Spinning up new predictors: introducing NMR and solubility

Csaba Fábri

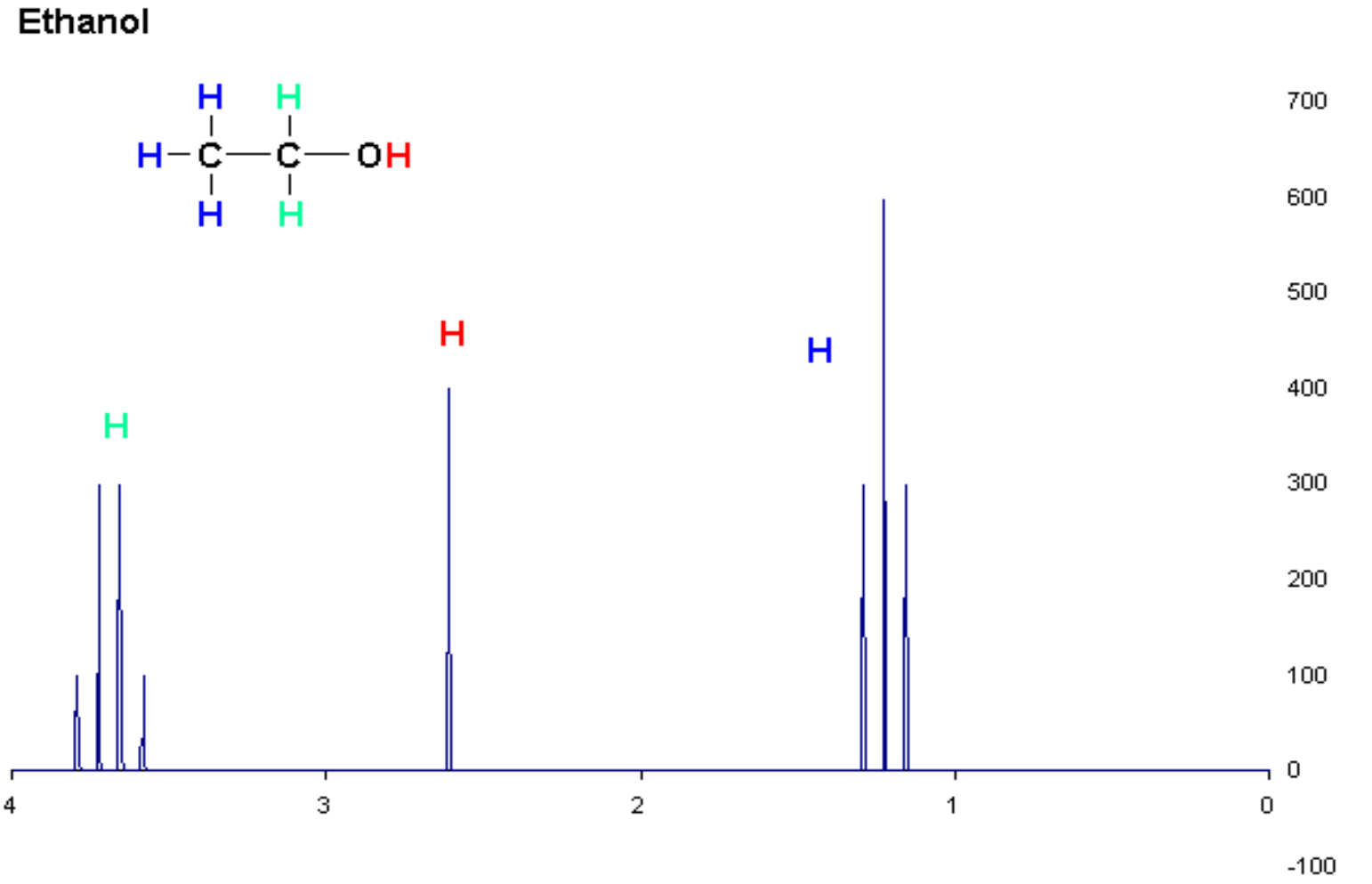
ChemAxon

Solutions for Cheminformatics

Agenda

- Nuclear magnetic resonance (NMR) spectroscopy
  - Introduction of NMR spectroscopy
  - Prediction of spectrum parameters
  - Validation of the chemical shift prediction model
  - NMR Predictor demonstration
  - NMR further plans
- Solubility
  - Brief summary of the solubility model
  - Solubility Predictor demonstration

Example: ^{1}H NMR spectrum of ethanol



NMR Predictor overview

- NMR is a very powerful tool in structure elucidation and validation
- Chemaxon's novel NMR Predictor
  - is capable of fast and accurate prediction of ^{1}H and ^{13}C NMR spectra for standard organic molecules
  - has an easy-to-use Graphical User Interface: all chemical information can be seen on one screen
  - can compare predicted and experimental NMR spectra
  - ready for web, can be used via MarvinSketch applet
  - is another useful predictor in Chemaxon's Discovery Toolkit
  - will be released in Marvin 5.10

Chemical shift prediction model

- Physicochemical and topological descriptors for the description of atomic environments (see next slide)
- Model fitting: multilinear least-squares regression (MLR) and support vector machine (SVM)
- ^{13}C NMR chemical shift model: 9 categories based on hybridization and the number of attached protons
- ^{1}H NMR chemical shift model: 2 categories (C-H protons and heteroatomic protons)
- Validation
- Training and test data were obtained from NMRShiftDB (http://nmrshiftdb.nmr.uni-koeln.de)

Chemical shift descriptor set

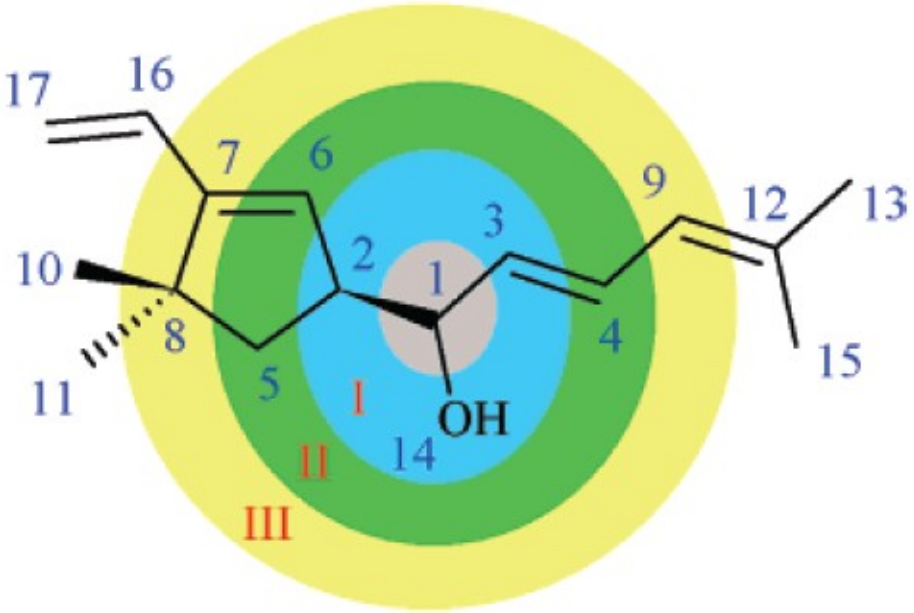


Figure 2. Encoding of the atom environment. The three nearest spheres are marked as cyan, green, and yellow circles, and the Roman numerals denoting the number of each sphere are shown. Atom no. 1 is the central atom. The blue numbers are assigned arbitrarily and serve as references (see the text for more details).

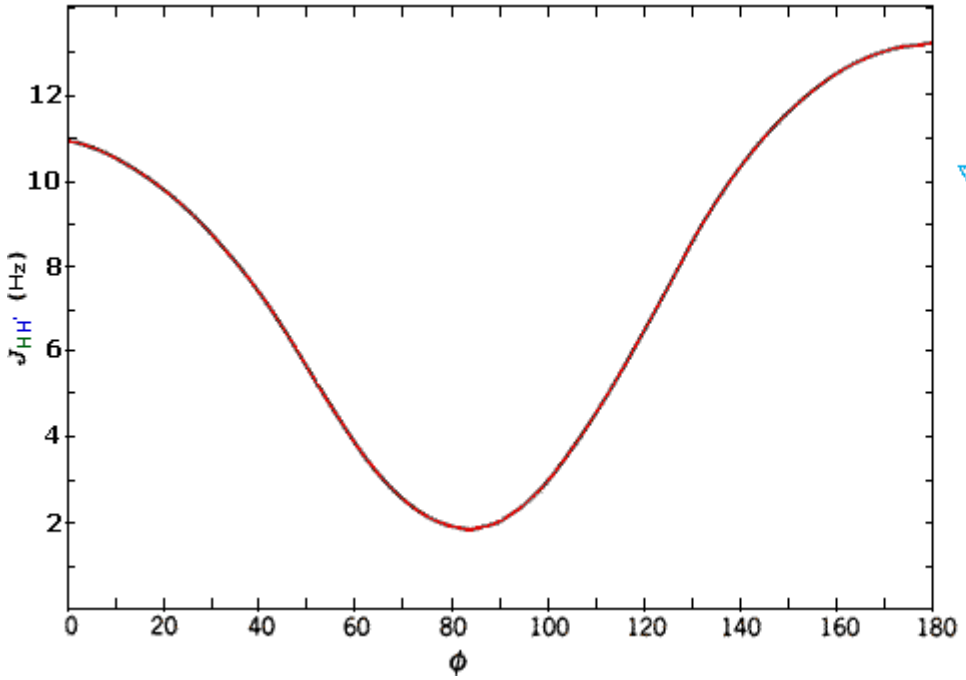
J. Chem. Inf. Comput. Sci. **40**, 1169 (2000)

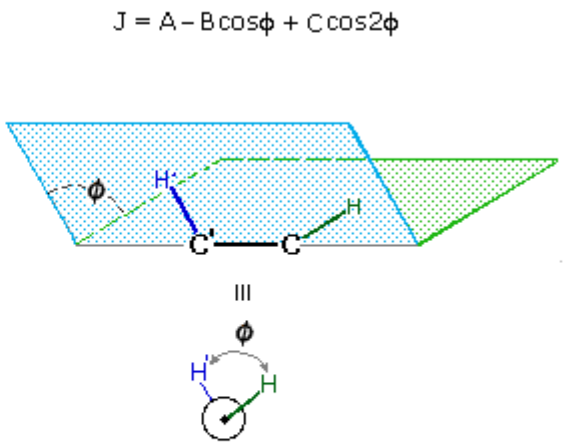
Coupling constant prediction

- Karplus equation for vicinal H-H couplings:

$J = 7.76 \cdot \cos^2\Phi - 1.10 \cdot \cos\Phi + 1.40$

- 3D structure generation and Boltzmann averaging for rotatable bonds





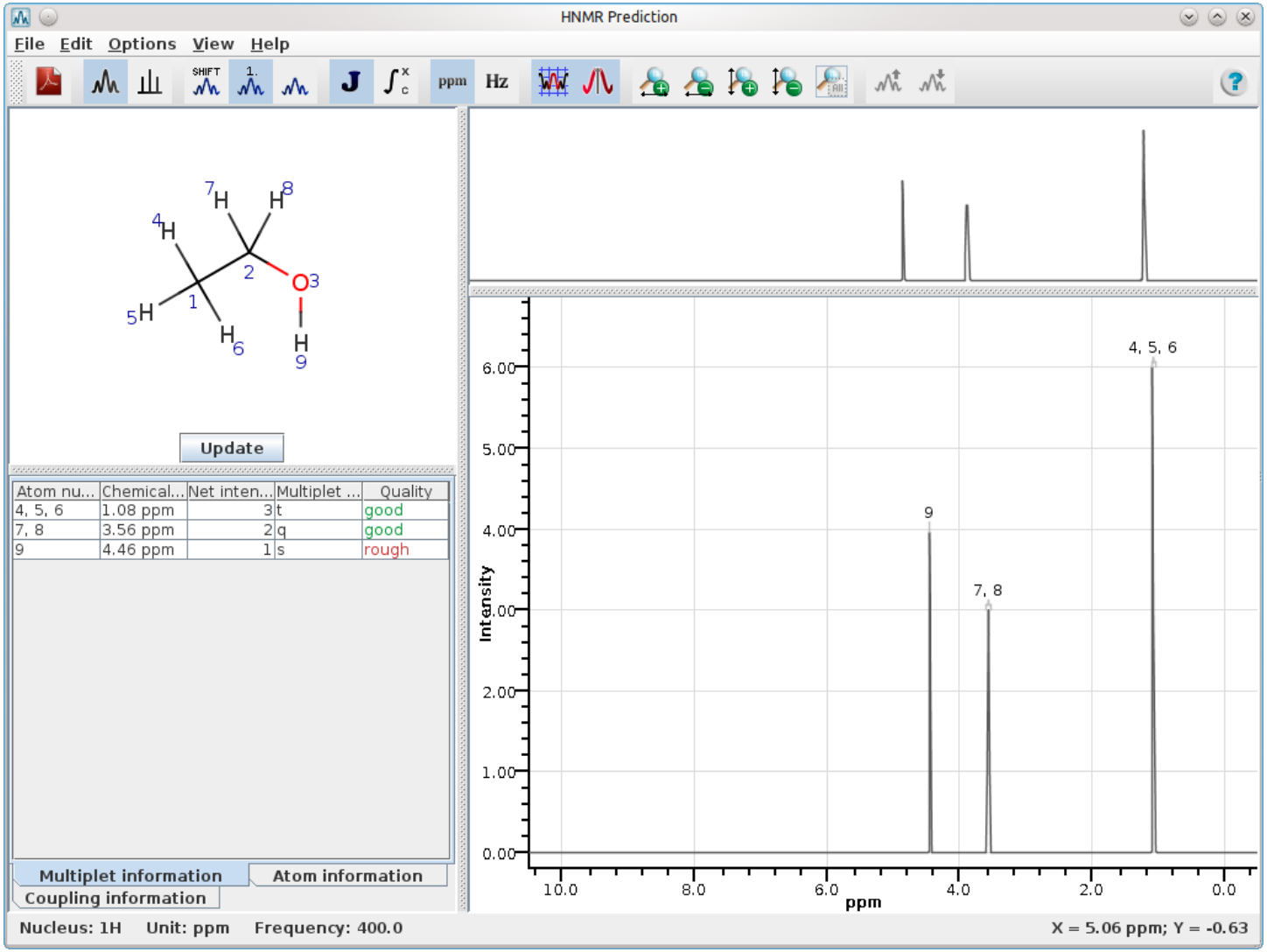
Validation of the chemical shift model

- Accuracy tests: ratio of residuals lower than a threshold

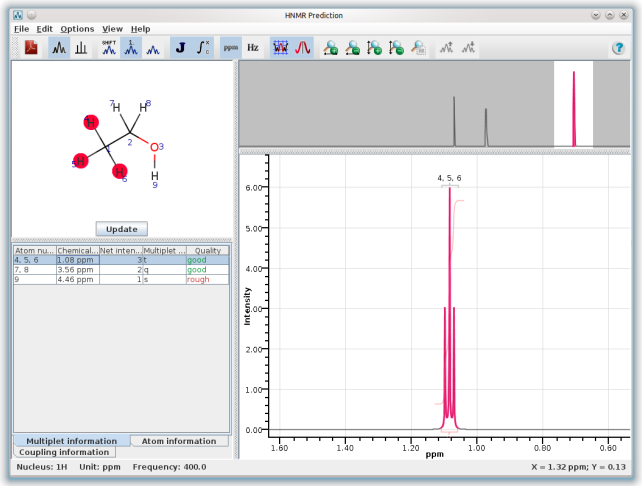
^{13}C NMR	
< 5 ppm	77%
< 10 ppm	93%

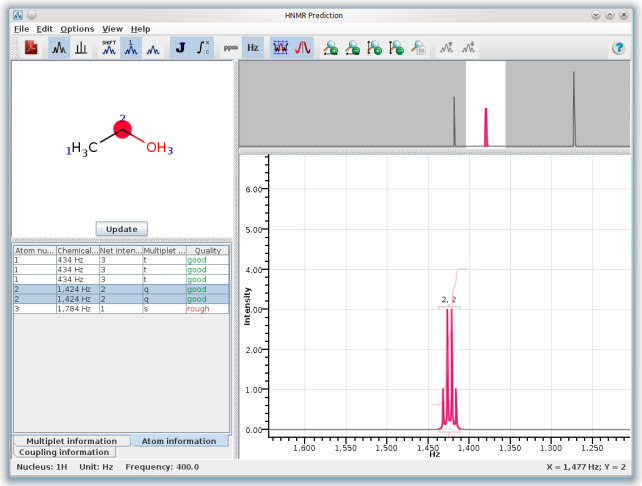
^{1}H NMR	
< 0.5 ppm	81%
< 1.0 ppm	95%

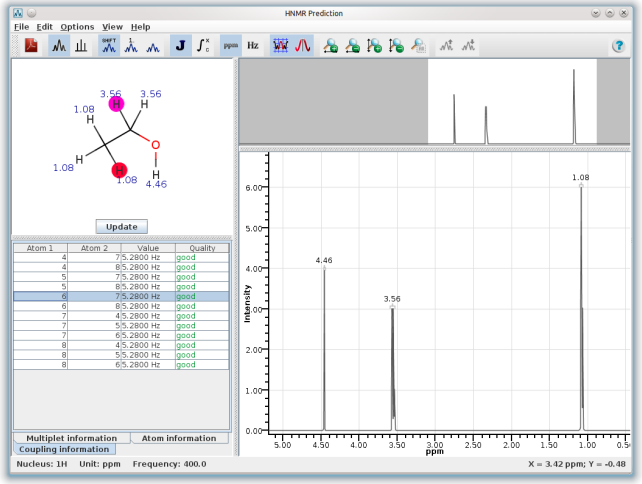
NMR Predictor features I

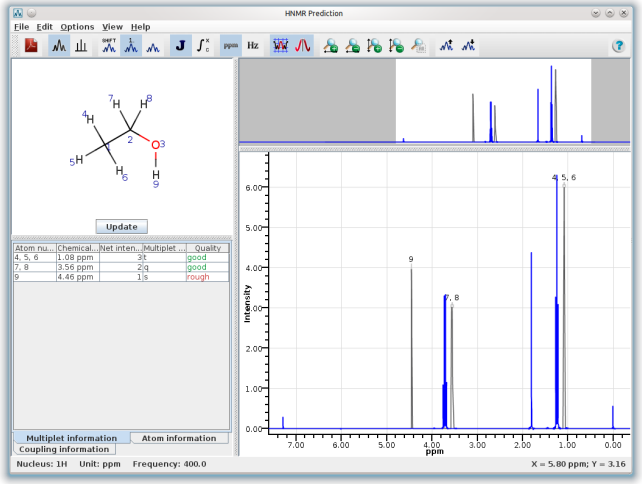


NMR Predictor features II

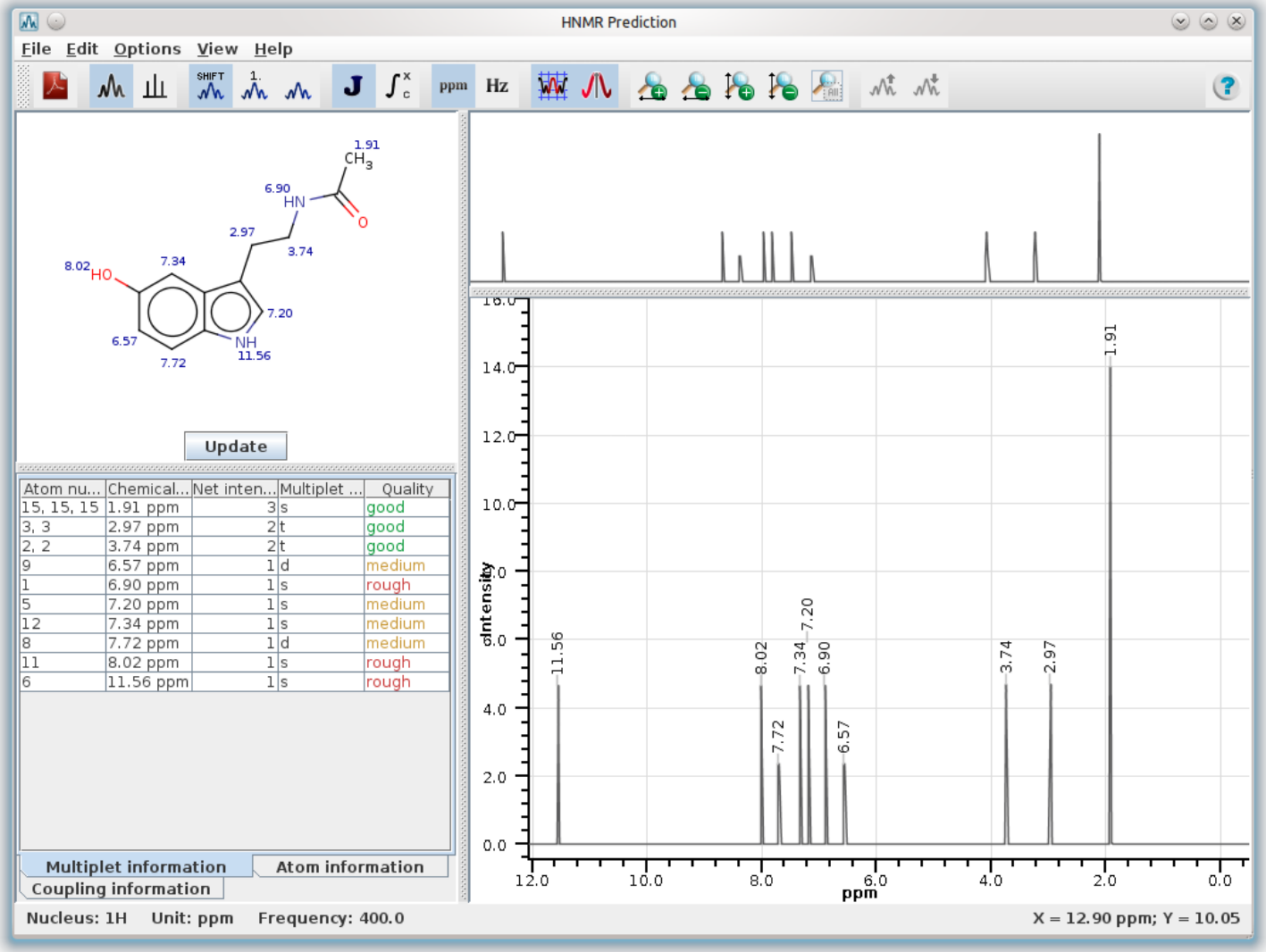








NMR Predictor features III



Example: N-acetyl-5-hydroxytryptamine

^{13}C NMR		^{1}H NMR	
Prediction / ppm	Experiment / ppm	Prediction / ppm	Experiment / ppm
172.25	168.94	11.56	10.49
152.57	150.12	8.02	8.62
129.20	127.88	7.72	7.13
126.44	130.80	7.34	6.84
122.74	122.89	7.20	7.04
118.41	110.92	6.90	7.94
112.10	111.52	6.57	6.60
110.05	102.21	3.74	3.29
108.04	111.23	2.97	2.72
41.92	-	1.91	1.82
26.28	25.30		
23.41	22.62		

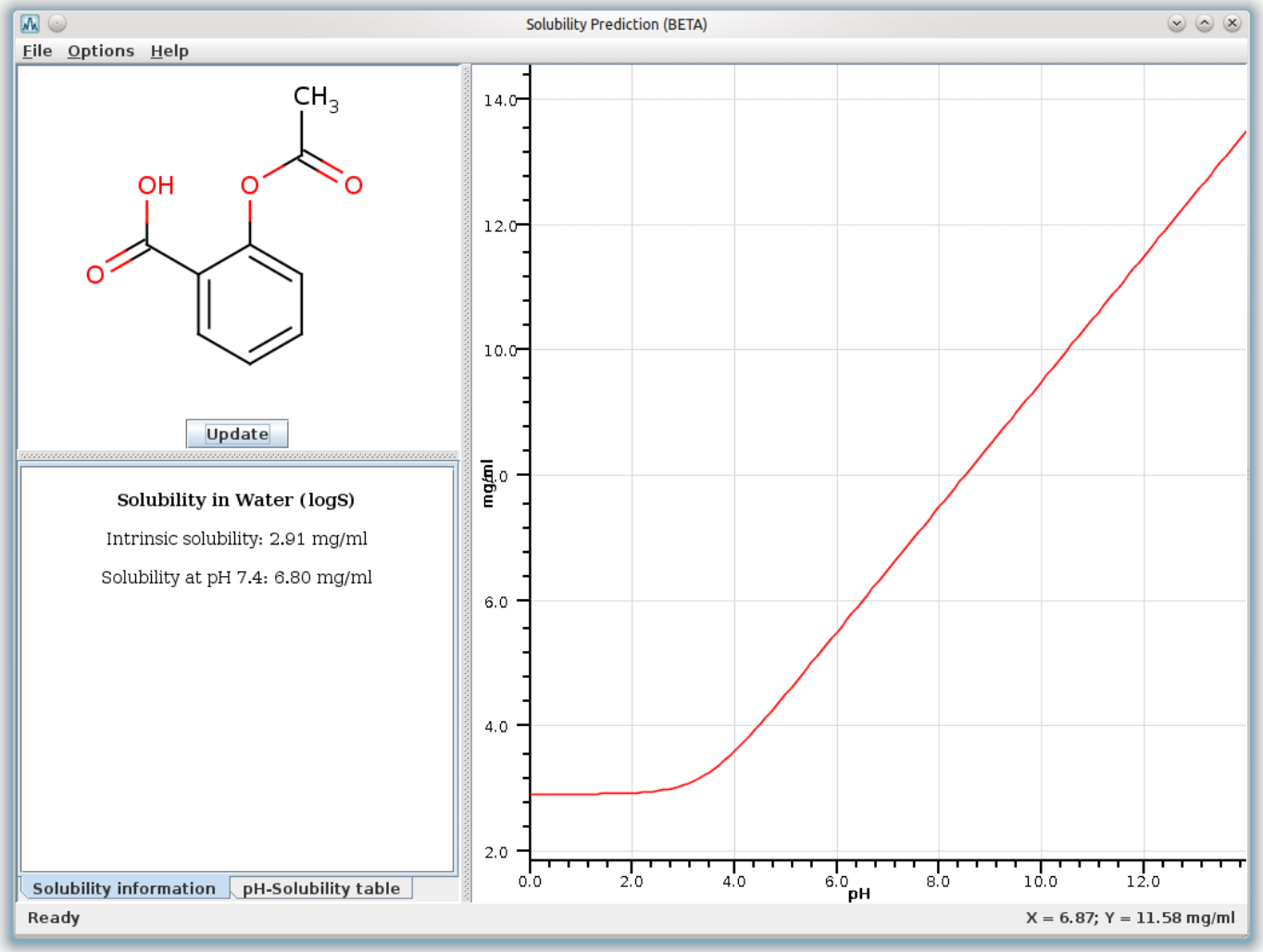
NMR further plans

- Calculate H-F and C-F coupling constants
- Handle diastereotopic protons
- Support more NMR data formats for import/export
- Train the NMR chemical shift model
- NMR prediction on www.chemicalize.org
- Extend the training set
- Consider additional fitting algorithms besides MLR and SVM

Solubility Predictor

- Brief summary:
  - QSPR model
  - Physicochemical and topological descriptors
  - Training for experimental intrinsic solubilities
- List of relevant features:
  - Prediction of intrinsic solubilities
  - Prediction of pH-dependent solubility curves
- Beta release in Marvin 5.11
- Demonstration

Solubility Predictor features



Summary

- Fast and accurate ^{1}H and ^{13}C NMR spectrum prediction for standard organic molecules (5.10 release)
- Development of the Solubility Predictor is in progress (beta release in 5.11)
- Any feedback is appreciated:

calculators-support@chemaxon.com

Acknowledgements

Calculations team members:

Zsolt Mohácsi, Imre Barna, Rita Veréb, István Rábel, József Szegezdi, Viktória Pálfi

NMR expert:

Csaba Peltz

Application scientists:

Iván Solt, Attila Szabó, Krisztián Niesz