Automatic Structure Verification: current status

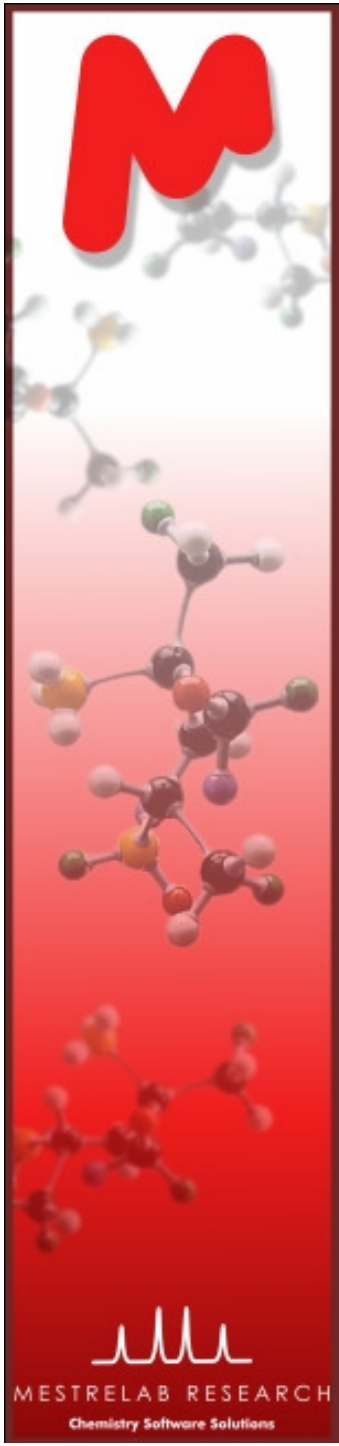
Stan Sykora, Carlos Cobas, Felipe Seoane, et al

SMASH 2010, Portland, OR, USA
DOI: 10.3247/SL3nmr10.008

MESTRELAB RESEARCH
Chemistry Software Solutions

www.mestrelab.com

www.ebyte.it



What is ASV? Part I

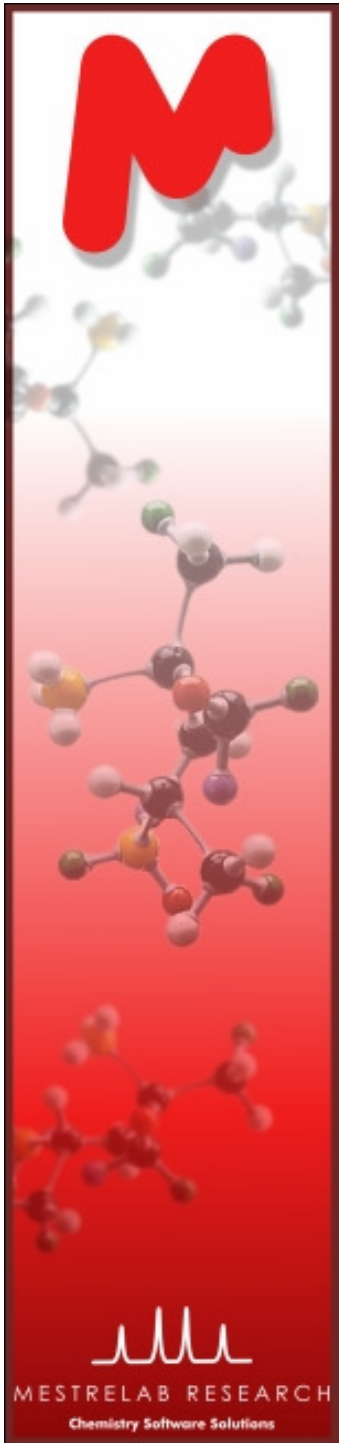
Molecule versus experimental spectra (1D, 2D, MS, …):
Questions to ask: *Might it fit? Is it definitely ruled out?*

Automatic (ASV) versus
Computer Assisted (CASV) structure verification
Necessity/possibility to cover both aspects

The NMR branch of Mnova ASV is based on
NMR predictions and
General rules of the NMR trade (quantitative aspects)

Practical issues to live with:
Real-life spectra: complete with
solvent peaks, artifact peaks, quantitation errors
Real-life predictions of parameters:
Error bounds, inversions, almost totally uncertain labiles

Target: Small and medium-sized molecules with
well-defined peaks & multiplets



What is ASV? Part II

First of all, it is a **scoring system** (a math concept)
based on a library of **ASV tests**.

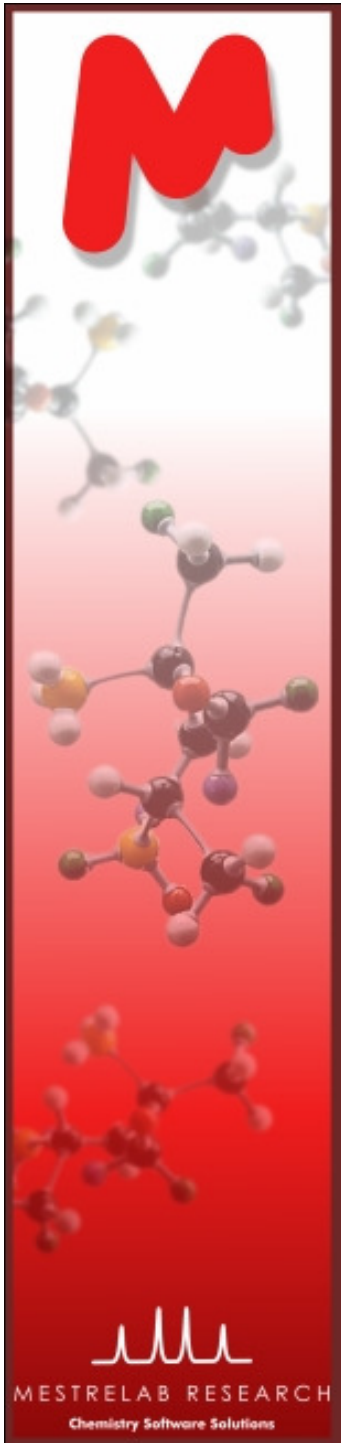
It is organized into **ASV tasks**,
each task comprising a number of ASV tests.

User Interface covers:
Simple tests (one molecule against one spectrum) and
various batch modes and scripts

It allows to combine 1D and 2D spectra (HSQC)

It also allows to comine NMR and MS data

Coming soon:
User scripting of custom ASV tasks



Exploitation of 1D spectra in Mnova ASV

Pre-requisites: GSD with Auto-Edit

- Solvent recognition
- Labile peaks recognition
- Impurities recognition
- Efficient J-coupling multiplets recognition

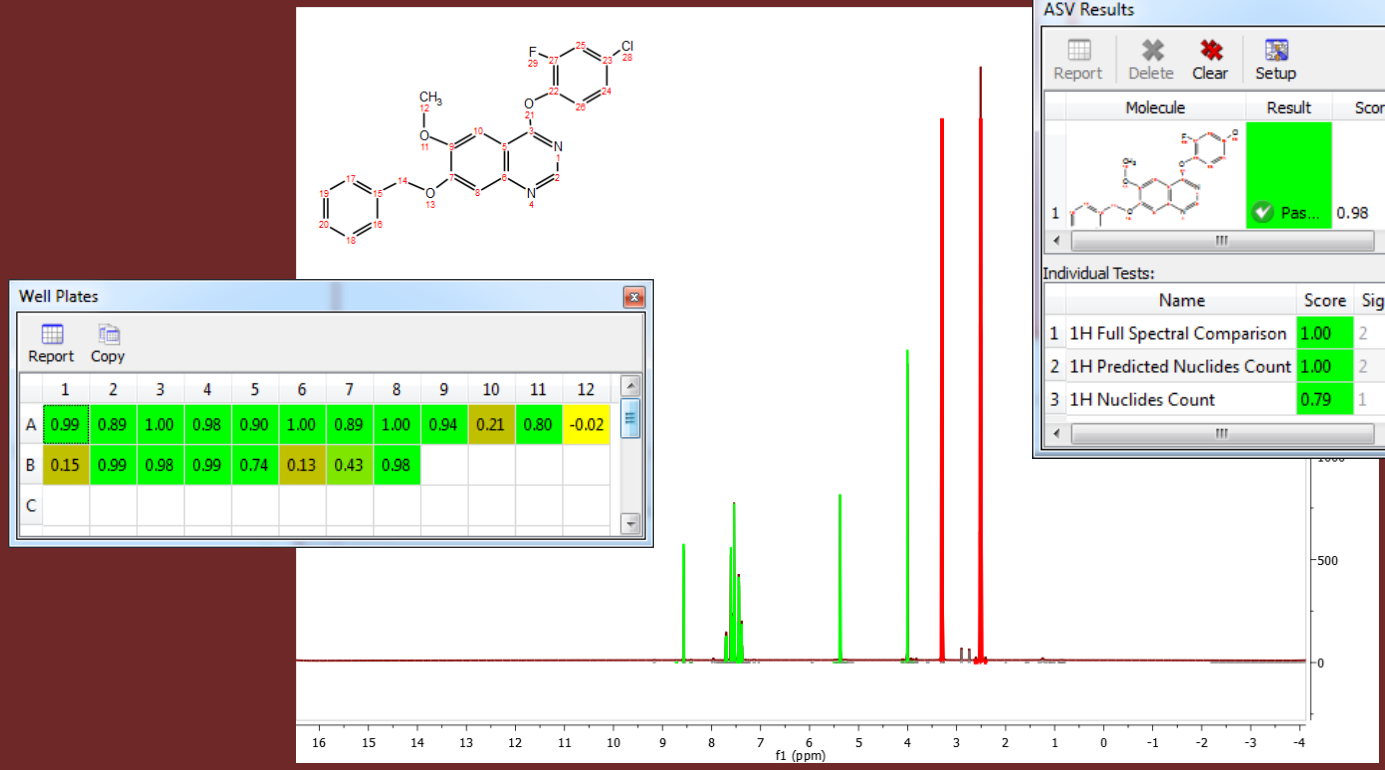
Currently implemented tools for ASV tests:

- Novel spectrum similarity metric
- Bayesian Number of Nuclides assessments
- Similarity metric extension for prediction error bounds

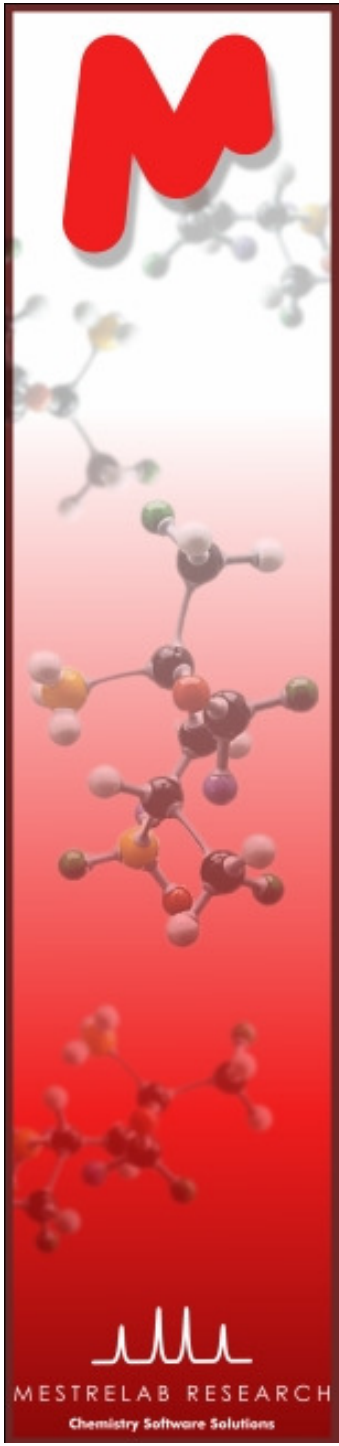
Present Status

The system is fully operative and already quite good

We now go as far as one can go
short of complete assignments
of nuclei to spectral multiplets



MESTRELAB RESEARCH
Chemistry Software Solutions



Perspectives

ASV has been so far rarely attempted
and it never reached maturity for widespread use

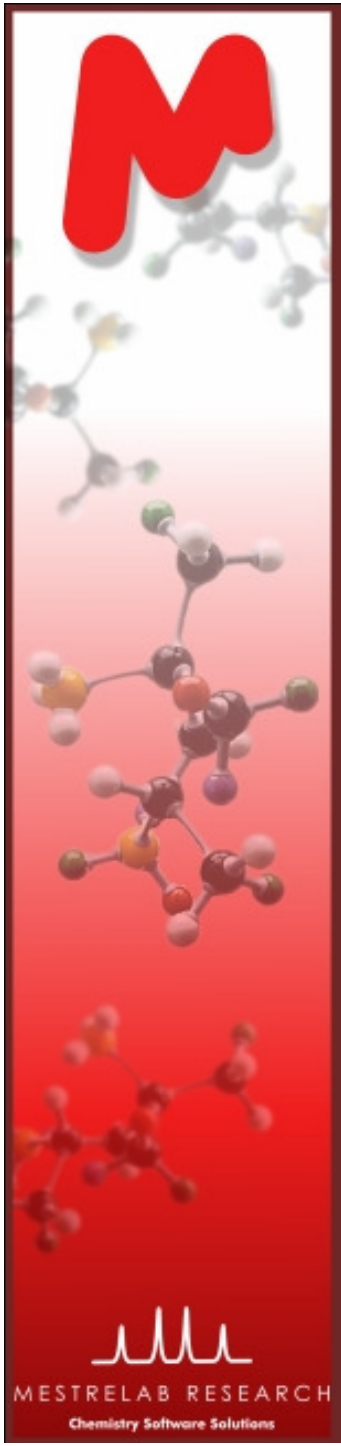
Its development will probably continue for decades:
it is a new category of NMR software

Planned short-term steps:

Ongoing improvements due to
better recognition of solvent and labile peaks

Enumeration of all possible assignments
and scoring on all of them.

This will involve the J-structure of the multiplets



Invitation

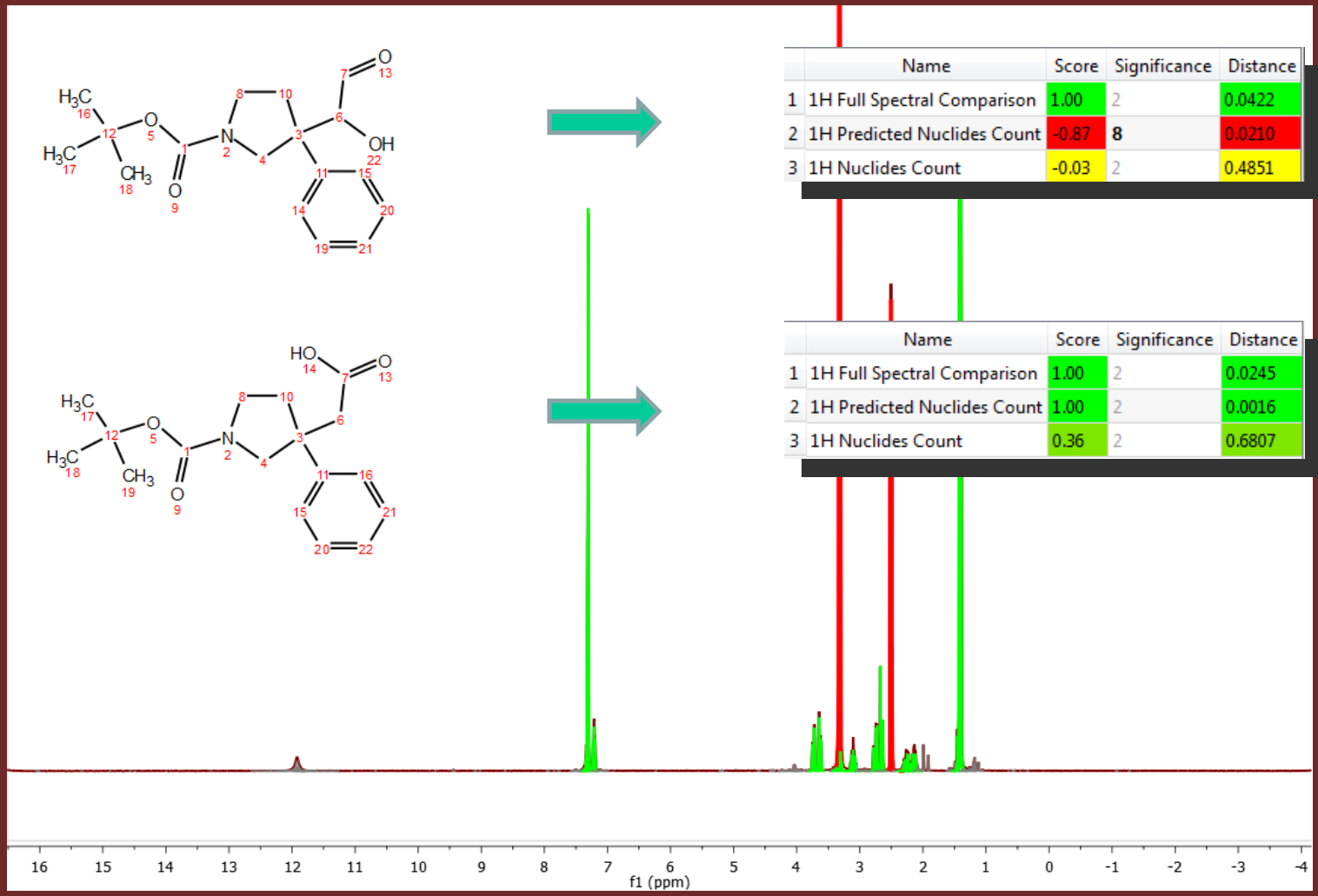
You are ALL invited to join the ASV bandwagon

You will get a useful, working product
becoming better and better

And YOU will be able to actively contribute
to its development

Example #1

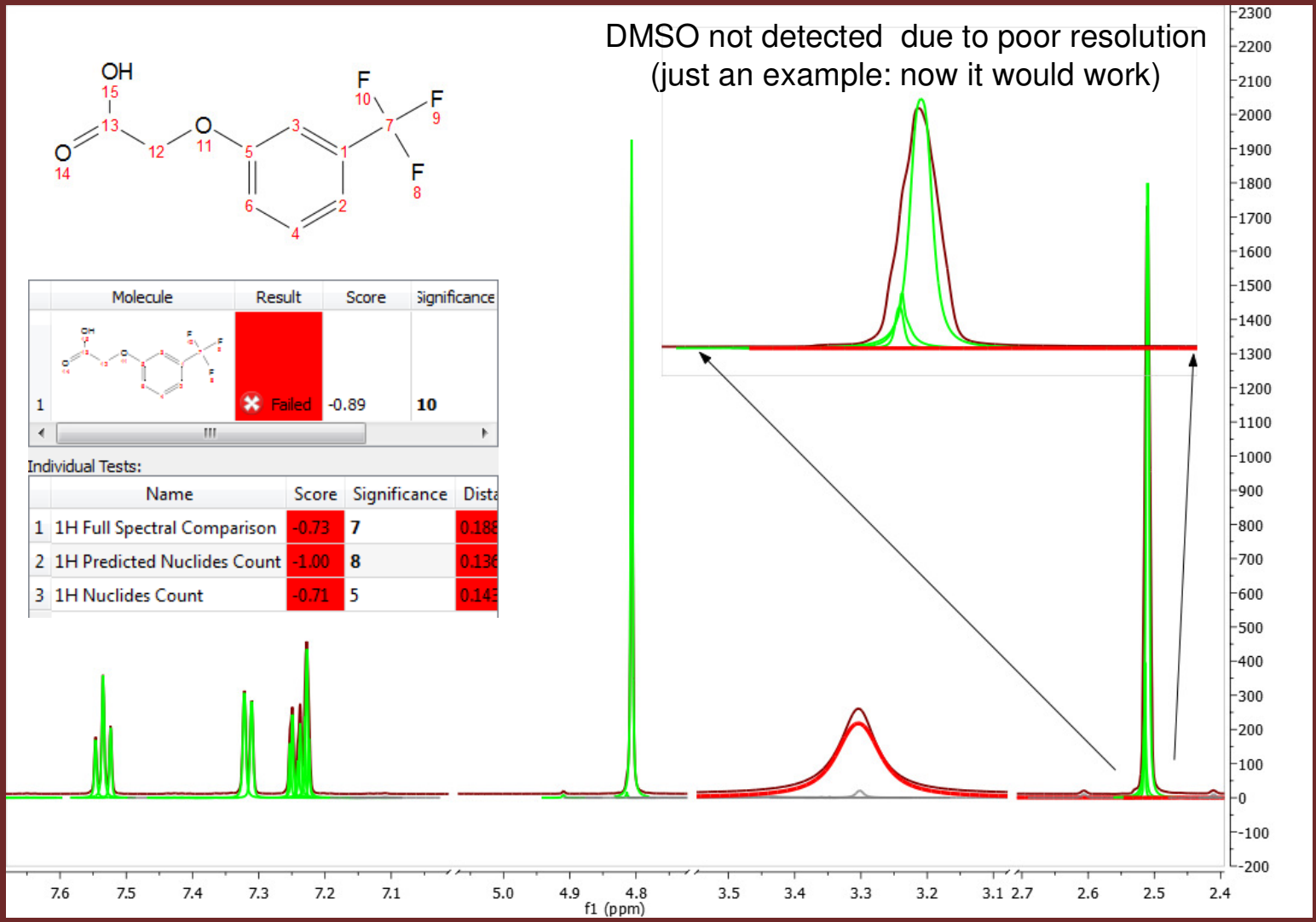
One spectrum, two (or more) molecules



MESTRELAB RESEARCH
Chemistry Software Solutions

Example #2a

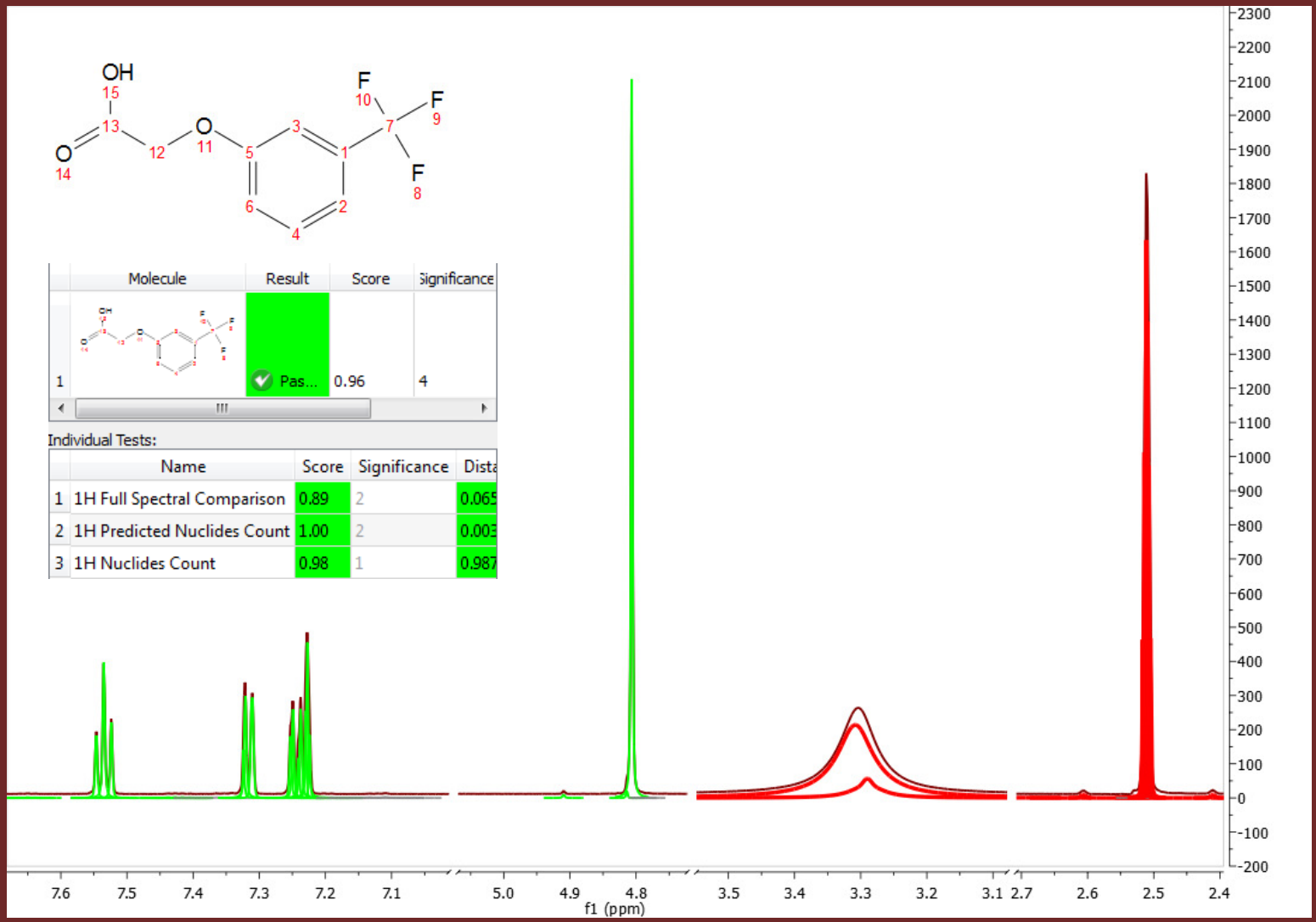
Failure due to bad solvent recognition



MESTRELAB RESEARCH
Chemistry Software Solutions

Example #2b

Success after manual marking of the solvent (CASV)

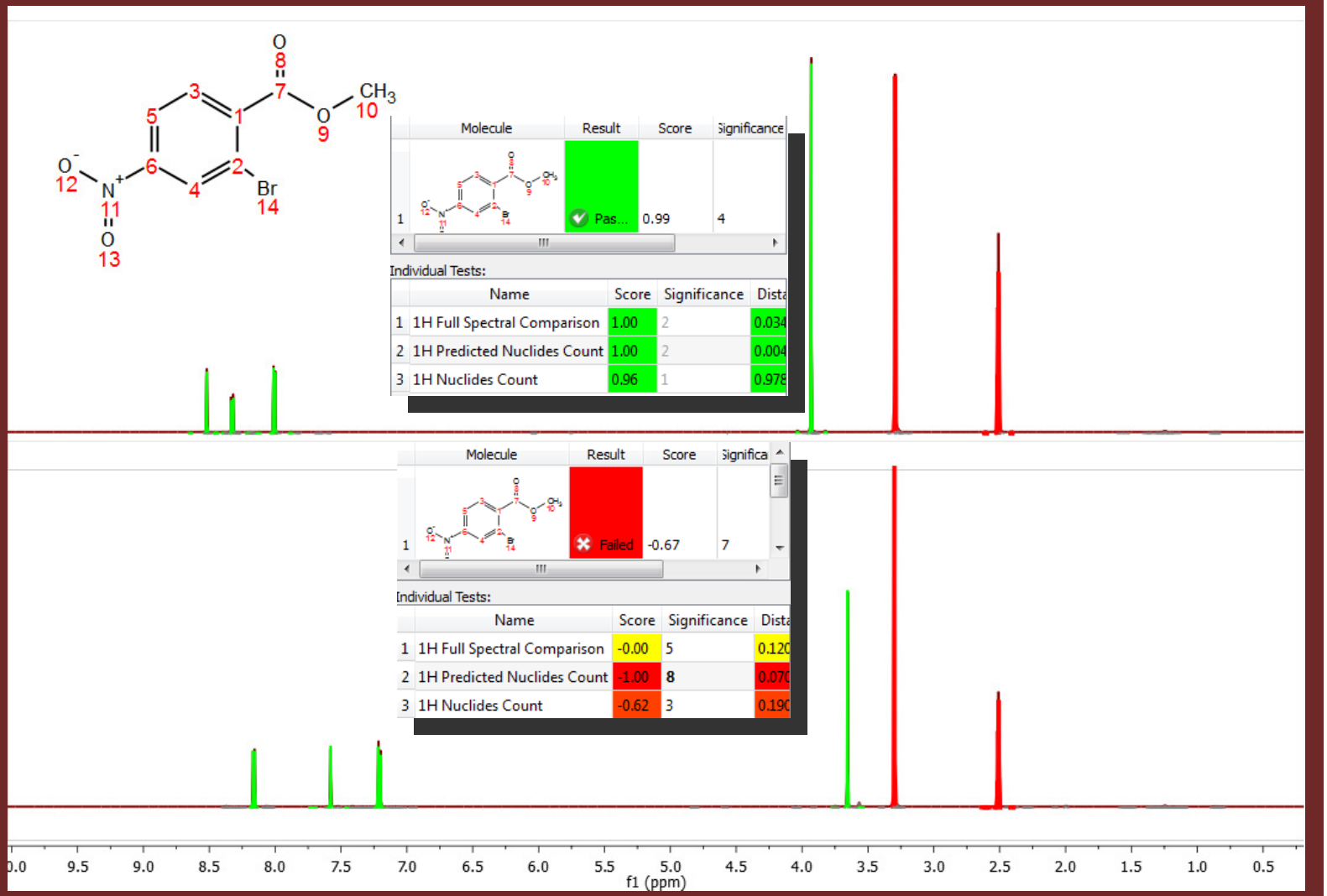


MESTRELAB RESEARCH

Chemistry Software Solutions

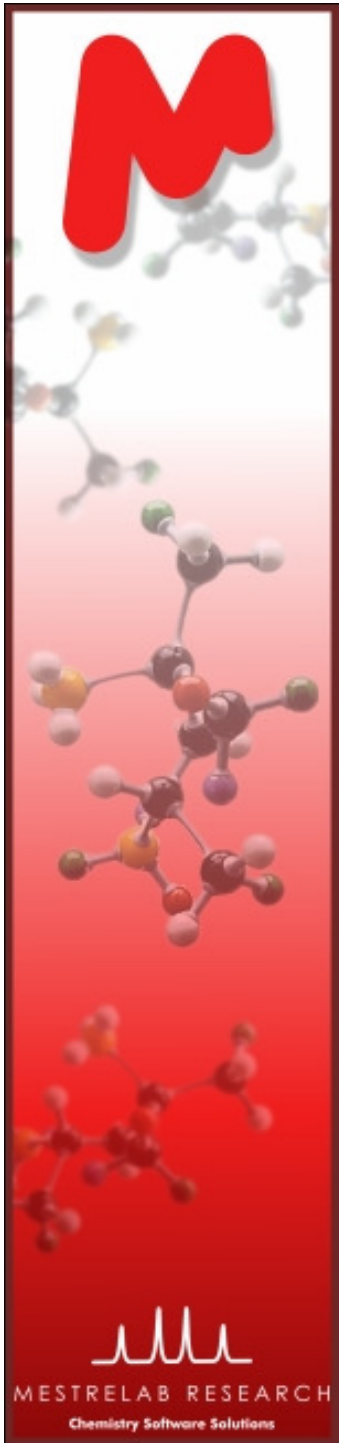
Example #3

One molecule, two (or more) spectra



MESTRELAB RESEARCH

Chemistry Software Solutions



Thank You for your Patience

Any Questions ?

SMASH 2010, Portland, OR, USA
DOI: 10.3247/SL3nmr10.008