The NMR-Purity and the new dimensions and standards of purity. qQMSA+CTLS

SAVE 90% !€

[1]Reino Laatikainen, Pekka Laatikainen and Henri Martonen
[2]USP Team: Sunil Paudel, Christine Castagna, Jana Brcek and Ben Shapiro
[1]Spin Discoveries Ltd., Kuopio, Finland
[2]USP: US Pharmacopeia, USA

€ in NMR spectrometer time, sample preparations, calibration, reference standards, calibrations and data handli

Integration vs. Total-Line-Shape Methods and qQMSA

1

Integration is sensitive to phase and baseline. When the spectrum presents isolated signals, with integration is possible to achieve 95% confidence intervals as low as 1.5% of the compound concentrations, and that the integration of well-isolated peaks can lead to relative uncertainties of 11% when there are even slight phase and baseline errors[1]. *This makes the integrals manipulable through phase,baseline and integration range, which are subjective parameters, but fortunately, easily from the original FID – in opposite to weighting and reference quality bias.*

The most serious of problem of integration is the peak overlap. To be precise, the signals to be integrated need to be sufficiently isolated. Lorentzian signals decay slowly to infinity: for a maximum error of 1%, integration limits of 25 times the line width (±25 Hz or 0.05 ppm with linewidth of 1 Hz at 500 MHz) are needed. For errors < 0.1%, the limits should be ±76 Hz (±0.15 ppm)[2].

The risk that an impurity signal is hiding under the integrated signal, increasing thus the integral – and purity – is serious with integration. The essential impurities can be expected to resemble to the target compound. The same problem but more serious, is with the MS and chromatographic methods.

1. Malz, F. and Jancke, H., *J.Pharm.Biomed.Analysis*, 2005, 38, 813–823. https://doi.org/10.1016/j.jpba.2005.01.043
2. Griffiths, L., *The Analyst*, 1998, 123, 1061–1068. https://doi.org/10.1039/a800625c

QMSA sounds as an ideal tool for pNMR. However, it to be perfect, the line-shape model should be perfect!

The line width and line shape depend on compound, spectrometer and weather.
In ChemAdder, the line shape can be described using asymmetrical Lorenzian, Gaussian and Dispersion functions, adding out-of-coil corrections, virtual long-range couplings and isotope shifts for ^{13}C and S, Cl.
Despite the most sophisticated line-shape functions, the fitting usually leads to typical observed-calculated difference spectrum (trace).

The *charasteric* trace arises from line-shape artefacts, which are compound- and proton-specific (see inserts, similar for the two multiplets). The artefacts cannot be totally removed with line-shape tools or the approaches are impractical.

If the observed-calculated area is not ZERO, as it should be, this leads to a small bias in concentrations.

Unfortunately, the area is sensitive to the line-shape model!

The bias can be minimized using the broadening (both windowing and/or IT-transforms), adding the ^{13}C couplings and isotope shifts to the model (clumsy) or by using ChemAdder pNMR Options (*'IT Supported Fitting , ITSUF'*)

The NMR Purity

2

NMR Spectral Purity*, shortly *NMR Purity* or *Spectral Purity

There are two principal ways to characterize chemical purity of a compound:

1. Determine the purity of the main component concentration. Even if the measurement is done very carefully, the accuracy is seldom better than 1-0.5%, and very accurate weighting (thus, in sufficient amounts and even *buoyancy* should be considered) and very pure references for calibration are needed.

2. Determine the impurities. For impurities, the demanded accuracy of the method is seldom an issue - it may be sufficient to say that the impurity% is <0.1% or 0.1-0.2%. Accurate weighting of small amounts and calibrations are not needed. The impurity concentrations can be determined with accuracy that is far better than demanded !
qQMSA+CTLS offers five quality parameters for the sample quality:

- **Purity%, a reference signal needed**
- **The NMR-Purity has dual meaning: *firstly, it means 'purity of the NMR spectrum'*, secondly, *'purity of sample based on NMR spectral purity' !***
- **eR2 = essential R2: measures goodness of observed-calculated fit)**
- **eRMS = essential RMS: measures spectral quality and analysis)**
- **Parameter MATCH: measures fit of spectral parameters to the default values of chemical shifts and coupling constants. The conditions (solvent, concentration, pH, temperature, ..) become critical with use of this parameter.**
- **In addition, the impurities can be identified or at least characterized.**

The NMR-purity parameters offer superior measures for chemical quality control, with minimal calibration and preparations! Also, the acquisition and relaxation times can be shortened to < 10s, because the relaxation effects can be treated through RESPONSE FACTORS.

SAVE 90% !€

RESPONSE FACTORS: the source of serious bias of P% (if ignored)

3

Definition of **RF = 1.0 if the intensity of NMR signal area/proton = that of the reference signal**

RF's of α-glucose in different experiments differ by up to **19%** from the default (1.0)[$]:

	qH[a] D_2O	*H[b] D_2O*	*qpresat[c] D_2O*	*presat[d] D_2O*	*qpresat[c] H_2O+D_2O*	*presat[d] H_2O+D_2O*
H1	**0.962**	**0.875**	0.960	0.880	0.950	0.924
H2	**0.974**	0.993	0.965	0.993	0.904	0.909
H3	**1.000**	0.910	**1.000**	0.920	0.969	**1.000**
H4	**0.978**	0.953	0.990	0.990	**1.000**	0.978
H5	**0.965**	0.997	0.975	**1.000**	0.850	0.885
H6A	**0.977**	0.997	0.953	0.994	0.884	0.868
H6B	**0.975**	**1.000**	0.955	0.981	**0.811**	**0.840**

[a] **Basic proton spectrum** (zg): 128k data points (td), ds=4, ns = 8, AQ=7.7s, **RD(d1)=52.3s** and 90° pulse.
[b] Basic proton spectrum (zg): td=128k, ds=4, ns=32, AQ=7.7s, d1=2.3s and 90° pulse.
[c] Noesypresat pulse sequence (noesygppr1d): 10 ms mixing, td=128k, ds=4, ns=8, AQ=7.7s, d1=3.0s, addnl. delay before suppression(d2)=49.3s and 90° pulse.
[d] As in c, but d2=0.
[$] *J.Magn.Reson., 242(2014)67.*

The *'principle of very details'*

4

The ^1H NMR spectral area of a compound can be modelled into ***very details*** - the rest of the area represents to impurities or the compounds which are not in the model:

NMR-impurity% = 100 * [Observed area-Simulated area]/[Observed area]

Then [**100 - NMR-Impurity**] represents the **NMR-purity** of the sample.

- **NMR-purity** of 100% means that there are no other compounds that give ^1H NMR signals, but not that there would be no salts or solvent trace (which must be defined in the model).
- Impurity signals hiding under crowded spectral regions can be revealed by QMSA!!
- NMR-impurity does not give the impurity weight%, if the molar weights corresponding to the impurity signals are not known, which is often the case. However, it is sufficient to know the type of proton (CH, CH_2 or CH_3) to get the molarity% of the corresponding compound.

Essential NMR Statistics

5

RMSE = RMS Error
R2 = The fraction of variance explained by model } depend on spectral width
Noise = RMSE for signal free parts of spectra!

Essential Range = percentage of regions with NMR signals (calculated intensity > threshold)
Essential RMSE = RMSE for the Essential Range,
Essential R2 = R2 for the Essential Range, } independent of spectral width !!
NMR-Purity = percentage of spectral area explained by QMSA
NMR-Impurity = percentage of spectral area NOT explained by QMSA

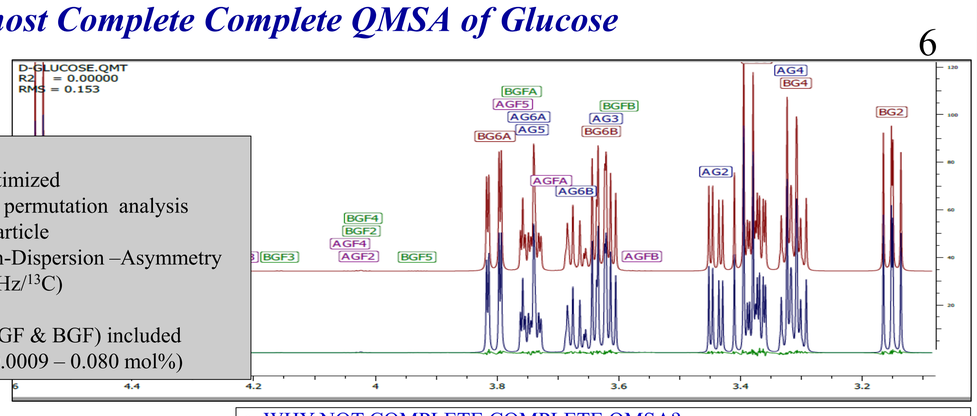
eRMSE is used in error analyses and a good measure of spectral quality !

Residual root mean square Error (RMSE or RMS) and R2-value, commonly used to describe the quality of QMSA, **are poor descriptors** because they depend on the spectral width, which can be 30 ppm while the spectrum occupies only 1 ppm!

The RMSE and R2 should be independent of the range and RMSE equal to noise, but usually RMSE > noise, because the line-shape model is never perfect!

Almost Complete Complete QMSA of Glucose

6



- ^{13}C Decoupled ^1H spectrum
- Chemical shifts & couplings optimized
- Long-range couplings after sign permutation analysis
- Different line-widths for each particle
- Line-shape: Lorenzian-Gaussian-Dispersion –Asymmetry
- Geminal ^{13}C isotope shifts (0.7 Hz/^{13}C)
- Response Factors optimized
- 2 Minor **furanose** tautomers (AGF & BGF) included
- Several other trace impurities (0.0009 – 0.080 mol%)

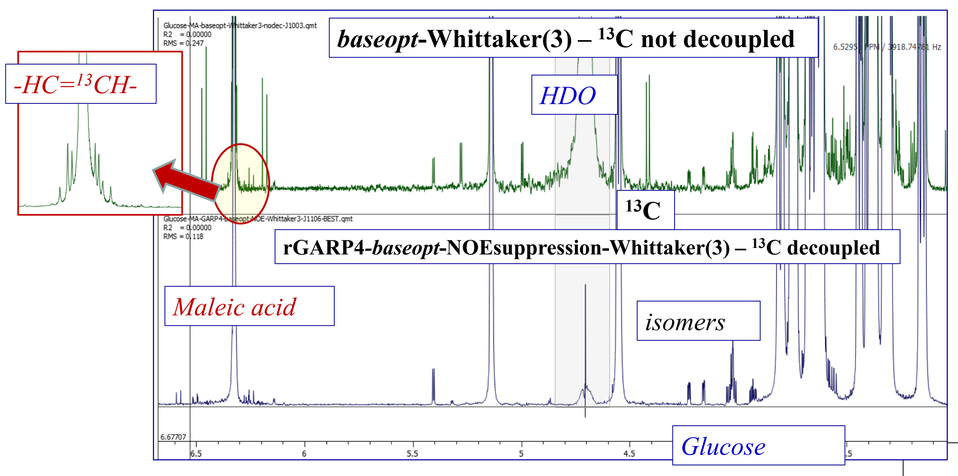
- Noise = 0.01%
- RMSE = 0.15%
- **Essential RMSE = 0.32%**
- Essential R2 = 99.96
- **NMR-purity = 99.86%**

- WHY NOT COMPLETE COMPLETE QMSA?
- Incomplete permutation of signs of the LR couplings
- All the geminal ^{13}C isotope shifts equal
- No vicinal and long-range ^{13}C isotope shifts
- Other isotope effects ignored ($^{17,18}O$)
- Dipolar couplings ignored
- Minor (for example, open-chain tautomers) impuritities not treated with QMSA
- …

The above effects are spin-particle specific, while the instrumental artefacts are similar for all the species !!

Removal of ^{13}C couplings using *'gentle GARP4 decoupling'*

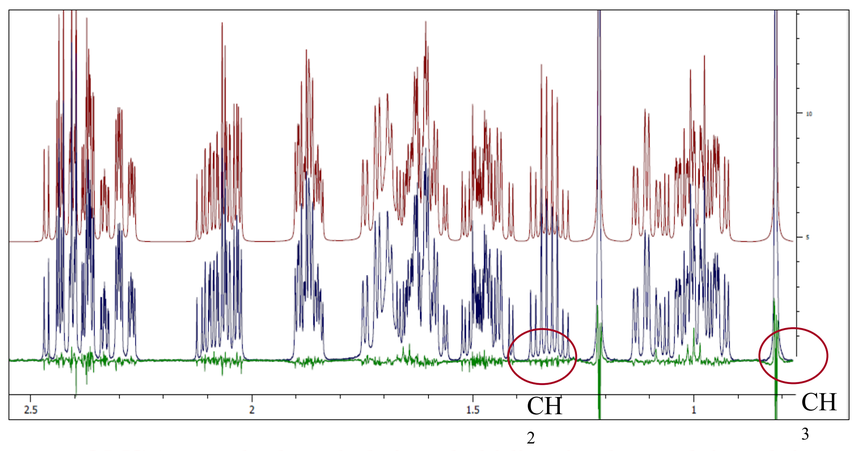
7



QMSA of Testosterone + ca. 0.40 wt% CH_3CH_2R-impurity

8

The spectrum was fitted by ignoring the impurity from the model:



QMSA reveals the ethyl signals hiding in the peak jungle!

Holistic quantitative QMSA(qQMSA) + CTLS

9

(CTLS = Constrained Total-Line-Shape)

A spectrum data may contain the 5 different type NMR signals :

1. *Quantum Mechanically modellable signals*
2. *Xtructures* (singlets, multiplets), like polymer and lipoprotein signals
3. *Xpectrals*, like albumin spectrum
4. *Xpurities*
5. *Integrals*

The common point is that the signal area/nucleus is the same:
Total area = QM + Xtructures + Xpectrals + Xpurities + Integrals
All the types can be handled in one model by ChemAdder !
See our poster on hQMSA!

https://www.chemadder.com

CA