

Advanced Prediction

PERCH NMR Software Course

Turku - 16- 17.5.2006



© 2001-2006 PERCH Solutions Ltd. - CSD/PERCH NMR Software Course, Turku 16-17.5.2006



Why Just NMR?

Information in very well arranged form:

- Chemical Shifts → Chemical Nature of Compound
- Coupling Constants → Connectivity & Stereochemistry
- Intensities → Quantitativity
- Relaxation Times → Kinetics



© 2001-2006 PERCH Solutions Ltd. - CSD/PERCH NMR Software Course, Turku 16-17.5.2006



What About Proton NMR?

- The most sensitive and useful NMR nucleus
- Chemical shifts sensitive to experimental conditions (solvent, temperature, concentration)
- Coupling patterns change with shift-changes in non-first-order spectra
- Couplings contain structural information and vice versa !!

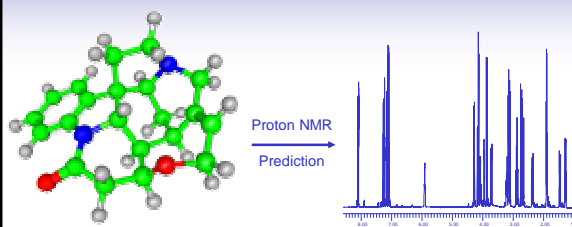
A simple data-base approach cannot work well



© 2001-2006 PERCH Solutions Ltd. - CSD/PERCH NMR Software Course, Turku 16-17.5.2006



NMR Spectral Prediction



Structure

Spectrum



© 2001-2006 PERCH Solutions Ltd. - CSD/PERCH NMR Software Course, Turku 16-17.5.2006



NMR Prediction - Methods

- Increments
- Spectral Libraries
- Hoes Code
- Neural Networks
- Quantum Chemical Calculations
- Model It HNMR



© 2001-2006 PERCH Solutions Ltd. - CSD/PERCH NMR Software Course, Turku 16-17.5.2006



Increments

- Simple
- Very fast
- Limited to classes of substances
- Substituent effects are NOT additive!

$$\sigma_{calc} = \sigma_0 + \sum_i \sigma_i + K$$

SpecTool

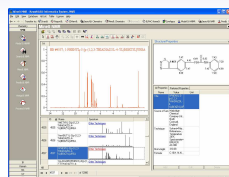


© 2001-2006 PERCH Solutions Ltd. - CSD/PERCH NMR Software Course, Turku 16-17.5.2006



Spectral Libraries

- Chemical (sub)structure
- Experimental data
- Manual comparison tedious
- Proton NMR spectra

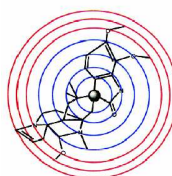


SpecInfo
HaveltAll NMR
SDBS
ACD

© 2001-2006 PERCH Solutions Ltd. - CSC/PERCH NMR Software Course, Turku 16-17.5.2006

The HOSE-Code

Hierarchically Ordered Spherical
description of Environment



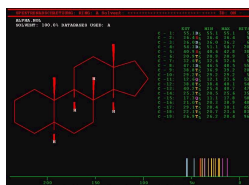
- Requires large database
- Slower
- Limited descriptors
- Stereochemistry?

SpecInfo
ACD
PredictIt NMR
CSEARCH

© 2001-2006 PERCH Solutions Ltd. - CSC/PERCH NMR Software Course, Turku 16-17.5.2006

3D HOSE-Code

- Stereochemistry represented by up/down-bonds
- Stereochemistry described by interactions over 3, 4 and 5 bonds
- Additionally descriptors for axial/equatorial substitution
- No 3D-coordinates needed
- Dynamics?
- Other than ax/eq?

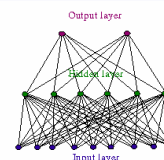


CSEARCH
ACD?

© 2001-2006 PERCH Solutions Ltd. - CSC/PERCH NMR Software Course, Turku 16-17.5.2006

Neural Networks

- Extremely fast
- Limited number of descriptors
- Time-consuming training required (optimization of weights)

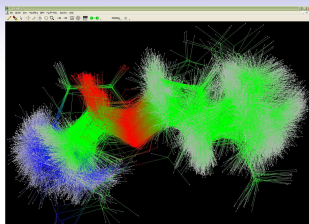


CSEARCH
ACD?

© 2001-2006 PERCH Solutions Ltd. - CSC/PERCH NMR Software Course, Turku 16-17.5.2006

PERCH NMR Software

- 3D-Structure
- Conformation
- Molecular flexibility
- Solvent
- Higher order effects

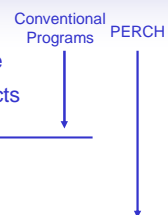


Most realistic descriptors !

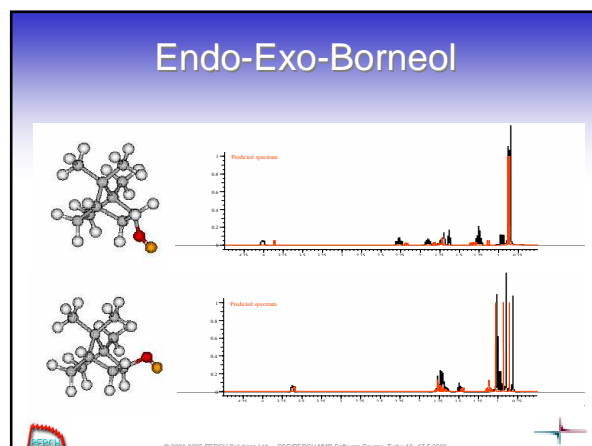
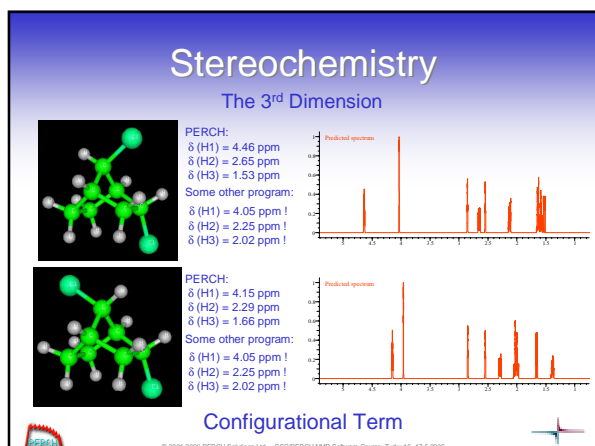
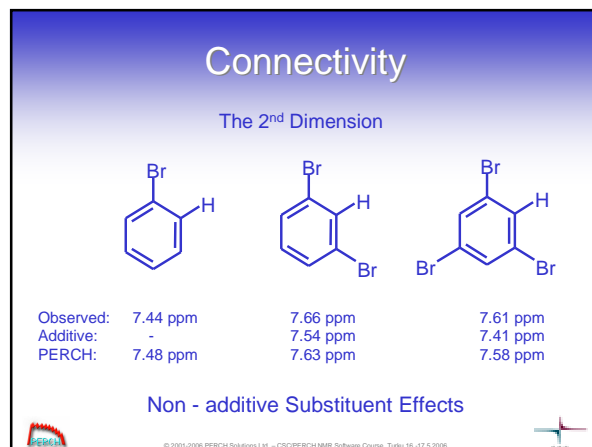
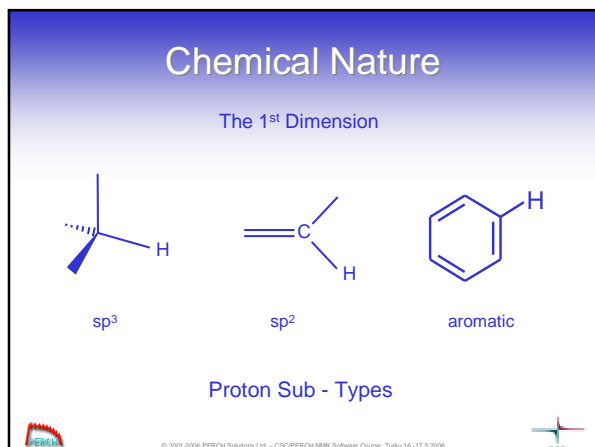
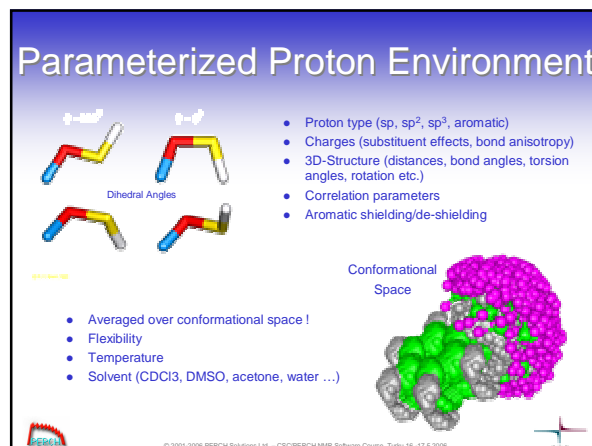
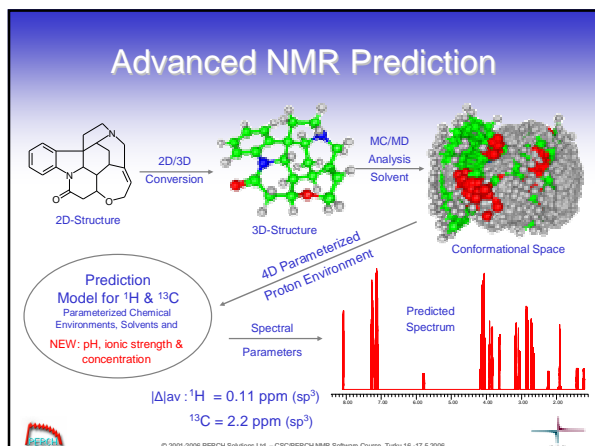
© 2001-2006 PERCH Solutions Ltd. - CSC/PERCH NMR Software Course, Turku 16-17.5.2006

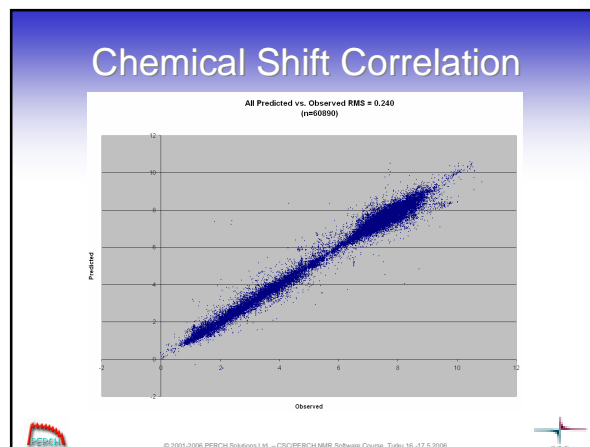
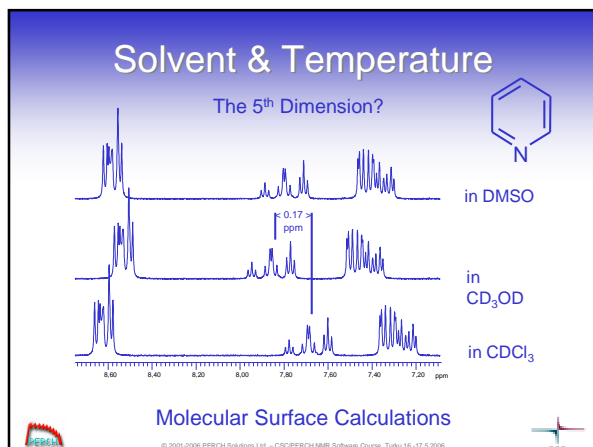
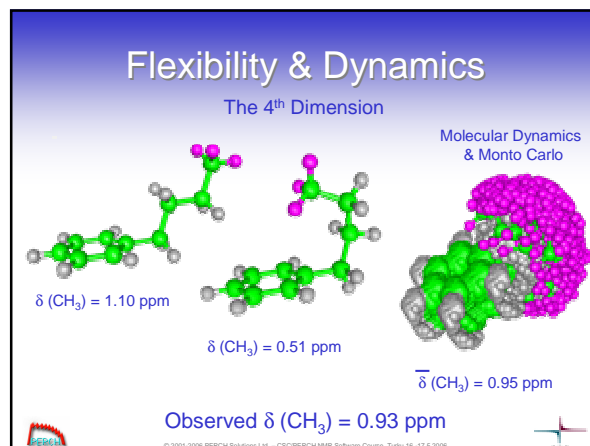
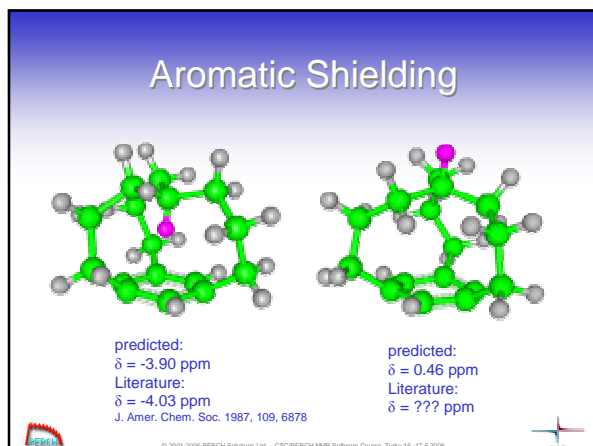
Proton NMR Prediction

1. Chemical Nature - Proton Type
2. Connectivity - Substituent Effects
3. Stereochemistry - 3D Structure
4. Time - Flexibility & Dynamics



© 2001-2006 PERCH Solutions Ltd. - CSC/PERCH NMR Software Course, Turku 16-17.5.2006





Reliability

All data, RMS = 0.240 (n=60890)
RMS = 0.208 (RMS < ±1, n=60451)

- "Hit" Predictions RMS = 0.068 (n=4178)
- "Good" Predictions RMS = 0.203 (n=44712)
- "Poor" Predictions RMS = 0.305 (n=7689)
- "Bad" Predictions RMS = 0.467 (n=3157)
- "Lack" Predictions RMS = 0.435 (n=1154)

© 2001-2008 PERCH Solutions Ltd. - CSD/PERCH NMR Software Course, Tulsa 16-17.5.2008

Comparison

Model It	Database / HOSE code
<ul style="list-style-type: none"> + More realistic descriptors + Enables conformational studies + Finds data used in model - Slow especially in adaptive mode - Sensitive to conformational errors 	<ul style="list-style-type: none"> + Very fast + Intrinsic link to data - Huge data base needed to find sufficient HOSE-spheres - Limited stereochemistry - Poor interpolation (Finds more than predicts)

© 2001-2008 PERCH Solutions Ltd. - CSD/PERCH NMR Software Course, Tulsa 16-17.5.2008

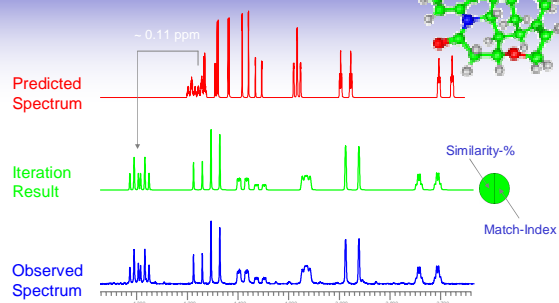
Limitations of NMR Databases

- Spectral Databases
 - Often unassigned
 - Limited parameters (field-strength, solvents)
 - Missing information (temperature, concentration)
- Non-Spectral Databases
 - Based on unverifiable literature data
 - Usually not completely assigned
 - Often based on rough first-order analysis only

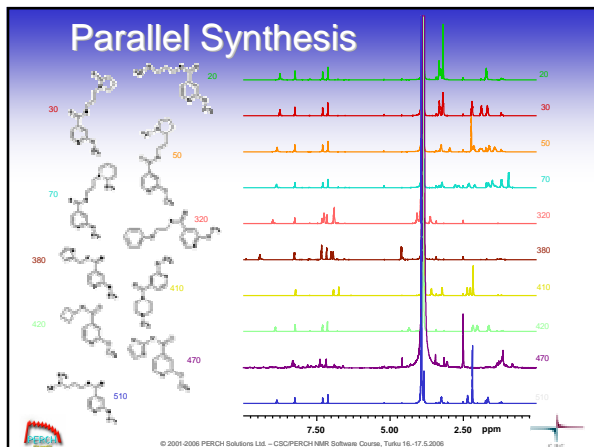
How to improve this?



Structure Verification



Parallel Synthesis



Cross-Verification

