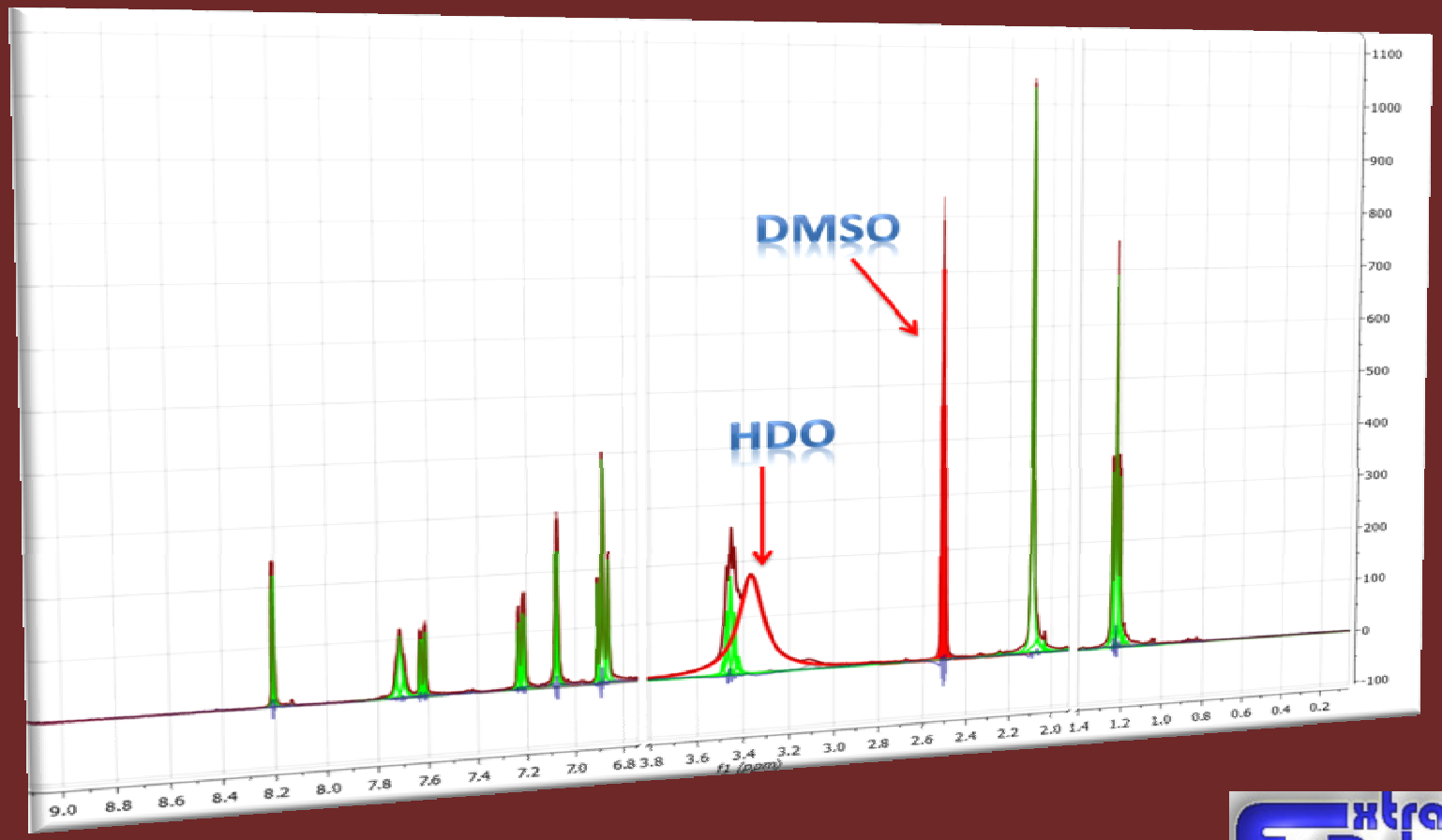
Automatic Solvent Recognition wizard:
example of an AI at work

Stan Sykora, Carlos Cobas,
Felipe Seoane, Pablo Monje, Esther Vaz, et al



MESTRELAB RESEARCH
Chemistry Software Solutions

www.mestrelab.com

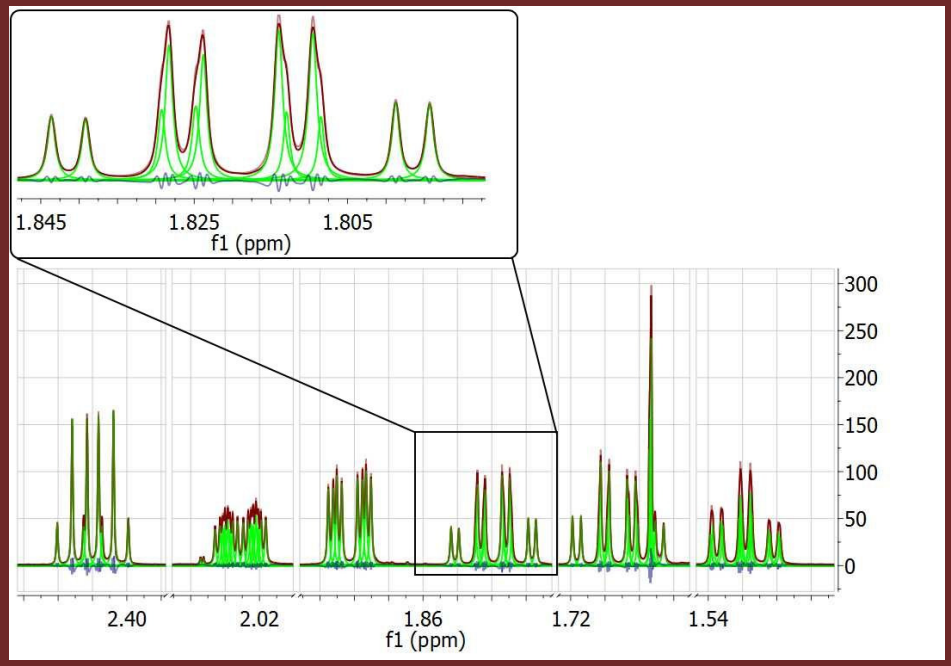
SMASH 2010, Portland, OR, USA
DOI: 10.3247/SL3nmr10.007

www.ebyte.it

GSD – Global Spectral Deconvolution

f-domain algorithm which automatically decomposes sets of superposed *Lorentzians* and *near-Lorentzians* and ends up in a **GSD Peaks List**

GSD is a standard feature of Mnova,
subject to continuous developmnet



GSD has been born in 2008
and presented in detail at several meetings in 2009 and 2010

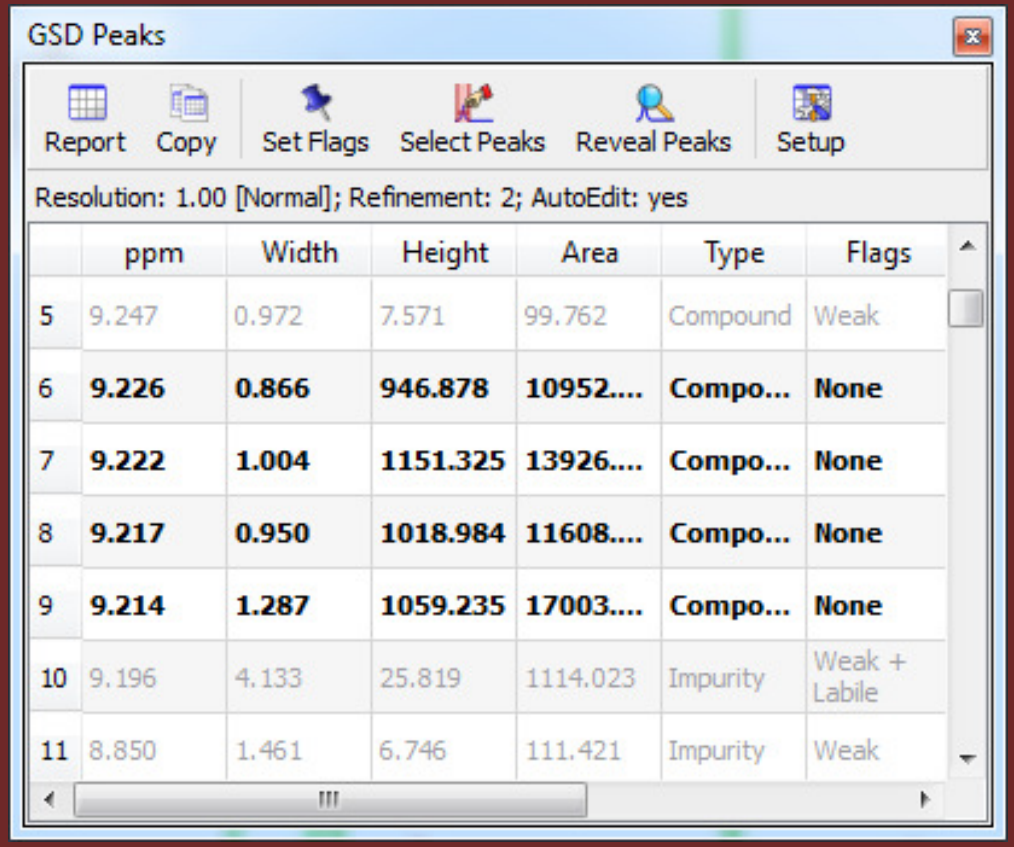
MESTRELAB RESEARCH
Chemistry Software Solutions

GSD Peaks Editing

Peaks editing brings in

specific NMR know-how

by **classifying** of the peaks in an GSD List

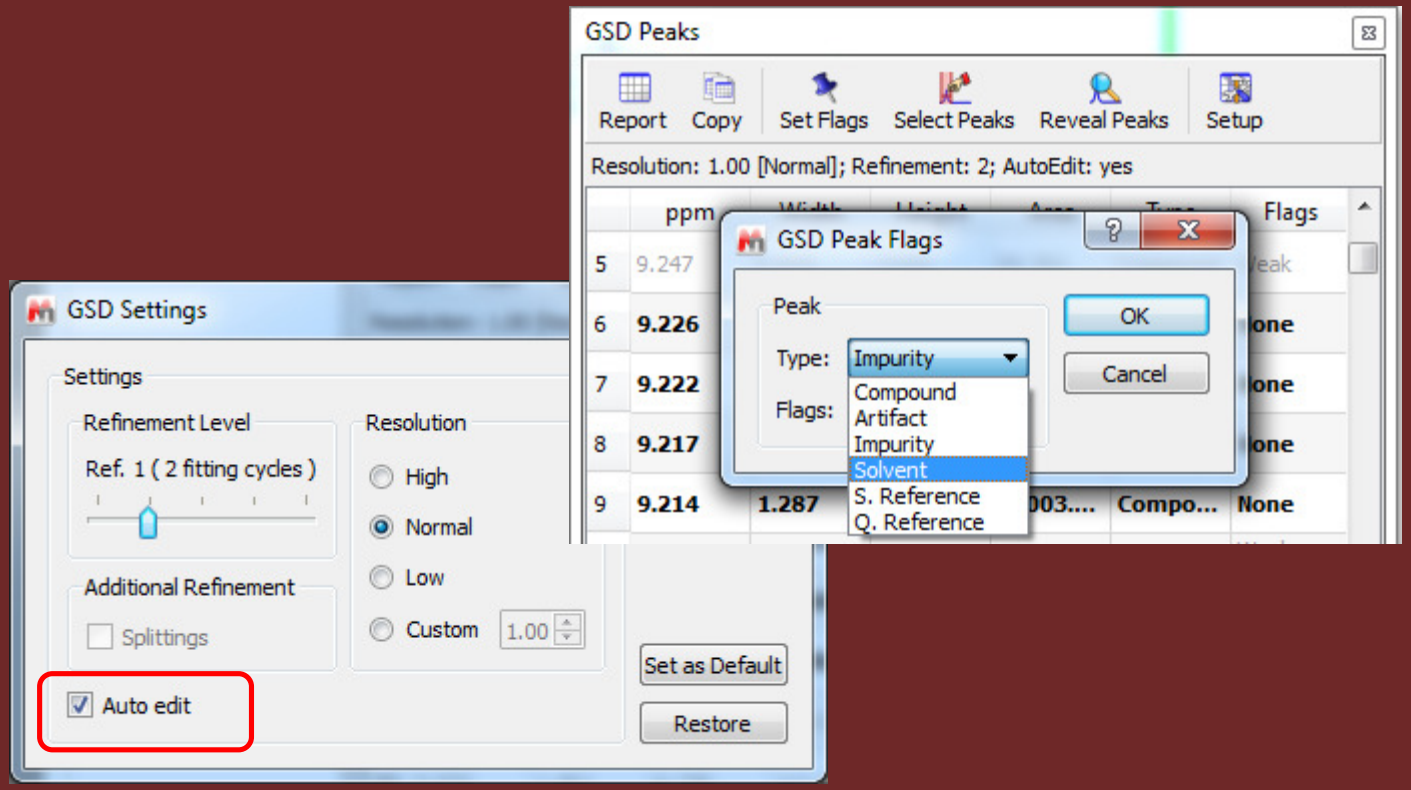


	ppm	Width	Height	Area	Type	Flags
5	9.247	0.972	7.571	99.762	Compound	Weak
6	**9.226**	**0.866**	**946.878**	**10952....**	**Compo...**	**None**
7	**9.222**	**1.004**	**1151.325**	**13926....**	**Compo...**	**None**
8	**9.217**	**0.950**	**1018.984**	**11608....**	**Compo...**	**None**
9	**9.214**	**1.287**	**1059.235**	**17003....**	**Compo...**	**None**
10	9.196	4.133	25.819	1114.023	Impurity	Weak + Labile
11	8.850	1.461	6.746	111.421	Impurity	Weak

MESTRELAB RESEARCH
Chemistry Software Solutions

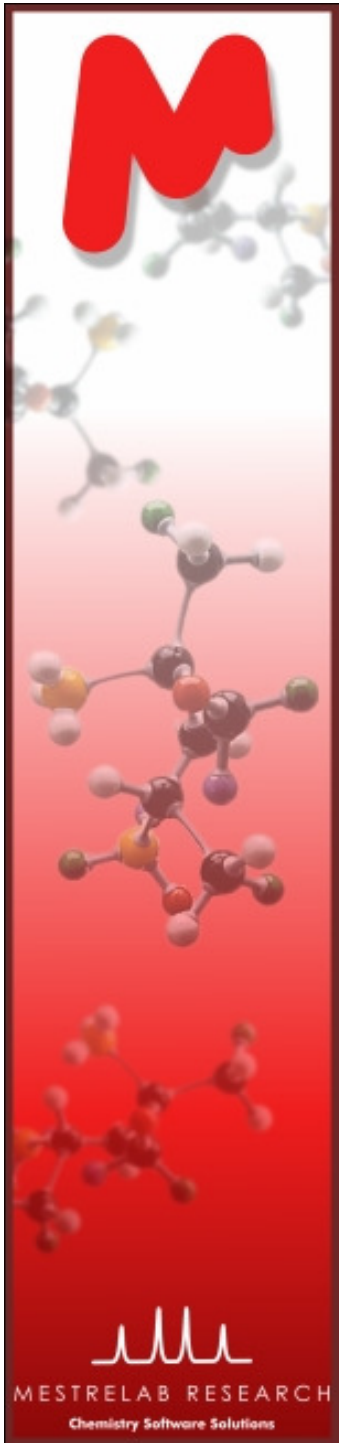
GSD Peaks Editing Modes

- ✓ Manual: Always accessible
- ✓ Automatic: An optional GSD feature always used in ASV



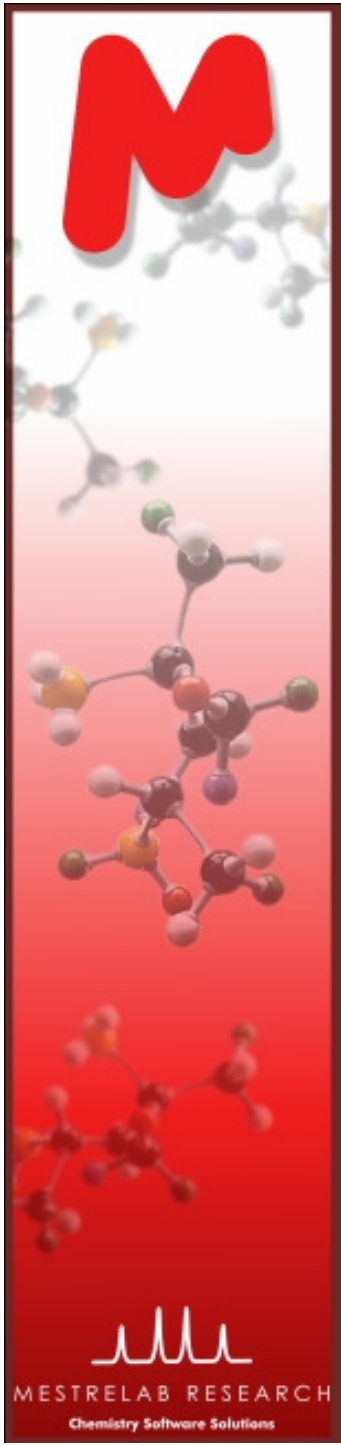
AutoEdit does not exclude subsequent manual peaks editing
(transparently followed by PostEdtit)

MESTRELAB RESEARCH
Chemistry Software Solutions



The GSD Auto-Edit algorithm: what is it up to

- ✓ Reference peaks recognition
- ✓ ^{13}C satellites recognition (where possible)
- ✓ Solvent peaks recognition (primary and secondary)
- ✓ Labile peaks recognition
- ✓ J-coupling multiplets recognition
- ✓ Impurity peaks recognition
- ✓ Weak peaks labeling



GSD Auto Editing results

Peak Types:

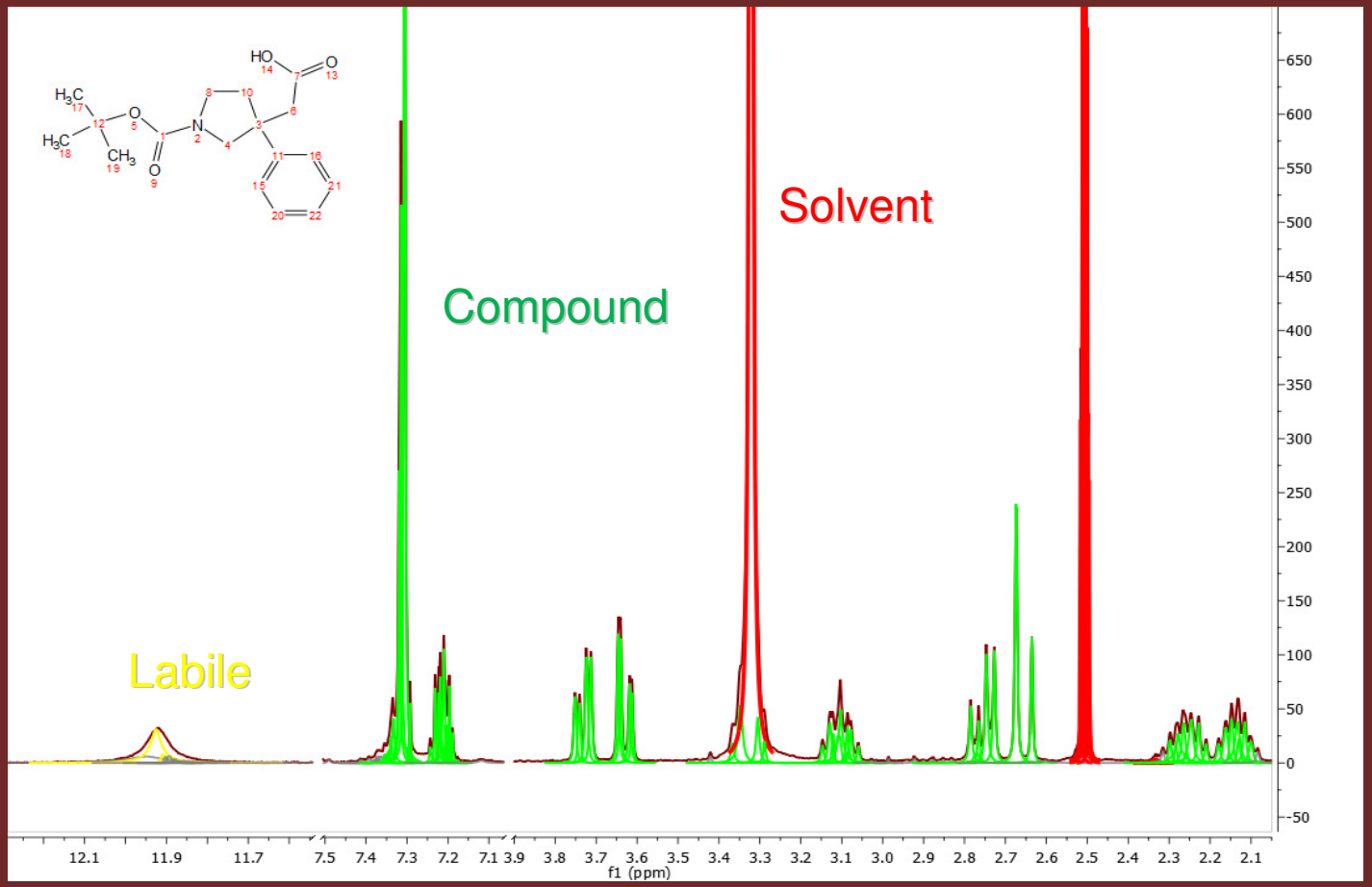
pt_Compound,
pt_Artifact,
pt_Impurity,
pt_Solvent,
pt_SReference,
pt_QReference

Peak Flags:

pf_SetByUser
pf_Hidden
pf_Weak
pf_C13Satellite
pf_Rotational
pf_Labile
pf_HasC13Sats

What can one do with the Edited Peaks List

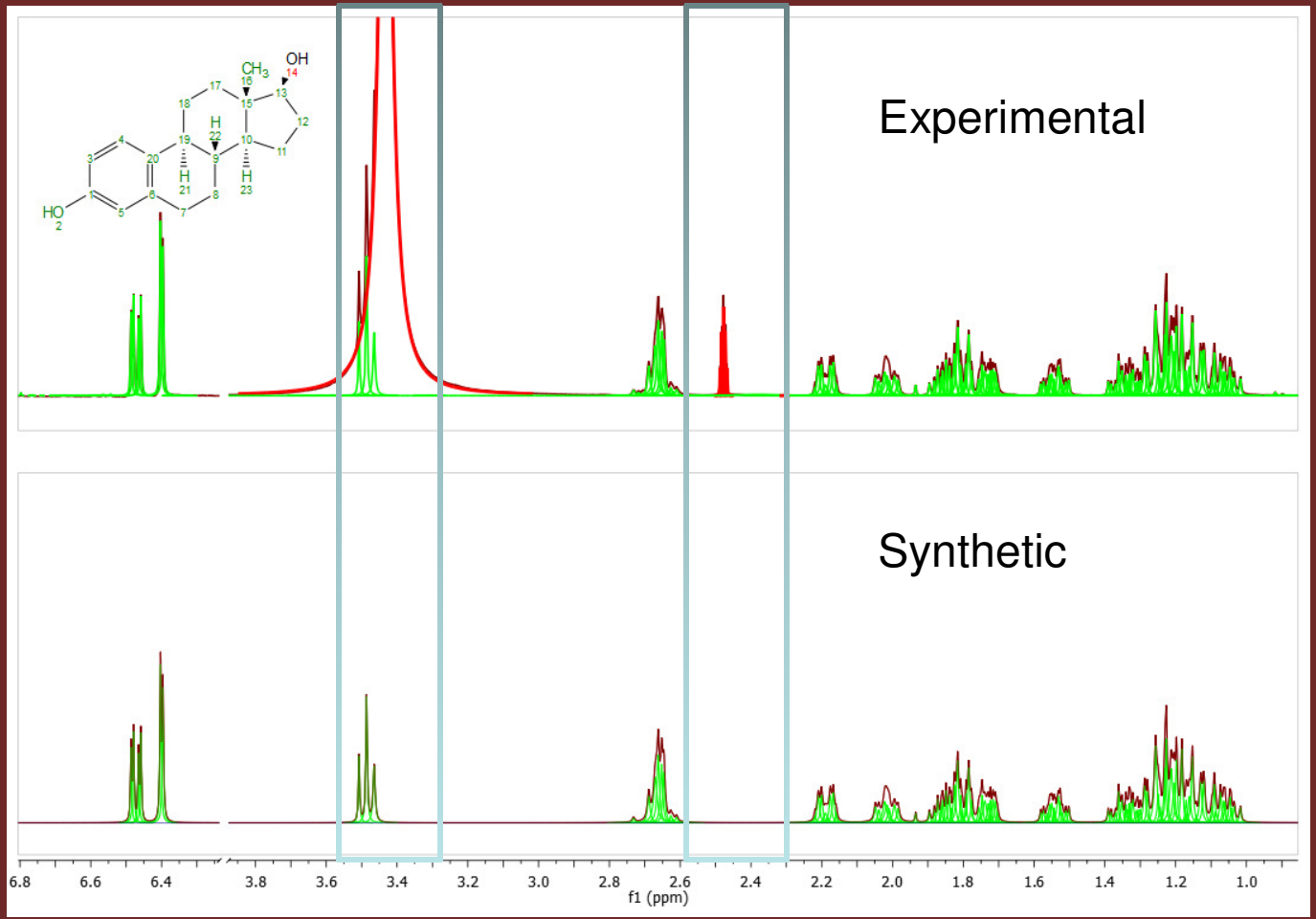
a) Color-labeled peak plots of the edited peaks



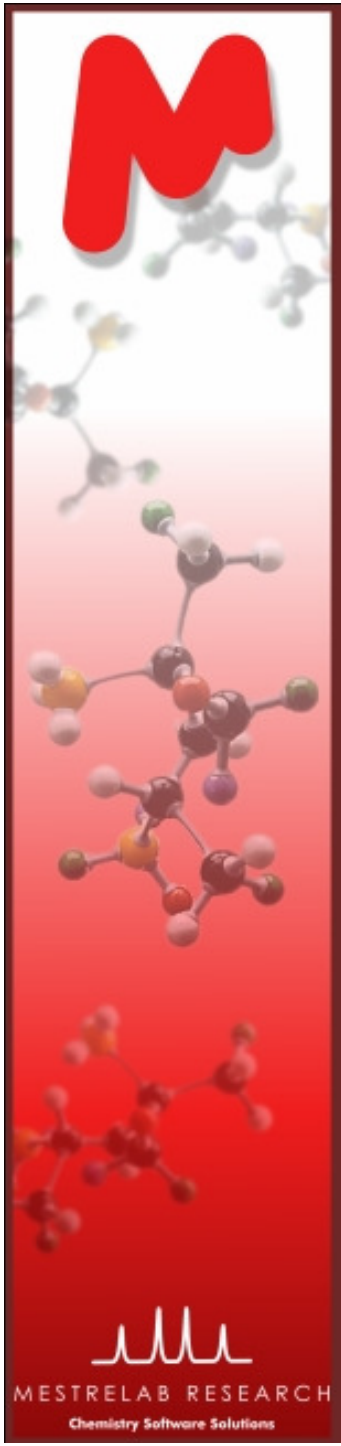
MESTRELAB RESEARCH
Chemistry Software Solutions

What can one do with the Edited Peaks List

b) Synthetic spectra of any desired category of peaks



MESTRELAB RESEARCH
Chemistry Software Solutions

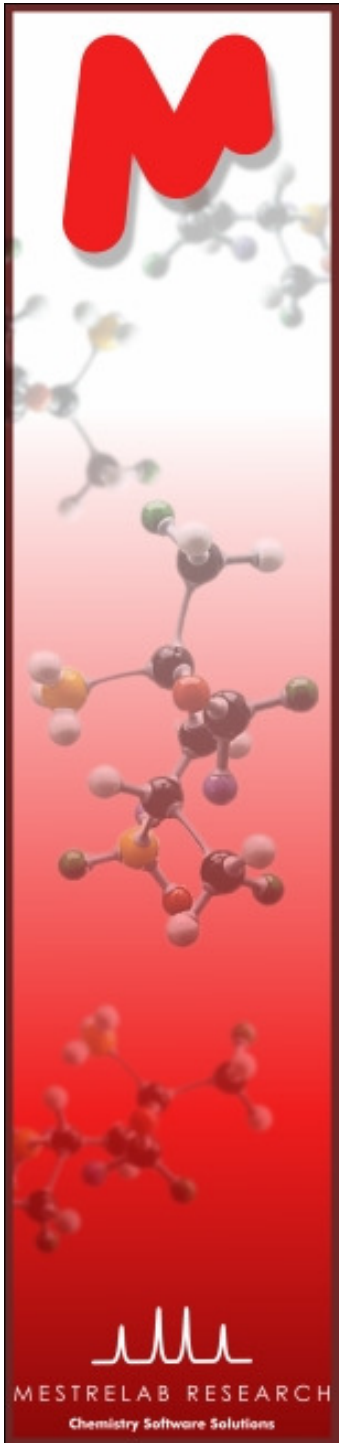


What can one do with the Edited Peaks List ?

c) Good editing is an essential pre-requisite of …

… ASV !

Automatic Structure Verification



Solvent recognition

Possibly the most important and tricky part of **AutoEdit**

Human solvent recognition:

Based on a few imperfect rules and lots of intuition

Automatic solvent recognition:

Based on a long list of imperfect scoring tests

Primary and secondary solvent signals:

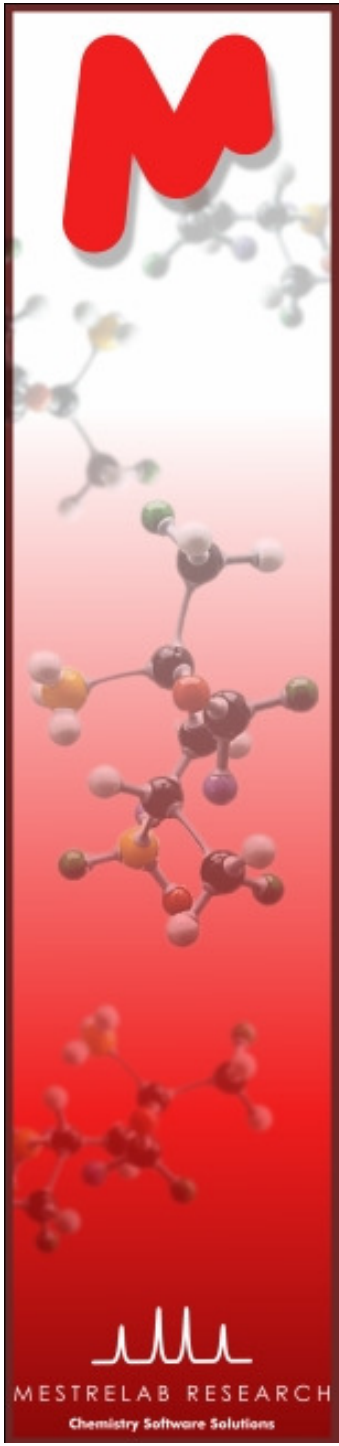
Primary solvent recognition is relatively easy

Secondary solvent signals (water) often present problems

Basic development rule:

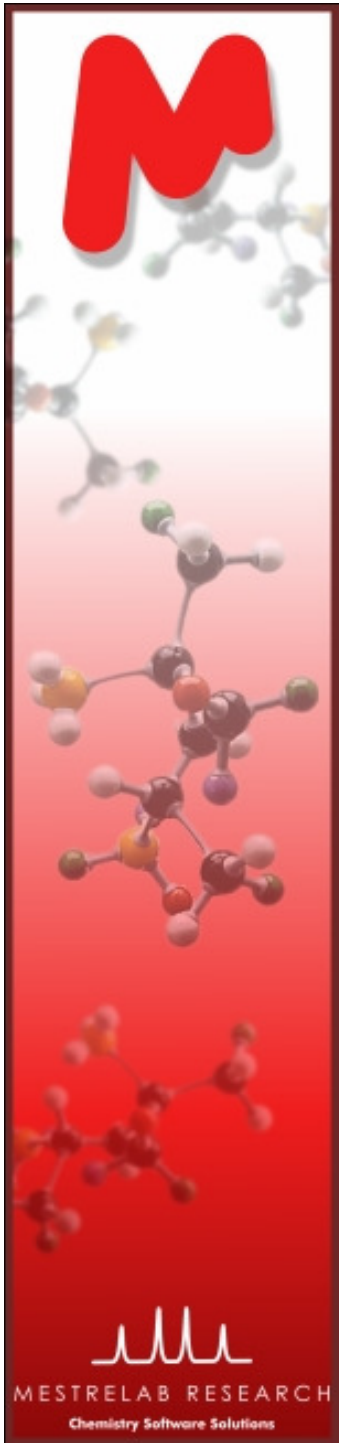
How the hell does Tony know that this peak belongs to solvent !!!

(if you ask him, he does not know)



Solvent description

```
DWORD SdFlags;      // See the FFP_Flags enum in EbPeaks.h
                    // Main Multiplet:
REAL  Shift;              // typical main multiplet shift value
REAL  ShiftLoPpm;         // low limit of main multiplet in ppm
REAL  ShiftHiPpm;         // high limit of main multiplet in ppm
DWORD Npeaks;             // number of main peaks (may not be 0)
REAL  JHz;                // coupling constant
REAL  Pweights[5];        // array of relative weights of the peaks
                    // Satellite Peaks:
REAL  SatWeight;          // satelites relative weight (0.011 or 0)
REAL  JSatHz;             // satelites coupling constant
REAL  SatShiftPpm;        // satelites relative isotope shift
                    // Associated Multiplet:
DWORD Npeaks2;            // number of peaks (may bee 0)
REAL  Weight2;            // relative weight (typical)
REAL  ShiftPpm2;          // relative isotope shift
REAL  JHz2;               // coupling constant
REAL  Pweights2[5];       // array of relative weights of the peaks
```



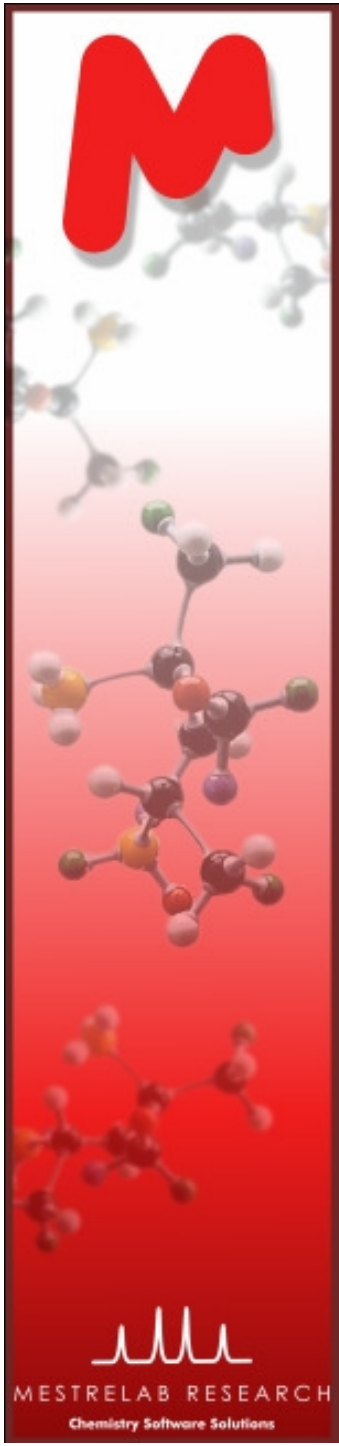
Solvent descriptors

Typically, a solvent entry in Mnova contains two Solvent Descriptors:

- One for the **primary solvent** peaks and
- One for the **secondary solvent** peaks (usually water)

However, it could contain any number of them (even a typical impurity might be nominally considered part of the solvent)

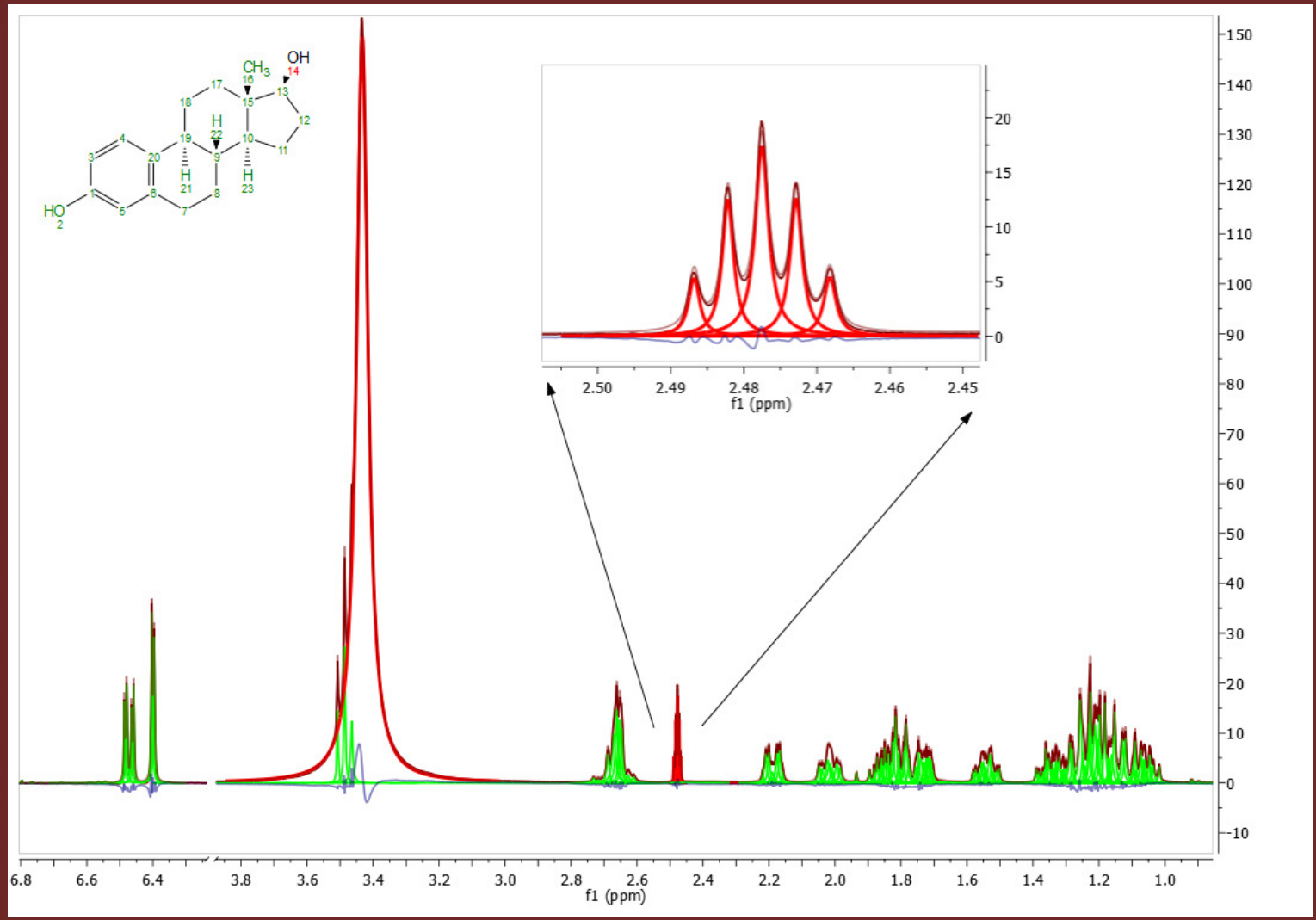
Coming soon:
Possibility to easily edit and add solvent description XML files.



Solvent recognition algorithm

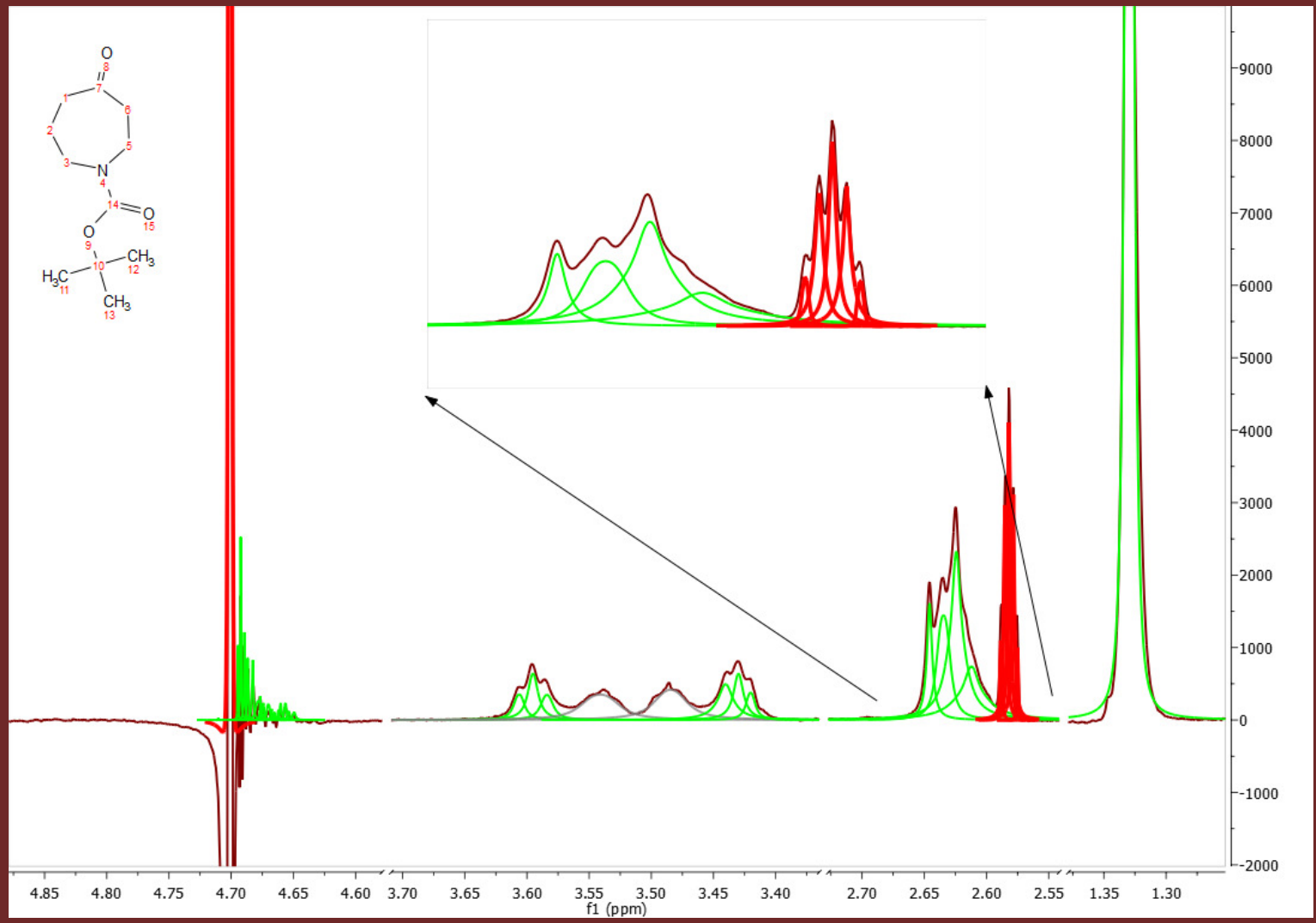
- ✓ Look for the central peak of a solvent multiplet
- ✓ Scan all peaks in the admissible range
- ✓ Associate a full-fledged scoring system with each peak
- ✓ Scan on all of these criteria (if pertinent):
  - Position
  - Width
  - Height
  - Kurtosis
  - Presence of J-coupling multiplet peaks
    - J-HD coupling constant value
    - Relative intensities
  - Presence of 13C satellite peaks
    - Relative intensites
    - J-CH coupling value
    - J-CH isotope shift
  - Presence of secondary J-coupling multiplet
    - Relative intensities
    - J-HD coupling in secondary multiplet

Some examples

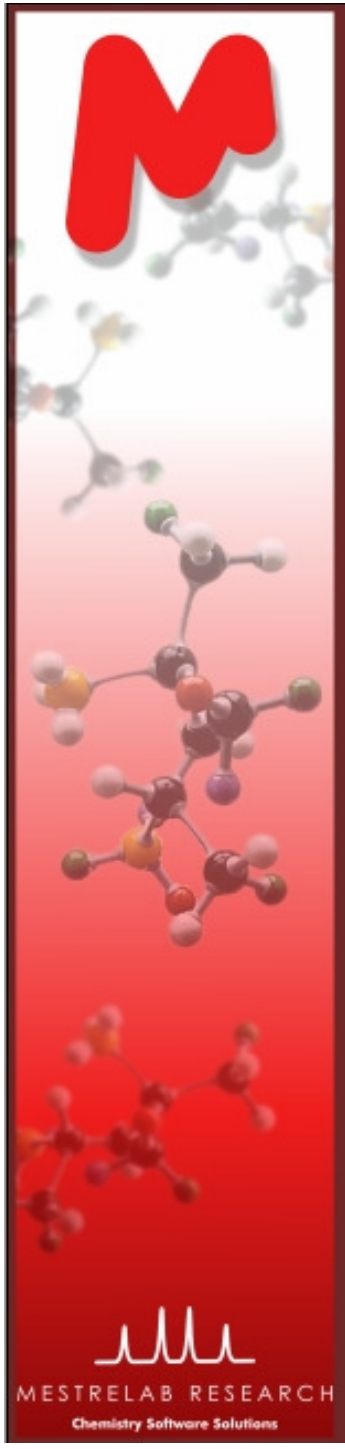


MESTRELAB RESEARCH
Chemistry Software Solutions

Some examples



MESTRELAB RESEARCH
Chemistry Software Solutions



Thank You for your Patience

Any Questions ?

SMASH 2010, Portland, OR, USA
DOI: 10.3247/SL3nmr10.007