1st International Electronic Conference on Metabolomics
1-30 November 2016
chaired by Dr. Peter Meikle
sponsored by metabolites

Quantitative Quantum Mechanical NMR Analysis: the Superior Tool for Analysis of Biofluids

Reino Laatikainen[1,*], Pekka Laatikainen[2], Henri Martonen[2], Mika Tiainen[1], and Elias Hakalehto[3]

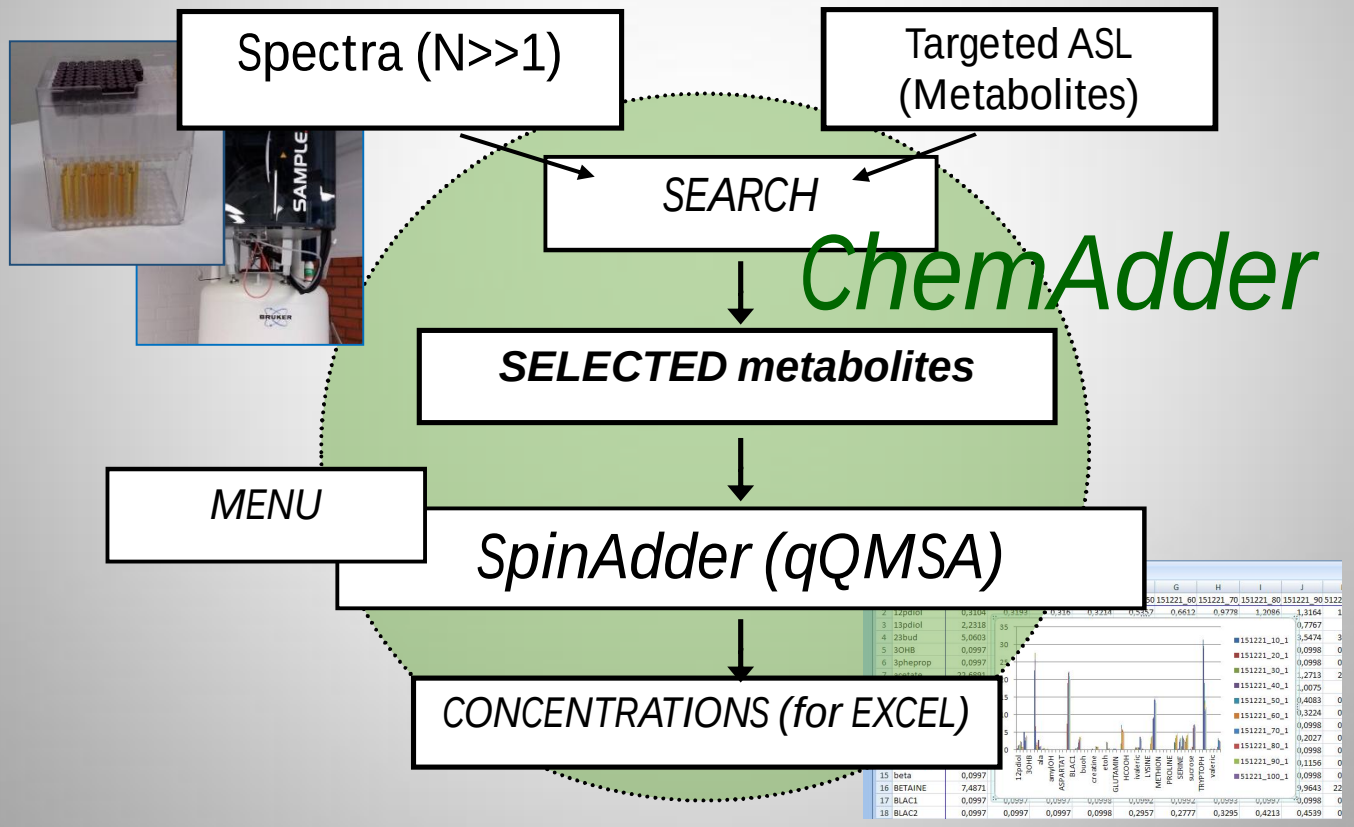
[1] School of Pharmacy, Univ. of Eastern Finland (UEF), P.O.Box 1627, FIN-70211Kuopio, Finland; [2] Dept. of Chemistry, Univ. of Jyväskylä, P.O.Box 35, FIN-40014 Jyväskylä, Finland; [3] Faculty of Science and Forestry, UEF, P.O.Box 111, FIN-80101 Joensuu, Finland.

UEF

* Corresponding author: reino.laatikainen@uef.fi

Quantitative Quantum Mechanical NMR Analysis: the Superior Tool for Analysis of Biofluids

Graphical abstract



Abstract:

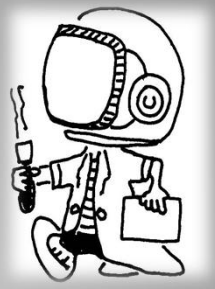
Almost automate quantitative analysis of biofluids is now behind a few clicks, from sample to EXCEL table after minimal sample preparation, without separations, calibration and reference materials, even for unknown compounds!

Each organic compound with protons gives a highly diagnostic and unique spectrum which is practically identical with any spectrometer operating at certain field. A distinctive feature of the 1D ^{1}H NMR spectra is that even the most complex spectrum of a compound can be described by a few spectral parameters within experimental accuracy, employing the quantum mechanical theory. The NMR spectral parameters offer also a very efficient way to store artefact free spectra in Adaptive Spectral Libraries (ASL), instead of variable quality experimental spectra. Once spectra have been measured and modelled in one magnetic field strength using Quantum Mechanical Spectral Analysis (QMSA), the spectra can be simulated in every detail in any other field and mixtures – to be used in quantification of the mixtures with ChemAdder software (see http://chemadder.com). The software is described and its application to analyses of serum, volatile fatty acids from biowaste and slaughterhouse waste are used as examples in our presentation.

Keywords: Metabolomics; Quantitative NMR; QMSA; ASL; ChemAdder

Introduction (Part 1)

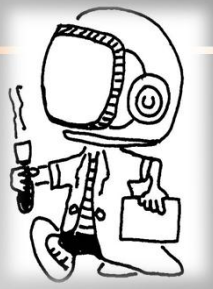
EXPERIMENTAL

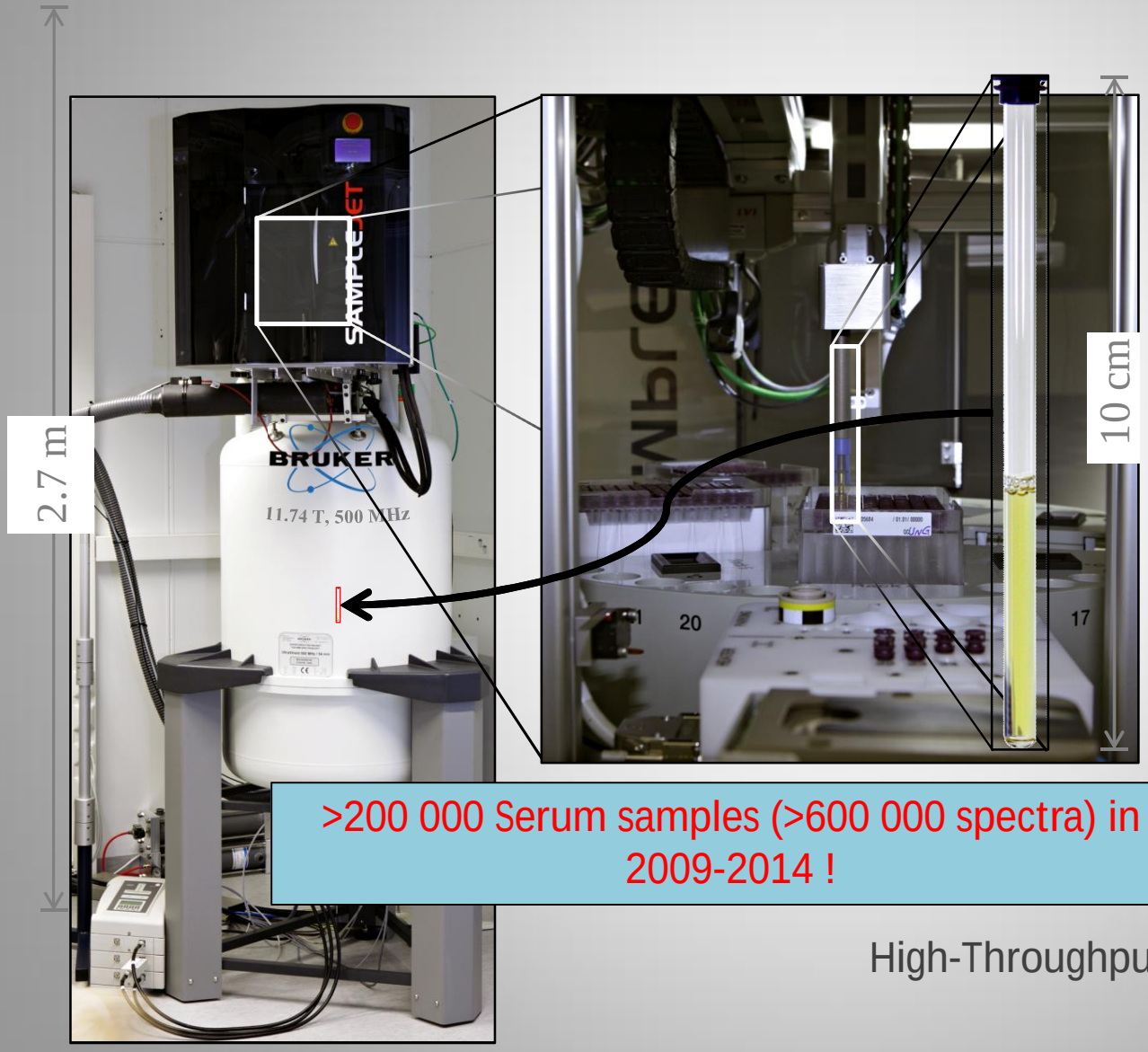


- The new NMR technology (automatic sample changer, autoshimming, autopreparing) allows almost automate measurement of 480 samples (> one weekend !) without break and operator !
- No baseline artefacts !
- No solvent suppression artefacts !
- No line-shape artefacts !
- Transfer to own computer of researcher ..to be analyzed using ChemAdder!
- < 20$€/sample (incl. amortization of instrument) !

NMR METABOLOMICS LABORATORY of UEF

High-throughput NMR metabolomics





- Sample into magnet
- Heat sample to +37ºC
- Tune & Homogenize magnetic field
- Measure data
- Analyze data
- Make conclusions

>200 000 Serum samples (>600 000 spectra) in 2009-2014 !

High-Throughput Serum NMR Metabolomics

Slice by Pasi Soininen

ABOWE project

Movable ABOWE Pilot biorefinery unit for industry wastes, in Poland for potato industry and restaurant biowaste, and in Sweden for slaughterhouse wastes. The unit was constructed in Savonia University of Applied Sciences, Kuopio, Finland, under supervision of Adjunct Professor Elias Hakalehto. Photo: Mika Ruotsalainen.

[1] den Boer E, Łukaszewska A, Kluczkiewicz W, Lewandowska D, King K, Reijonen T, Kuhmonen T, Suhonen A, Jääskeläinen A, Heitto A, Laatikainen R, Hakalehto E, Volatile fatty acids as an added value from biowaste, *Waste Management*, http://dx.doi.org/10.1016/j.wasman.2016.08.006.
[2] Schwede S, Thorin E, Lindmark J, Klintenberg P, Jääskeläinen A, Suhonen A, Laatikainen R, Hakalehto E, Using slaughterhouse waste in a biochemical based biorefinery -results from pilot scale tests. *Environmental Technology*, http://dx.doi.org/10.1080/09593330.2016.1225128

NMR Spectrum (600 MHz) of an ABOWE sample

Aliphatic Region: aliphatic acids are easily identified

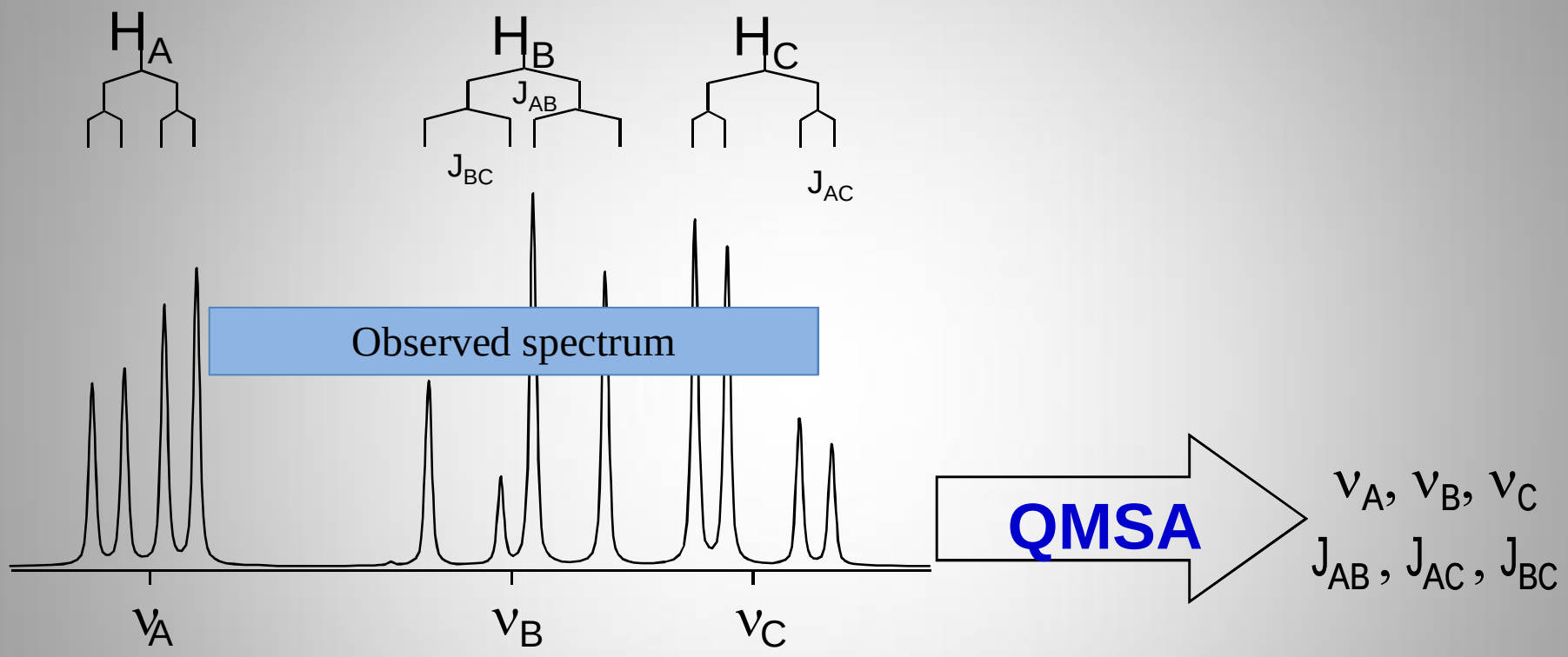
2,3-Butanediol (RS & RR) have a very unique signal but it often overlaps with valerate in HPLC

A structure (or a part of it) be also identified from splittings (coupling constants) of multiplets: the couplings do not depend on instrument or sample.

Principles of QMSA and qQMSA

- qQMSA: Tiainen M, Soininen P, Laatikainen R, Quantitative Quantum Mechanical Spectral Analysis (qQMSA) of ^{1}H NMR Spectra of Complex Mixtures and Biofluids, *J.Magn.Reson.*, 242, 67 (2014).
- A review: Laatikainen R, Tiainen M, Korhonen S-P, Niemitz, M, "Computerized Analysis of High-resolution Solution-state Spectra" in Encyclopedia of Magnetic Resonance, eds R. K. Harris and R. E. Wasylishen, John Wiley: Chichester. Published 15th December 2011. (DOI: 10.1002/9780470034590.emrstm1226).
- QMSA Iterator: Laatikainen R, Niemitz M, Weber U, Sundelin J, Hassinen T, and Vepsäläinen J, General Strategies for Total-Line-Shape Type Spectral Analysis of NMR Spectra Using Integral Transform Iterator, *J.Magn.Reson.* A120, 1-10 (1996).

Quantum Mechanical Spectral Analysis (QMSA)



Chemical shift (n) = weight point of multiplet
Coupling constant (J) ≈ difference of two lines => fine structure

THE PARAMETERS ARE INDEPENDENT OF INSTRUMENTATION ..The problem with signals in MS, GC and HPLC !!

Quantum Mechanical NMR Spectral Analysis: Math

NMR intensity spectrum $I(\nu)$ is sum of *spectra* of chemical components $S(\nu)$, *background* $B(\nu)$ & *noise*(ν)

$$I(\nu) = \Sigma x_n S_n(\nu) + B(\nu) + noise(\nu)$$

where ν is frequency.

Each spectrum $S(\nu)$ is a function of spectral parameters

$$S_n(\nu) = F_n(\nu, \underline{w}, \underline{J}, \underline{R}, \underline{\Delta}, Line\text{-}shape)$$

Where $\underline{w}$ = chemical shifts, $\underline{J}$ = coupling constants, $\underline{R}$ =response factors (≈1.0), $\underline{\Delta}$ = line-widths and line-shape.

Structure analysis: $I(\nu)$ => $\underline{w}$ & $\underline{J}$ => *structure*

Quantitative NMR: $I(\nu)$ => x_n *(populations)*

A non-linear mathematical inverse problem – solved iteratively !!

(F_n is a non-explicit function the values of which can be calculated and differentiated by using matrix formalism)

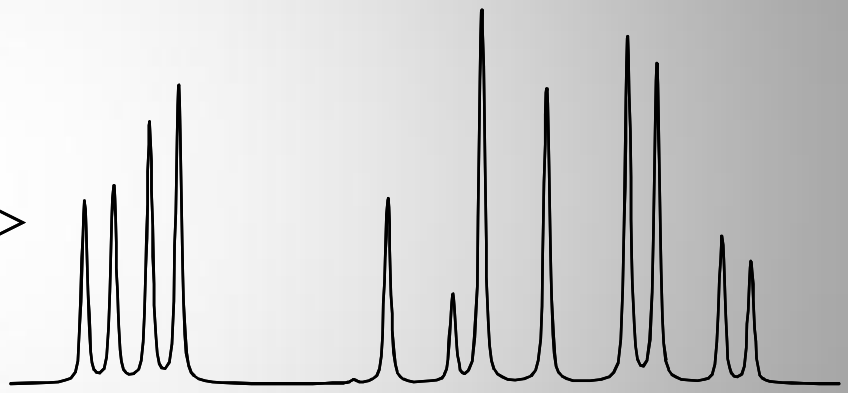
If chemical shifts, coupling constants & line-shape are given, spectrum can be simulated quantum mechanically !

Simulated spectrum

ν_A, ν_B, ν_C

J_{AB}, J_{AC}, J_{BC}

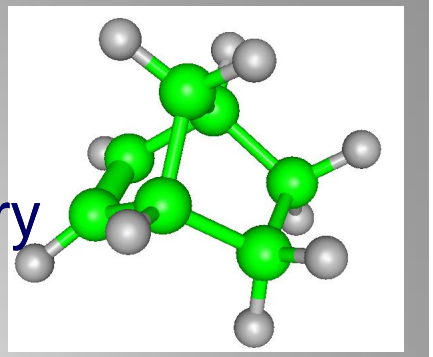
QMSA

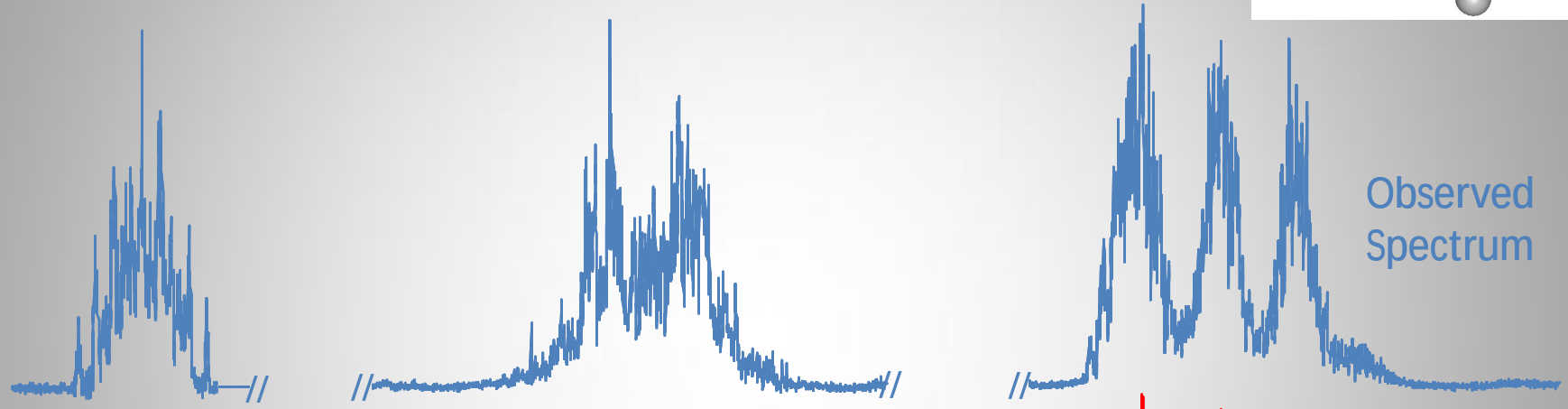


=> Model spectra for quantitative analysis - and ASL

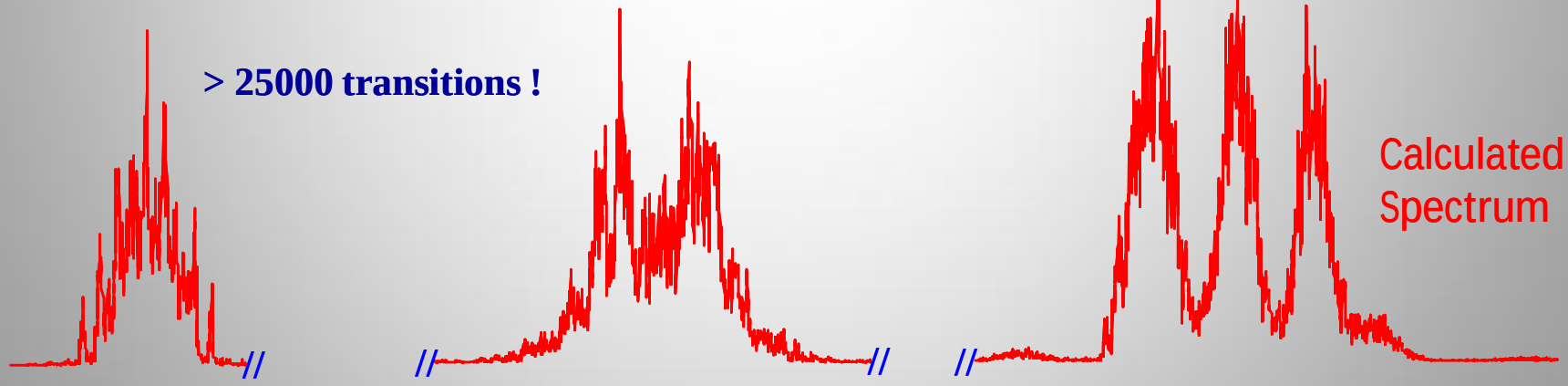
The chemical shifts depend slightly (0.001-0.05 ppm) on sample, but in qQMSA they can be recognized effectively from their coupling patterns ...this forms a problem in the (non-QM) methods based on experimental model spectra.

Even the most complex NMR spectra obey strict quantum mechanical rules and can be simulated in very details



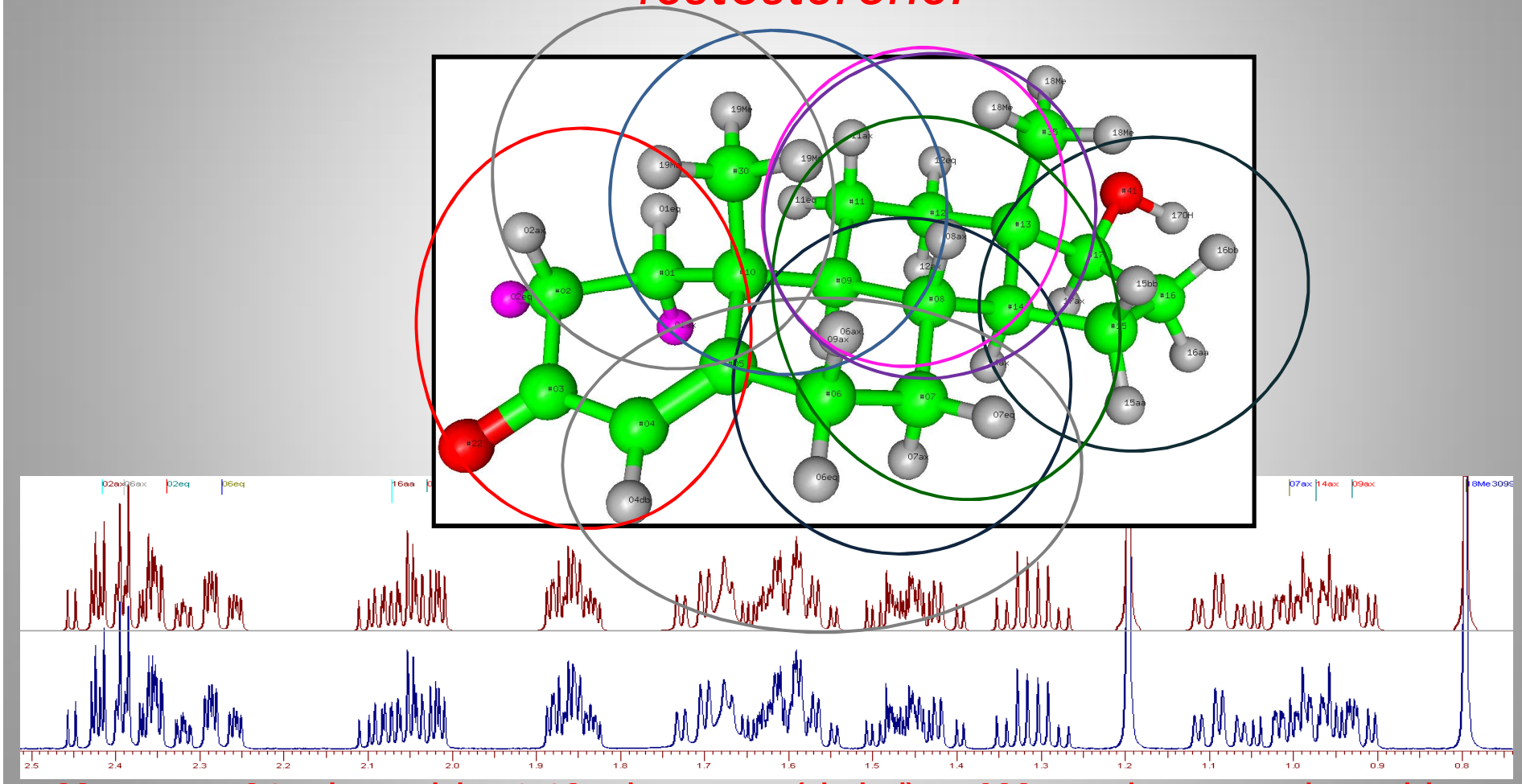


> 25000 transitions !



Large Spin-networks can be now simulated (by automate splitting into sub-systems)

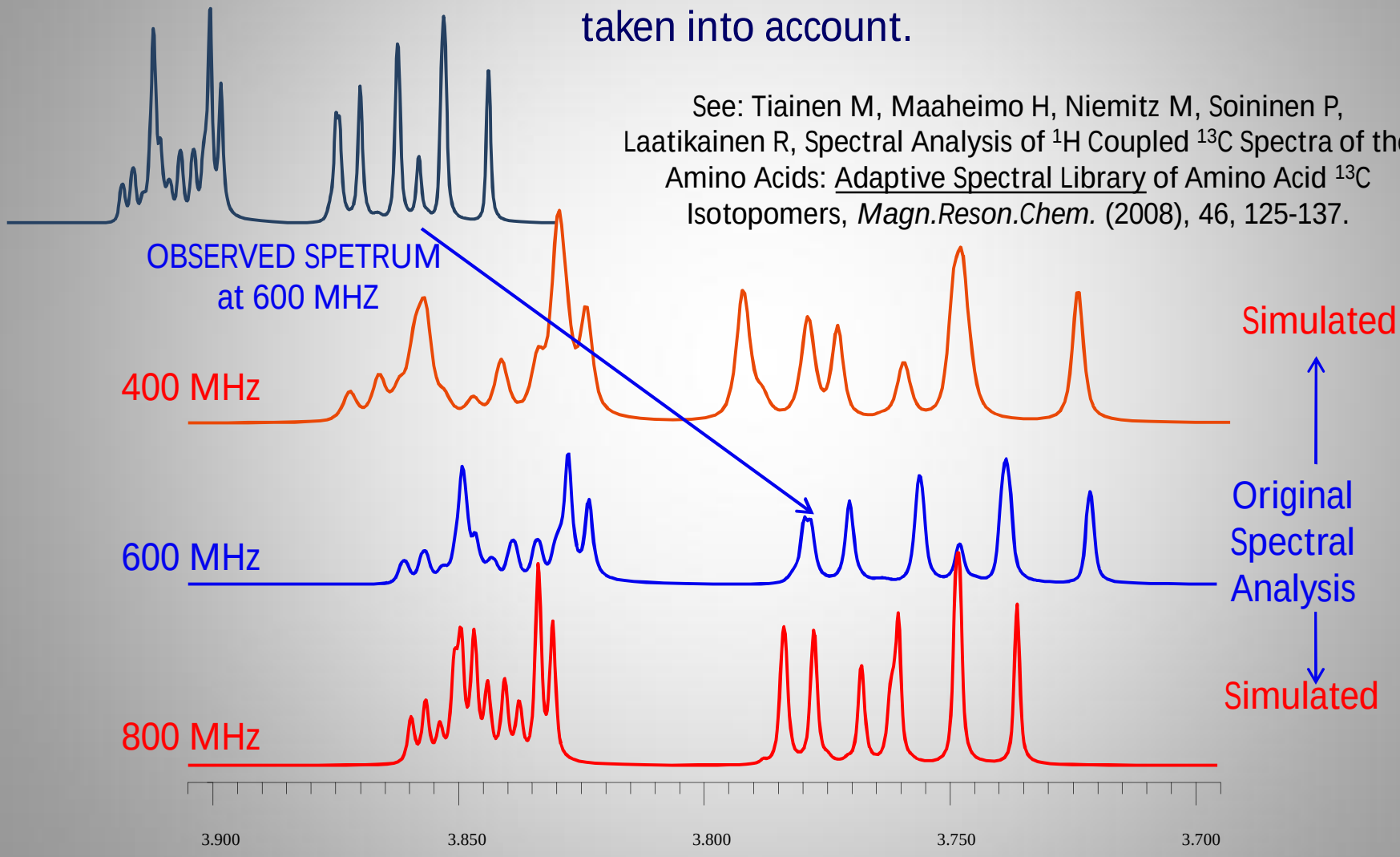
Testosterone:



28 protons, 24-spin particles & 13 sub-systems (circled) => 688 non-degenerated transitions, only ! Simulation time ca. 5 s.

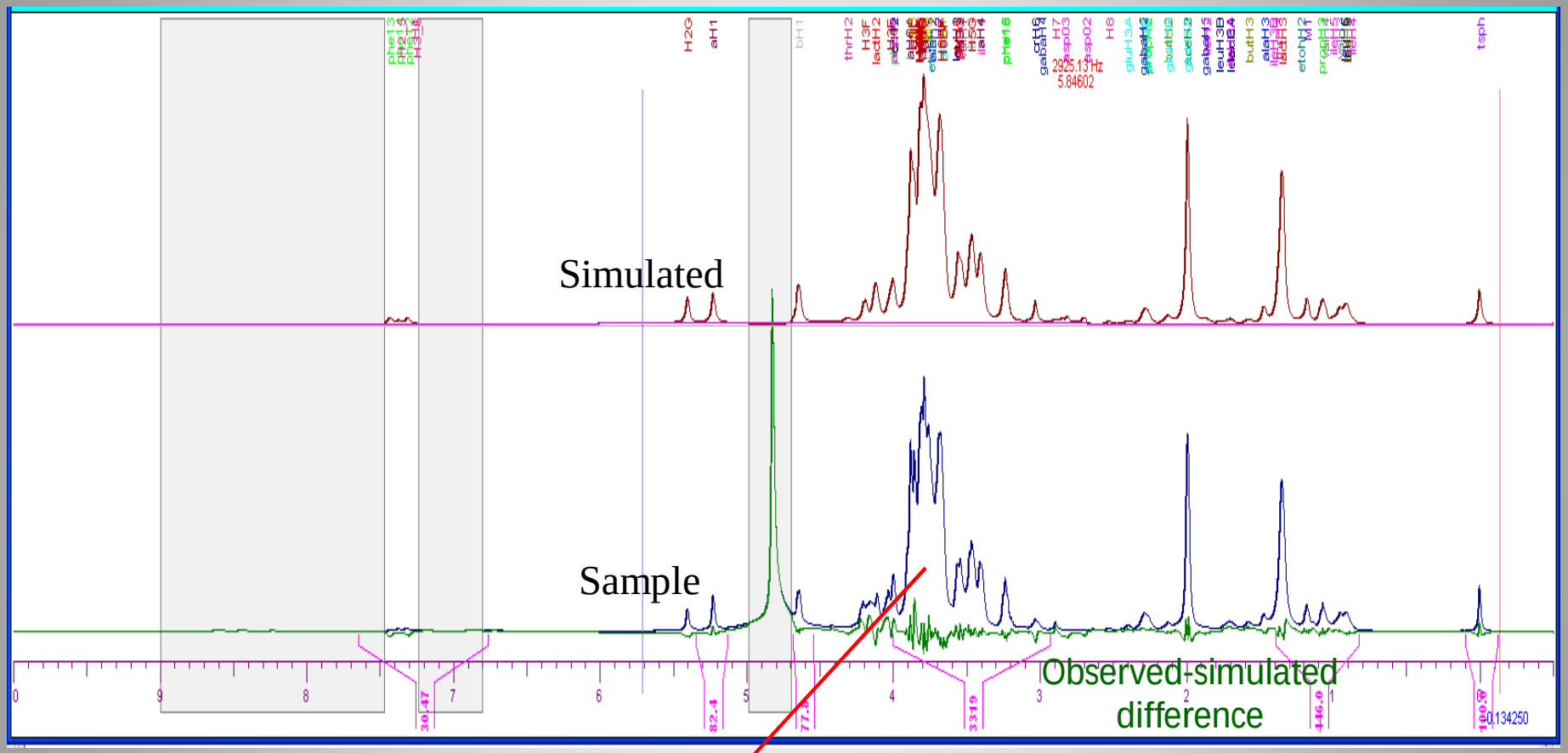
Adaptive Spectrum Libraries (ASL):

Analyze spectrum at one (magnetic) field, then the spectrum at any other field and line-shape can be then simulated ! Also variations in the chemical shifts can be taken into account.



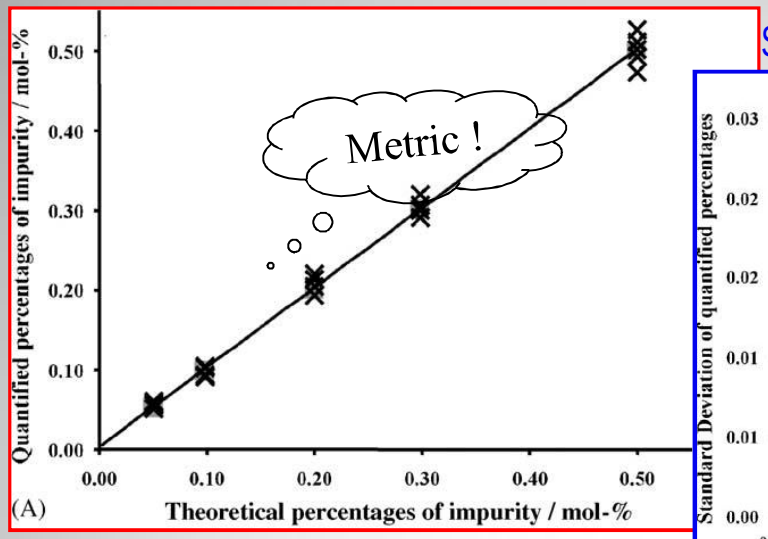
See: Tiainen M, Maaheimo H, Niemitz M, Soininen P, Laatikainen R, Spectral Analysis of ^{1}H Coupled ^{13}C Spectra of the Amino Acids: Adaptive Spectral Library of Amino Acid ^{13}C Isotopomers, *Magn.Reson.Chem.* (2008), 46, 125-137.

Quantitative QMSA of an ABOWE Swedish slaughterhouse sample using 23 metabolites:

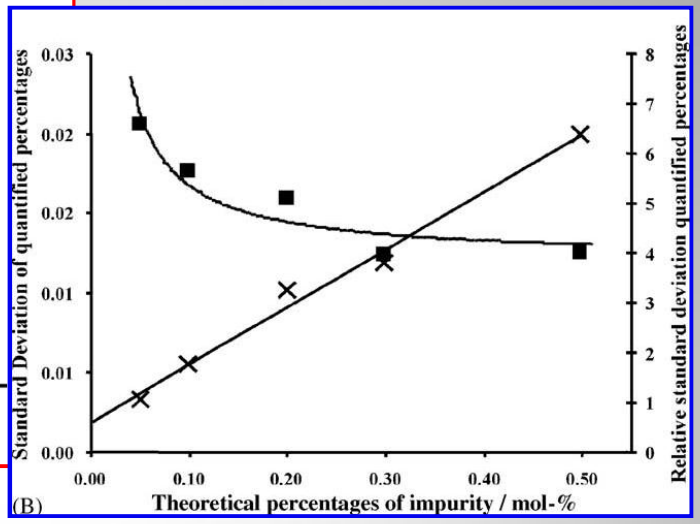


Sometimes spectral lines are broadened by Fe & Mn-ions, like above. It forms no problem for qQMSA - but how to manage it with the methods based on experimental model spectra !?

Linearity & accuracy



Standard deviation vs. mol% %



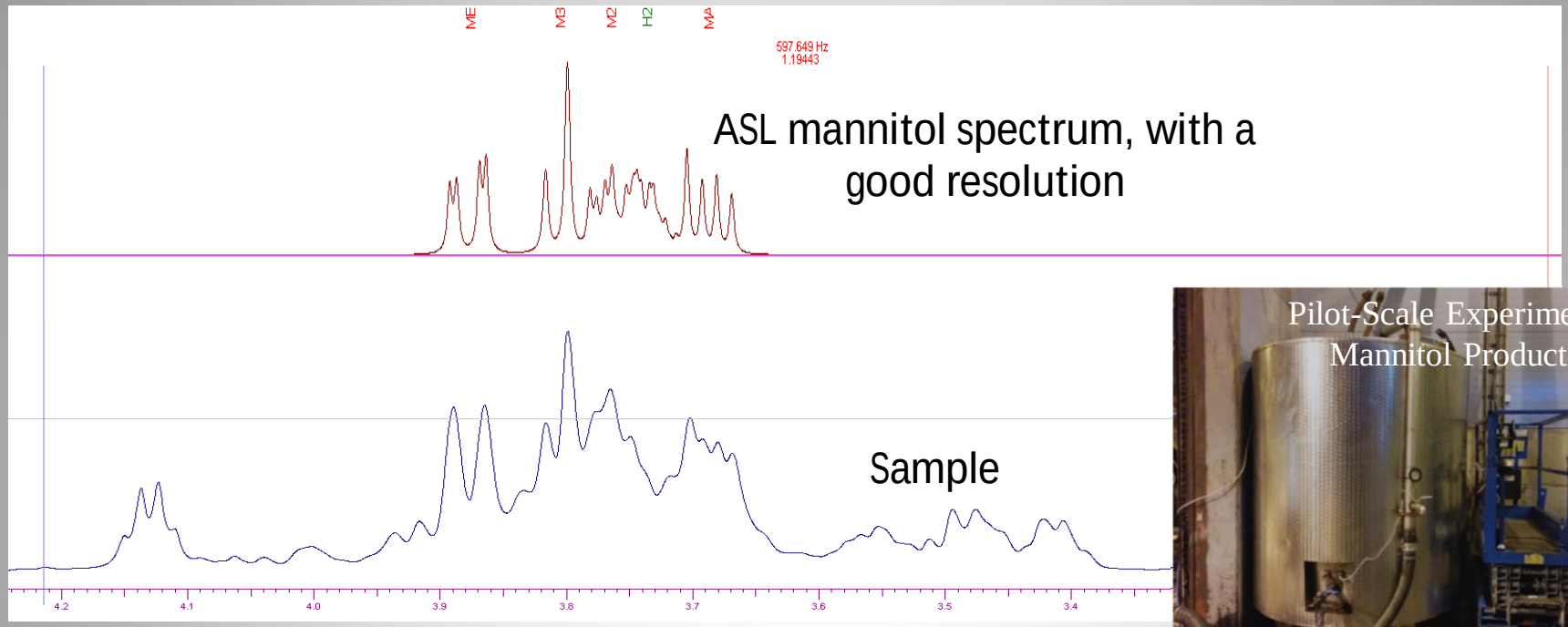
Calculated vs. real impurity concentrations
(in mol%) R^2 = 0.995 (from a drug impurity analysis[1])

NO CALIBRATION ...if systematic errors of < 5%, or 10% for T2 edited spectra, are tolerated - the bias may arise from water suppression and T2 editing but they can assumed to be constant for same type experiments and samples [2] !!

[1] Soininen P, Haarala J, Vepsäläinen J, Niemitz M, Laatikainen R, Strategies for Organic Impurity Quantification by ^{1}H NMR spectroscopy: Constrained Total-Line-Shape Fitting, *Anal.Chim.Acta* (2005) 542, 178-185.

[2] Tiainen M, Soininen P, Laatikainen R, Quantitative Quantum Mechanical Spectral Analysis (qQMSA) of ^{1}H NMR Spectra of Complex Mixtures and Biofluids, *J.Magn.Reson.*, 242, 67 (2014).

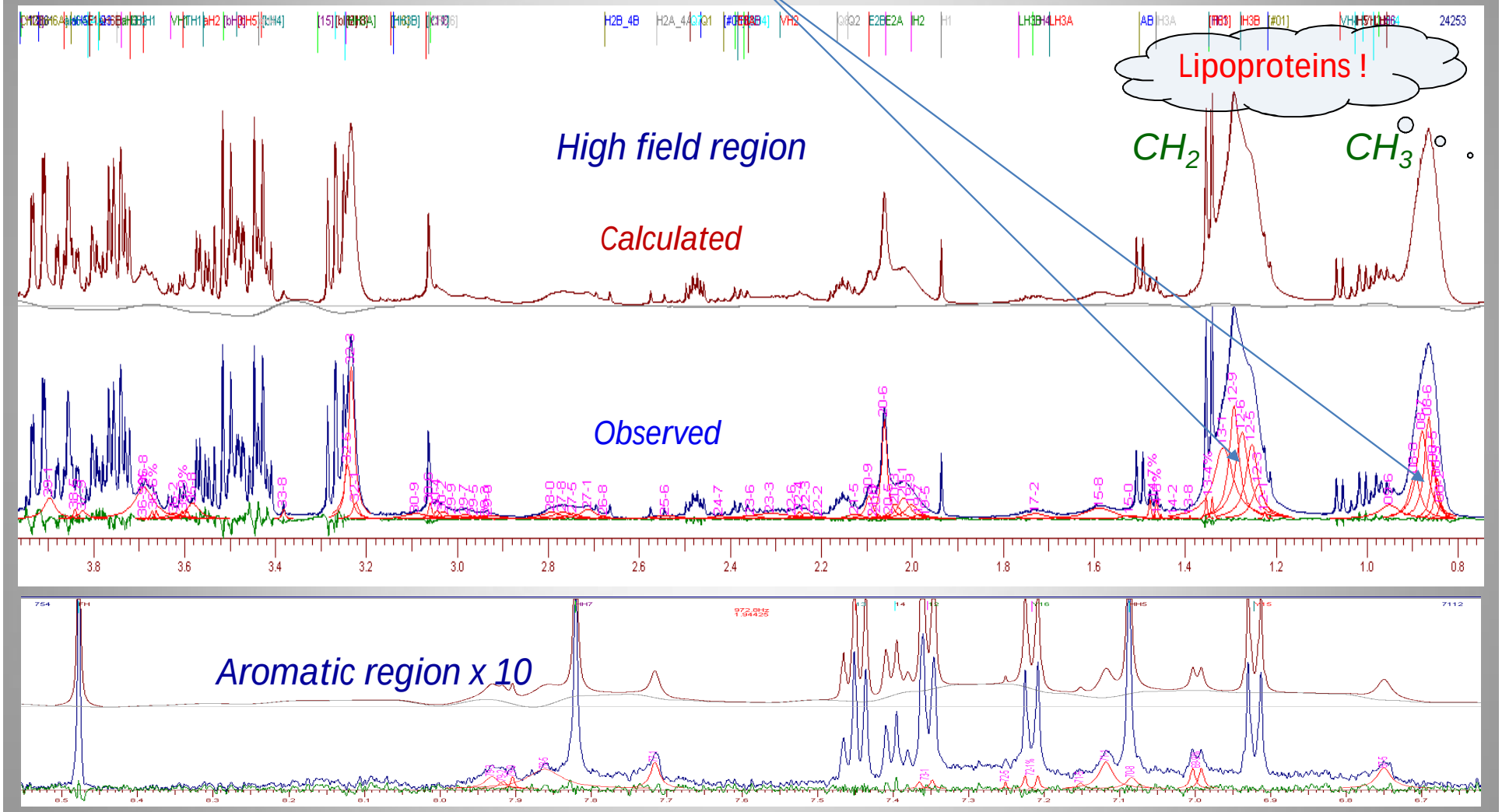
Detection of mannitol: spiking



Spiking is sometimes used to ensure identification of a component in complex samples or crowded range (like above).

Hakalehto E, Heitto A, Kivelä J, Laatikainen R, (2016). Meat industry hygiene, outlines of safety and material recycling of biotechnological means. In: Hakalehto E (Ed.). Microbiological Industrial Hygiene. Nova Science Publishers. New York, USA. Pp 249-270.

T2 edited spectra of serum: the lipoprotein signals are described by *structures* !



One can use also prior knowledge (= information that can be written for SpinAdder in form of linear equations between the model parameters).

qQMSA of SERUM

- T2 edited spectra – with H_2O suppression : minor but tolerable artefacts for some compounds, response factors (R, should be ≈ 1.0) vary for signals and compounds => systematic errors usually < 10%.
- Macromolecules (proteins removed) – only H_2O suppression: systematic errors of < 5%, standard deviations in concentrations < 2% - recommended !

Tiainen M, Soininen P, Laatikainen R, Quantitative Quantum Mechanical Spectral Analysis (qQMSA) of ^{1}H NMR Spectra of Complex Mixtures and Biofluids, *J.Magn.Reson.*, 2014, 242, 67-78.

Software: ChemAdder & SpinAdder (Part 3)

- ChemAdder: the multispectra user interface for QMSA and qQMSA.
- SpinAdder: the new generation of QM spin engine.
- 2-5 min/spectrum (in multispectral mode).

Features

- Novel Qt technology & graphics and support (C++).
- Spectral input in JDX-format.
- Large spin systems (no limit met, yet, ..except long $-(CH_2)_n$- systems).
- Smart chemical shift permutator for complex spectra.
- Graphics & data: No limitation in number of spectra treated simultaneously.
- Fast essential metabolite search from ASL's using FZZY tool: takes advantage from the multispectral data.
- Up to 100 or more (?) metabolites.
- Targeted ASL (*Adaptive Spectral Library*): metabolite libraries for each sample type – any field – any line-shape.
- Output in TXT or CSV (EXCEL) format, in mg/ml or, *also for unknown compounds*, in mmol/ml.
- Wizarded protocols.
- Tailored protocols (*MENUs*) and default profiles for sample types.
- Maximal information using combination of QM spectra, structures and prior knowledge (= any information that can be written for iterator in form of linear equation).
- Tools for preparation of ASL spectra even from very poor spectra (with bad baseline and solvent suppression artefacts) or even from peak lists.
- QMSA oriented tools for examination of 1D and 2D spectra.

ChemAdder user interface:

Simultaneous analysis of a set of spectra: extra information for metabolite search from ASL's !

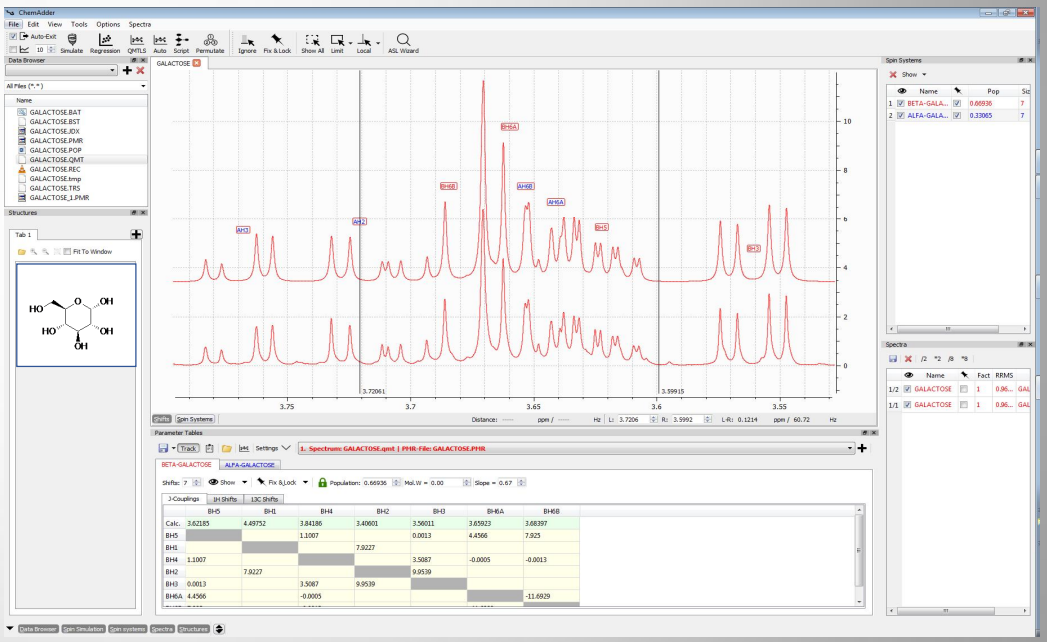
REPORT (in TXT or CSV format)

&QM	NAME	N	PROTONS	POPULATION	MOL%	mMOL	Weight(mg/ml)
%Q	**lactate**	**1**	**4**	**0.9004E+01**	**9.0918**	**86.3675**	**7.7731**
%Q	**acetate**	**2**	**3**	**0.1680E+02**	**16.9621**	**161.1321**	**9.6679**
%Q	ala	3	4	0.6009E+01	6.0681	57.6442	5.1303
%Q	valine	4	8	0.2383E+01	2.4059	22.8546	2.5140
%Q	leu	5	10	0.2505E+01	2.5291	24.0256	3.1474
%Q	ile	6	10	0.1930E+01	1.9491	18.5155	2.4255
%Q	**etoh**	**7**	**5**	**0.7797E+01**	**7.8735**	**74.7941**	**3.4405**
%Q	butyrate	8	7	0.7870E+00	0.7947	7.5494	0.6643
%Q	propio	9	5	0.1874E+01	1.8922	17.9750	1.3661
%Q	glu	10	5	0.9932E-02	0.0100	0.0953	0.0141
%Q	**b-gluc**	**11**	**7**	**0.8427E+01**	**8.5091**	**80.8322**	**14.5498**
%Q	**a-gluc**	**12**	**7**	**0.5282E+01**	**5.3338**	**50.6687**	**9.1204**
%Q	gly	13	2	0.1311E+01	1.3236	12.5734	0.9430
%Q	thr	14	5	0.9269E+00	0.9360	8.8912	1.0581
%Q	phe	15	8	0.1850E+01	1.8678	17.7434	2.6793
%Q	3pheprop	16	9	0.2500E+00	0.2524	2.3980	0.3597
%Q	creatine	17	5	0.1144E+01	1.1552	10.9735	0.9766
%Q	gaba	18	6	0.9932E-02	0.0100	0.0953	0.0098
%Q	asp	19	3	0.1473E+01	1.4870	14.1258	1.8787
%Q	**mannitol**	**20**	**8**	**0.9783E+01**	**9.8790**	**93.8454**	**17.0799**
%Q	**23bud**	**21**	**8**	**0.1504E+02**	**15.1904**	**144.3013**	**12.9871**
%Q	**sucrose**	**22**	**14**	**0.4436E+01**	**4.4791**	**42.5497**	**16.0838**
%Q	tsp	23	9	0.9685E+00	0.9779	9.2900	1.3573
TOTAL(excl. reference) =				**99.0315**	**100.0000**	**949.9510**	**113.8694**

SMART PERMUTATOR and preparation of ASLs

GALACTOSE (α and β forms) gives a very crowded spectrum of 5 protons within 0.1 ppm (60 Hz): the prediction of the chemical shifts is impossible with such an accuracy and the 2D spectra are almost useless because of strong couplings which also make the multiplets of individual protons unrecognizable and very sensitive to the shifts !

Solution is the smart shift permutator, which goes through the combinations of shift order by a try-and-learn algorithm:



- ChemAdder & SpinAdder allow analysis of very poor spectra (with bad baseline, solvent suppression artefacts and impurities) or even from peak lists.

SpinAdder: the new generation QMSA engine: non-QM signals can be described by **structures**

Even the smallest details of spectrum can be described by the combination of QM spectra and the *structures* !

Analysis of a time series of spectra from a bioreactor experiment, as guided by ChemAdder WIZARD

SPECTRA (N>>1)
JDX-format

Targeted ASL
Metabolites (M)

FZZY SEARCH ENGINE

WIZARD

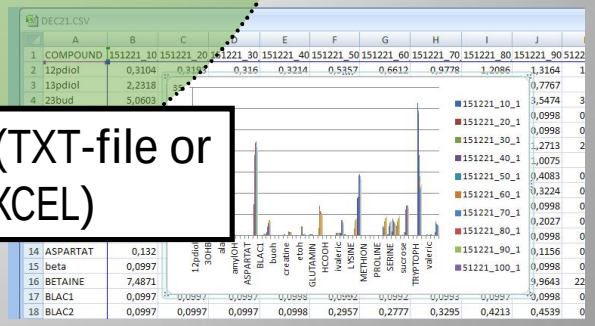
SELECTED metabolites
ASL files & average spectrum

ChemAdder

```
SIMULATE 5000 1000
BROADENIN
REGRESSIO
QMTLS 5000 1000
REM BROADENING
QMTLS 5000 1000
#LINEWIDTH OPTIMIZED = F
QMTLS 5000 1000
```

MENU (script)

SpinAdder (qQMSA)



CONCENTRATIONS (TXT-file or
CSV-file for EXCEL)

Conclusions (Part 4)

- Up to 100 metabolites in one sample?
- Dynamic range of 0.01-100 mol% (<0.01 mol% for drug impurities)
- qQMSA applications:
  - Any mixture and impurity analysis.
  - Biofluids: serum, CSF, lipid extracts of serum, urine, ...
  - Bioextracts, juices, ...
  - HSQC slices: ^{13}C Isotopomers in metabolic flux-analysis (to be published).

WHY qQMSA is SUPERIOR

Numerous protocols have been proposed for analysis of mixture NMR spectra, as based on (A) model experimental spectra of pure compounds, (TLS = Total-Line-Shape) presenting the signals as group of spectral lines or (BIN) binning.

Why qQMSA is superior:

- Signals can be very complex (see above norbornene and 2,3-butadiol) - not easily described in the TLS methods and they depend on field.
- Line shapes and widths may vary greatly.
- ASL spectra are free of artefacts and impurity signals !
- Overlay problem: chemical shifts (multiplet positions) vary 0.001 – 0.05 ppm (1 – 30 Hz at 600 MHz), which is 5 - 30 x line-width. Binning destroys information (for example, CH_3 signals of butyrate and valerate are amalgamated, see page 9). - A challenge also for the qQMSA iterators: solved in SpinAdder by FZZY algorithm !
- QMSA yields also couplings – chemical confidence or even identification of a new compound or recognition of known structural moiety.

CONS & PROS

+ Sample preparation, ..just filter and add buffer.
+ No calibration (or only once) - no pure reference compounds.
+ Instrument (field) independence.
+ Semiquantitative analysis of sample at one glance.
+ Chemical confidence (identification of components directly from spectrum), ..also carbohydrates (a problem with MS).
+ ASL (Adaptive Spectrum Libraries): NO EXPERIMENTAL ARTEFACTS !
+ Almost automate analysis - from sample to EXCEL.

- Not one-line.
- Some expertise needed.
- Not very sensitive, sample size > 0.3 ml.
- Expensive instrumentation, liquid helium and nitrogen - demands high number of samples to be economically feasible.
- Availability of ASL's.

RECOMMENDATIONS

- NMR is invaluable in checking composition and detecting metabolites (especially sugars) of fermentation products, whenever starting materials, conditions or protocols are modified.
- NMR suits to calibration of the one-line methods like GC and HPLC; not necessary to prepare the calibration samples containing accurate concentrations - which may not be available in pure form and in measurable amounts. (In inert solvents and with sufficient relaxation delays the response factors are very close to 1.0.)
- Remove high molecular weight components, instead of T2-editing.
- 1D TOCSY useful in recognition of unknown spin-systems.

ChemAdder Project

- University of Eastern Finland (UEF): Prof. Reino Laatikainen

  Univ. of Jyväskylä: Pekka Laatikainen & Henri Martonen

- For more about the project and software

See http://chemadder.com