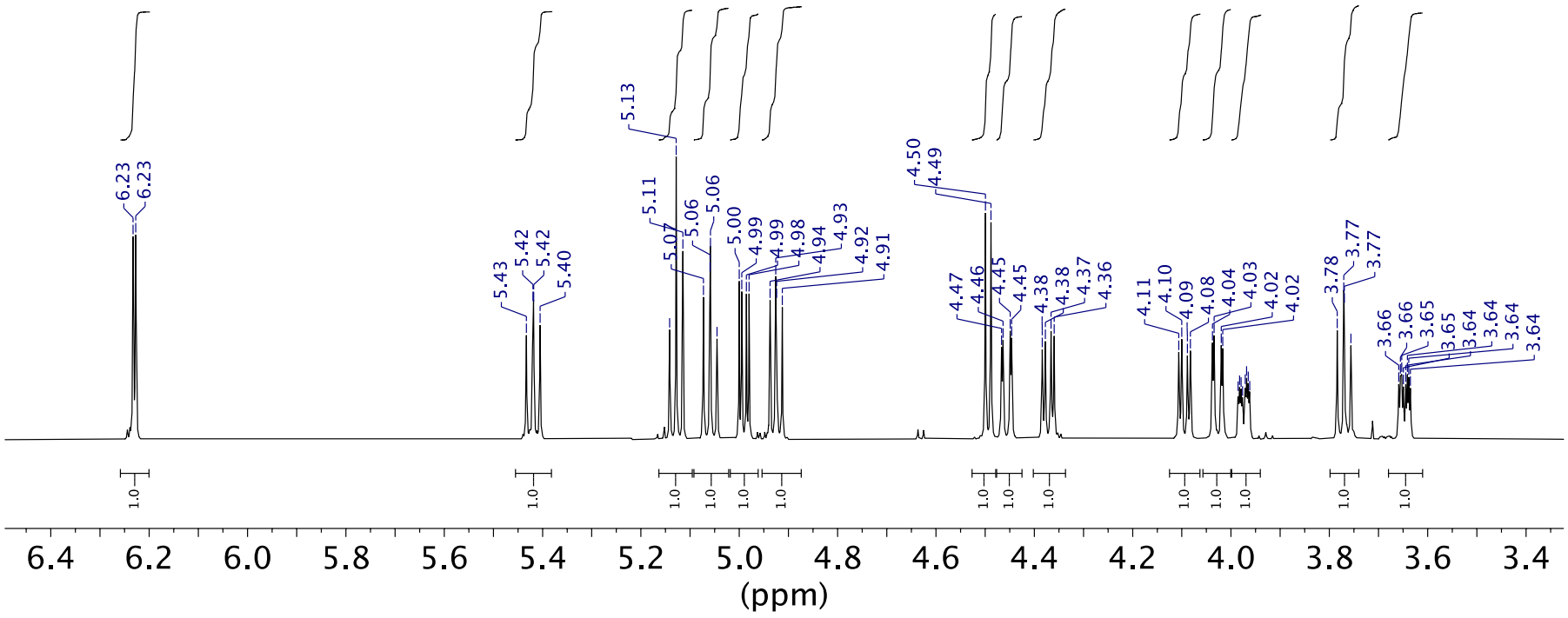


In order to simplify/speed the process, what information should we make use of?

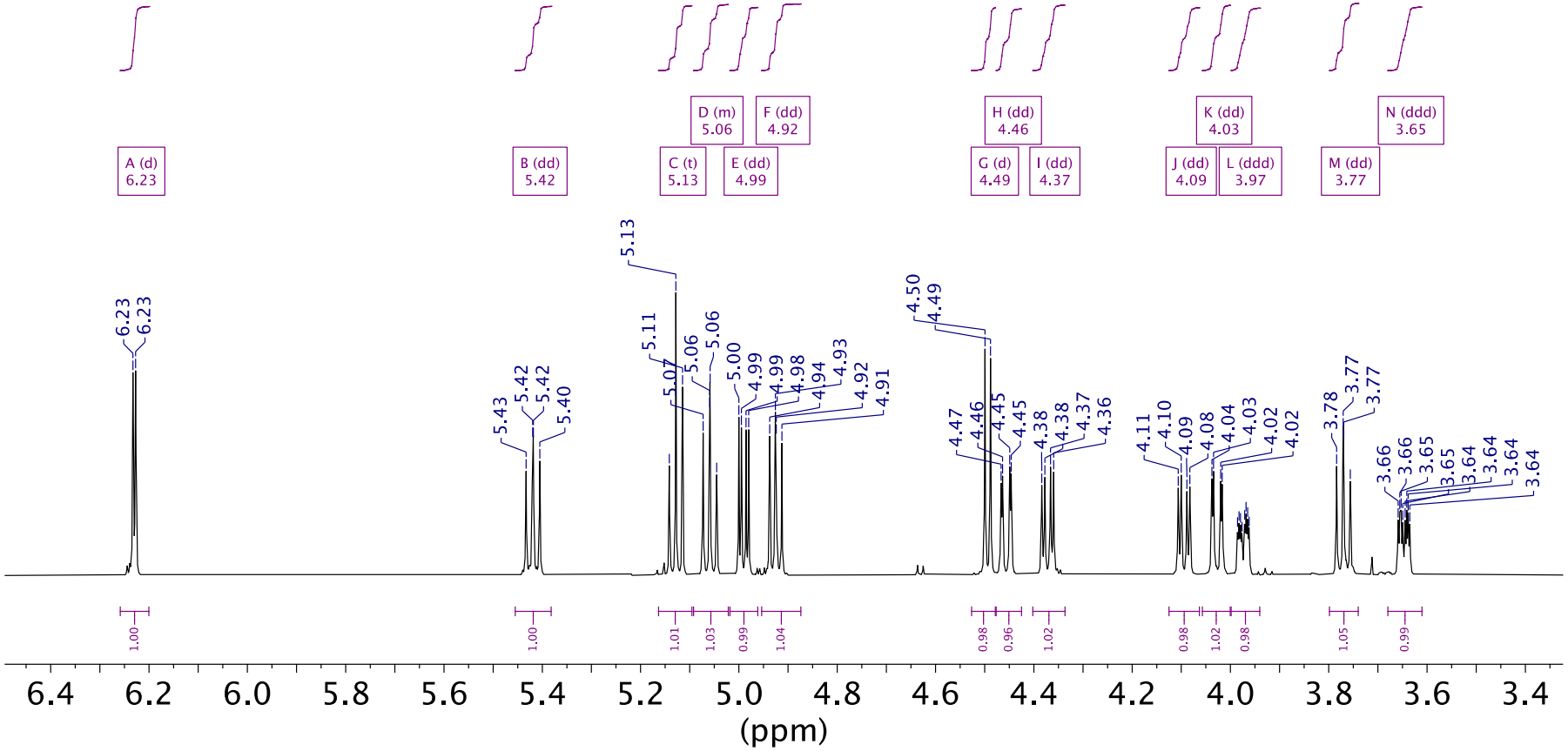
Information that can be measured directly from the spectra without the need for interpretation / analysis at this point.

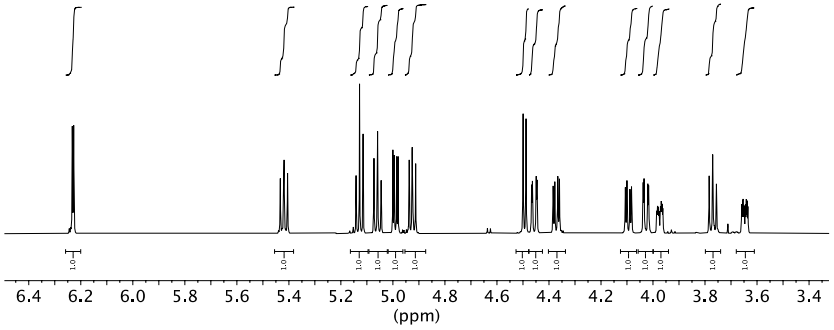

100.91
88.92
76.06
72.92
71.97
71.60
70.71
69.29
69.28
67.73
61.53
61.29
104
100
96
92
88
84
80
76
72
68
64
60
(ppm)

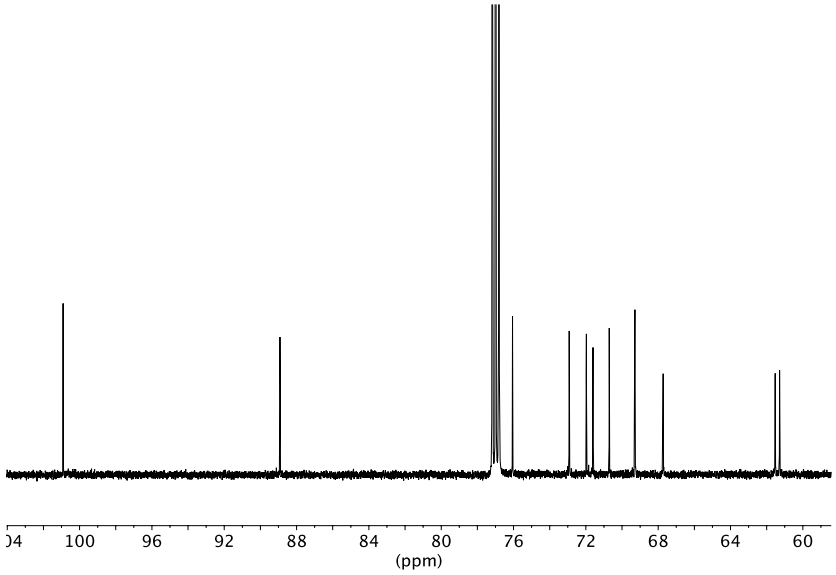

1.0
1.0
1.0
1.0
1.0
1.0
1.0
1.0
1.0
1.0
1.0
1.0
6.4
6.2
6.0
5.8
5.6
5.4
5.2
5.0
4.8
4.6
4.4
4.2
4.0
3.8
3.6
3.4
(ppm)

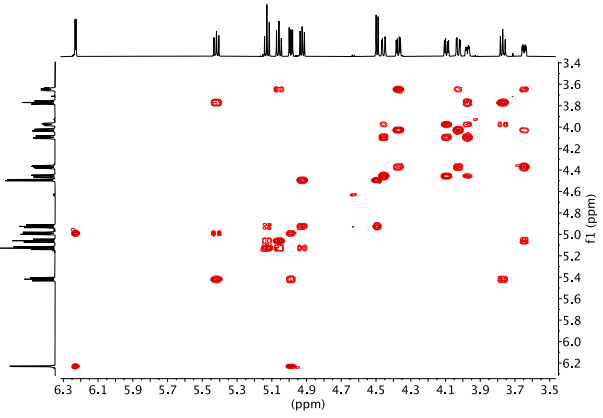

6.23
6.23
5.43
5.42
5.42
5.40
5.13
5.11
5.07
5.06
5.06
5.00
4.99
4.99
4.98
4.94
4.93
4.92
4.91
4.50
4.49
4.47
4.46
4.45
4.45
4.38
4.38
4.37
4.36
4.11
4.10
4.09
4.08
4.04
4.03
4.02
4.02
3.78
3.77
3.77
3.66
3.66
3.65
3.65
3.64
3.64
3.64
3.64
1.0
6.4
6.2
6.0
5.8
5.6
5.4
5.2
5.0
4.8
4.6
4.4
4.2
4.0
3.8
3.6
3.4
(ppm)

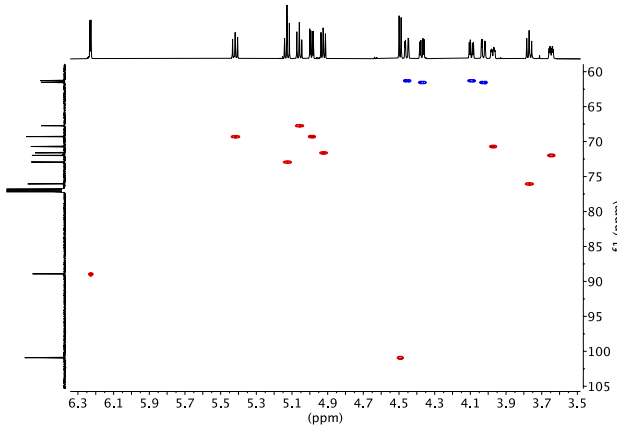
^{1}H NMR (700 MHz, cdcl$_3$) δ 6.23 (d, *J* = 3.7 Hz, 1H), 5.42 (dd, *J* = 10.3, 9.3 Hz, 1H), 5.13 (t, *J* = 9.4 Hz, 1H), 5.09 – 5.02 (m, 1H), 4.99 (dd, *J* = 10.3, 3.8 Hz, 1H), 4.92 (dd, *J* = 9.4, 7.9 Hz, 1H), 4.49 (d, *J* = 8.0 Hz, 1H), 4.46 (dd, *J* = 12.3, 2.1 Hz, 1H), 4.37 (dd, *J* = 12.5, 4.3 Hz, 1H), 4.09 (dd, *J* = 12.3, 4.3 Hz, 1H), 4.03 (dd, *J* = 12.5, 2.3 Hz, 1H), 3.97 (ddd, *J* = 10.1, 4.3, 2.1 Hz, 1H), 3.77 (dd, *J* = 10.1, 9.3 Hz, 1H), 3.65 (ddd, *J* = 10.0, 4.3, 2.3 Hz, 1H), 2.16 (s, 3H), 2.11 (s, 3H), 2.07 (s, 3H), 2.02 (s, 3H), 2.01 (s, 3H), 1.99 (s, 3H), 1.98 (s, 3H), 1.97 (s, 3H).

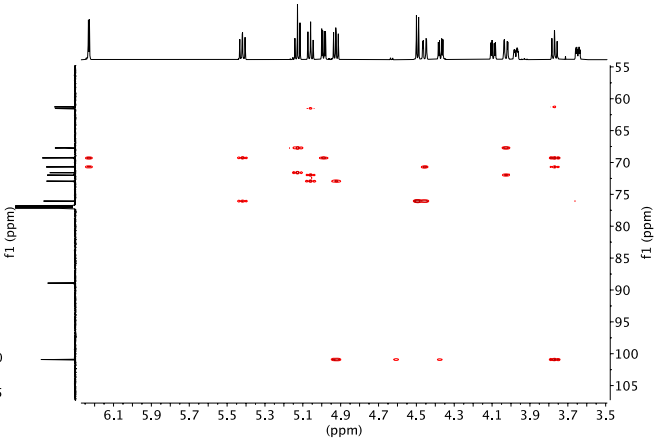



6.4 6.2 6.0 5.8 5.6 5.4 5.2 5.0 4.8 4.6 4.4 4.2 4.0 3.8 3.6 3.4
(ppm)


100 96 92 88 84 80 76 72 68 64 60
(ppm)


6.3 6.1 5.9 5.7 5.5 5.3 5.1 4.9 4.7 4.5 4.3 4.1 3.9 3.7 3.5
(ppm)
f1 (ppm)


6.3 6.1 5.9 5.7 5.5 5.3 5.1 4.9 4.7 4.5 4.3 4.1 3.9 3.7 3.5
(ppm)
f1 (ppm)


6.1 5.9 5.7 5.5 5.3 5.1 4.9 4.7 4.5 4.3 4.1 3.9 3.7 3.5
(ppm)
f1 (ppm)

How do we extract data from the spectra?

Check consistent referencing

“Peak Picking”

Identify chemical shift correlations

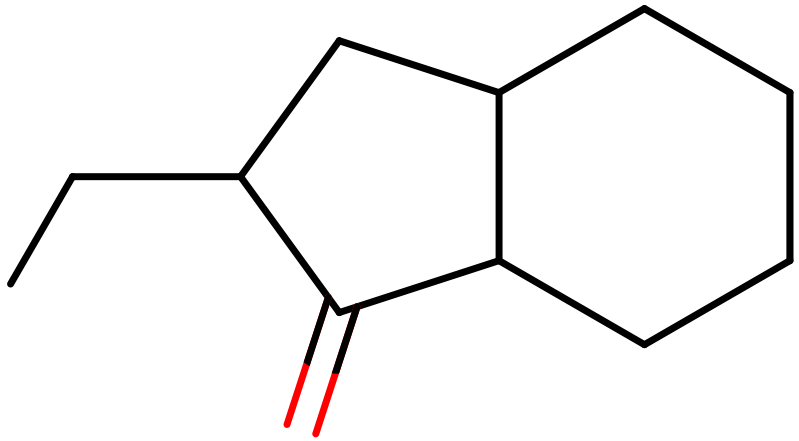
The Concept

Use extracted data to identify CH_x fragments, where x = 0, 1, 2, or 3 (ie quaternary, methine, methylene, methyl)

Position fragments over atoms in the structure (from smiles string)

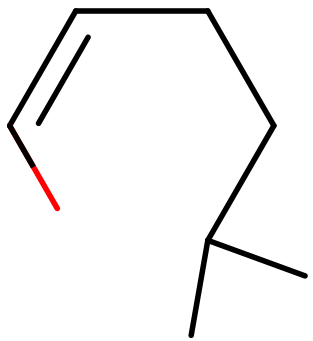
Display connectivity between fragments (COSY, HMBC) as well as displaying 1D ”spectra”

2-ethylindanone

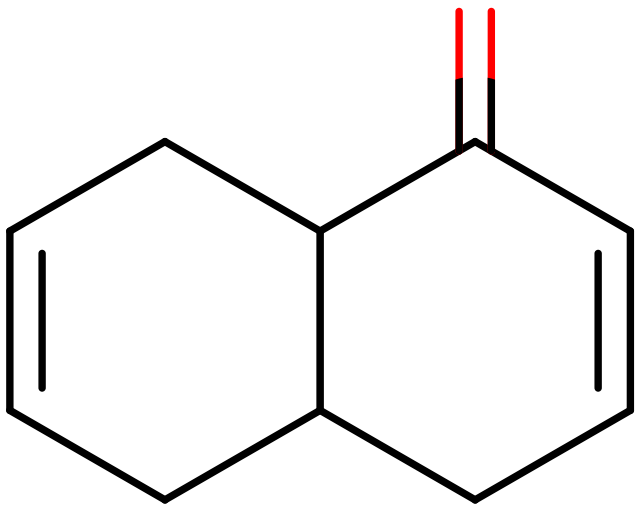


The problem with predictions

beta-lapachone

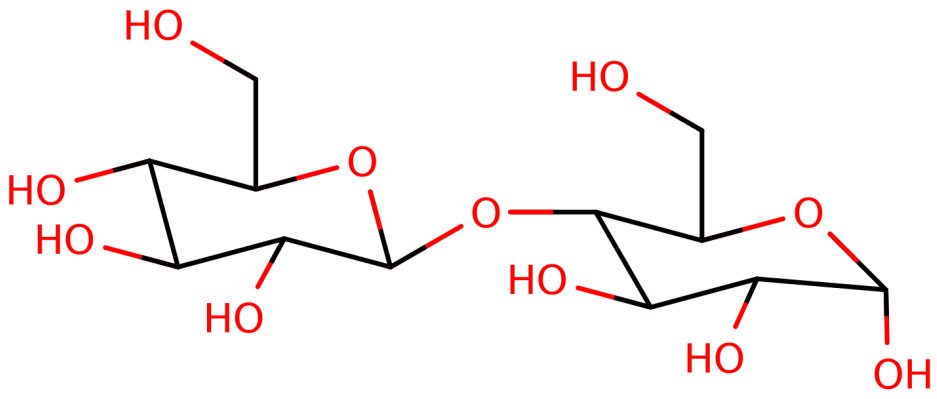


REF5



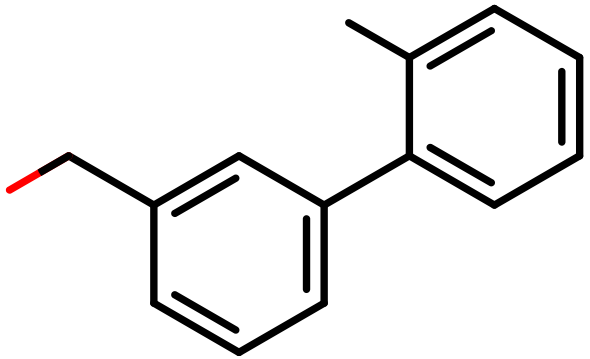
Cellobiose

Not really, I lied!



If you want to fully confirm the structure, rather than just the constitution, you need additional information from ^{1}H-^{1}H coupling constants and / or NOE's. But that task is made much easier by knowing where in the molecule each proton signal belongs.

Fasiglifam*



* Thanks to Christopher Sleigh of Sygnature Discovery

Advantages:

Relatively simple extraction of data – little or no interpretation required at the point of data extraction

Assignment involves positioning CH_x fragments on structure – concepts that chemists like

Possible to use 13C NMR prediction software to make initial placements

Can save assignment (and supporting evidence!)

Easy to check assignments and spot inconsistencies

Easy to change assignments and try an alternative

Areas most needing improvement:

Resolution in 2D spectra – particularly COSY (A)

Improved methods for identifying “x” in CH_x (A)

Improvements to “peak picking” to reliably identify chemical shift correlations (B)

A = NMR people problem! (a lot already underway)

B = Software problem (some work already in the literature)

SimpleNMR

Get simpleNMR from github.com:

https://github.com/EricHughesABC?tab=repositories

simpleNMR

Read the README file (bottom of page)

Releases v0.0.8

Select the appropriate executable