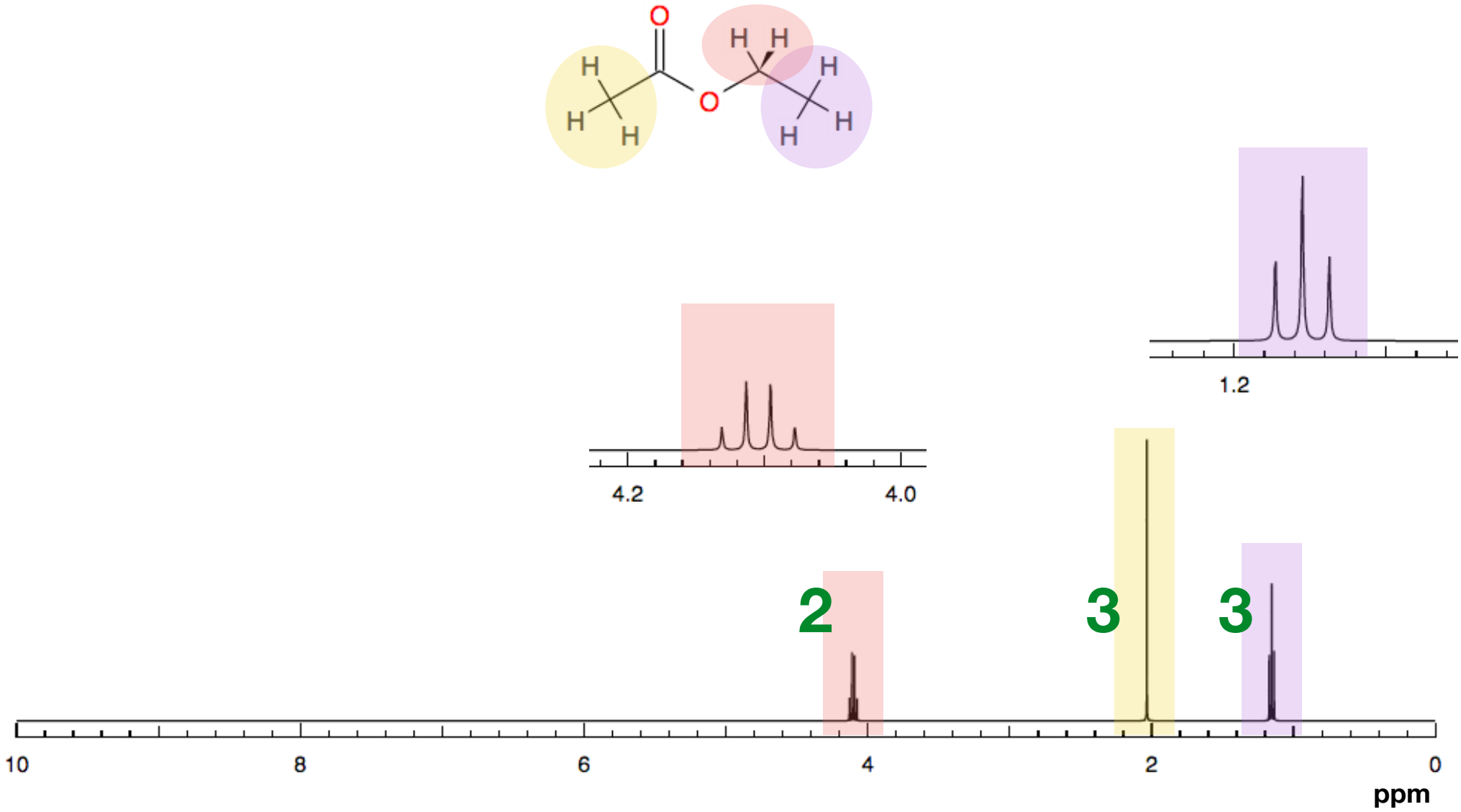
Ask Ernö

Self learning of NMR prediction tool

Service of cheminformatics
luc.patiny@epfl.ch

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

^{1}H NMR spectroscopy

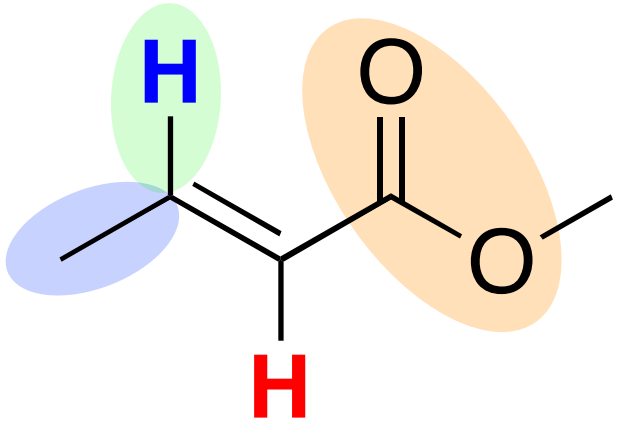


How to predict chemical shifts ?

Prediction using “increments”

$$\delta_{C=CH} = 5.25 + Z_{gem} + Z_{cis} + Z_{trans}$$

$$\delta_{C=CH} = 5.25 + 0.8 - 0.22 + 0 = 5.83$$

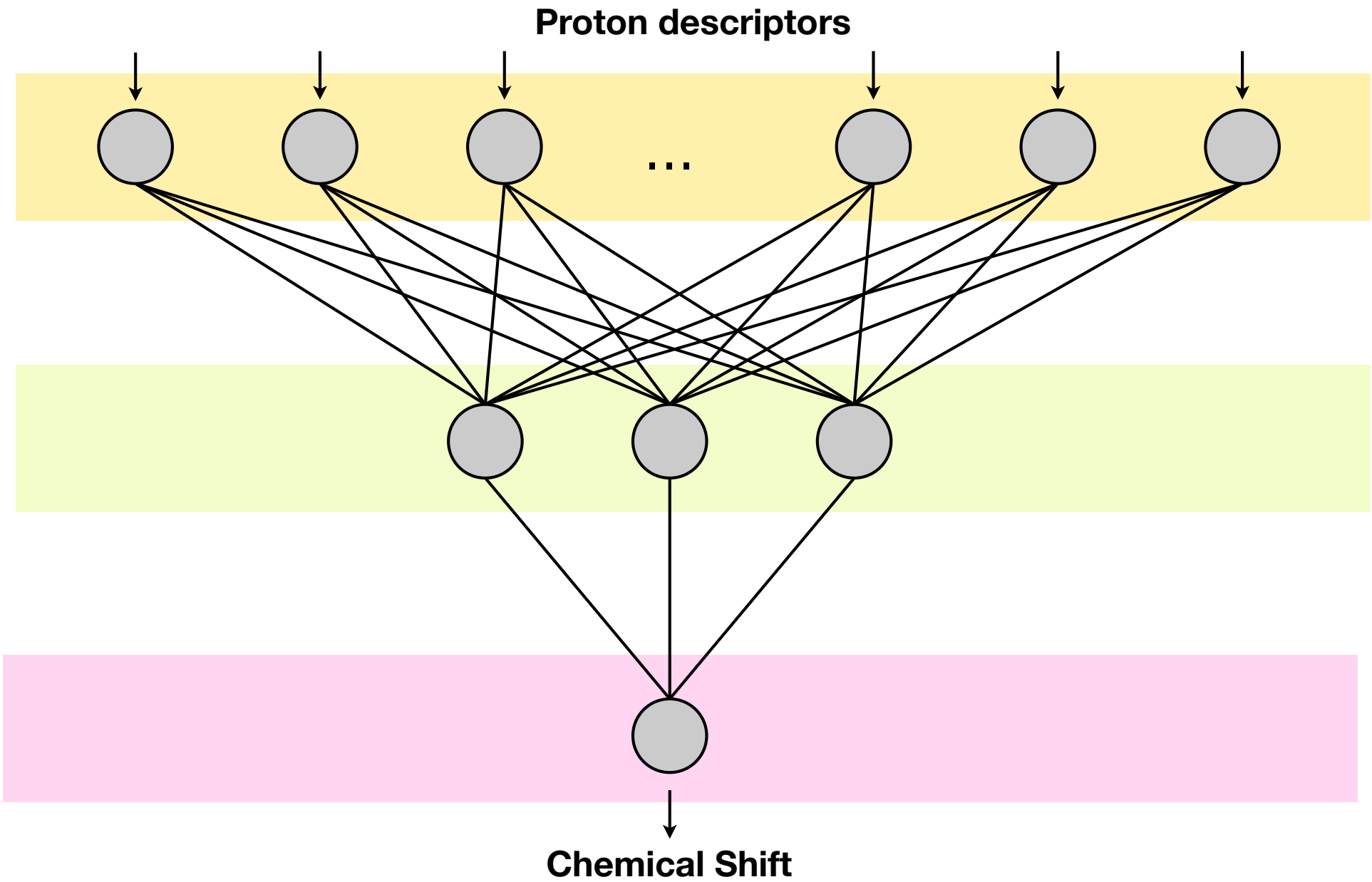


$$\delta_{C=CH} = 5.25 + 0.45 + 1.18 + 0 = 6.88$$

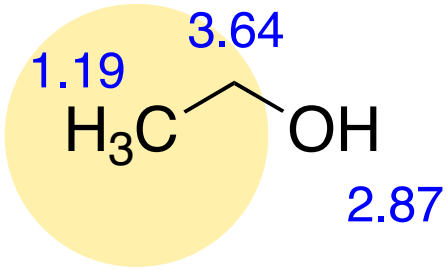
Substituent R	Z gem	Z cis	Z trans
-H	0	0	0
-alkyl	0.45	-0.22	-0.28
-F	1.54	-0.4	-1.02
-Cl	1.08	0.18	0.13
-COOR isolated	0.8	1.18	0.55

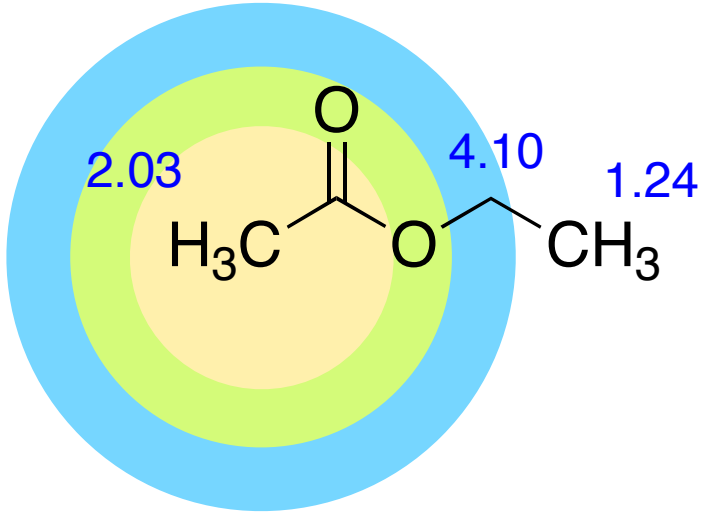
Pretsch, “Spectral data for Structure determination or Organic compounds”, 1989

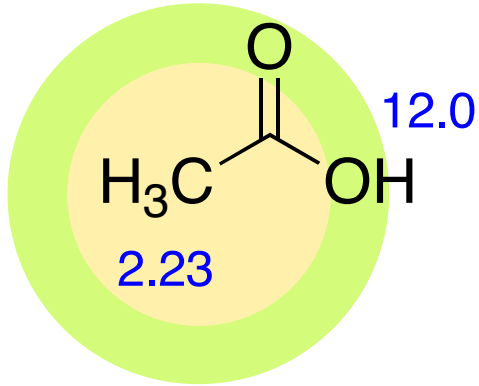
Neural Network

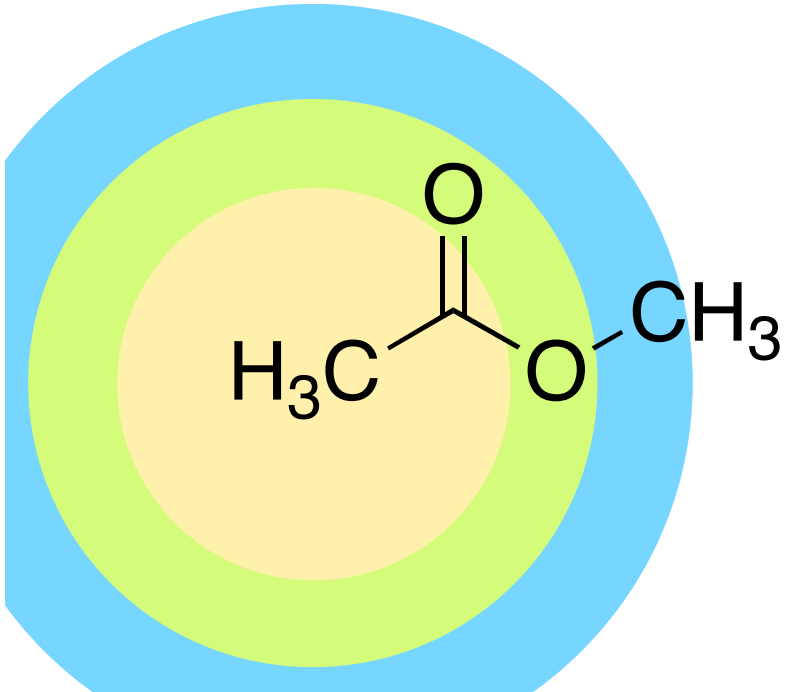


HOSE code









$$\delta = \frac{1.19+2.03+2.23}{3} = 1.81$$

$$\delta = \frac{2.03+2.23}{2} = 2.13$$

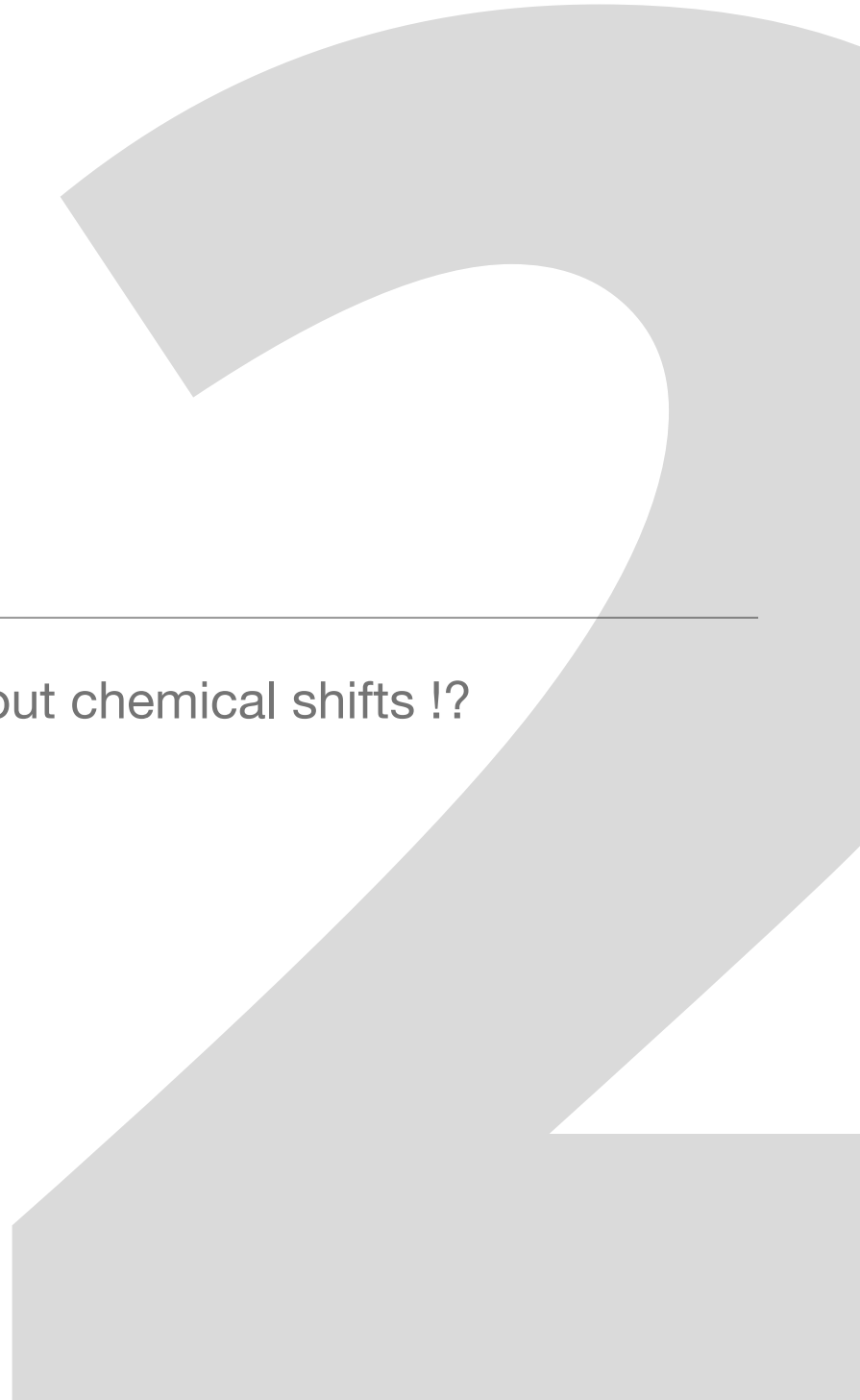
$$\delta = 2.03$$

But …

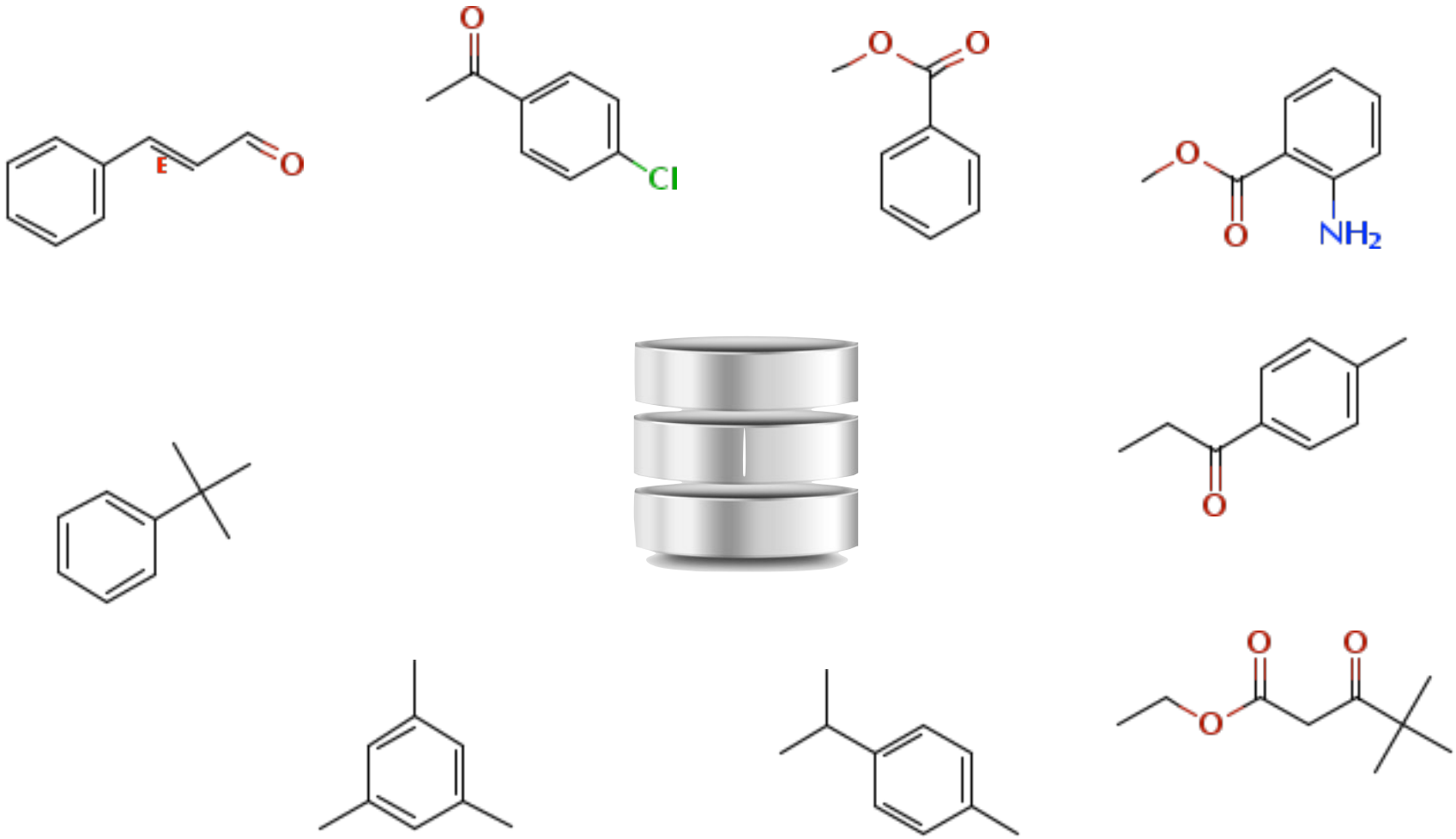
All those methods rely on assigned NMR spectra

Let’s skip the human …

Creating a NMR chemical shifts predictor without chemical shifts !?

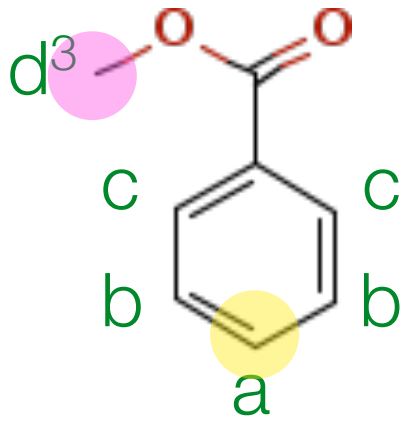


1. Creating a dataset: molfile / spectrum

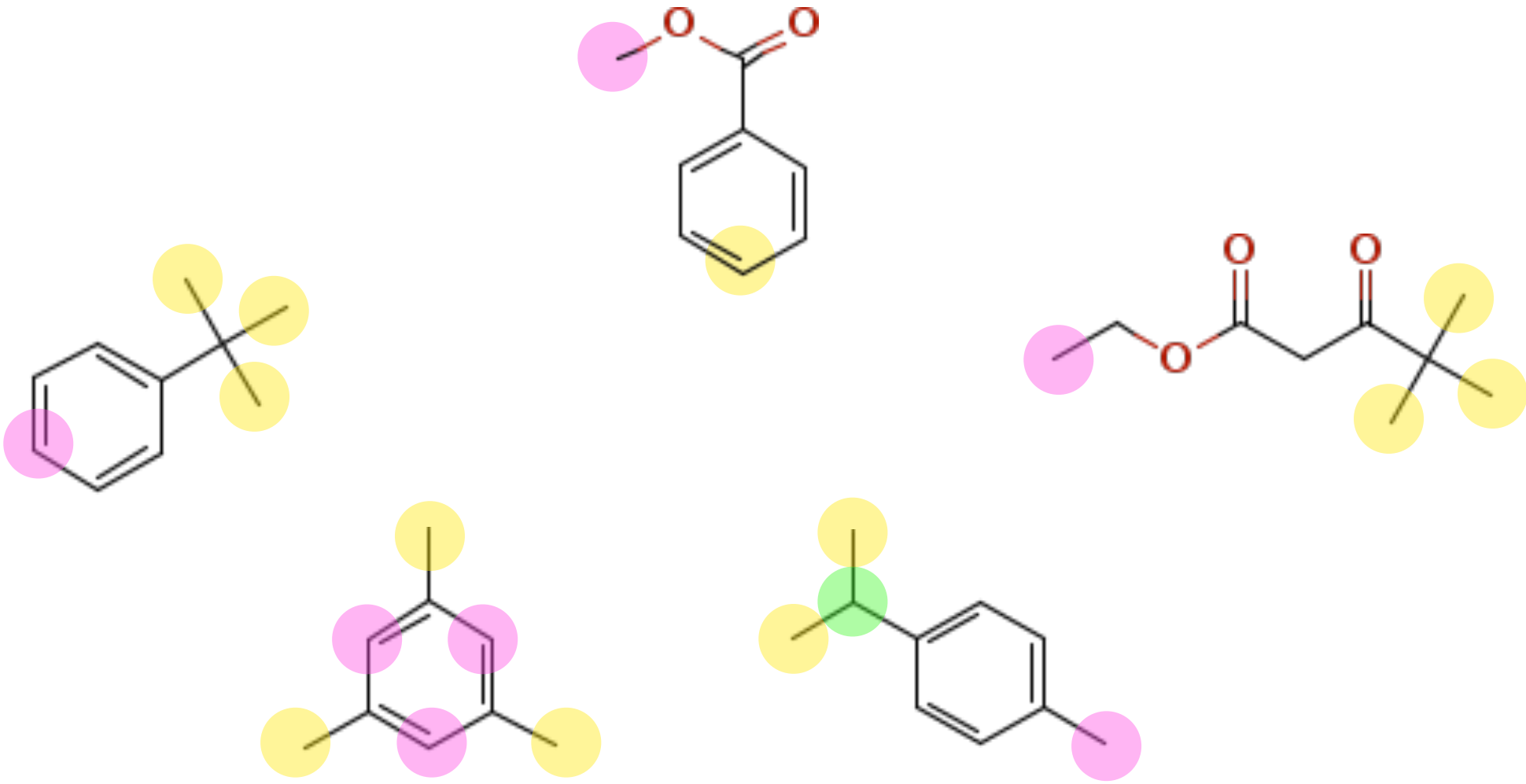


2341 molecules / NMR spectra

2. Number and kind of hydrogens



3. Unambiguous number of protons



Demo

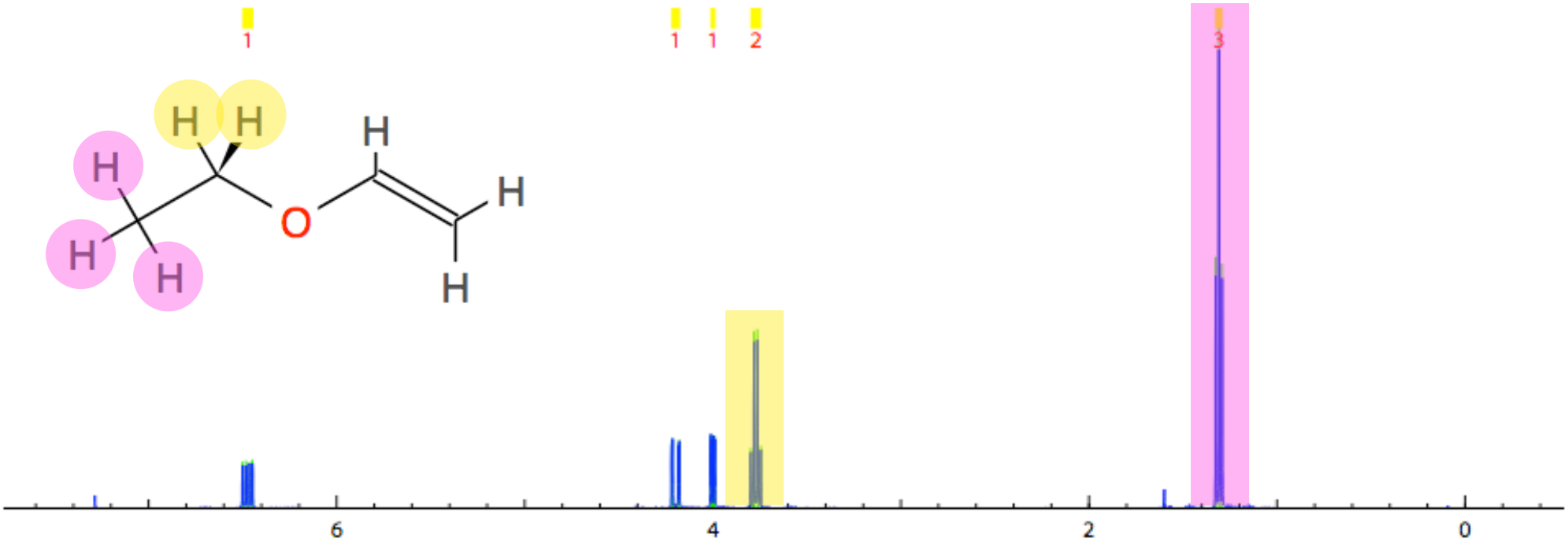
3. Peak picking and automatic integration

- Integration of the zones

- Removal of the NMR solvent

Demo

4. Assign non ambiguous protons

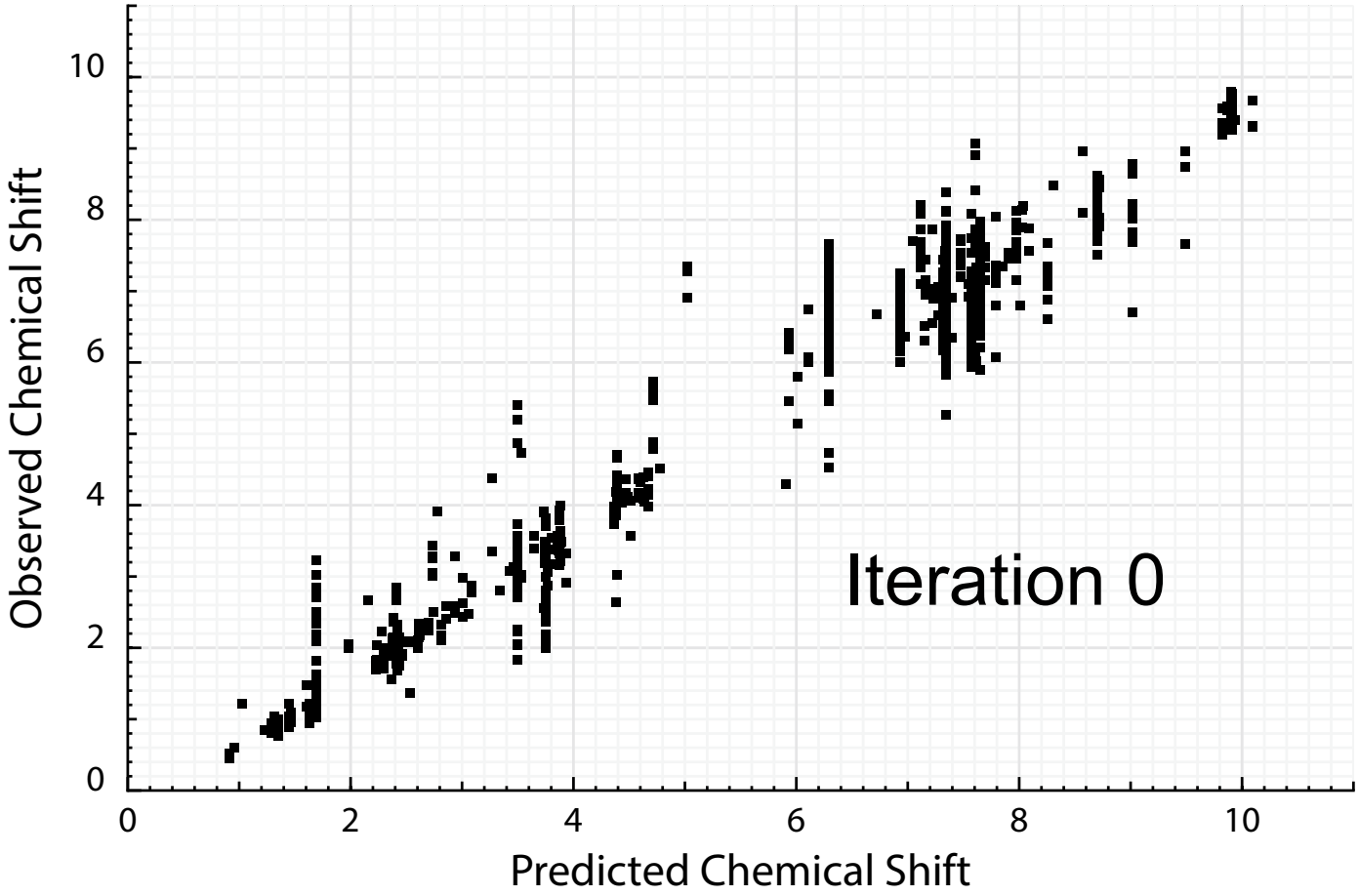


… and create corresponding HOSE codes

The results

3

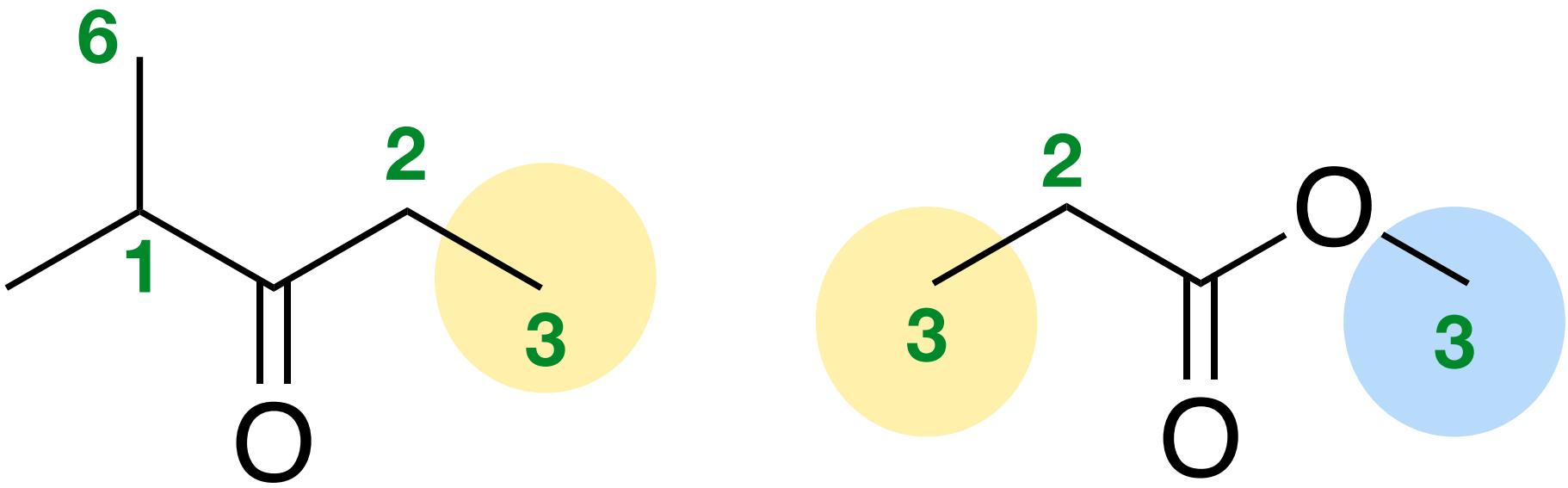
Analysis of the 2341 molecule / spectra set



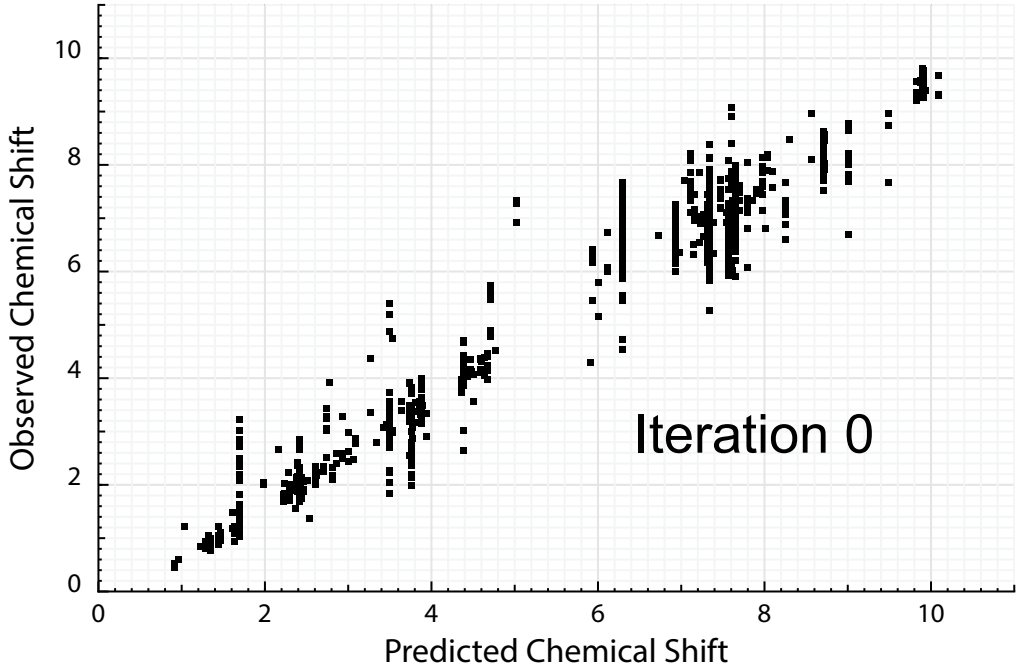
HOSE 3 : 62
HOSE 4 : 254
HOSE 5 : 461

Test set of 298 molecules

Using what we learned

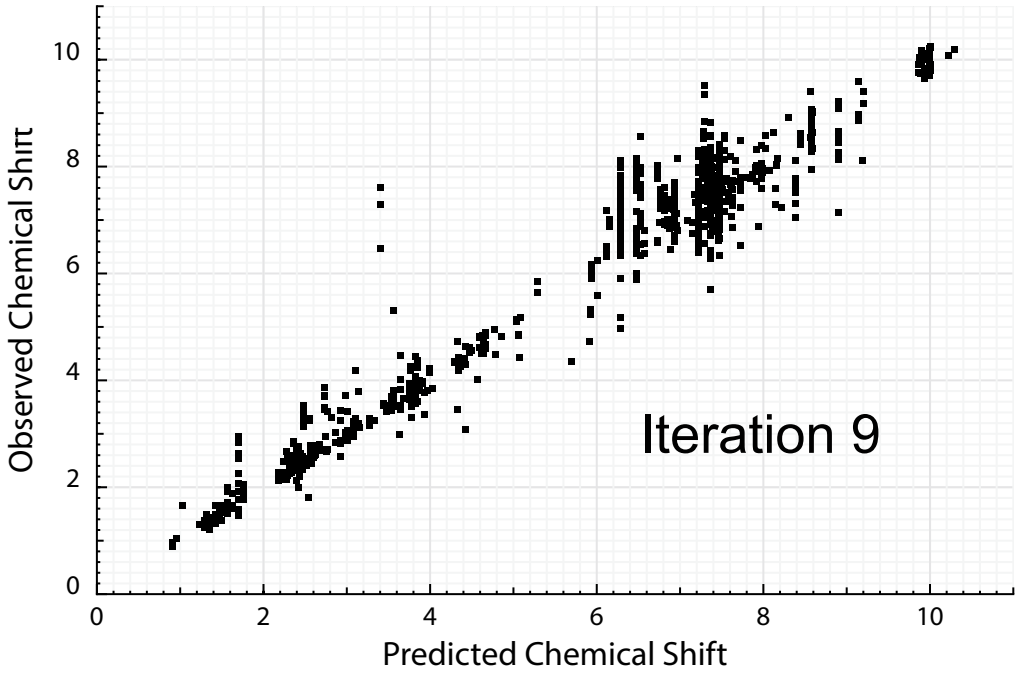


9 iterations further …



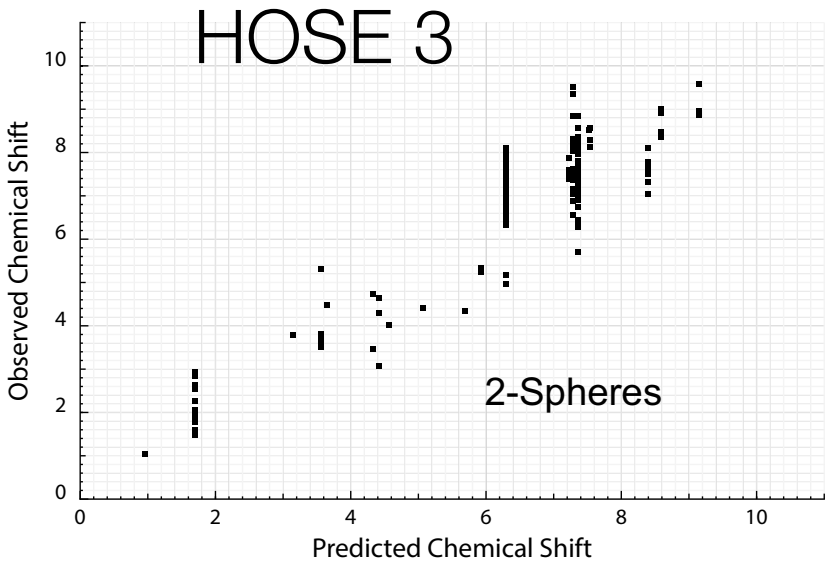
HOSE 3 : 62
HOSE 4 : 254
HOSE 5 : 461

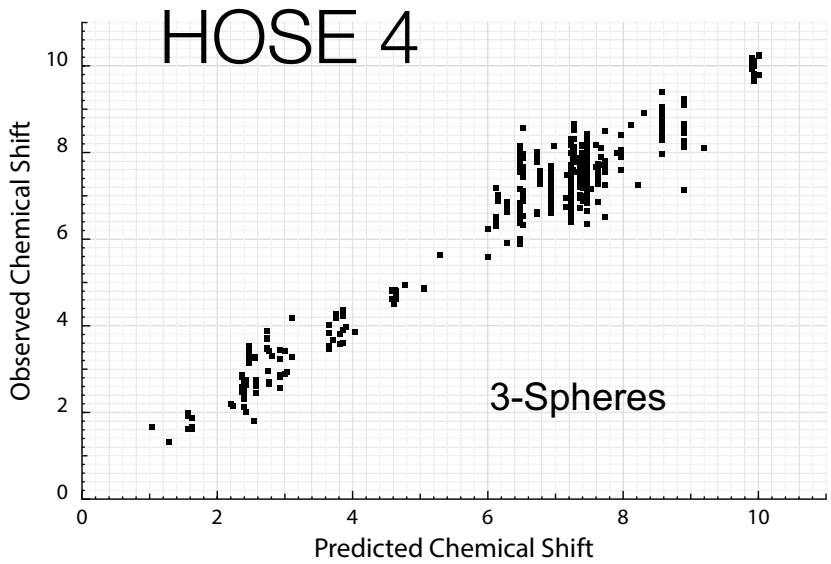
HOSE 3 : 63
HOSE 4 : 382
HOSE 5 : 916

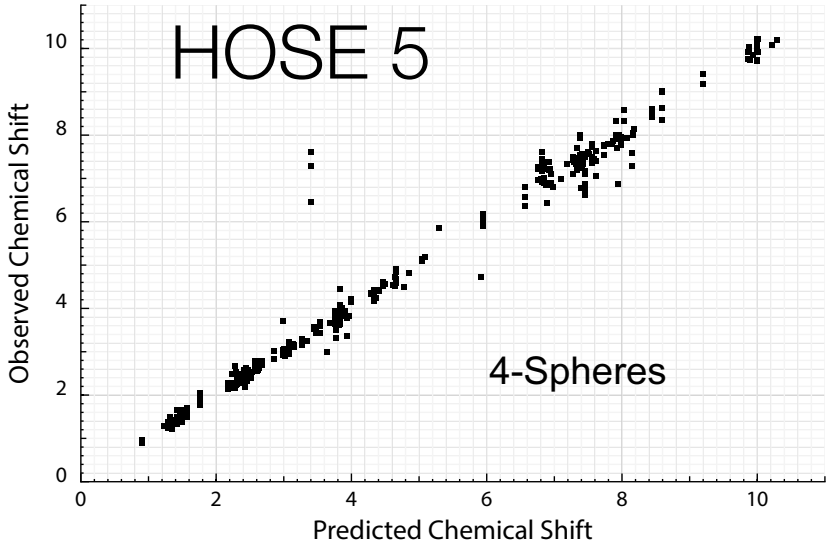


Test set of 298 molecules

Quality of prediction based on HOSE code level





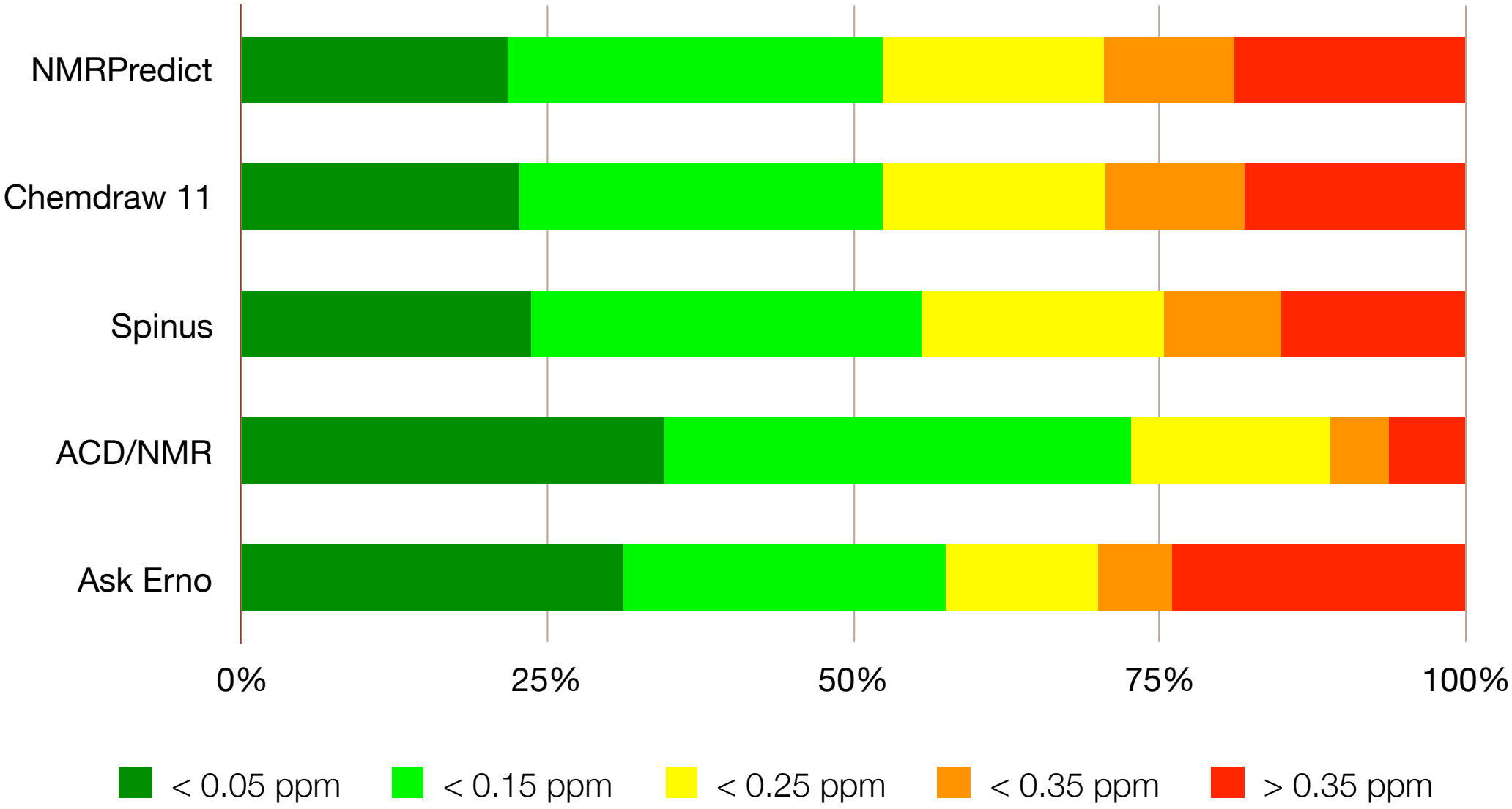


Demo

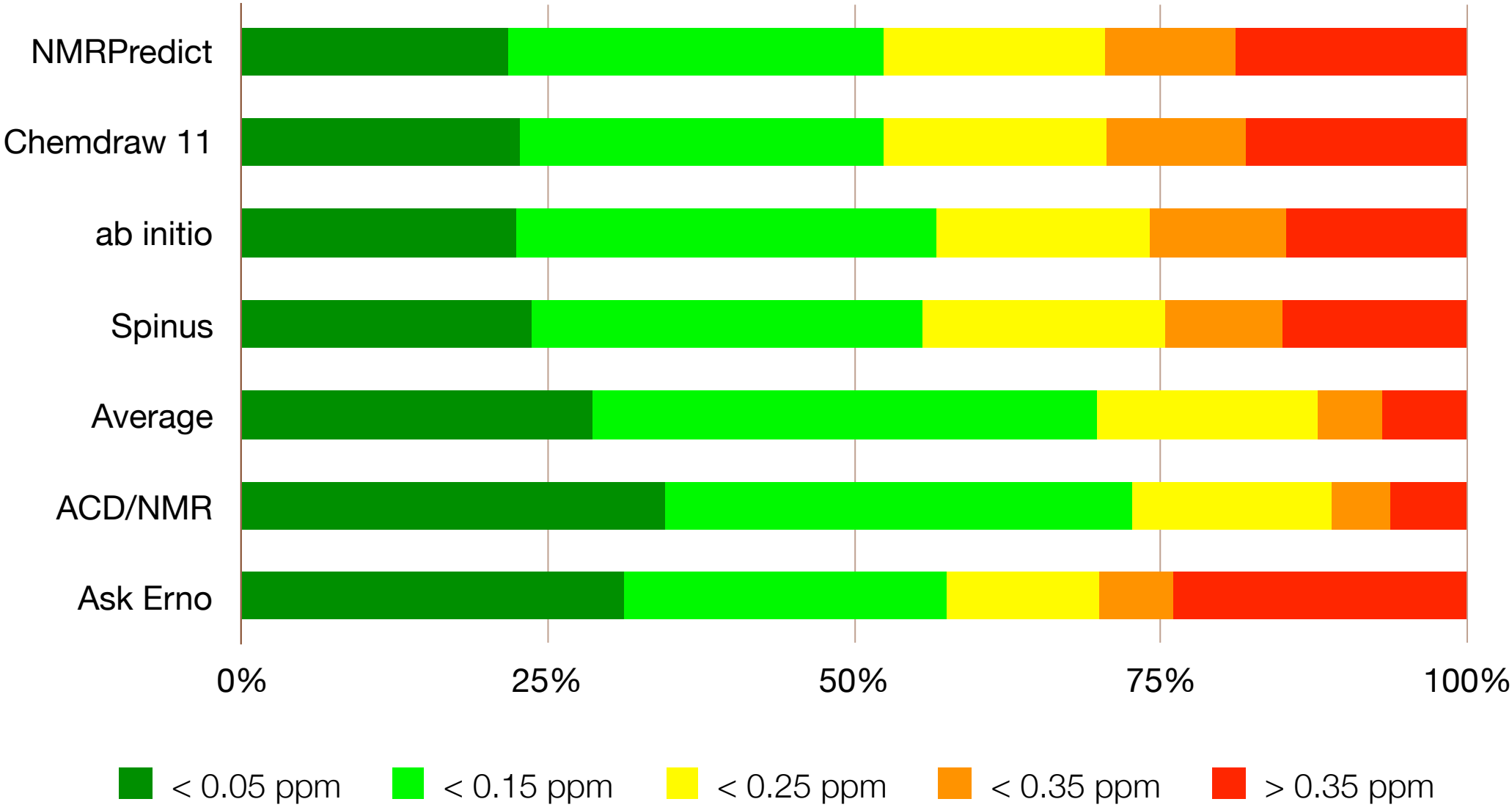
Conclusions

4

Prediction average errors



Prediction average errors



Thanks !

NMR project

- Andrès Castillo
- Julien Wist
- Andrès Bernal
- Reiner Dieden

visualizer

- Norman Pellet
- Michaël Zasso
- Daniel Kostro

openchemlib

- Thomas Sander