QMSA Letters 2025

The effects of C, Si, Cl and S isotopes in ^{1}H NMR spectra – not that simple – for perfectionists

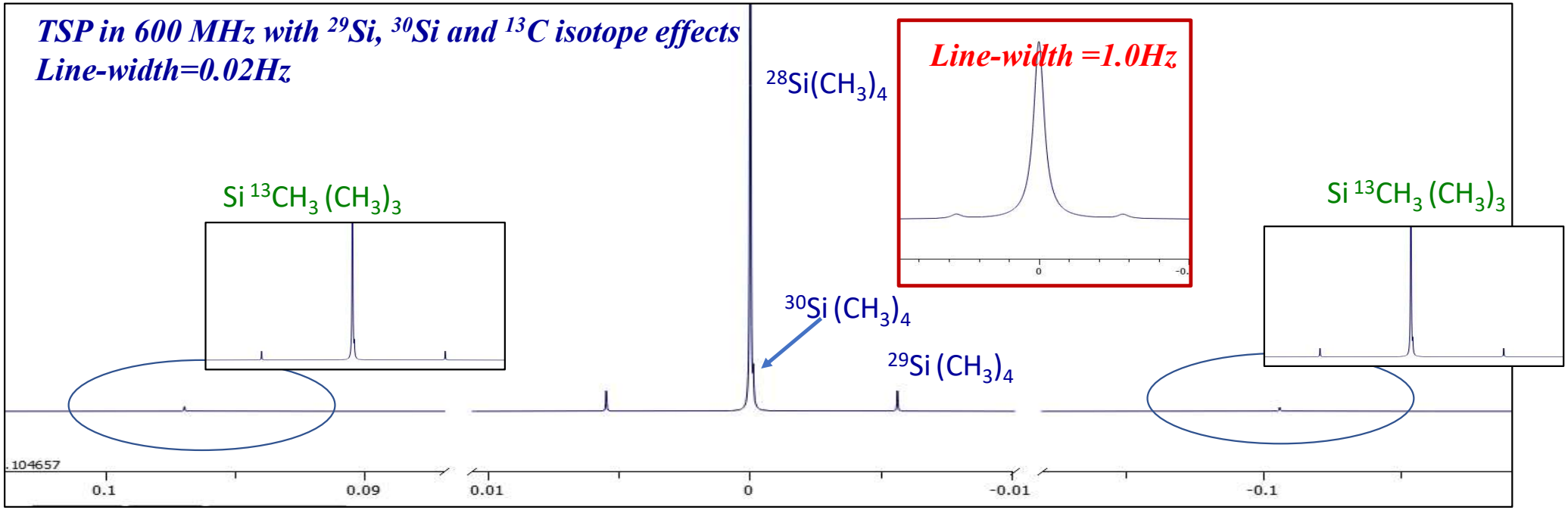
Reino Laatikainen[a], Pekka Laatikainen[a] and Henri Martonen[a]

[a] Spin Discoveries Ltd., Kuopio, Finland

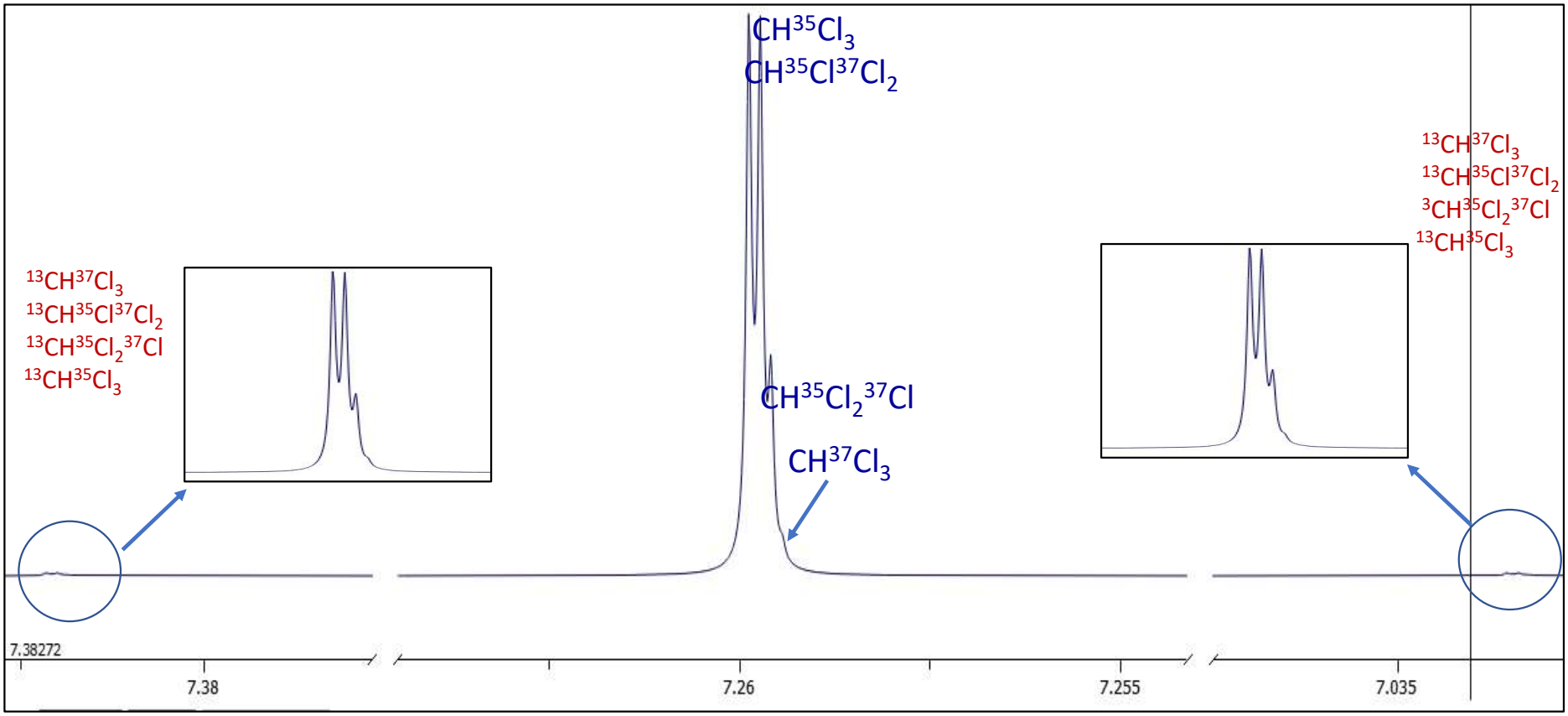
[b] University of Turku, Turku, Finland

Isotope effects are not insignificant !

- The NMR standards TMS, TSP and DSS have three Si-isotopomers: two spin=0 isomers (with 92.22% of ^{28}Si and 3.09% of ^{30}Si), ^{29}Si with spin ½ yielding the easily notable 4.685% satellites to the major signal.
- The chemical shift of ^{29}Si-doublet is not the same as that of the major ^{28}Si-isotopomer. The ^{30}Si-isotopomer has also a different chemical shift (its isotope effect is assumed to be double to that of the ^{29}Si isotope). - **If TMS, TSP and DSS are used as quantitative reference these effects need to be considered**! In qQMSA these isotope signals can be added to the line shape.



$CHCl_3$-signal with a high resolution (Line-Width = 0.05Hz); 4 Cl-isotopomers:



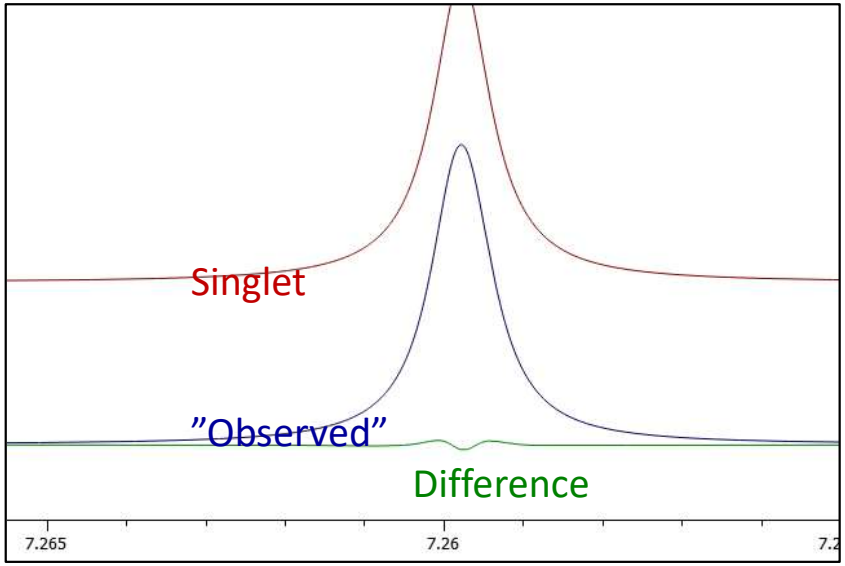
The $CHCl_3$-signal is a poor line-shape reference, instead of the complex isotopomer mixture signal, (in ChemAdder) one can use its integral!

$CHCl_3$ Isotopomers

- Chlorine has two major isotopes ^{35}Cl and ^{37}Cl, the normal ratio of which is ca. 0.7576:0.2424.
- The ratio yields four isotopomers $CH^{35}Cl_3$, $CH^{35}Cl_2{}^{37}Cl$, $CH^{35}Cl^{37}Cl_2$, $CH^{37}Cl_3$. Their theoretical percentages are 43.1 : 41.8 : 13.8 : 0.013. In addition, there are the ^{13}C satellites (1.1%) for each isotopomer (see previous page)!
- The isotopomer signals are resolved at high field and resolution: the isotope shifts are 0.00013 ppm (or 0.04 Hz at 600 MHz) for each ^{37}Cl (*JACS, 109, 6508 (1987)).*

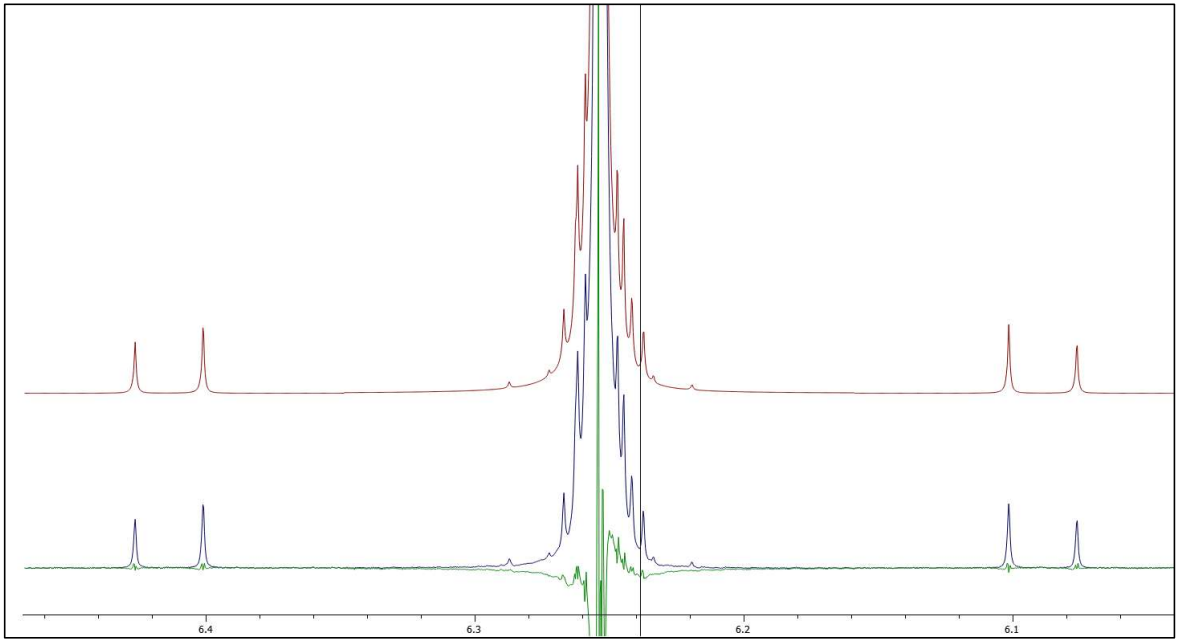
The $CHCl_3$ signal at 600 MHz, with line-width of 0.50 Hz, as fitted as a singlet (ignoring the isotope shifts) but optimizing the line-shape (Lorentzian, Gaussian and Asymmetry contributions + "*out-of-coil*" line shape correction).

Only a small bias is caused by the isotope shifts!

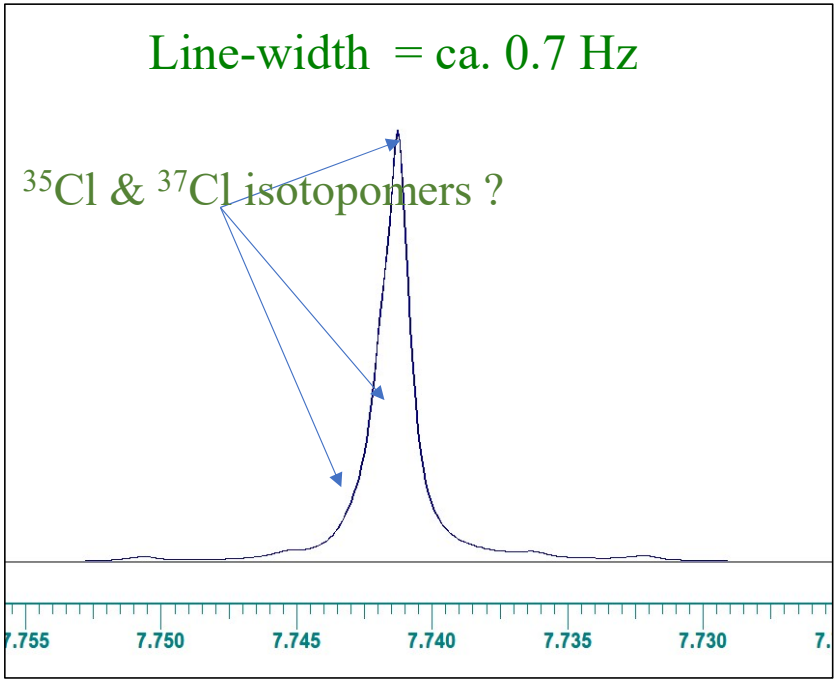


Good news: the effects can be fitted by line-shape, and they are insignificant in **benchtop** spectra! In new ChemAdder one can use its integral for quantitative referencing!

Maleic acid



^{35}Cl & ^{37}Cl and ^{13}C isotope effects in 1,2,4,5-chlorobenzene



^{13}C Isotope effects and prior knowledge

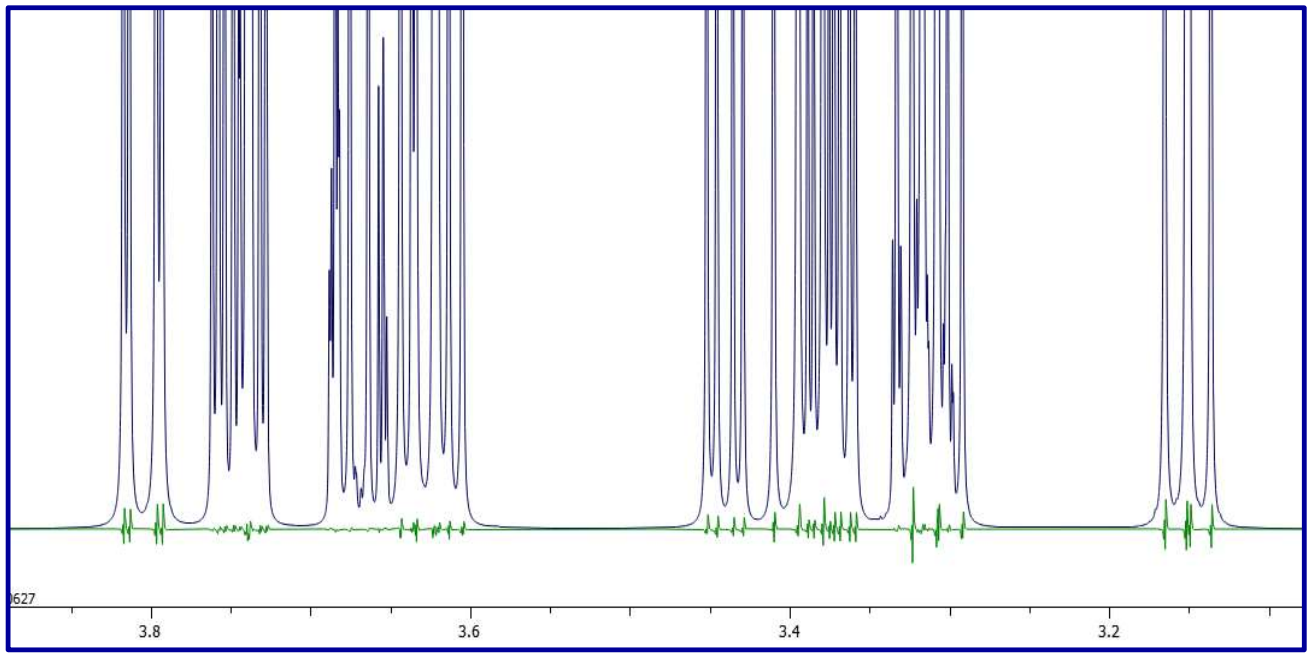
- The isotope effects are small and the ^{13}C satellites are weak so that their independent fitting is impossible. However, one can assume that:
  - The population of the ^{13}C satellites is 1.1%, 2.2% or 3.3% of the total area. The Si-isotopomers with ^{13}C can be ignored, due to their low populations.
  - The line-widths and line-shapes of all the isotopomers are the same.
  - The couplings can also be kept fixed, although in very different conditions they may be changed.
  - The relative shifts do not vary and can be fixed.

All the above constraints are included in ChemAdder ASL-files. For example, the isotopomer populations are fixed by using Response Factors and setting them fixed. The relative isotopomer shifts are fixed by naming the isotopomer shifts, for example, as TSP@Si28, TSP@Si29 and TSP@Si30.

Glucose

The **effective*** ^{13}**C contributions** to glucose calculated spectrum at 600 Mz spectrum, with line-width = 0.5 Hz:

*Means that after adding the isotope effects to a synthetic spectrum, the default line-shape parameters were optimized - so that the ***essential RMSE*** dropped from 0.25 to 0.14%.



Conclusion: the ^{13}C isotopes effects cannot be completely described with Lorenzian-Gaussian-Asymmetry-Dispersion line-shape - ignoring of the effects leads to variance of 0.14% (in the essential RMSE).

The effects can be significant in impurity analyses in high field, but they can be ignored at the benchtop.

ChemAdder offers no graphical interface for adding the isotopes, yet …but

The following ASL-files are given in REFERENCES-folder
TMS.ASL, TMS_Si.ASL, TMS_SiC.ASL,
TSP.ASL, TSP_Si.ASL, TSP_SiC.ASL,
DSS.ASL, DSS_Si.ASL, DSS_SiC.ASL,
Acetate_13C.ASL

Also, the following isotope signals can be useful:
Creatinine_13C.ASL
Creatine_13C.ASL
Maleic acid_13C.ASL
Butyric, lactic and propionic acids in some biological samples.

Sulphur has two significant isotopes (^{32}S 95%, ^{34}S 4.2%), the ^{33}S 0.75% can be ignored here. The isotope effect of ^{34}S can be assumed to be the same as with ^{29}Si (0.00012 ppm, 0.72 Hz at 600 MHz). The effect can be visible, for example, for dimethyl sulfoxide and dimethyl sulfone, with a high resolution and field. The isotope effects may be compensated by using our "*out-of-coil*" line shape correction.

How to define ^{13}C-isotope shifts in PMR-file

For example, glucose

```
&CHEMICAL SHIFTS:
GLUCOSE@B    2*SPIN=1  SPECIES=1H  MOL%(Y)=62.8092 [ DEF=62.4535 MIN=1.000 MAX=100.0 OCC=10 ] SE=0.3486 MW=180.16
 BG1      4.552439  1*1*1  STAT=Y   PRED= 4.5530 RANGE(0)= 0.033 WIDTH(Y)= 0.511 RESP(Y)= 0.9111 C13=-0.70
 BG2      3.151073  1*1*1  STAT=Y   PRED= 3.1518 RANGE(0)= 0.026 WIDTH(Y)= 0.517 RESP(N)= 1.0000 C13=-0.70,-0.70
 BG3      3.393743  1*1*1  STAT=Y   PRED= 3.3945 RANGE(0)= 0.027 WIDTH(Y)= 0.519 RESP(Y)= 0.9934 C13=-0.70,-0.70
 BG4      3.309449  1*1*1  STAT=Y   PRED= 3.3103 RANGE(0)= 0.027 WIDTH(Y)= 0.515 RESP(Y)= 1.0024 C13=-0.70,-0.70
 BG5      3.372964  1*1*1  STAT=Y   PRED= 3.3736 RANGE(0)= 0.027 WIDTH(Y)= 0.518 RESP(Y)= 1.0112 C13=-0.70,-0.70
 BG6A     3.804547  1*1*1  STAT=Y   PRED= 3.8051 RANGE(0)= 0.029 WIDTH(Y)= 0.517 RESP(Y)= 0.9876 C13=-0.70
 BG6B     3.629129  1*1*1  STAT=Y   PRED= 3.6298 RANGE(0)= 0.028 WIDTH(Y)= 0.511 RESP(Y)= 0.9593 C13=-0.70
GLUCOSE@A    2*SPIN=1  SPECIES=1H  MOL%(Y)=37.1908 [ DEF=37.5226 MIN=1.000 MAX=100.0 OCC=10 ] SE=0.2064 MW=180.16
 AG1      5.139082  1*1*1  STAT=Y   PRED= 5.1397 RANGE(0)= 0.036 WIDTH(Y)= 0.514 RESP(N)= 1.0000 C13=-0.70
 AG2      ....
```

In addition, activate the ^{13}C correction by setting

```
LINESHAPE CORRECTION = I    ! N = NO | I = ISOTOPE
```

in the profile file.

Line-shape

- The theoretically, NMR line-shape obeys Lorentzian function L(v). Its essential property is that the function decays slowly.
- FID manipulations (resolution enhancement) may change line-shape so that it closes to Gaussian function G(v), which decays quickly.
- The third basic NMR line-shape function is dispersion D(v), which decays very slowly. We have found that in practice the first derivative of the G(v) is more useful is better in describing asymmetry of line-shape.
- Sometimes (obviously as an instrumental artefact), a better fit to the line-shape may demand that the left and right line-widths are different.

The ChemAdder default line-shape is composed of three terms:

$$\text{Intensity}(v) = \textit{Constant} * I_i * [a*L(v, \Delta_L,\Delta_R) + b*G(v, \Delta_L,\Delta_R) + c *D(v, \Delta_L,\Delta_R)]$$

Where *Constant* relates the observed and calculated intensity (depends on concentration), I_i = the intensity of transition i (from QM simulation) Δ_L and Δ_R = the line-widths for the left and right sides, and optimizable parameters. Also, the coefficients a, b and c are optimizable. The line-widths depend on the relaxation time, but then one usually assume that $a + b + c = 1$. However, for strong (solvent) or important signals (reference) nuclei specific line-shape can be used. It a spectrum is strongly resolution enhanced, it may lead to negative 'wings' in the line-shape, which can be described with $b>1$ and $a<0$. We have not been found the dispersion contribution useful in normal spectra.

All the experimental line-shapes cannot be completely described by the above model. ChemAdder offers also corrections to some special line-shape defects.

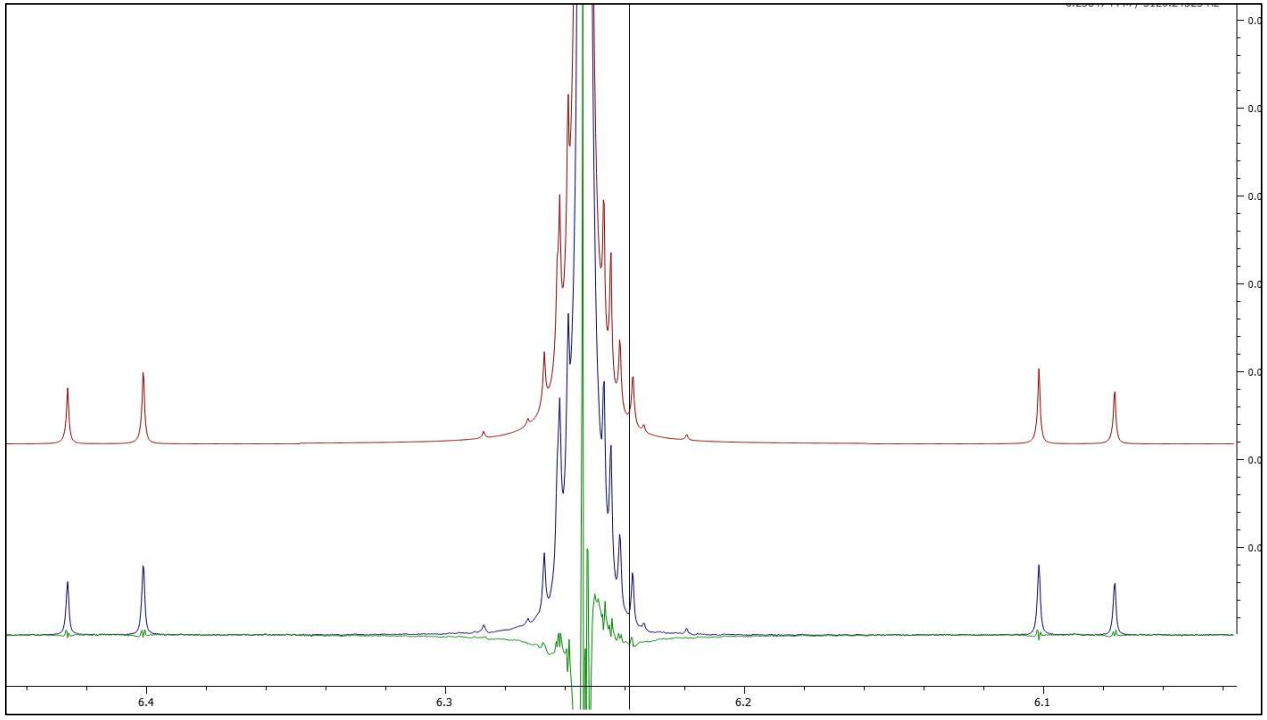
“Out of Coil” correction, virtual couplings and isotope shifts

- For strong and broadened signals, can be added a correction, which can be thought to describe signals coming from out of the coil sample, locating in less homogenous magnetic field. In ChemAdder, this effect is described by 5 additional asymmetric Lorentzians, the line-widths of which vary between 0.5 to 32 times the line-width of the species. The couple of the narrow terms can be used to fix isotope shifts. The broadest terms act as an efficient descriptor for the baseline!
- Small long-range couplings cause effects which appear as broadening, which cannot be completely modelled by the Lorentzian - Gaussian line-shape. In ChemAdder one can add non-specific ‘virtual’ couplings to multiplets. Thus, every line of a multiplet is split to a doublet, triplet or quartet. The splitting is optimized as coupling.
- ^{13}C isotopes lead to ca. 0.00012 ppm effect on the shifts of the protons in geminal position. Thus, the effect of three methyl's on C**H**-proton shift of $(CH_3)_3$-C**H,** at 600 MHz is 3 * 600 * 0.0012 = 2.16 Hz and intensity of the shoulder is 3*1.1 = 3.3%. The effect is seldom visible by eye, but may cause a small extra RMS-error, which is usually insignificant. Anyhow, the ^{13}C isotope shifts can be easily added to simulated spectra.

All the above corrections can be added simultaneously. They increase the computational efforts, but they can be added only to selected protons, for example, to the reference and solvent signals.

^{13}C Effects

- The ^{13}C satellites may form a nuisance, when the concentration range of a sample is large and there is a few major components. For example, glucose in biofluids.
- If the satellites are removed, there remains ^{13}C isotope shift that moves the decoupled isotopomer multiplets (which are similar to the main multiplets) by ca. 0.0012 ppm (0.72 Hz) at 600 MHz.
- Also, the geminal and vicinal ^{13}C,H-couplings may have odd effects, for example, maleic acid:



Conclusion

- The **^{29}Si couplings** in TMS, TSP and DSS are visible in any field, and should be added to qQMSA, especially when the signal is used as a quantitative or line-shape reference – which is not a good idea in some samples.
- Also, the **chlorine isotope effects** can be normally ignored. Although the broadening in $CHCl_3$ at 600 MHz is only ca. 0.08 Hz, it cannot be completely compensated by optimizing the line-shape – our analysis shows that the isotope effect is in this case an origin of the *mysterious* observed-calculated difference.
- Because the model for glucose (which dominates many biological samples) with the ^{13}C satellites would be very complex, **we recommend that the samples for qQMSA are always measured removing the ^{13}C satellites and decoupling the long-range ^{13}C,H couplings.**
- The geminal ^{13}C isotope shifts vary (0.7 ppb is a fair estimate) and can be added to some strong signals, like methyls.
- Also, the vicinal and even long-range shifts should be considered – **impossible in practice!** However, the isotope shifts are proton-specific (the same for every line of a proton), as also are some other effects like long-range couplings => **The SOLUTION: a proton specific line-shape correction, see QMSA Letters ?(2022).**
- **Forget the isotope shifts for benchtop spectra !**