

# Tools for Structure Elucidation



Sandra Groscurth



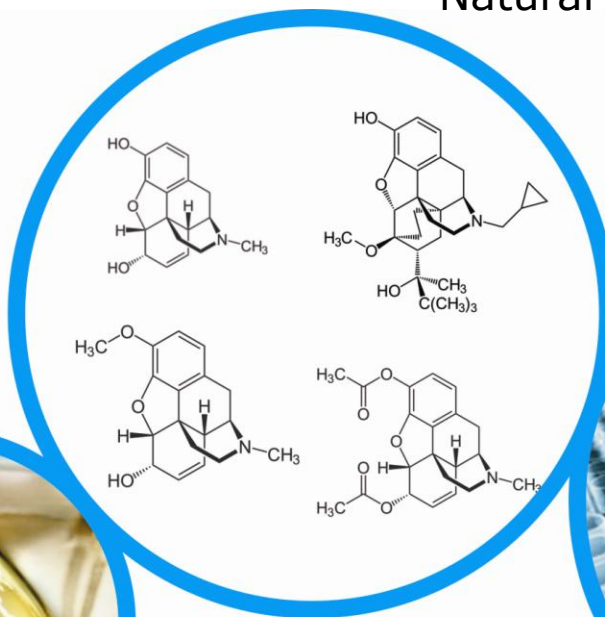
# Workflow in Natural Product Research



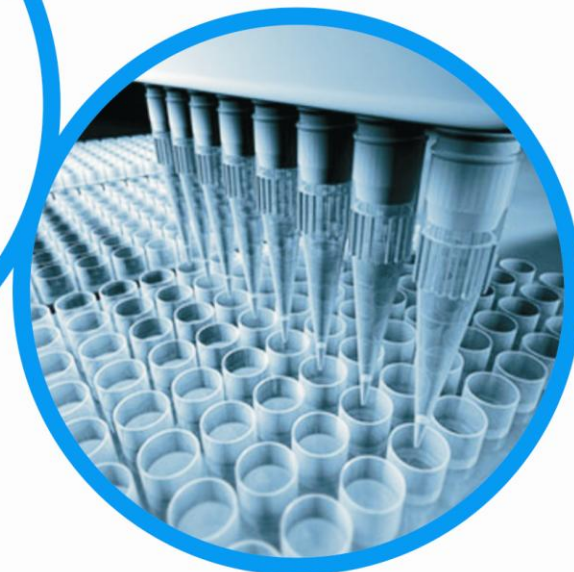
Organism



Natural Products

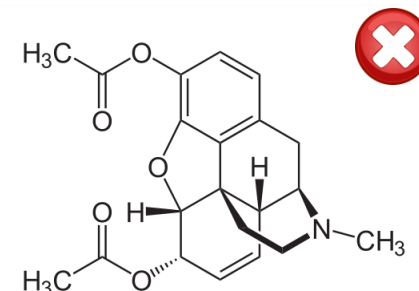
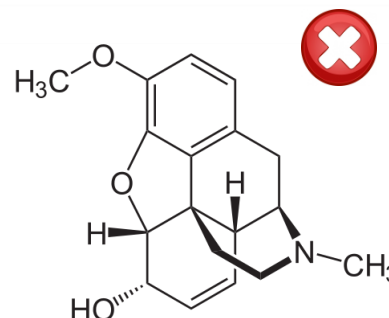
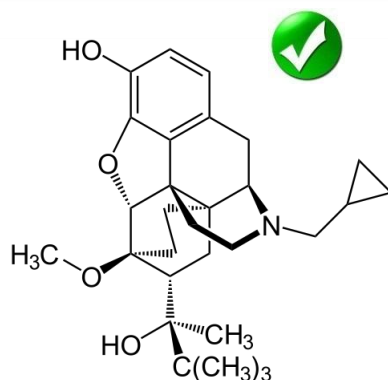
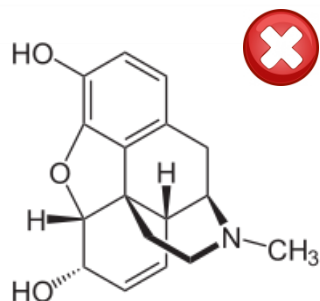


Extraction



Activity Assay

# Characterization of Unknown Structures



characterization of extracted compounds that show activity:

first of all, is it a new compound, is the structure unknown?

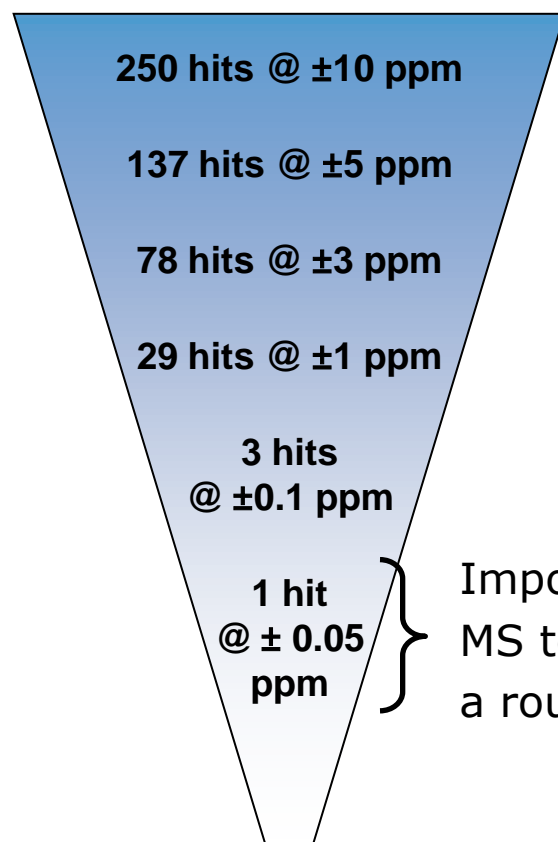
⇒ dereplication by MS/MS

⇒ molecular formula determination by mass spectrometry

# Structural Characterization by MS



Number of hits depending on mass accuracy



$m/z = 609.28$ :

10 ppm  $\Rightarrow$  0.006 Da

0.05 ppm  $\Rightarrow$  0.00003 Da  
0.03 mDa

Impossible for any  
MS technology on  
a routine basis

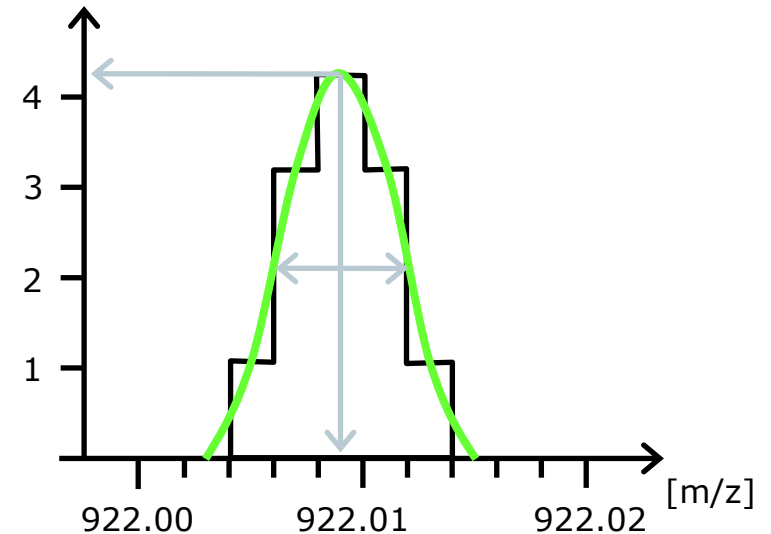
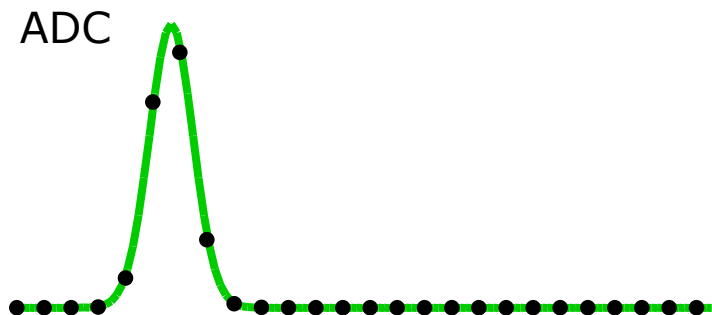
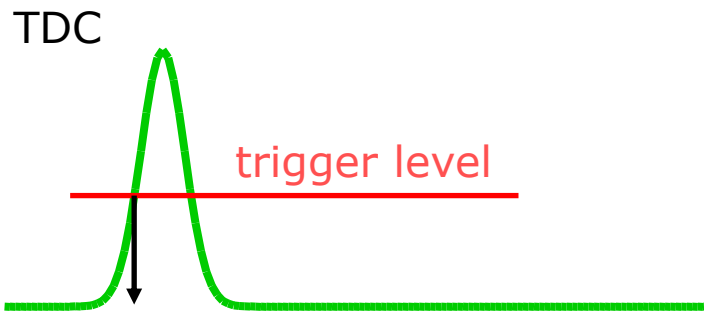


- resolution 20000 FWHM
- accuracy 1–2 ppm

# Accurate Mass Determination



detection of single ion signal:



2 GHz Digitizer = sampling rate 0.5 ns

# Time to Digital Converter



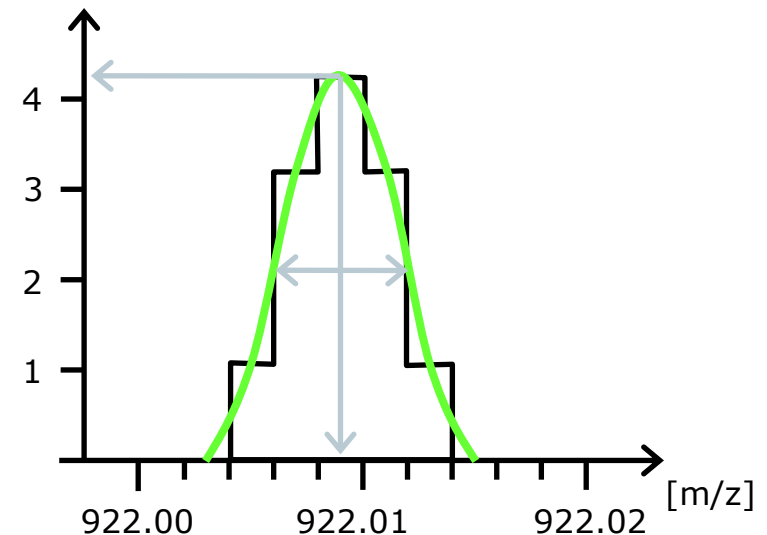
time information of the histogram peak is converted by interpolation and calibration into mass information

mass position = 922.009 m/z

intensity = 4 counts

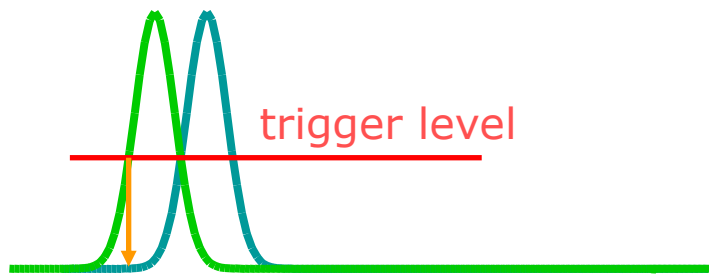
peak width (FWMH) = 0.06 Da

resolution =  $\frac{\text{mass position}}{\text{peak width}} = 15400$

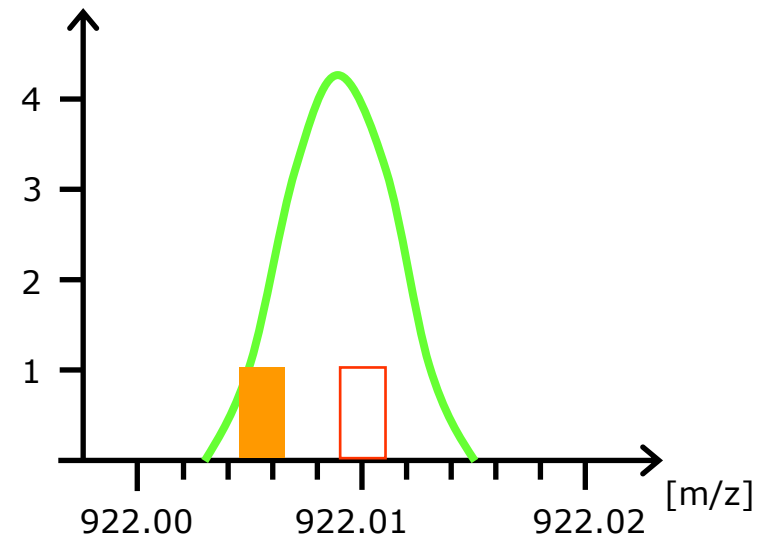


# Time to Digital Converter

with higher sample concentration two or more ions per acceleration pulse may reach the MCP detector



- ⇒ only the first ion causes a trigger event
- ⇒ the second ion is not converted



# Time to Digital Converter

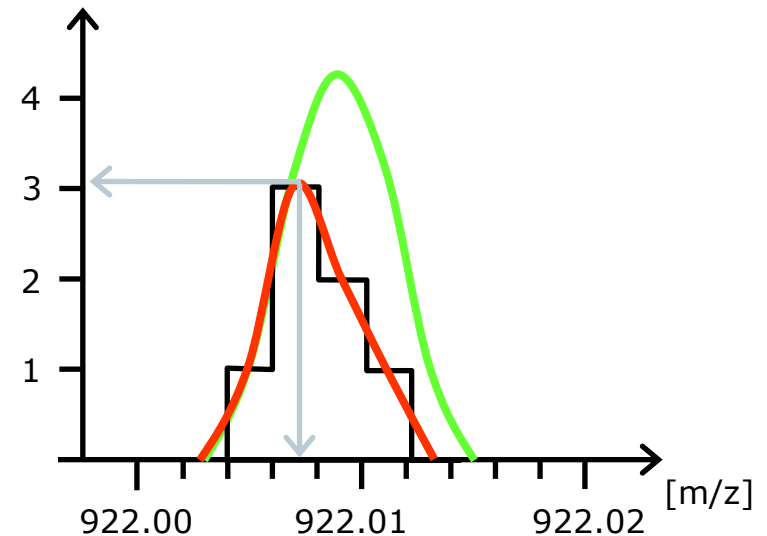


with higher sample concentration two or more ions per acceleration pulse may reach the MCP detector

in this example with the same number of ions reaching the detector, only 60% of these ions were converted

⇒ wrong intensity

⇒ wrong mass position



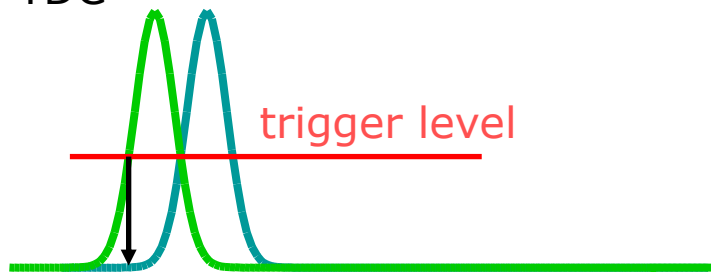


# Accurate Mass Determination

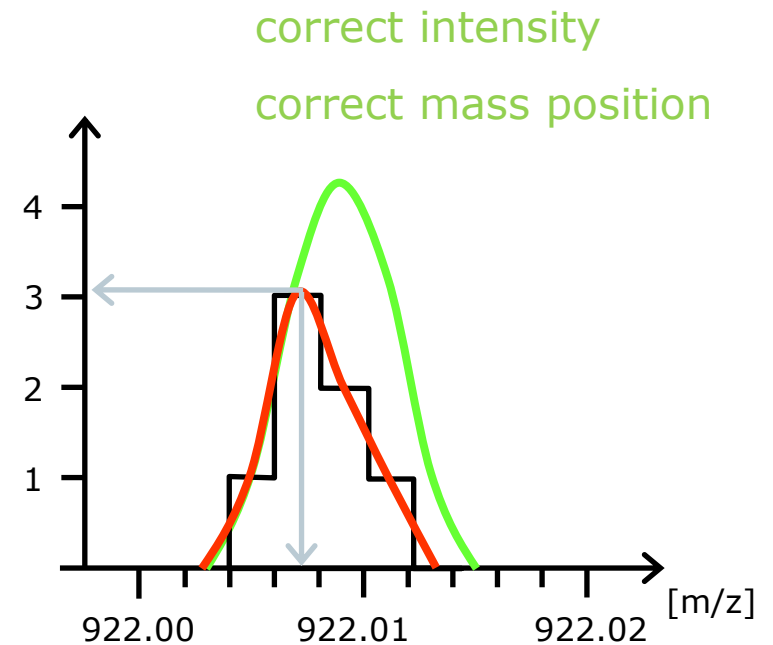
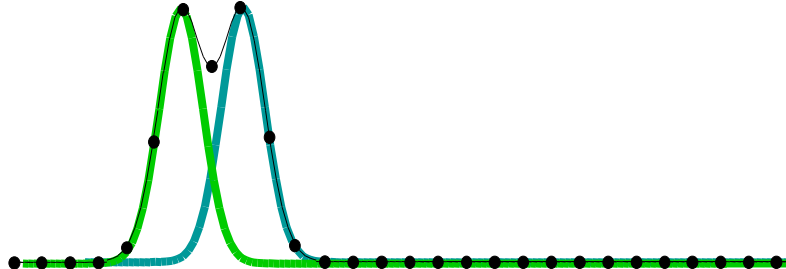


to avoid dead time effects: different digitizer technology

TDC



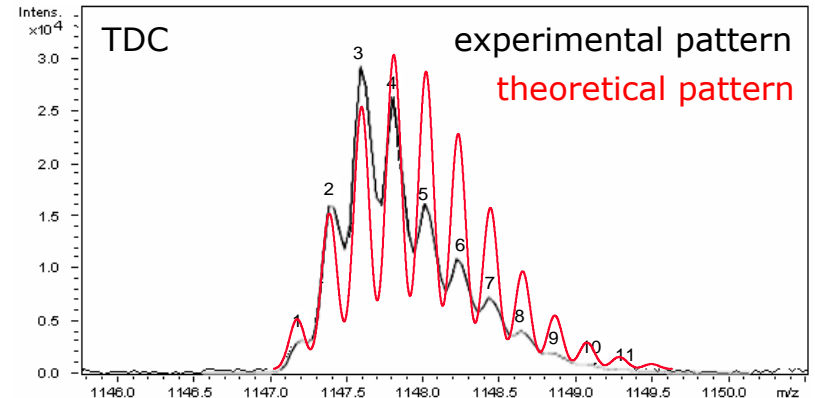
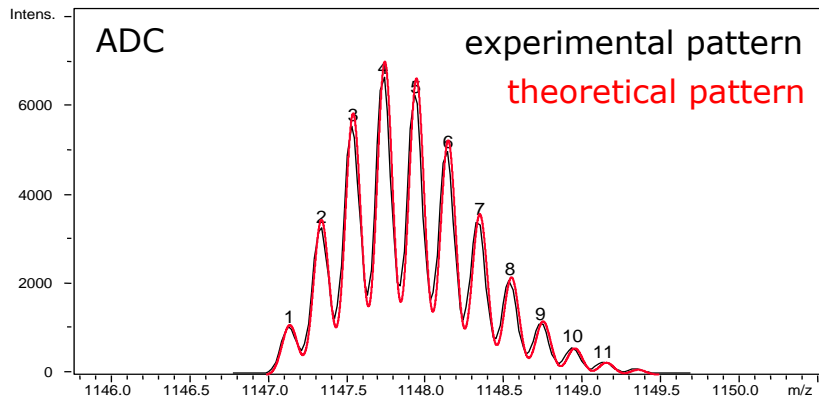
ADC



# Accurate Mass Determination

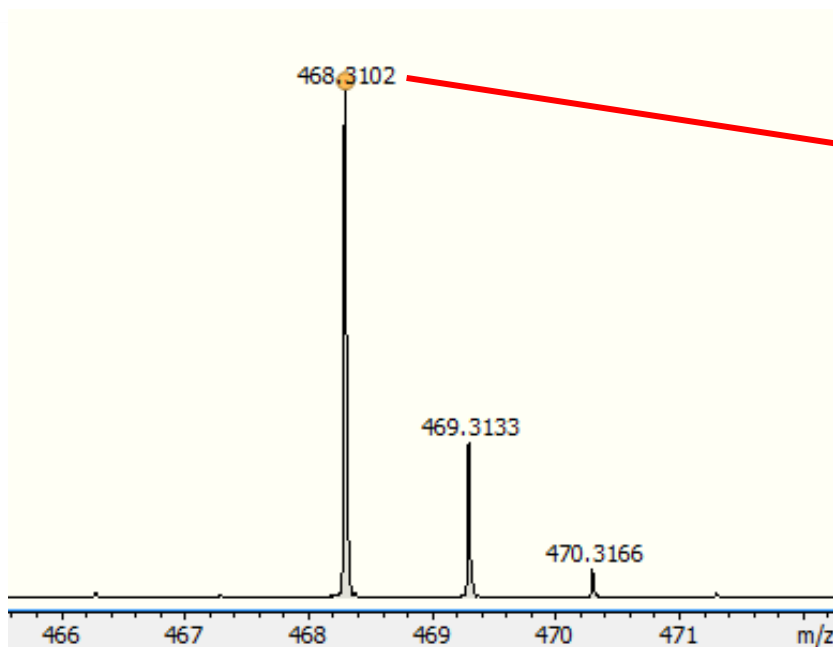


comparing true isotopic pattern of insulin



⇒ ADC digitizer technology preferred for any MS instrument

# SmartFormula – MS Data Interpretation



SmartFormula Manually

Lower formula:

Upper formula:

Note: for  $m < 2000$  the elements C, H, N, and O are considered implicitly.

Adducts, pos.  ☐ Collect adducts

Adducts, neg.

Measured  $m/z$   Tolerance:  mDa Charge:

Meas. $m/z$	#	Ion Formula	Score	$m/z$	err [mDa]	err [ppm]	mSigma	rdl
468.3102	1	<b>C<sub>29</sub>H<sub>42</sub>NO<sub>4</sub></b>	100.00	468.3108	0.6	1.3	9.4	9.
	2	C <sub>30</sub> H <sub>38</sub> N <sub>5</sub>	34.74	468.3122	2.0	4.2	20.4	14.
	3	C <sub>14</sub> H <sub>38</sub> N <sub>13</sub> O <sub>5</sub>	20.20	468.3113	1.1	2.4	60.5	2.

☐ Automatically locate monoisotopic peak Maximum number of formulae

☒ Check rings plus double bonds Minimum  Maximum

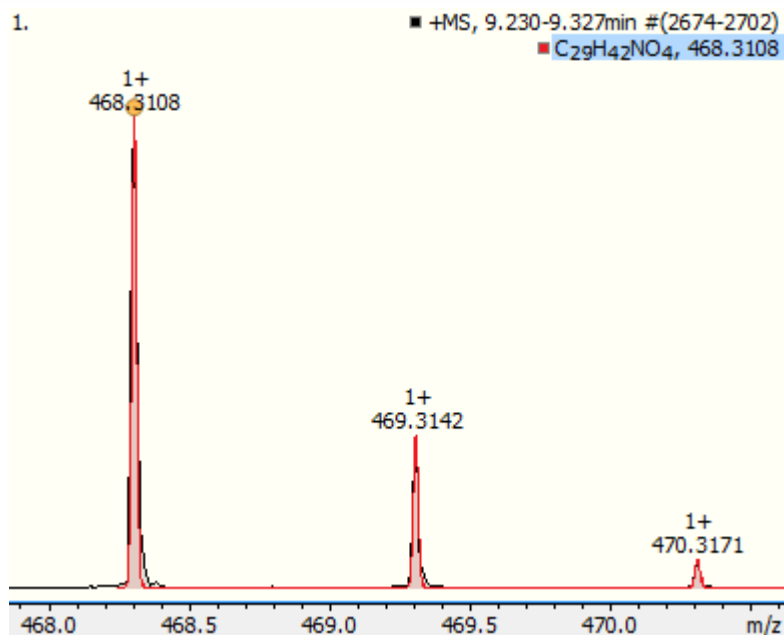
Electron configuration

☒ Filter H/C element ratio Minimum H/C:  Maximum H/C:

☒ Estimate carbon number ☒ Generate immediately

⇒ unambiguous elemental composition determination by combined accurate mass and isotopic pattern information

# SmartFormula – MS Data Interpretation



SmartFormula Manually

Lower formula:  Generate  
Upper formula:  Help

Note: for  $m < 2000$  the elements C, H, N, and O are considered implicitly.

Adducts, pos. M+H Collect adducts  
Adducts, neg. M-H

Measured m/z 468.3102 Tolerance: 2 mDa Charge: 1

Meas. m/z	#	Ion Formula	Score	m/z	err [mDa]	err [ppm]	mSigma	rdl
468.3102	1	$C_{29}H_{42}NO_4$	100.00	468.3108	0.6	1.3	9.4	9.
	2	$C_{30}H_{38}N_5$	34.74	468.3122	2.0	4.2	20.4	14.
	3	$C_{14}H_{38}N_{13}O_5$	20.20	468.3113	1.1	2.4	60.5	2.

☐ Automatically locate monoisotopic peak Maximum number of formulae 500  
☒ Check rings plus double bonds Minimum -0.5 Maximum 40  
Electron configuration even  
☒ Filter H/C element ratio Minimum H/C: 0 Maximum H/C: 3  
☒ Estimate carbon number ☒ Generate immediately  
Copy to SmartFormula Parameters Show Pattern

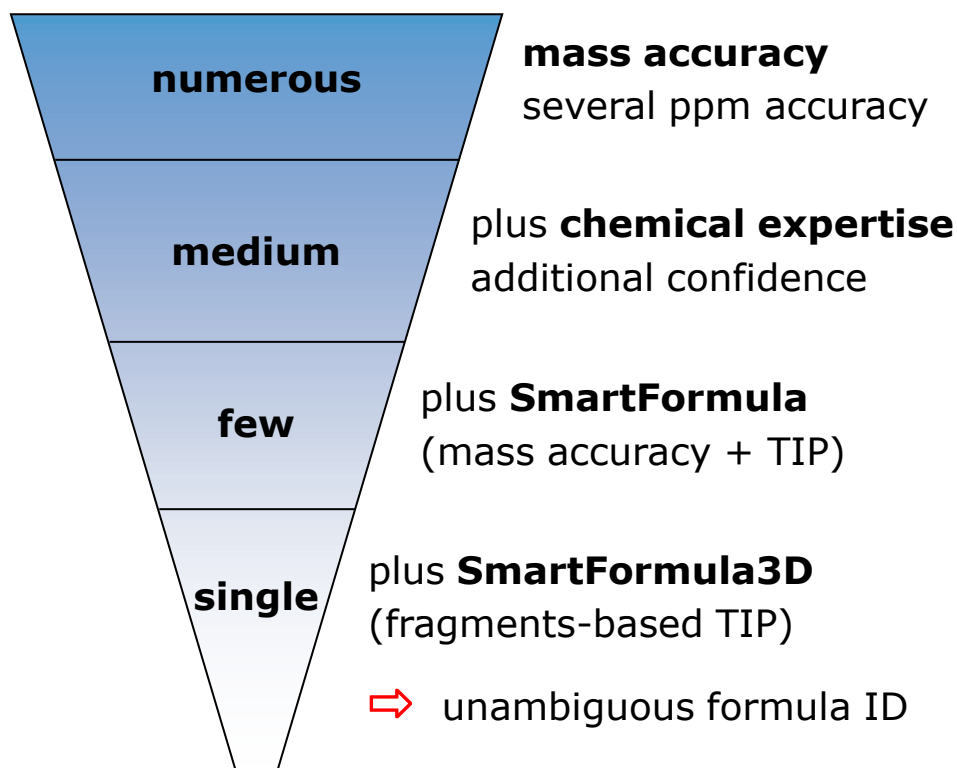
Overlay of **measured** and **simulated** spectrum:

⇒ sum formula  $C_{29}H_{42}NO_4$  shows perfect match for the isotopic patterns

# Accurate molecular formula from MS/MS



# possible formulae

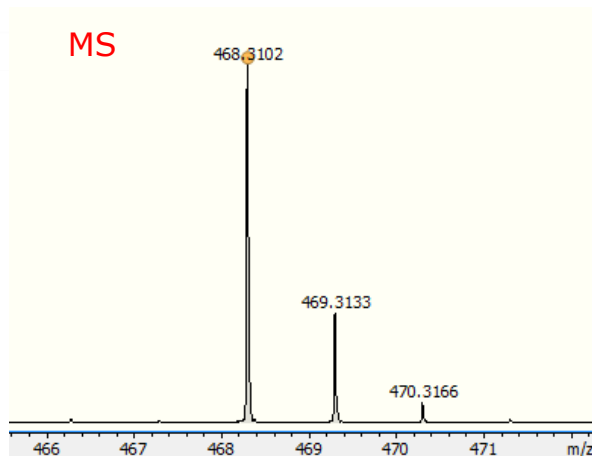


⇒ definitive molecular formulae

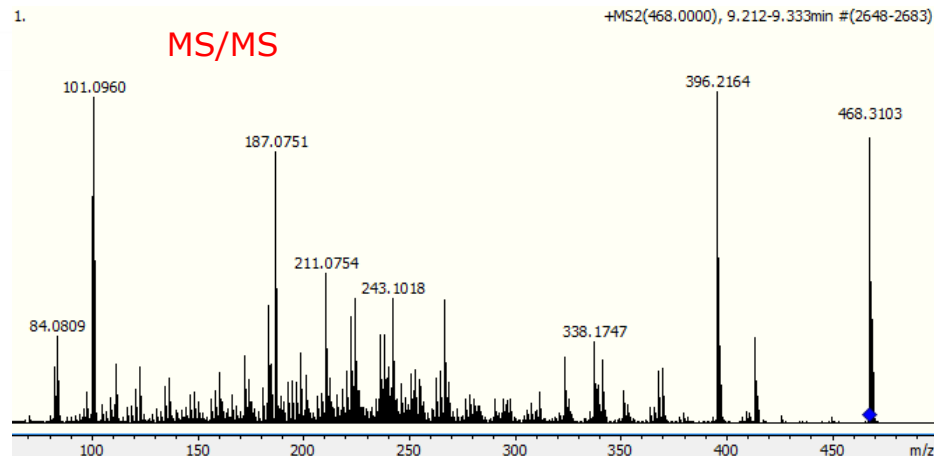


- resolution 40-60000 FWHM
- accuracy 1 ppm (MS & MS/MS)
- 20 full spectra/sec

# Accurate molecular formula from MS/MS



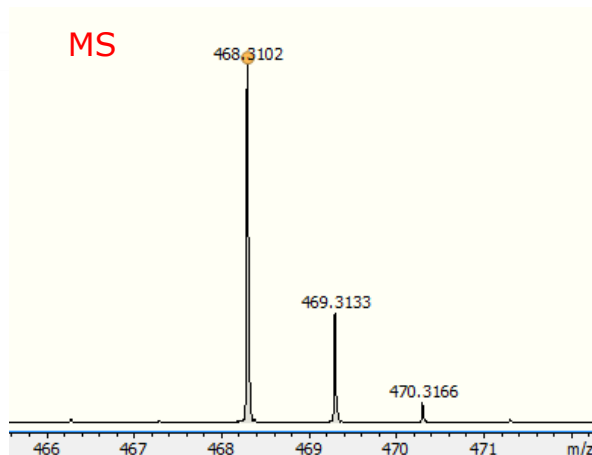
Parent



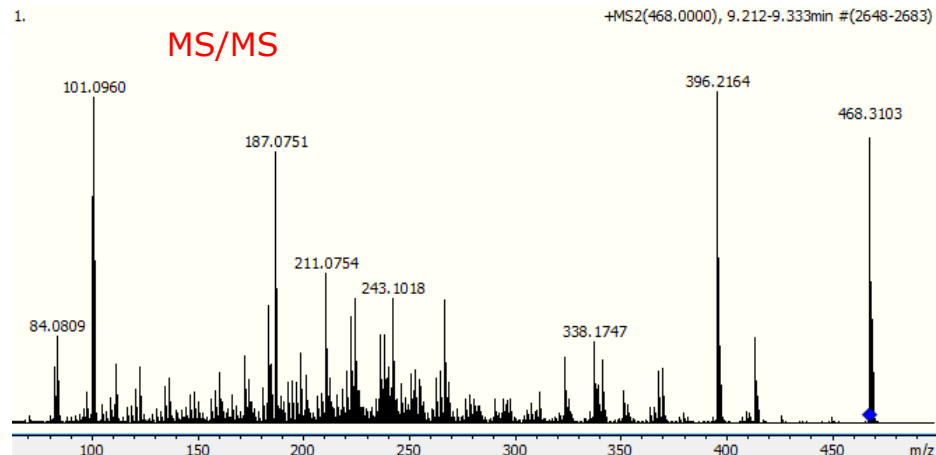
Fragments

- elemental formulae for parent ions by SmartFormula
  - elemental formulae for fragment ions by SmartFormula
- ⇒ every 'true' fragment formula must be a sub-set of the 'true' parent formula
- ⇒ SmartFormula3D

# Accurate molecular formula from MS/MS



Parent



Fragments

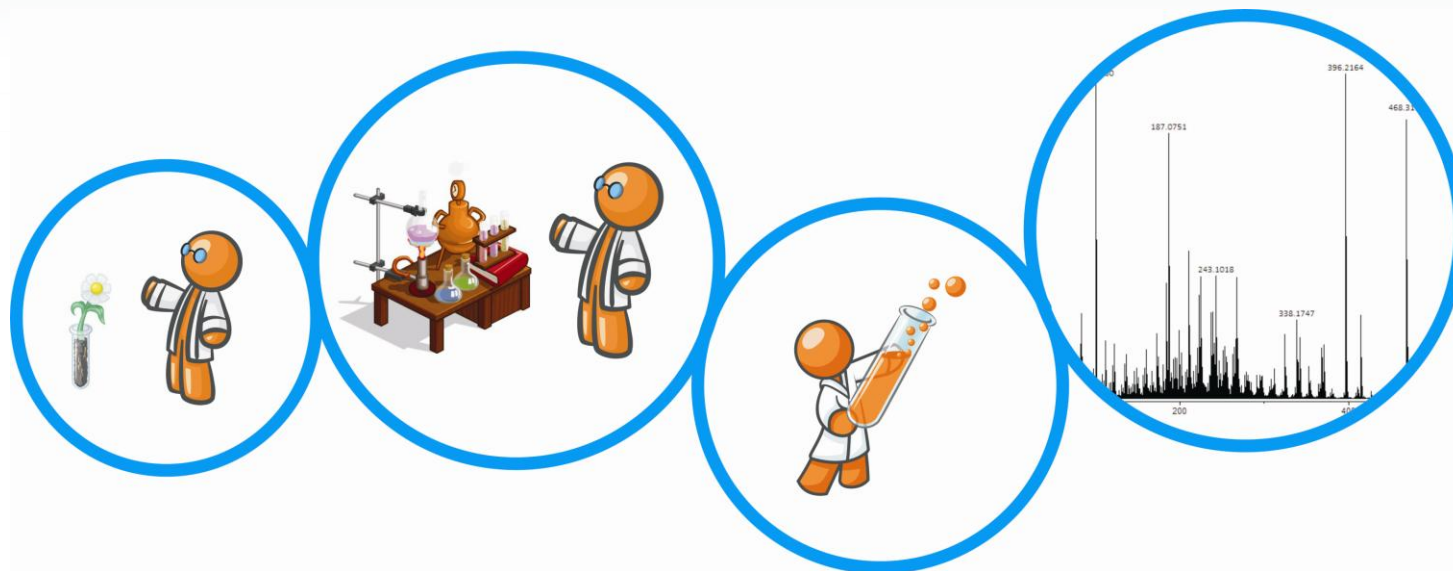
	SumFormula	m/z calc	err[mDa]	err[ppm]	mSigma	eConf
<input checked="" type="checkbox"/>	C <sub>29</sub> H <sub>42</sub> NO <sub>4</sub>	468.3108	0.6	1.2	21.2	even

proposed molecular formulae  
of precursor ion after  
combination of MS and  
MS/MS information

	SumFormula	SumFormula Loss	m/z Loss	err[mDa] Loss
<input type="checkbox"/>	C <sub>24</sub> H <sub>30</sub> NO <sub>4</sub>	C <sub>5</sub> H <sub>12</sub>	72.0938	0.1
<input type="checkbox"/>	C <sub>24</sub> H <sub>30</sub> NO <sub>4</sub>	C <sub>5</sub> H <sub>12</sub>	72.0938	0.1
<input type="checkbox"/>	C <sub>22</sub> H <sub>26</sub> NO <sub>3</sub>	C <sub>7</sub> H <sub>16</sub> O	116.1189	1.2
<input type="checkbox"/>	C <sub>22</sub> H <sub>26</sub> NO <sub>3</sub>			1.2
<input type="checkbox"/>	C <sub>18</sub> H <sub>15</sub> O			0.1
<input type="checkbox"/>	C <sub>12</sub> H <sub>11</sub> O <sub>2</sub>			0.3
<input type="checkbox"/>	C <sub>6</sub> H <sub>13</sub> O			0.5

molecular formulae of  
fragment ions

# Accurate molecular formula from MS/MS



characterization of unknown compound by MS/MS:

molecular formula based on TIP of MS and MS/MS spectra

⇒  $C_{29}H_{41}NO_4$

⇒ structure elucidation by NMR



# NMR Spectra for Structure Elucidation



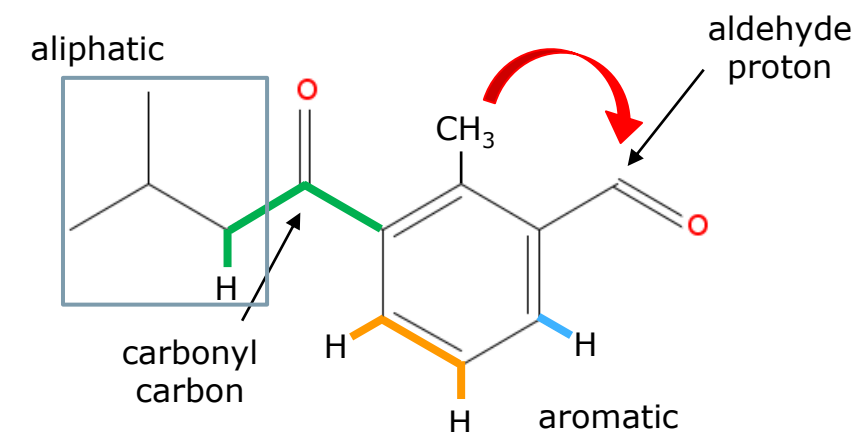
- 1D NMR spectra

- 1D Proton
- 1D Carbon

chemical shifts  $\Rightarrow$  functional groups

- 2D NMR spectra

- **HSQC**  $\Rightarrow$   $^1J_{CH}$
- **COSY**  $\Rightarrow$   $^3J_{HH}$
- **HMBC**  $\Rightarrow$   $^2J_{CH}/^3J_{CH}$
- **TOCSY**  $\Rightarrow$  spin system
- **ROESY**  $\Rightarrow$  through space



correlations  $\Rightarrow$  neighboring atoms

experiment	rel. sensitivity
------------	------------------

$^1\text{H}$	1.00
--------------	------

$^{13}\text{C}$	0.01
-----------------	------

HSQC	0.25
------	------

COSY	0.20
------	------

HMBC	0.07
------	------

TOCSY	0.12
-------	------

ROESY	0.07
-------	------



- NMR spectra for structure elucidation are less sensitive
- sample amounts are typically very little

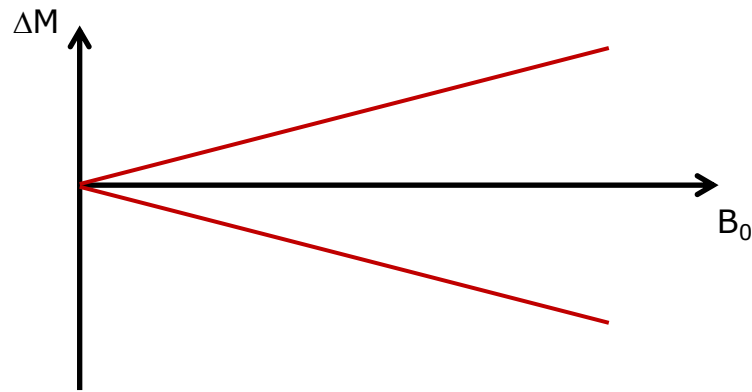
in order to reduce the experimental time

⇒ increase NMR sensitivity

# Increase NMR Sensitivity



- increase sensitivity by increasing the magnetic field strength

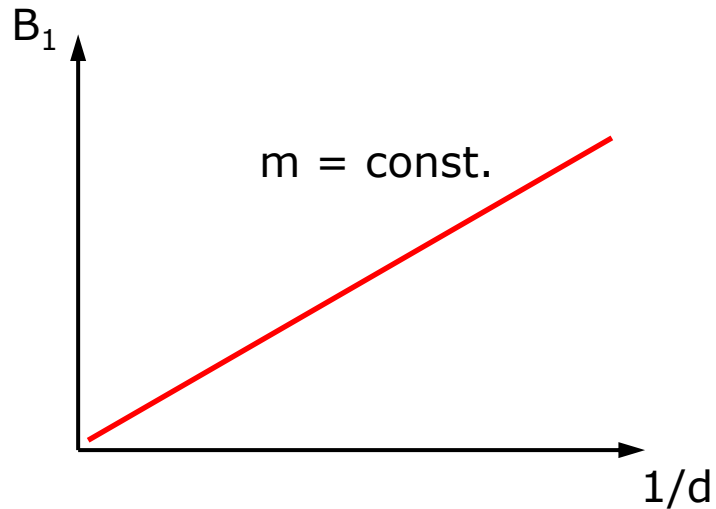


- increase (mass) sensitivity through probe technology
  - small volume probes: increase mass sensitivity
  - CryoProbes: decrease noise and thus increase signal to noise
  - combine both: best signal to noise and highest mass sensitivity

# Small Volume NMR Probes

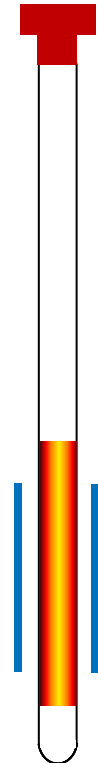


benefit of small volume NMR probes: increase mass sensitivity



relevant parameters:

- distance  $^1\text{H}$  coil to NMR tube
- diameter of NMR tube



⇒ same number of spins increases sensitivity for small volume probes

- no problems with solvent impurities (high analyte concentration even with very small sample amounts)
- no problems with solvent suppression (no radiation damping, etc.)
- advantage of working with small sample amounts
  - often easier to get clean samples in small amounts
  - no danger to overload HPLC columns
  - possibility to use analytical HPLC columns and fraction collect (better separation with small sample amounts)

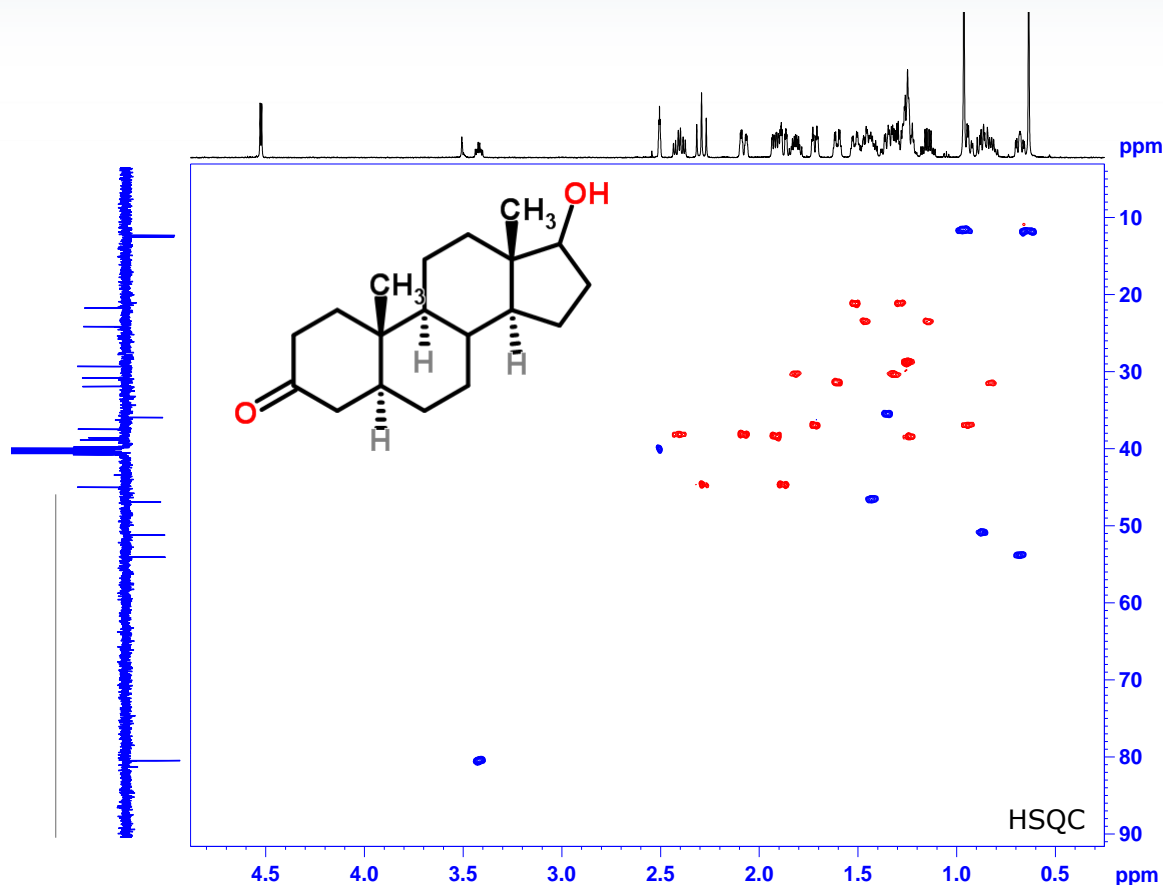
# Small Volume NMR Probes



Probehead Diameter [mm]	1	1.7	3	5
Sample Volume [μl]	5	30	180	500
Recommended Concentration [mM]	30	10	2.2	1
Sample Amount [μmol]	0.15	0.30	0.40	0.50
Mass Sensitivity	~ 4	~ 2.8	~ 1.4	1
Experiment Time	1	2	8	16

⇒ **0.6 mg** of compound is enough to acquire all required 1D and 2D NMR spectra on a 1.7mm RT @ 400 MHz over the weekend

# Small Volume NMR Probes: 1.7mm RT



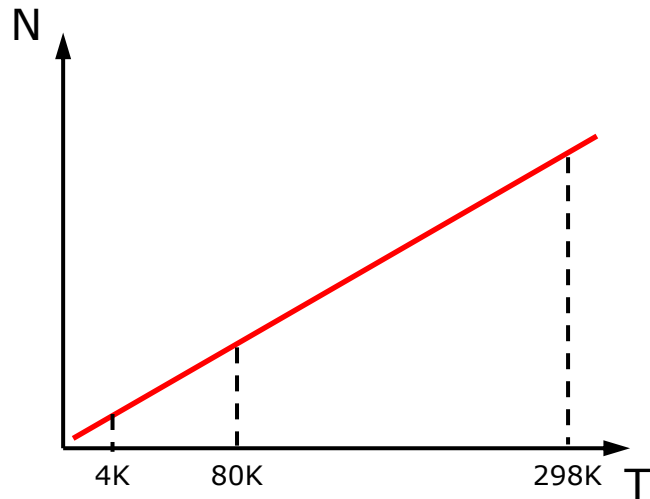
1D $^1\text{H}$	1 scan
COSY	20 min
TOCSY	45 min
NOESY	4 h
HSQC	10 min
HMBC	3.5 h
H2BC	4 h
1D $^{13}\text{C}$	8 h

120 $\mu\text{g}$  of a Steroid ( $\approx 350\text{g/mol}$ ) in 30 $\mu\text{l}$  ( $\approx 11\text{mM}$ ) < 24 h on 600MHz

# Cryogenic NMR Probes



benefit of cryogenic NMR probes: decrease noise



$$S/N \sim \frac{M \cdot B_0}{\sqrt{R_c(T_c + TPA) + R_s(T_s + TPA)}}$$

M: magnetization  
B<sub>0</sub>: static field  
R<sub>c</sub>: 'resistance in coil'  
R<sub>s</sub>: 'resistance in sample'  
T<sub>c</sub>: coil temperature  
T<sub>s</sub>: sample temperature  
TPA: preamplifiers equivalent noise temperature

cooling NMR coil reduces electronic noise

⇒ increases signal/noise



# Cryogenic NMR Probes: huge selection



He cooled

<b>CryoProbes</b>	<b>400</b>	<b>500</b>	<b>600</b>	<b>700</b>	<b>800</b>	<b>850</b>	<b>900</b>	<b>950</b>	<b>1000</b>
<b><i>DCH C-H-D</i></b>	✓	✓	✓	✓					
<b><i>TCI H-C/N-D</i></b>	✓	✓	✓	✓	✓	✓	✓	✓	✓
<b><i>TCI H-C/N-D 1.7mm</i></b>		✓	✓	✓	✓	✓			
<b><i>QCI H/P-C/N/D</i></b>		✓	✓	✓	✓	✓			
<b><i>QCI H/F-C/N/D</i></b>		✓	✓	✓					
<b><i>TXO C/H-N-D</i></b>		✓	✓	✓	✓				
<b><i>DUX 2H</i></b>		✓	✓	✓	✓	✓			
<b><i>DUL-C-H-D 10mm</i></b>	✓	✓	✓						
<b><i>BBFO</i></b>		✓							
<b><i>BBO H&amp;F</i></b>	✓	✓	✓	✓					

N<sub>2</sub> cooled

<b>CryoProbe Prodigy</b>	<b>400</b>	<b>500</b>	<b>600</b>	<b>700</b>
<b><i>Prodigy BBO</i></b>	✓	✓	✓	
<b><i>Prodigy TCI</i></b>		✓	✓	✓

# Highest Versatility: BBO CryoProbe



- $^{13}\text{C}$  sensitivity 4 \* BBO RT
- $^{15}\text{N}$  sensitivity 4 \* BBO RT
- $^1\text{H}$  sensitivity 3 \* BBO RT
- observe and inverse detection
- cold preamplifiers for BB/ $^1\text{H}/^2\text{H}$
- z-gradient
- ATM compatible
- 0 ° – 80 ° C (standard sample temperature range)
- available for 400 – 700 MHz

# Maximum Carbon Sensitivity: DCH



1D  $^1\text{H}$

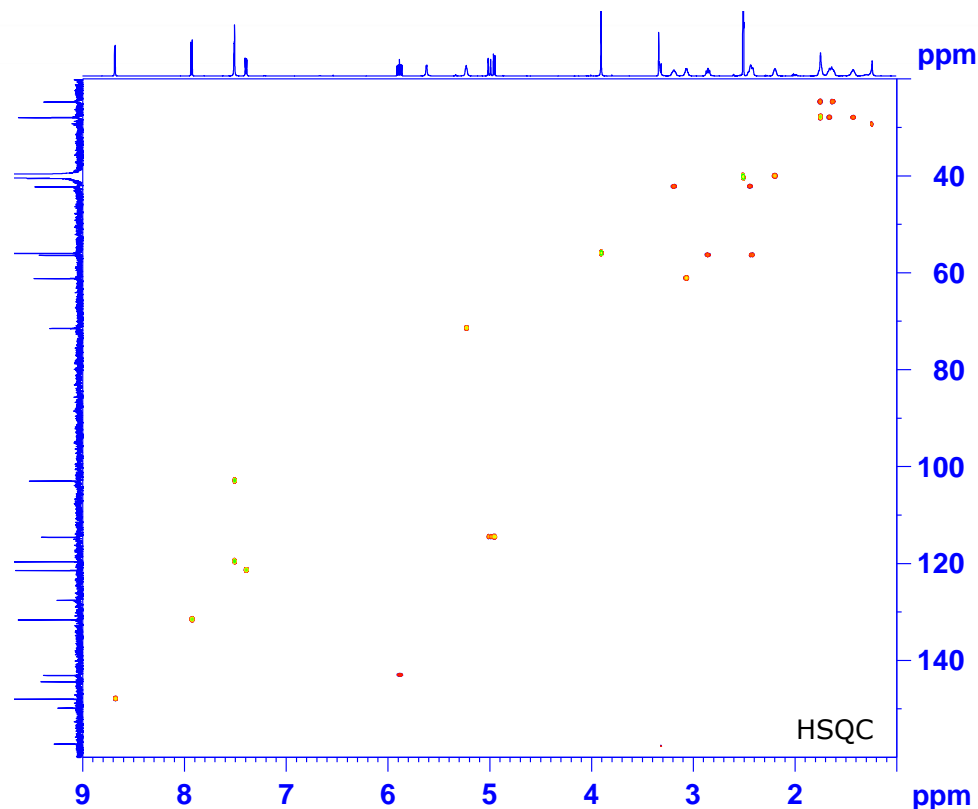
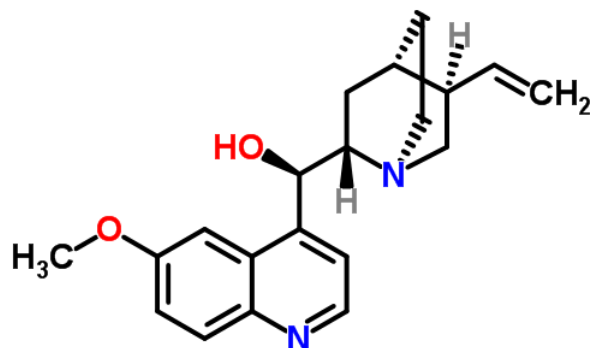
1 min

HSQC

10 min

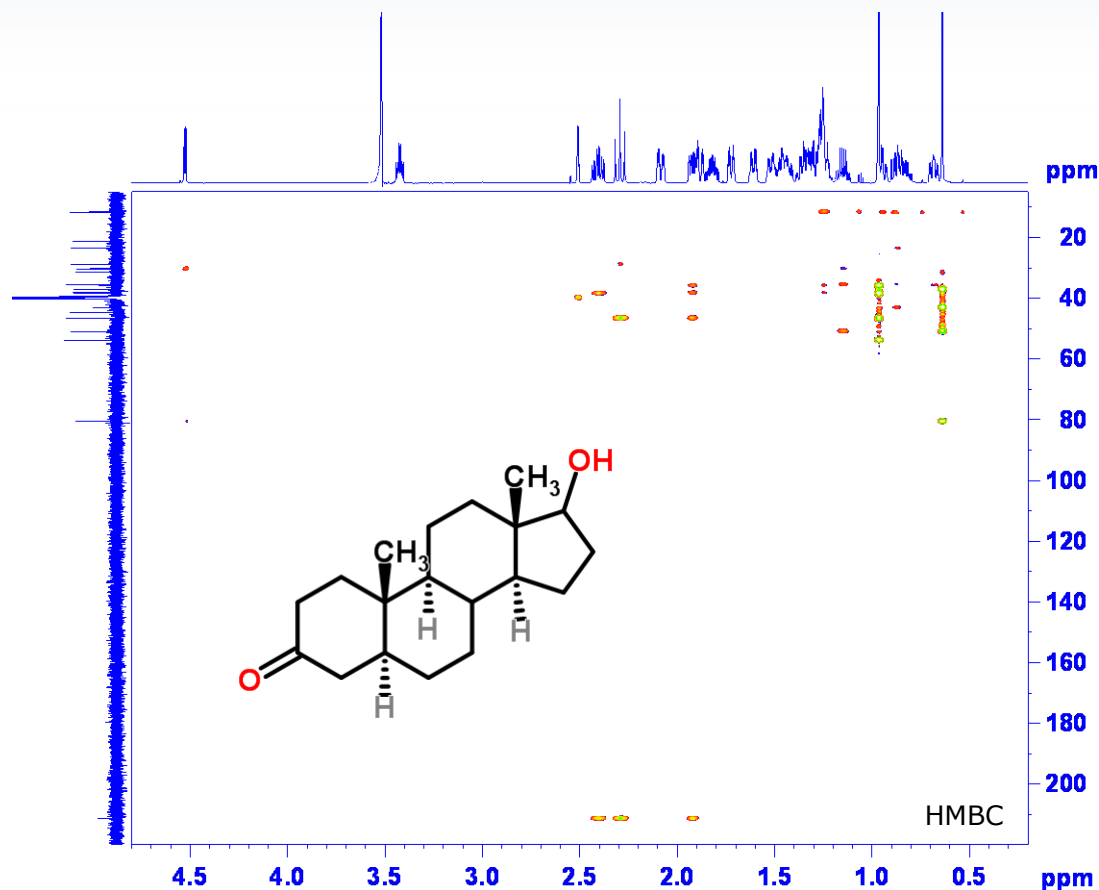
1D  $^{13}\text{C}$

30 min



0.5mg of Quinine ( $\approx 225\text{mol/l}$ ) in 600 $\mu\text{l}$  ( $\approx 2.4\text{mM}$ ) on 700MHz

# Maximum Mass Sensitivity: 1.7 CryoProbe



1D  $^1\text{H}$

1 scan

HSQC

5 min

HMBC

10 min

1D  $^{13}\text{C}$

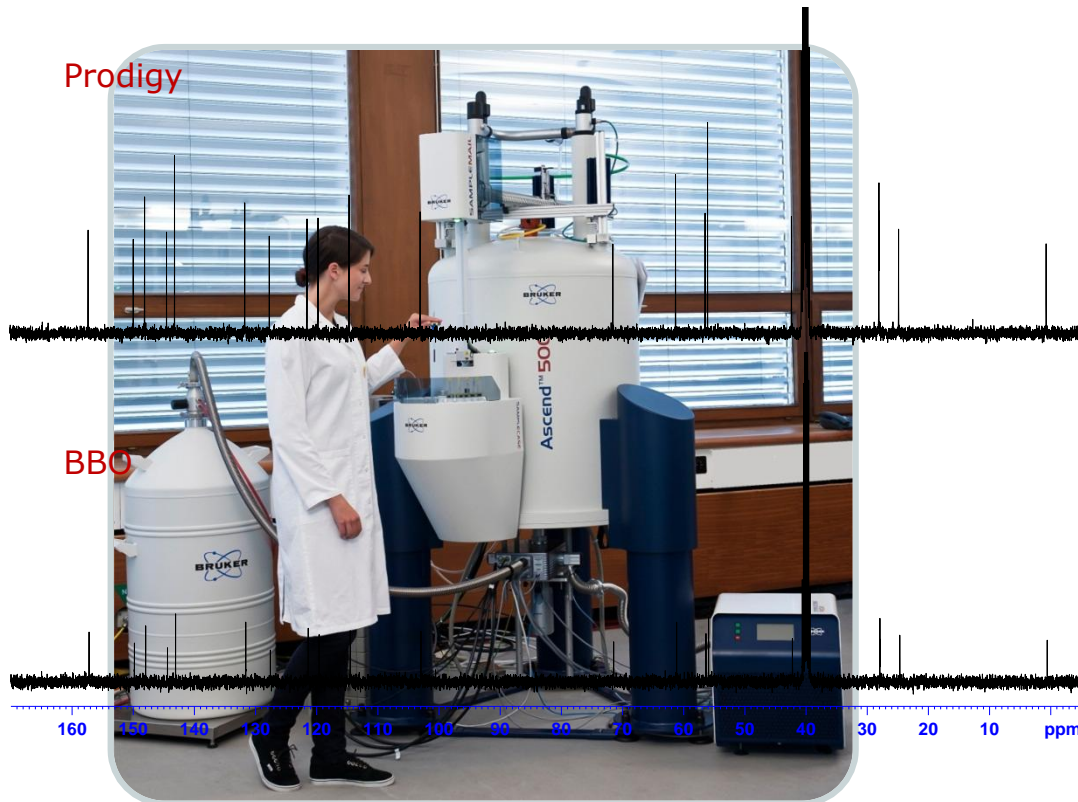
40 min

120 $\mu\text{g}$  of a Steroid ( $\approx 350\text{g/mol}$ ) in 30 $\mu\text{l}$  ( $\approx 11\text{mM}$ ) < 2 h on 600MHz

# CryoProbe Prodigy



## CryoProbe Prodigy: alternative at medium field

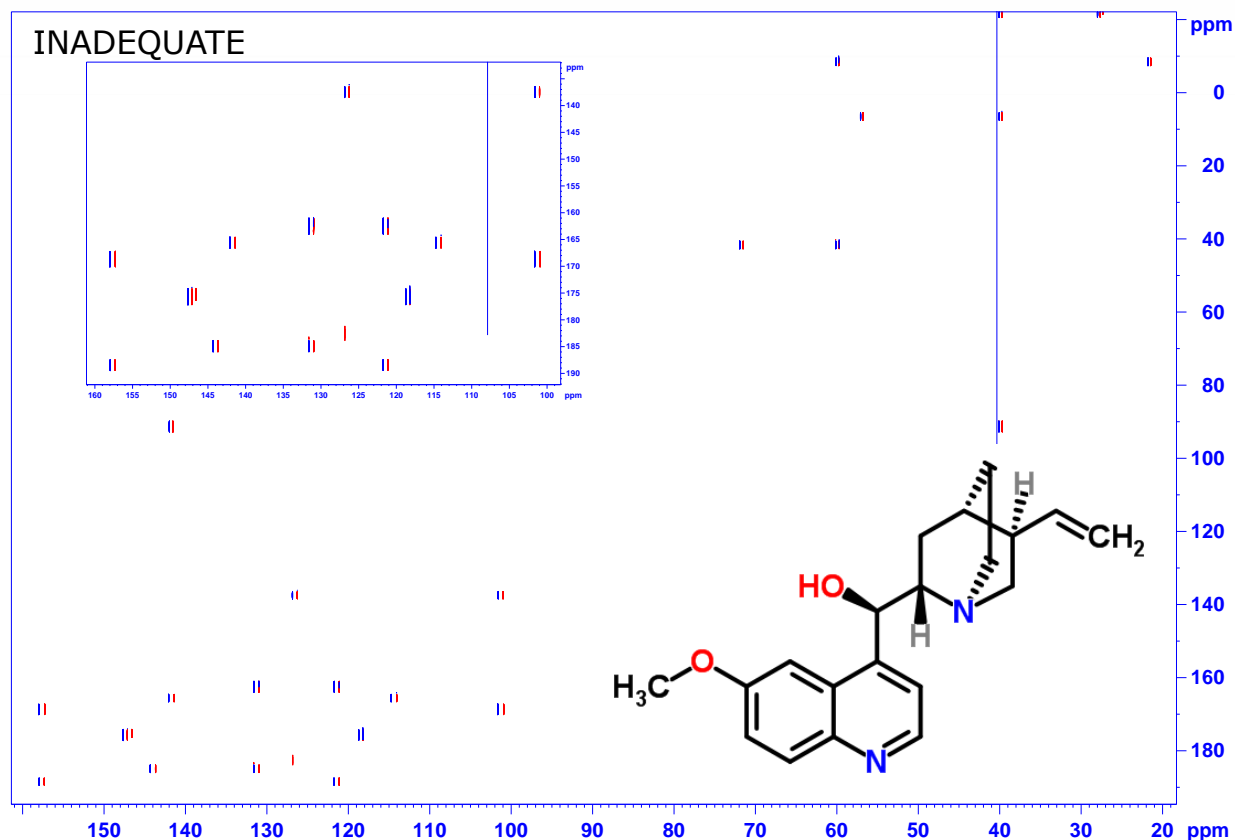


50mM Quinine, 32 scans

- liquid nitrogen for cryo-cooling
- small footprint
- Prodigy BBO
- Prodigy TCI

⇒  $^{13}\text{C}$  sensitivity 2-3 \* BBO

# CryoProbe Prodigy: $^{13}\text{C}$ sensitivity



17.5mg of Quinine ( $\approx 225\text{g/mol}$ ) in 600 $\mu\text{l}$  ( $\approx 100\text{mM}$ ): expt. time 16 h

# Summary NMR Probes



## small volume probes

⇒ increasing mass sensitivity

- 1.7mm RT
- 1mm RT

## cryogenic probes

⇒ decrease noise and increase signal/noise

### N<sub>2</sub> cooled

- Prodigy ⇒ sensitivity boost at affordable price

### He cooled

- BBO ⇒ highest versatility
- DCH ⇒ highest carbon sensitivity
- 1.7mm ⇒ highest mass sensitivity



1D Proton, 1D Carbon  
functional groups



2D HSQC

$^1J_{\text{CH}}$  correlation (CH, CH<sub>2</sub>, CH<sub>3</sub>)



2D COSY

$^3J_{\text{HH}}$  correlation

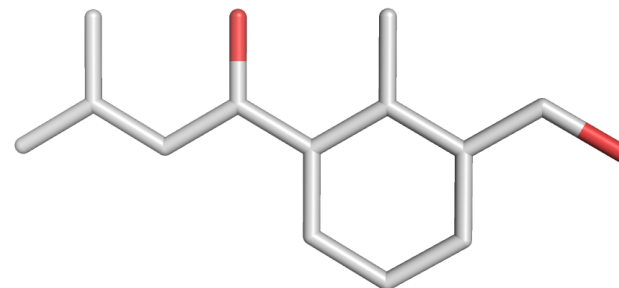


2D HMBC, 2D TOCSY

$^2J_{\text{CH}}/^3J_{\text{CH}}$  correlation, spin system

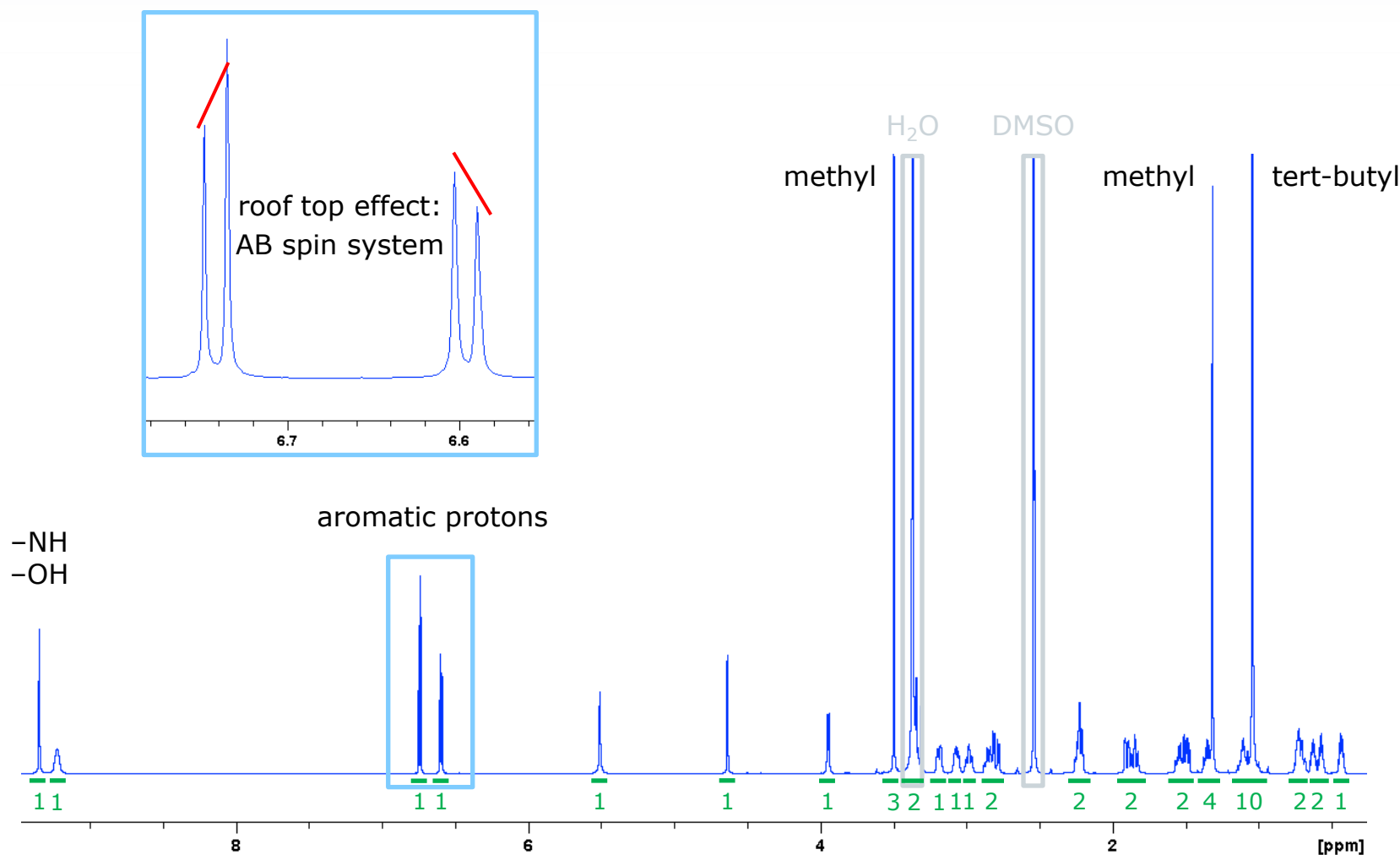


2D ROESY, 2D NOESY  
through space correlation





# NMR Spectroscopy – 1D Proton

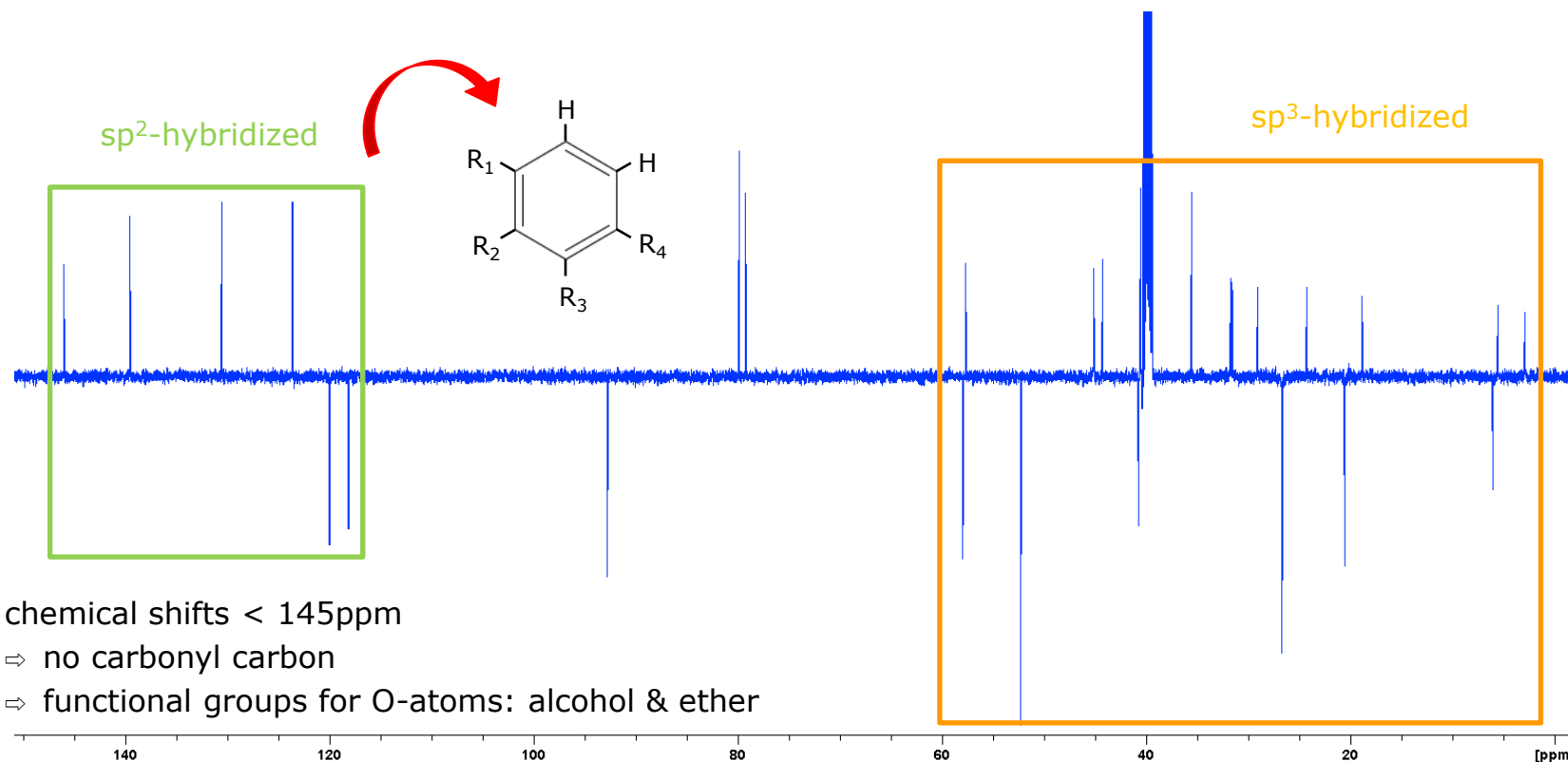


# NMR Spectroscopy – 1D Carbon



APT (**A**ttached **P**roton **T**est) spectrum for assigning multiplicities in  $^{13}\text{C}$  spectra

- ⇒ positive signals:  $\text{CH}_2$ -groups & quaternary C
- ⇒ negative signals: CH-groups &  $\text{CH}_3$ -groups

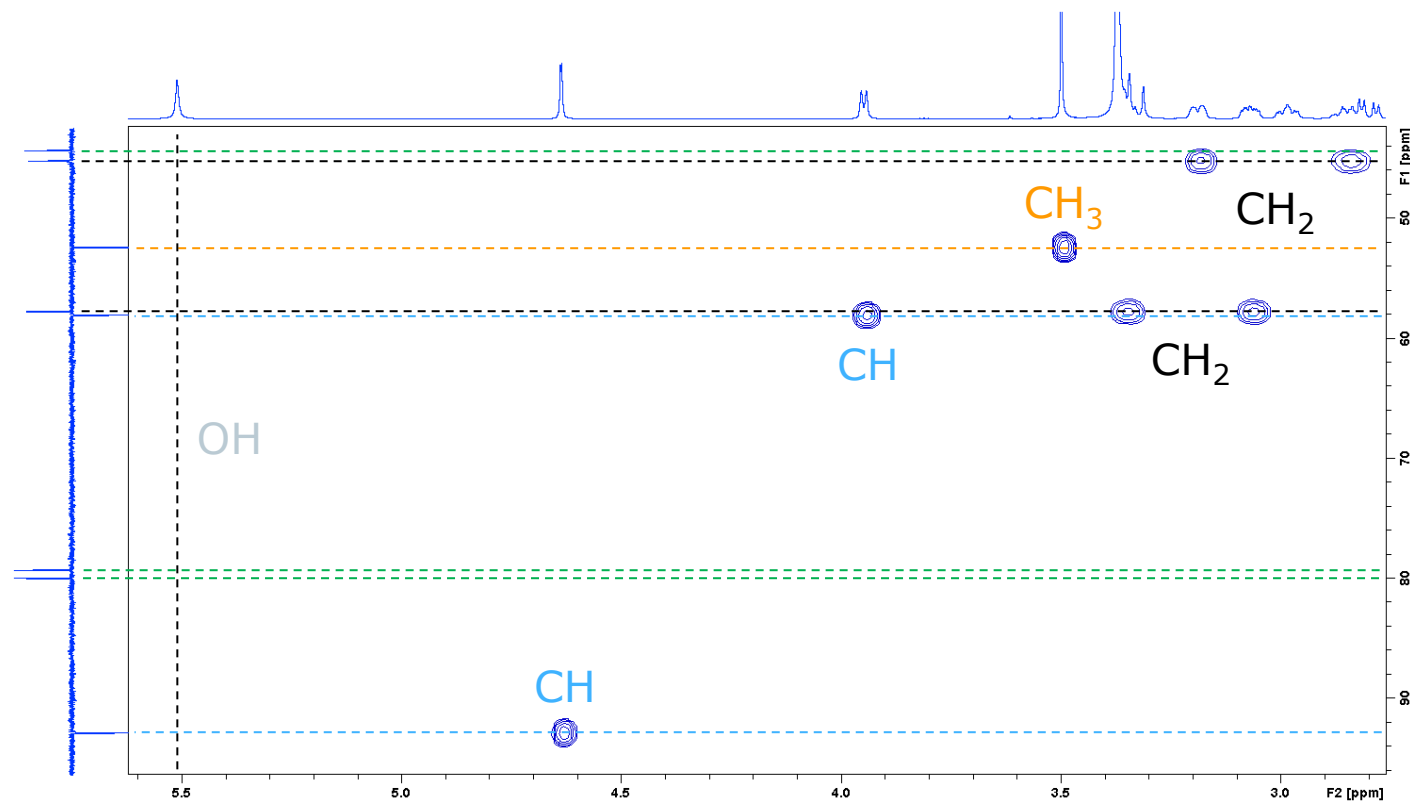
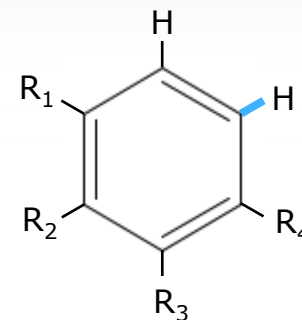


# NMR Spectroscopy – 2D HSQC



HSQC spectrum  $\Rightarrow$   $^1J_{CH}$  correlation

(identifying protons & carbons that are bound to each other)

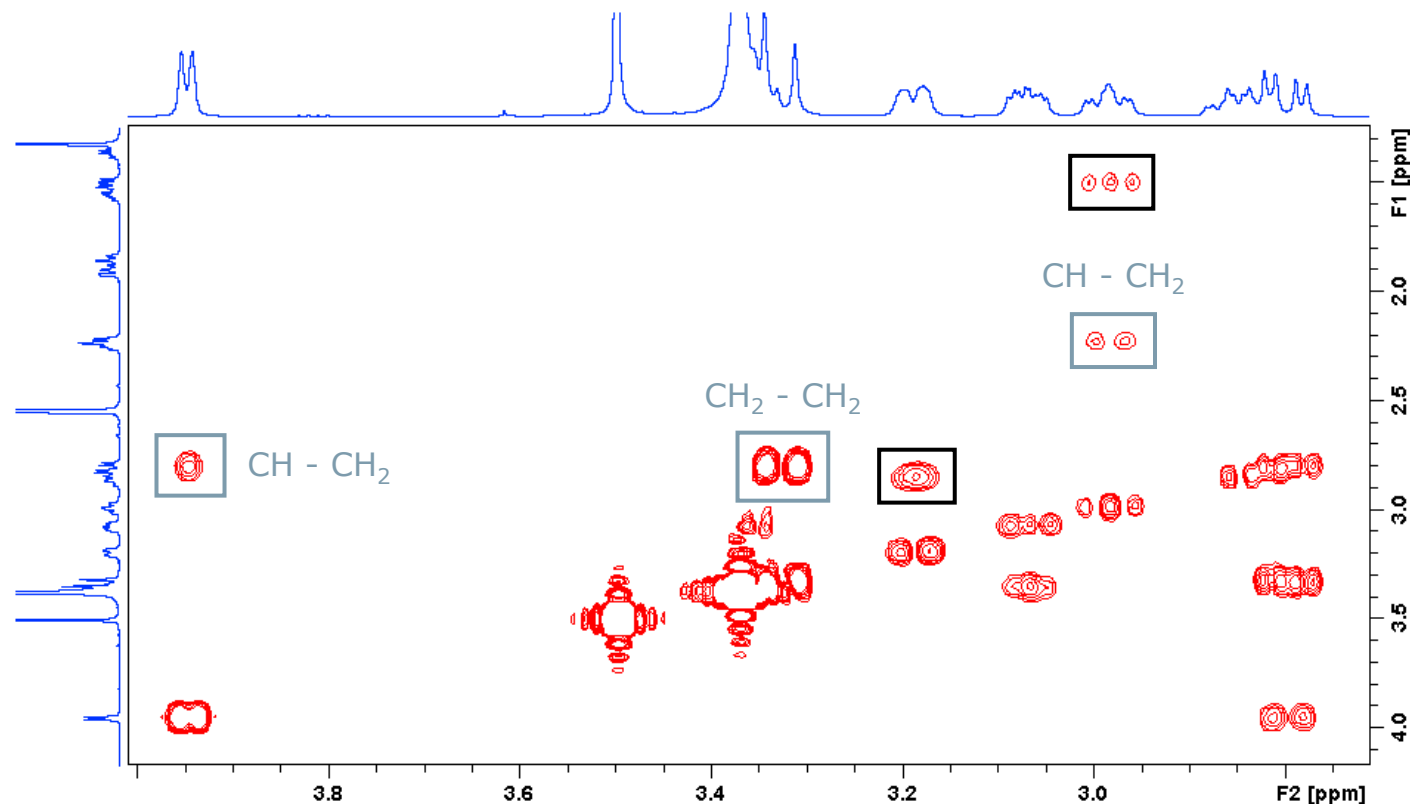
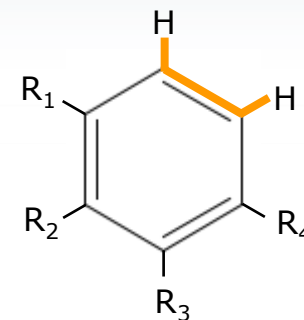


$\Rightarrow$  5x CH<sub>3</sub>  
9x CH<sub>2</sub>  
4x CH sp<sup>3</sup>  
2x CH sp<sup>2</sup>  
5x C sp<sup>3</sup>  
4x C sp<sup>2</sup>  
2x OH/NH

# NMR Spectroscopy – 2D COSY



COSY spectrum  $\Rightarrow$   $^3J_{HH}$  correlation  
(identifying  $\text{CH}_x$ -fragments that are covalently bound)



geminal correlations

$\Rightarrow$  no additional  
information

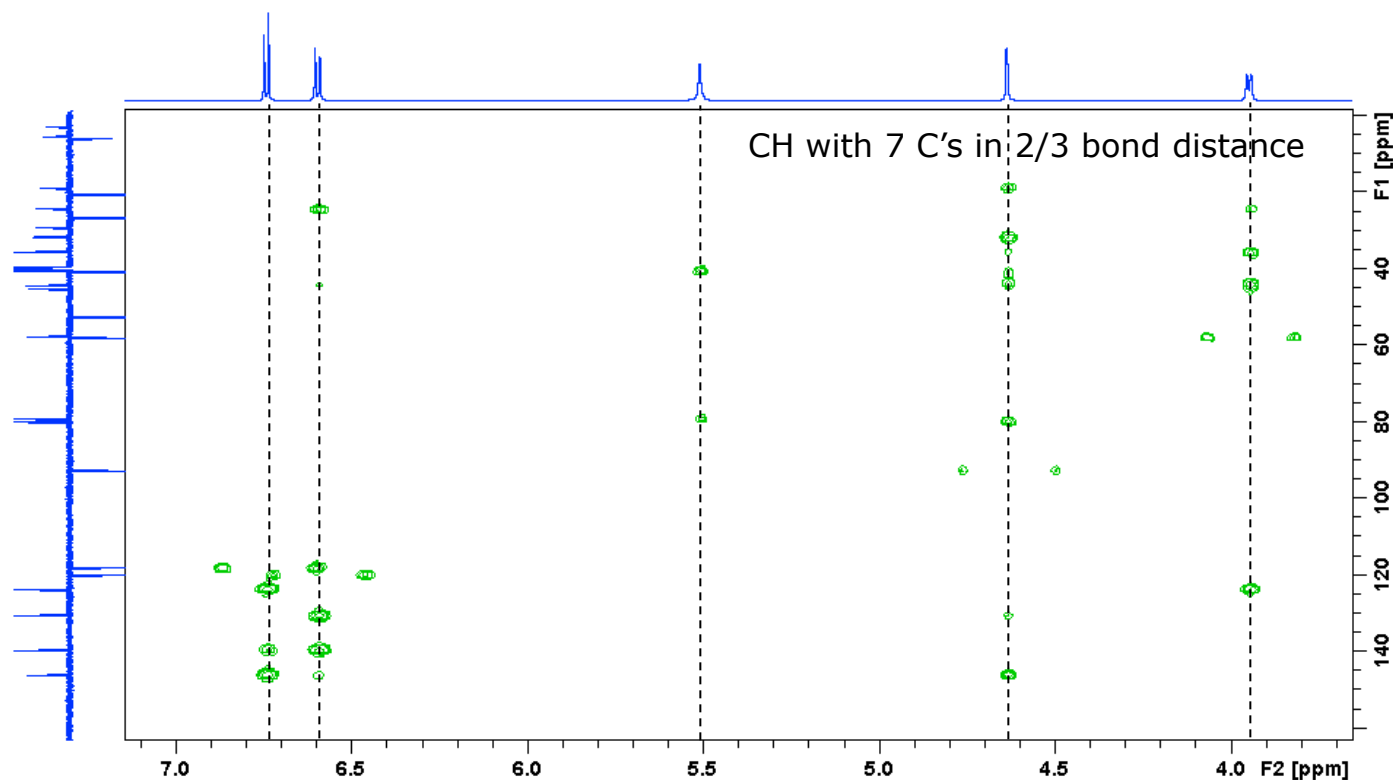
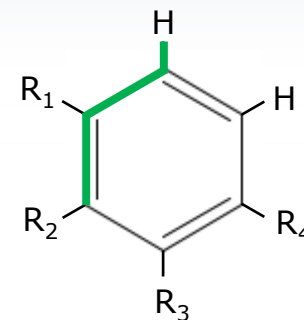
vicinal correlations

$\Rightarrow$  covalently bound  
 $\text{CH}_x$ -fragments

# NMR Spectroscopy – 2D HMBC



HMBC spectrum  $\Rightarrow$   $^2J_{\text{CH}}/^3J_{\text{CH}}$  correlation  
(connecting  $\text{CH}_x$ - fragments with neighbored carbons)



$\Rightarrow$  information  
about  
quaternary  
carbons

# Putting together the fragments



1D Proton, 1D Carbon  
functional groups



2D HSQC  
 $^1J_{\text{CH}}$  correlation (distinguish CH, CH<sub>2</sub>, CH<sub>3</sub>)



2D COSY  
 $^3J_{\text{HH}}$  correlation



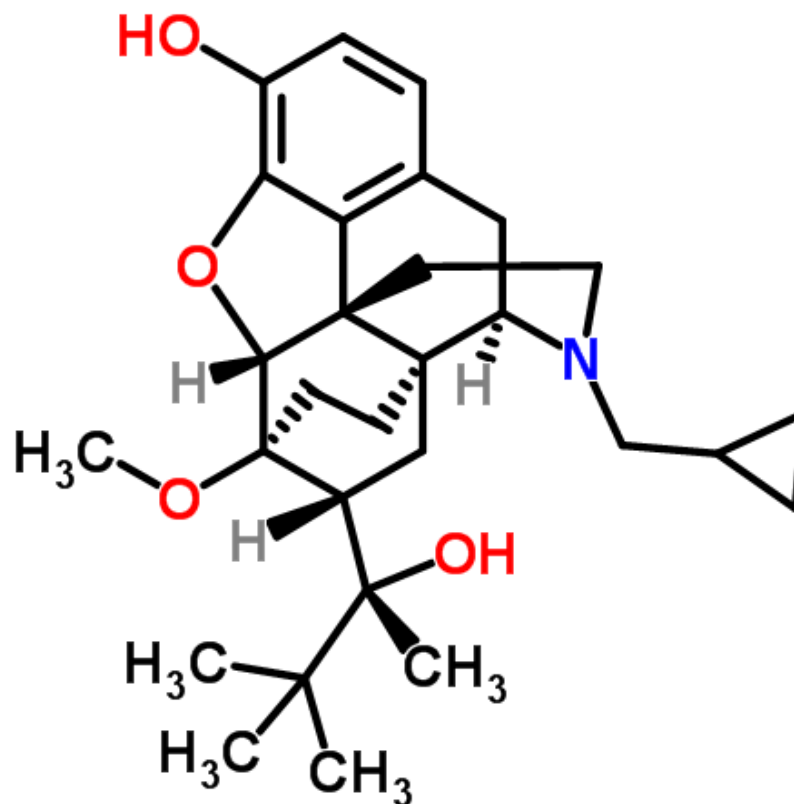
2D HMBC, 2D TOCSY  
 $^2J_{\text{CH}}/^3J_{\text{CH}}$  correlation, spin system

application of the rule of double-bond equivalents (DBE):

- 10 DBE for C<sub>29</sub>H<sub>41</sub>NO<sub>4</sub>
- 3 DBEs C=C

⇒ 7 ring closure left for C<sub>29</sub>H<sub>41</sub>NO<sub>4</sub>

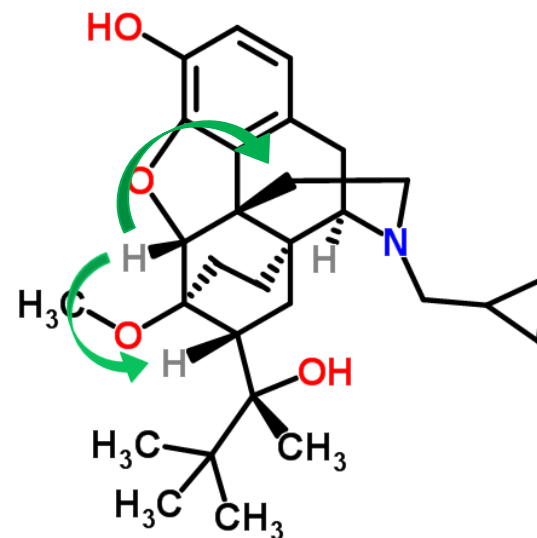
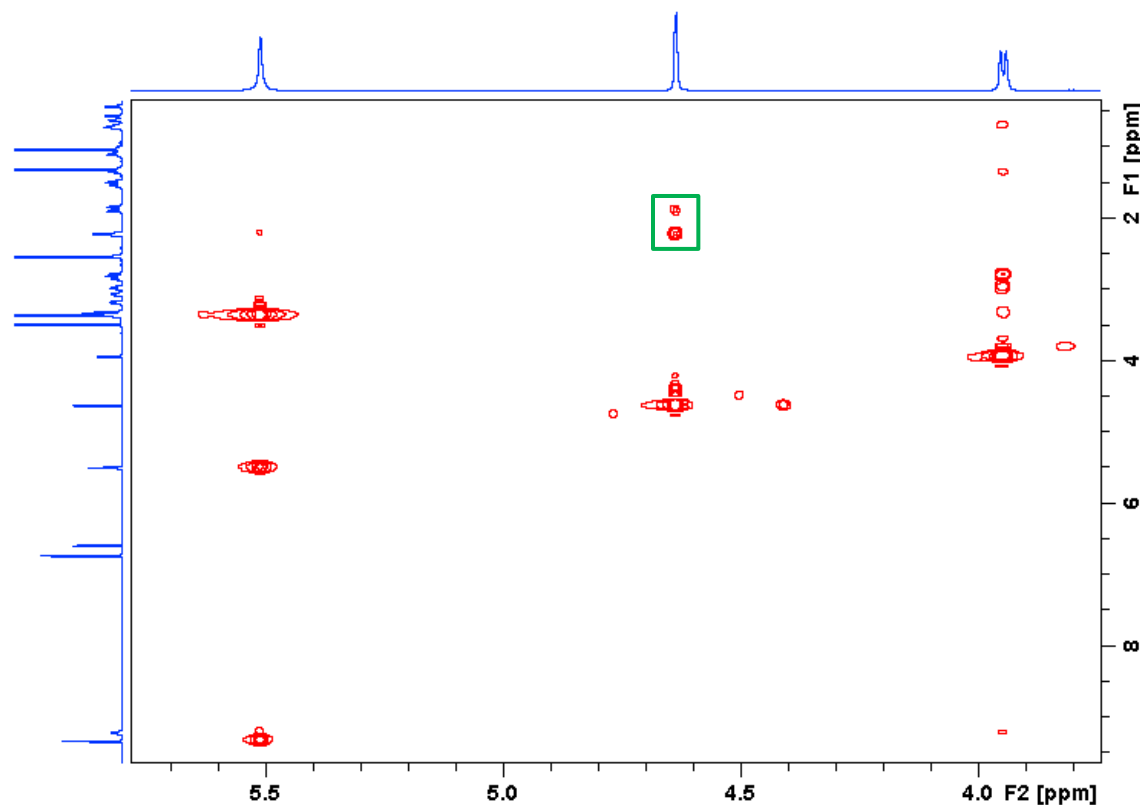
... after quite some detective work ...



# NMR Spectroscopy – 2D NOESY

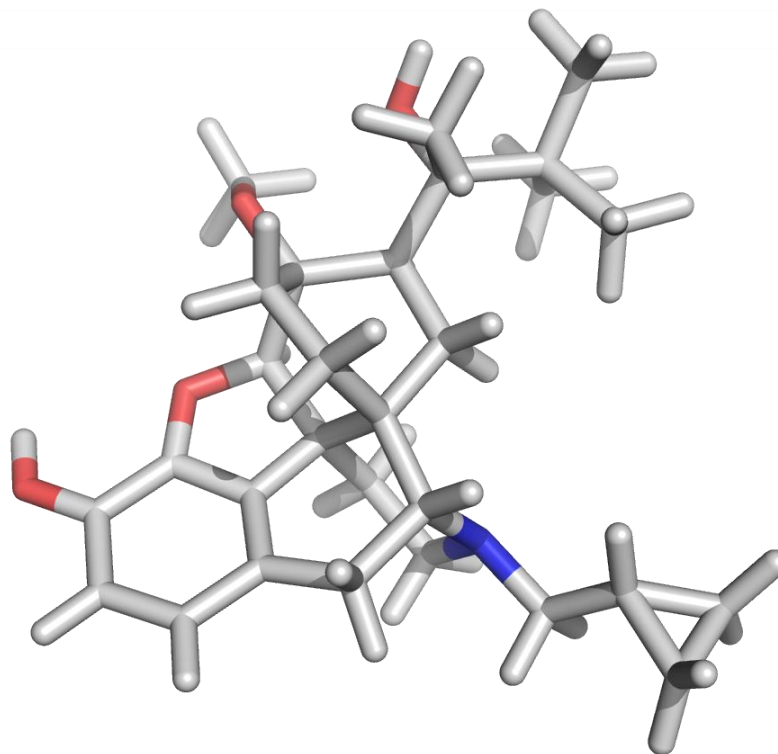


ROESY / NOESY spectrum  $\Rightarrow$  **through space** correlation  
(stereochemistry and 3D structure)



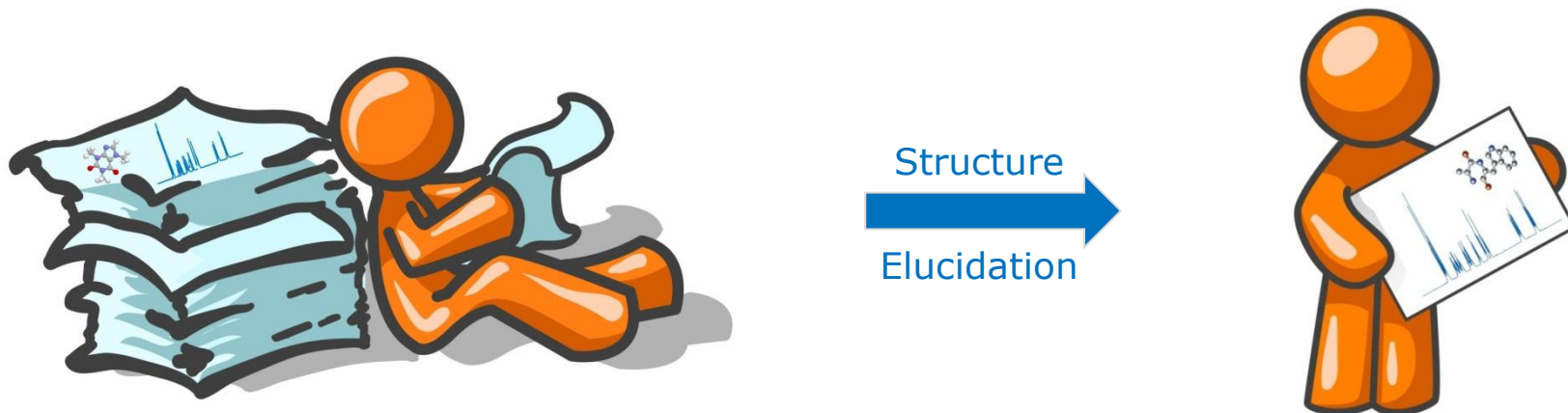


## 3D structure of Buprenorphine



- ⇒ small volume NMR probes allow acquisition of all 1D & 2D spectra required for structure elucidation on only **0.6** mg of compound within one weekend

structure elucidation is a time-consuming and challenging task



a lot of analytical data have to be analyzed in order to find the right structure

⇒ computer assisted data interpretation

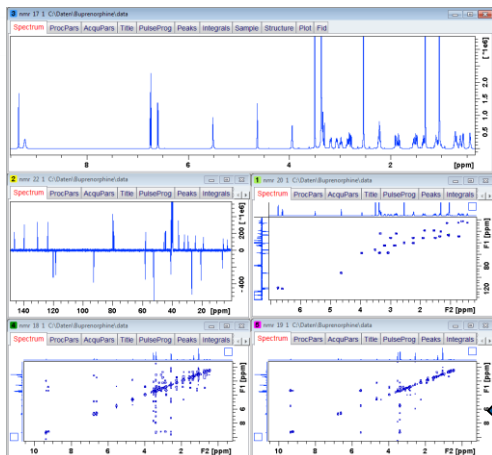
**for small organic molecule NMR:**

**CMC-se**

# Manual Structure Elucidation



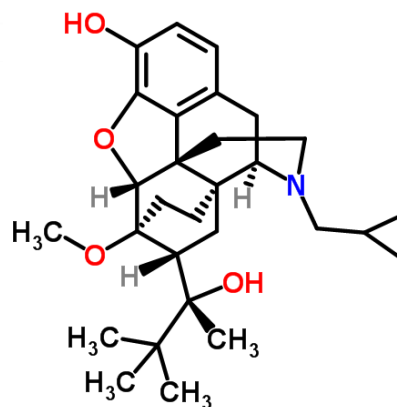
## Acquired Data



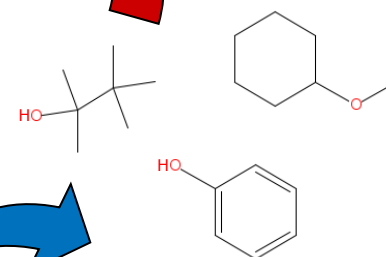
## Extract Information

#	<sup>13</sup> C chem. shift	<sup>1</sup> H chem. shift	Equivalent	Hybridization	Func. group	HMBC	COSY
21	26.69	1.04	3	sp <sup>3</sup>	tert-butyl	16, 9	
12	52.31	3.49		sp <sup>3</sup>	methoxy	8	
5	120.01	6.59		sp <sup>2</sup>	aromatic	2, 3, 22	
6	118.15	6.74		sp <sup>2</sup>	aromatic	1, 2, 4	
18	31.73	2.23		sp <sup>2</sup>		3, 7, 14	13
18'	31.73	1.90		sp <sup>2</sup>		3, 7, 14	13
15	40.78	2.22		sp <sup>3</sup>		7, 8, 9, 23, 24	19

## Generate Structure



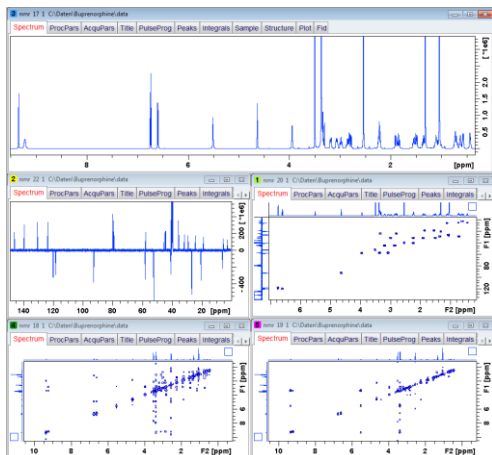
## Identify Fragments



# CMC-se – Assisted Structure Elucidation



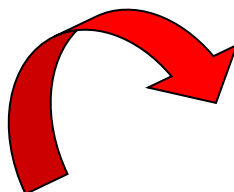
## Acquired Data



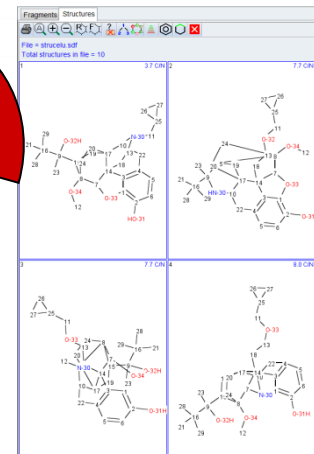
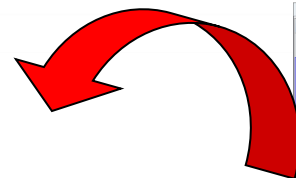
## Populates Table

Table with 10 columns: Name, DBS, Exp, Equiv, F1, F2, F3, F4, F5, F6, F7, F8, F9, F10, F11, F12, F13, F14, F15, F16, F17, F18, F19, F20, F21, F22, F23, F24, F25, F26, F27, F28, F29, F30, F31, F32, F33, F34, F35, F36, F37, F38, F39, F40, F41, F42, F43, F44, F45, F46, F47, F48, F49, F50, F51, F52, F53, F54, F55, F56, F57, F58, F59, F60, F61, F62, F63, F64, F65, F66, F67, F68, F69, F70, F71, F72, F73, F74, F75, F76, F77, F78, F79, F80, F81, F82, F83, F84, F85, F86, F87, F88, F89, F90, F91, F92, F93, F94, F95, F96, F97, F98, F99, F100. The table contains data for various NMR experiments, including 1D and 2D spectra.

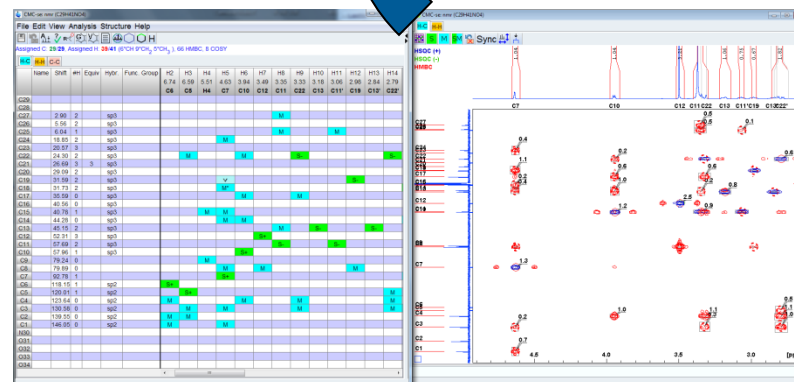
## Table Refinement



## Structure Evaluation



## Structure Generation



## The Correlation Table

CMC-se: nmr (C29H41NO4)

File Edit View Analysis Structure Help

Assigned C: 29/29, Assigned H: 39/41 (6<sup>°</sup>CH 9<sup>°</sup>CH<sub>2</sub> 5<sup>°</sup>CH<sub>3</sub>), 66 HMBC, 8 COSY

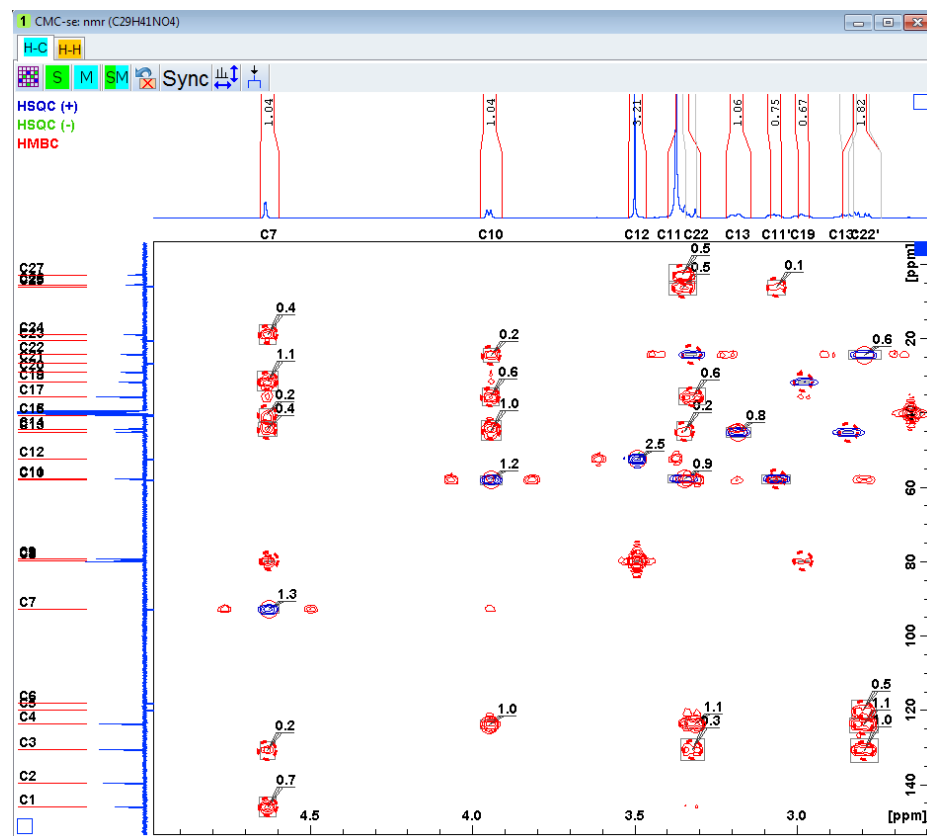
H-C H-H C-C

	Name	Shift	#H	Equiv	Hybr.	Func. Group	H1 9.34 H1	H2 6.74 C6	H3 6.59 C5	H4 5.51 H4	H5 4.63 C7	H6 3.94 C10	H7 3.49 C12	H8 3.35 C11	H9 3.33 C22	H10 3.18 C13	H11 3.06 C11'	H12 2.98 C19	H13 2.84 C13'	H14 2.79 C22'
C29																				
C28																				
C27		2.90	2		sp3									M						
C26		5.56	2		sp3												M			
C25		6.04	1		sp3															
C24		18.85	2		sp3						M									
C23		20.57	3		sp3															
C22		24.30	2		sp3				M						S-					S-
C21		26.69	3	3	sp3															
C20		29.09	2		sp3															
C19		31.59	2		sp3													S-		
C18		31.73	2		sp3															
C17		35.59	0		sp3															
C16		40.56	0		sp3															
C15		40.78	1		sp3															
C14		44.28	0		sp3															
C13		45.15	2		sp3															
C12		52.31	3		sp3															
C11		57.69	2		sp3															
C10		57.96	1		sp3															
C9		79.24	0																	
C8		79.89	0																	
C7		92.78	1																	
C6		118.15	1		sp2															
C5		120.01	1		sp2															
C4		123.64	0		sp2															
C3		130.58	0		sp2															
C2		139.55	0		sp2															
C1		146.05	0		sp2															
N30																				
O31																				
O32																				
O33																				
O34																				

- automatically populated
- all atoms from the molecule included as rows or columns
- filled in cells indicated the presence of correlations
- fully editable by the user
- convenient way for organizing the data for structure elucidation

## The Combined Spectra Display

- all data from the project included in a single screen
- all automatically picked peaks are displayed
- manual corrections to the automatic data analysis are possible through this window
- correlated cursor between all windows for data evaluation



# CMC-se – Assisted Structure Elucidation



CMC-se: nmr (C29H41NO4)

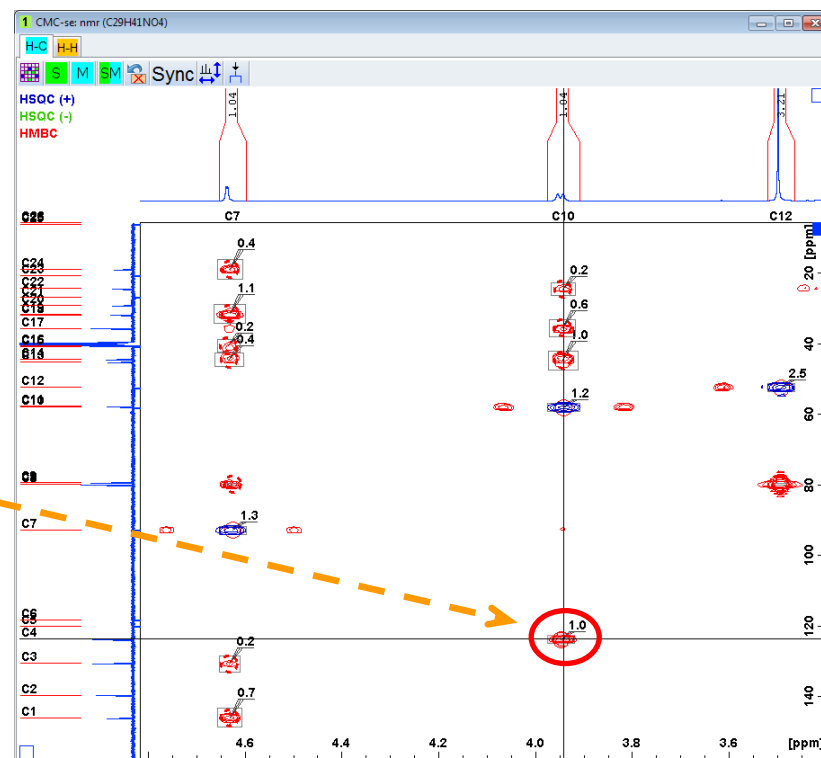
File Edit View Analysis Structure Help

Assigned C: 29/29, Assigned H: 39/41 (6\*CH 9\*CH<sub>2</sub> 5\*CH<sub>3</sub>), 66 HMBC, 8 COSY

H-C H-H C-C

	Name	Shift	#H	Equiv	Hybr.	Func. Group	H1 9.34 H1	H2 6.74 C6	H3 6.59 C5	H4 5.51 H4	H5 4.63 C7	H6 3.94 C10	H7 3.49 C12	H8 3.35 C11	H9 3.33 C22	H10 3.18 C13
C26		5.56	2		sp3											
C25		6.04	1		sp3									M		
C24		18.85	2		sp3											
C23		20.57	3		sp3											
C22		24.30	2		sp3			M			M				S-	
C21		26.69	3	3	sp3											
C20		29.09	2		sp3											
C19		31.59	2		sp3											
C18		31.73	2		sp3											
C17		35.59	0		sp3						M				M	
C16		40.56	0		sp3											
C15		40.78	1		sp3				M	M						
C14		44.28	0		sp3						M					
C13		45.15	2		sp3									M		S-
C12		52.31	3		sp3						S+					
C11		57.69	2		sp3								S-			
C10		57.96	1		sp3						S+					
C9		79.24	0						M							
C8		79.89	0							M						
C7		92.78	1													
C6		118.15	1		sp2		M	S+								
C5		120.01	1		sp2				S+							
C4		123.64	0		sp2			M								
C3		130.58	0		sp2				M							
C2		139.55	0		sp2		M	M	M							
C1		146.05	0		sp2		M	M								

- clicking on a peak in the data results in that cell being highlighted in the table and vice versa



- modifications in the spectra display and correlation table are synchronized

# CMC-se – Assisted Structure Elucidation



## NMR information

CMC-se v1.0 (C2H41NO4) H-COSY

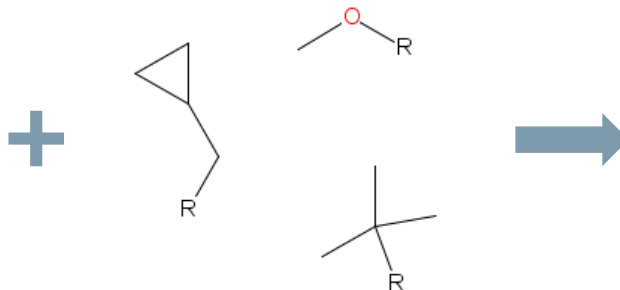
File Edit View Analysis Structure Help

Assigned C: 29/29 Assigned H: 39/41 (6°CH 9°CH<sub>2</sub> 5°CH<sub>3</sub>), 66 HMBC, 8 COSY

	Name	Shift	#H	Equiv	Hybr.	Func. Group	H1 9.34 H1	H2 6.74 C8	H3 6.59 C8	H4 5.51 H4	H5 4.63 C7	H6 3.94 C10	H7 3.49 C12	H8 3.35 C11	H9 3.33 C22
C29															
C27		2.90	2		sp3										M
C28		5.56	2		sp3										M
C25		6.04	1		sp3										M
C24		18.85	2		sp3						M				
C23		20.57	3		sp3						M				S
C22		24.30	2		sp3										
C21		26.69	3		sp3										
C20		29.09	2		sp3										
C19		31.59	2		sp3										
C18		31.73	2		sp3										
C17		35.59	0		sp3										M
C16		40.56	0		sp3										
C15		40.78	1		sp3						M	M			
C14		44.28	0		sp3						M	M			
C13		45.15	2		sp3										M
C12		52.31	3		sp3										S
C11		57.69	2		sp3										
C10		57.96	1		sp3										
C9		79.24	0												
C8		79.89	0												
C7		92.76	1												
C6		118.15	1		sp2		M	S							
C5		120.01	1		sp2										
C4		123.64	0		sp2										M
C3		130.58	0		sp2										M
C2		139.55	0		sp2		M	M							
C1		146.05	0		sp2		M	M							
N30															
O31															
O32															
O33															
O34															

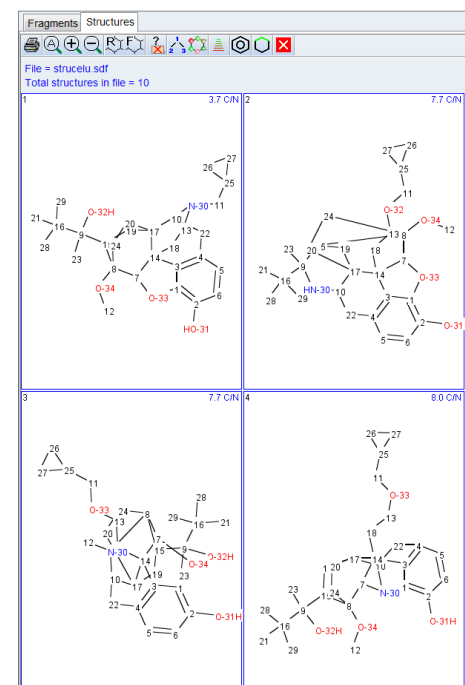
correlation table

## MS information



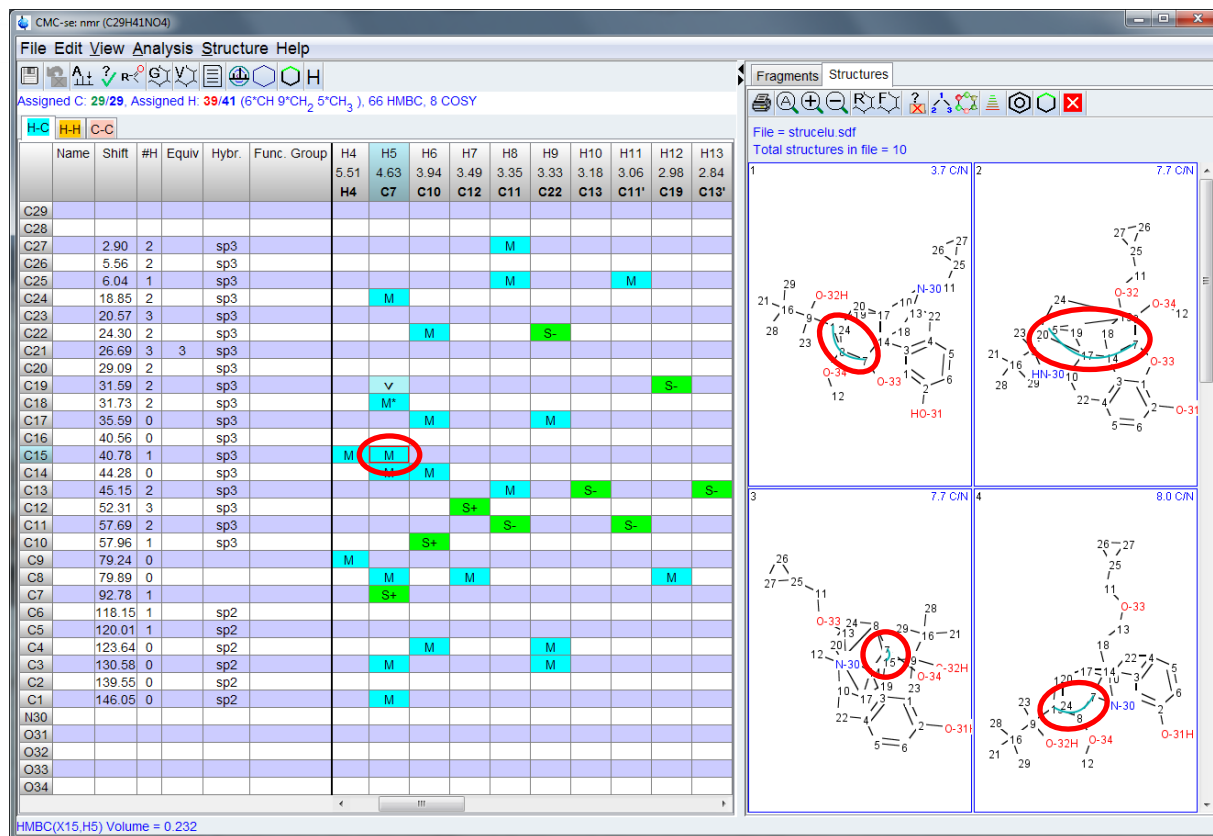
fragments (optional)

## structure proposals





# CMC-se – Assisted Structure Elucidation



- generated structures are linked to the correlation table
- individual or multiple correlations can be viewed on all structure proposals

# CMC-se – Assisted Structure Elucidation

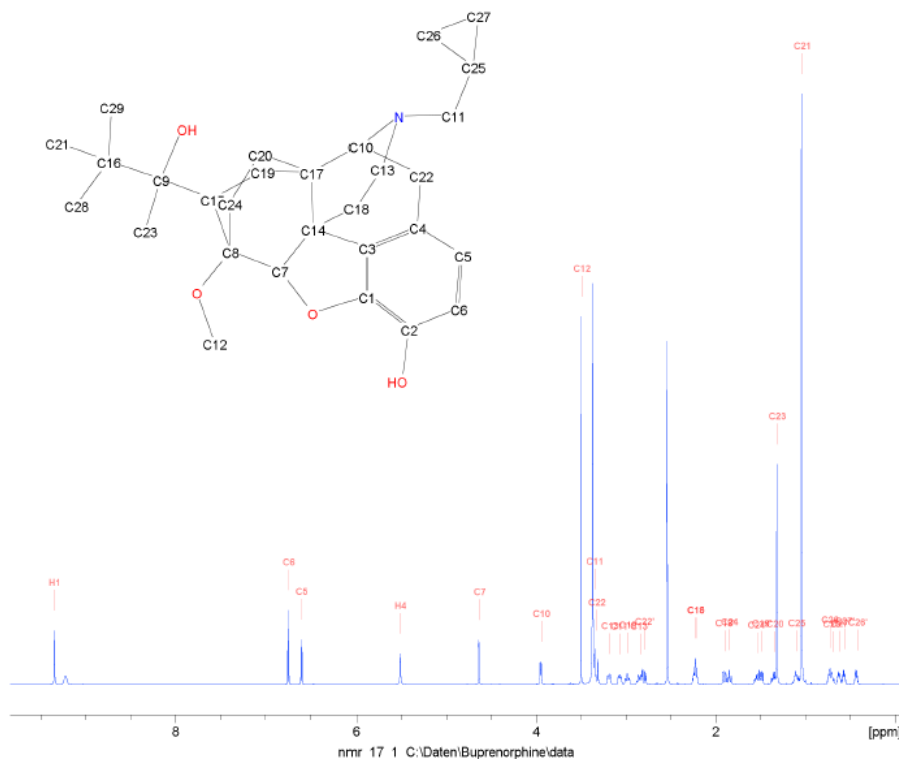


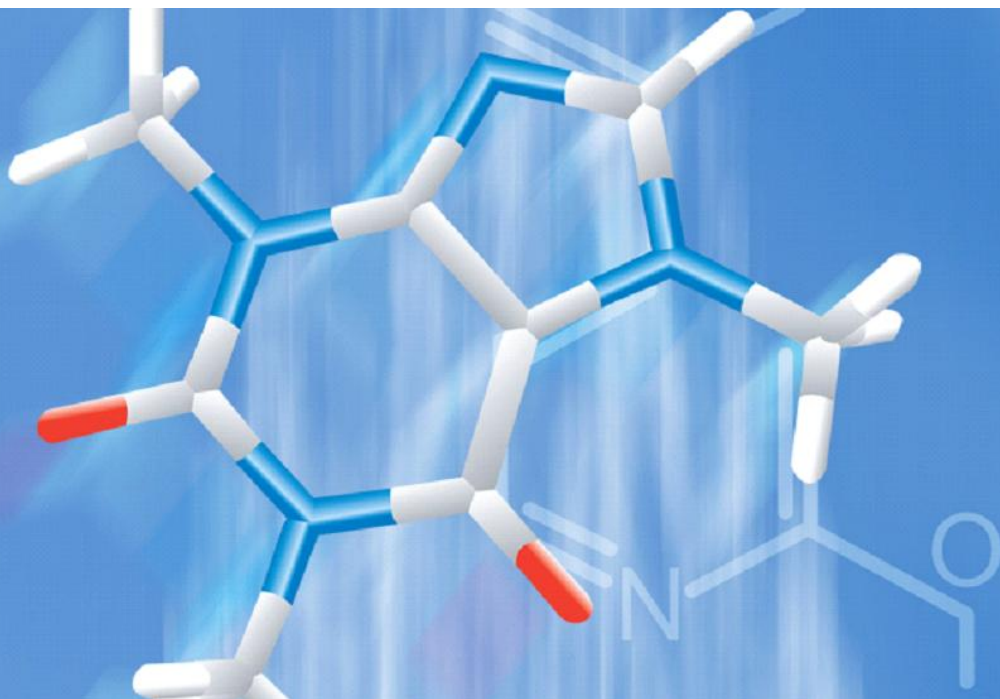
## Bruker Structure Elucidation Report nmr (C:\Daten\Buprenorphine\data\nmr\strucelu)



<sup>1</sup>H table of assignments

Atom	Shift [ppm]	Multiplicity	Bound to	Correlation table
C26'	0.42		(C26)	H29
C27'	0.56		(C27)	H28
C27	0.62		(C27)	H27
C20'	0.70		(C20)	H26
C26	0.73		(C26)	H25
C21	1.04	s	(C21)	H24
C25	1.10		(C25)	H23
C23	1.31	s	(C23)	H22
C20	1.34		(C20)	H21
C19'	1.49		(C19)	H20
C24'	1.54		(C24)	H19
C24	1.85		(C24)	H18
C18'	1.90		(C18)	H17
C15	2.22		(C15)	H16
C18	2.23		(C18)	H15
C22'	2.79		(C22)	H14
C13'	2.84		(C13)	H13
C19	2.98		(C19)	H12
C11'	3.06		(C11)	H11
C13	3.18		(C13)	H10
C22	3.33		(C22)	H9
C11	3.35		(C11)	H8
C12	3.49	s	(C12)	H7
C10	3.94		(C10)	H6
C7	4.63		(C7)	H5
H4	5.51	s		H4
C5	6.59	d	(C5)	H3
C6	6.74	d	(C6)	H2
H1	9.34	s		H1





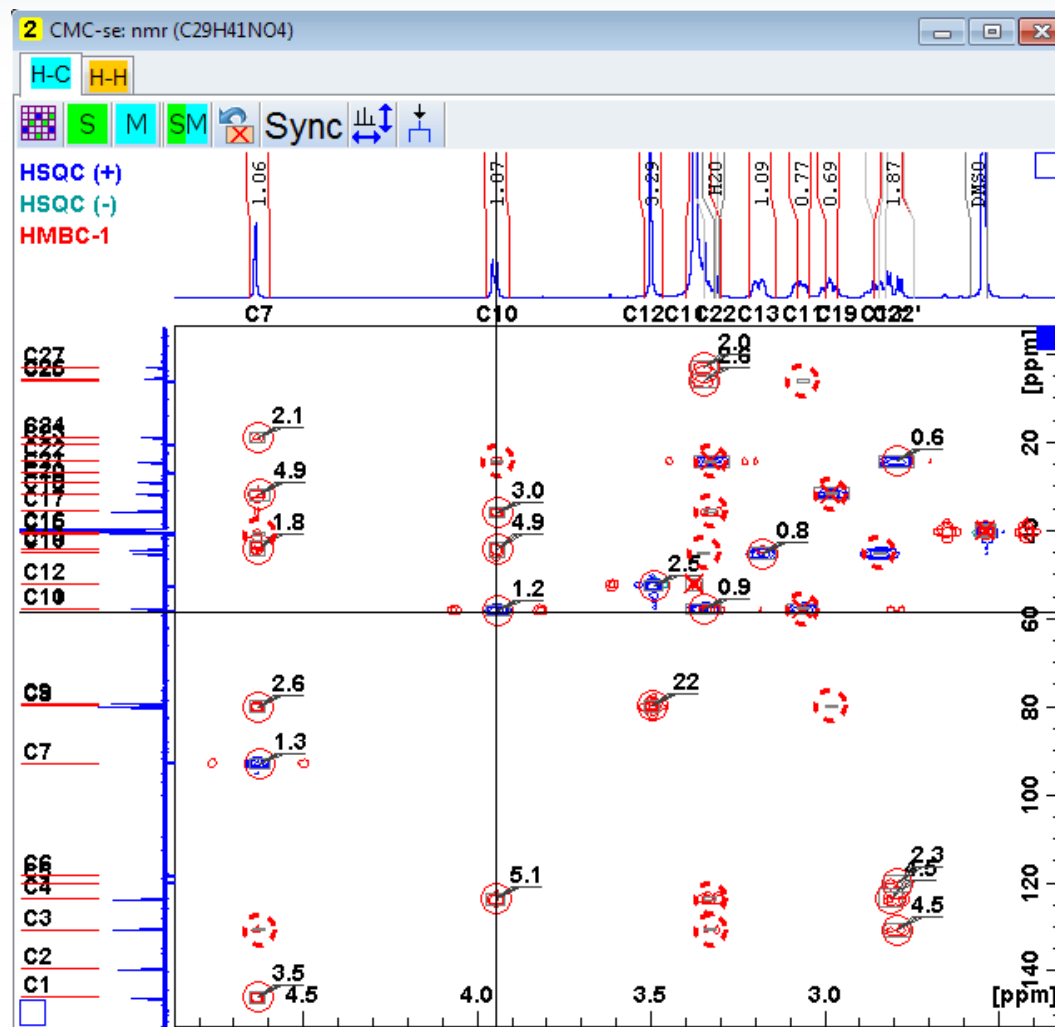
- **CMC-se**

Behind the Scenes:  
Interpretation Engine

# Automated Data Analysis



- S/N estimation
- peak picking
- artifact removal
- solvent signal search
- build correlation lists
- **find common frequencies**



## $^1\text{H}$ - $^{13}\text{C}$ HSQC spectrum

- multiplicity edited
  - negative signals  $\Rightarrow$   $\text{CH}_2$ -group
  - positive signals  $\Rightarrow$   $\text{CH}$ - or  $\text{CH}_3$ -group
- quantitative
  - integration of peaks  $\Rightarrow$  distinguish  $\text{CH}$ - from  $\text{CH}_3$ -groups
  - integral check with 1D  $^1\text{H}$  spectrum
- $^{13}\text{C}$  chemical shift
  - chemical shift  $< 60$  ppm  $\text{sp}^3$  hybridized
  - chemical shift  $> 100$  ppm  $\text{sp}^2$  hybridized

# Interpretation of NMR Constraints



HSQC

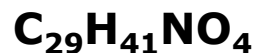
1D Proton

1D Carbon



proton distribution  
carbon hybridization

molecular formula



- 5\* CH<sub>3</sub> ⇒ sp<sup>3</sup> hybridized
- 9\* CH<sub>2</sub> ⇒ 9\* sp<sup>3</sup> / 0\* sp<sup>2</sup> hybridized
- 6\* CH ⇒ 4\* sp<sup>3</sup> / 2\* sp<sup>2</sup> hybridized
- 9\* C ⇒ 5\* sp<sup>3</sup> / 4\* sp<sup>2</sup> hybridized

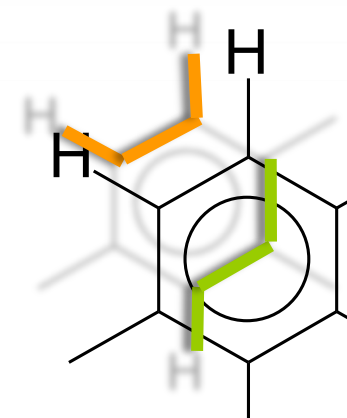
# Interpretation of NMR Constraints



HMBC and COSY cross peaks for connection of heavy atoms

**HMBC (required)**  $^2J - ^3J$

**COSY (optional)**  $^3J$

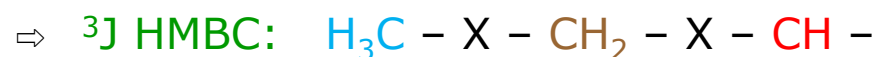
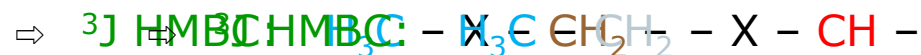
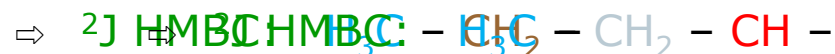


- $\text{CH}_3 \Rightarrow$  one neighboring carbon atom  $\Rightarrow$  0-1 COSY, 0-2 HMBC correlations
- $\text{CH}_2 \Rightarrow$  two neighboring carbon atoms  $\Rightarrow$  0-2 COSY, 0-4 HMBC correlations
- ...
- $\text{sp}^2$  hybridized carbon has to have a  $\text{sp}/\text{sp}^2$  hybridized neighbor

# Interpretation of NMR Constraints



- ~~first bond identification:~~



too few correlations:  
only fragments build up

too many correlations:  
even more possibilities



# Interpretation of NMR Constraints



**HSQC**

**1D Proton**

**1D Carbon**



proton distribution  
carbon hybridization

**HMBC**

**COSY**



atom connection



- HMBC correlations are essential to assemble a molecular structure from individual atoms
- number of structure proposals depends on the number of (ambiguous/unambiguous) correlations

# Adjustments for Structure Generation

A screenshot of the "Structure Generation Options" dialog box. The window has a title bar with a blue icon and the text "Structure Generation Options". The main area is divided into several sections with blue headers: "Execution Control", "Substructures", "Ring rules", "Correlations", and "Chemistry rules". Under "Execution Control", there are input fields for "Terminate when this many structures were generated (0=no limit)" (value 0) and "Terminate after this many seconds (0=no limit)" (value 300), and a checked checkbox for "Use multiple processors". Under "Substructures", there is an unchecked checkbox for "Enable substructure filtering". Under "Ring rules", there are checkboxes for "Structure does not contain any rings" (unchecked), "Maximum ring length (0=no limit)" (value 0), "Forbidden rings lengths (Comma separated e.g. '3,4')", and "Required rings lengths (Comma separated e.g. '5,6')". Under "Correlations", there are checkboxes for "Use COSY correlations" (checked), "Use HMBC correlations" (checked), "Auto-eliminate invalid or long range COSY correlations" (unchecked), and "Auto-eliminate invalid or long range HMBC correlations" (unchecked). There is also an input field for "Maximum number of eliminated correlations (COSY+HMBC)" (value 0) and a dropdown menu for "HMBC/COSY autoelimination policy" (set to "Optimal"). At the bottom, there is a "Chemistry rules" section with a dropdown arrow. At the very bottom of the dialog are two buttons: "Generate Structures" and "Cancel".

- long range correlations have to be taken into account
  - increasing number of correlations that are auto-eliminated by the software step by step
- information about rings length
- information about substructures (MS fragments)

# Evaluating Structure Proposals



CMC-se: nmr (C29H41NO4)

File Edit View Analysis Structure Help

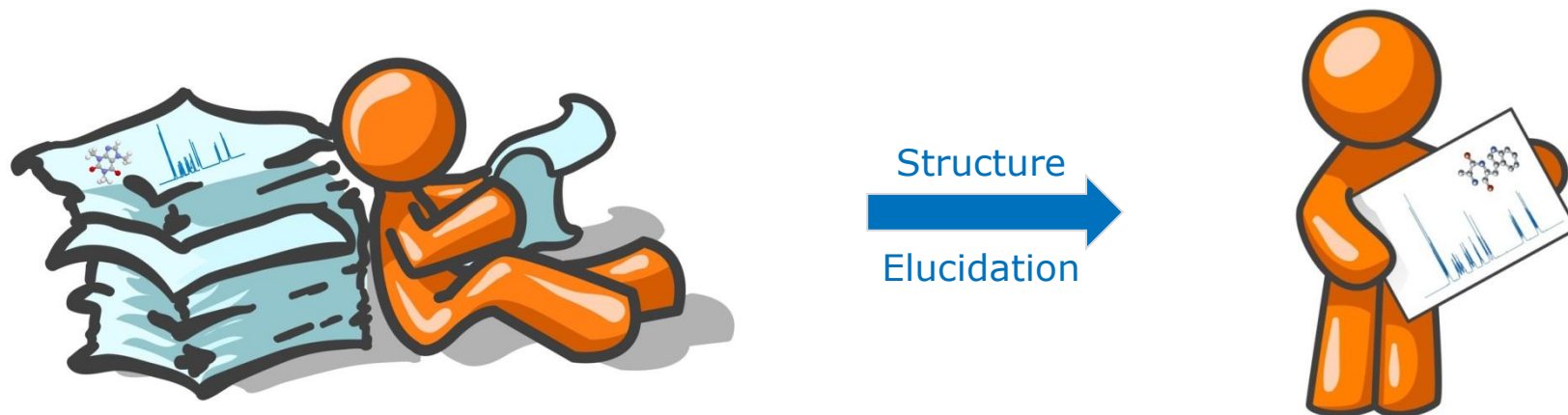
Assigned C: 29/29, Assigned H: 39/41 (6°CH 9°CH<sub>2</sub> 5°CH<sub>3</sub>), 66 HMBC, 8 COSY

H-C H-H C-C

Name	Shift	#H	Equiv	Hybr.	Func. Group	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12	H13	H14	H15	H16	
						H1	C6	C5	H4	C7	C10	C12	C11	C22	C13	C11'	C19	C13'	C22'	C18	C15	
C29																						
C28																						
C27	2.90	2		sp3									M									
C26	5.56	2		sp3																		
C25	6.04	1		sp3										M								
C24	18.85	2		sp3						M											M	
C23	20.57	3		sp3																	M	
C22	24.30	2		sp3				M			M			S-					S-			
C21	26.69	3	3	sp3																		
C20	29.09	2		sp3																		
C19	31.59	2		sp3													S-					
C18	31.73	2		sp3						M*										S-		
C17	35.59	0		sp3							M											
C16	40.56	0		sp3																		
C15	40.78	1		sp3				M	M												S+	
C14	44.28	0		sp3					M	M											M	
C13	45.15	2		sp3									M		S-				S-			
C12	52.31	3		sp3								S+										
C11	57.69	2		sp3									S-			S-						
C10	57.96	1		sp3								S+										
C9	79.24	0							M												M	
C8	79.89	0								M		M									M	
C7	92.78	1																			M	
C6	118.15	1		sp2		M	S+															
C5	120.01	1		sp2				S+														
C4	123.64	0		sp2					M		M										M	
C3	130.58	0		sp2					M												M	
C2	139.55	0		sp2		M	M	M														
C1	146.05	0		sp2		M	M				M											
N30																						
O31																						
O32																						
O33																						
O34																						

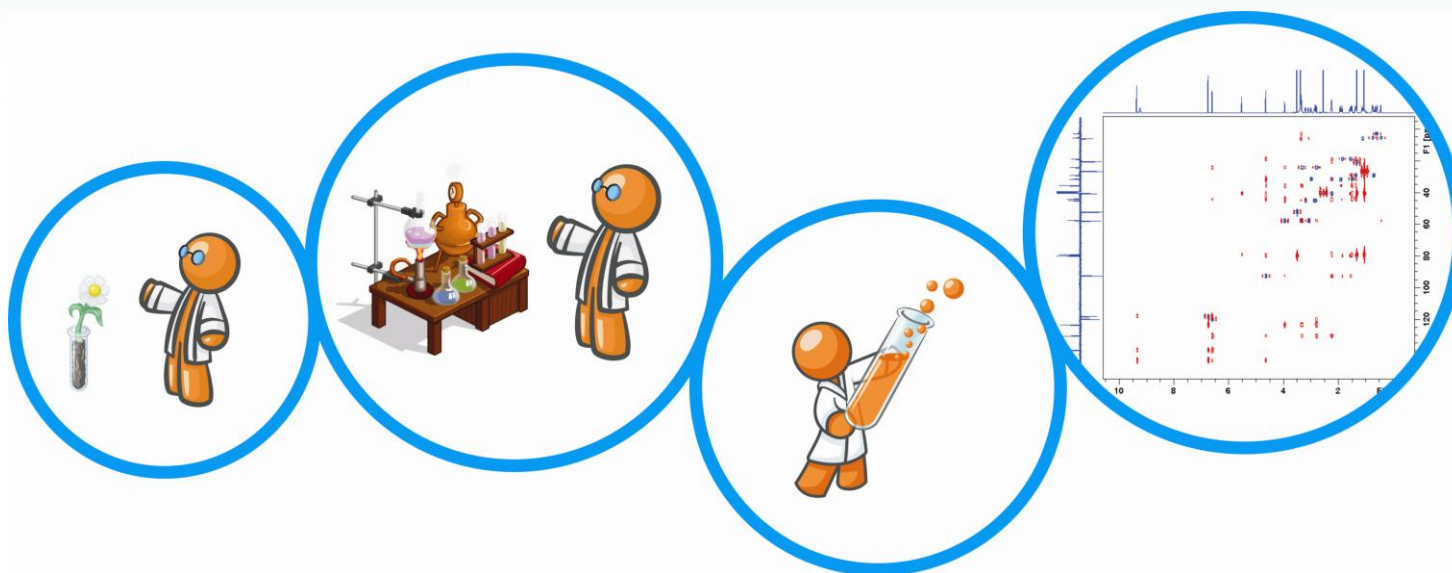
File = structu1.sdf  
Total structures in file = 10

as structure elucidation is a time-consuming and challenging task ...



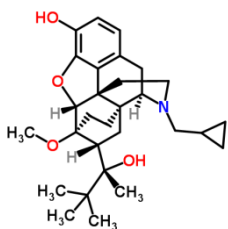
- CMC-se significantly speeds up the process of data analysis
- CMC-se organizes the workflow
- CMC-se generates structure proposals that all fit the experimental data

# Structure Elucidation of Natural Products



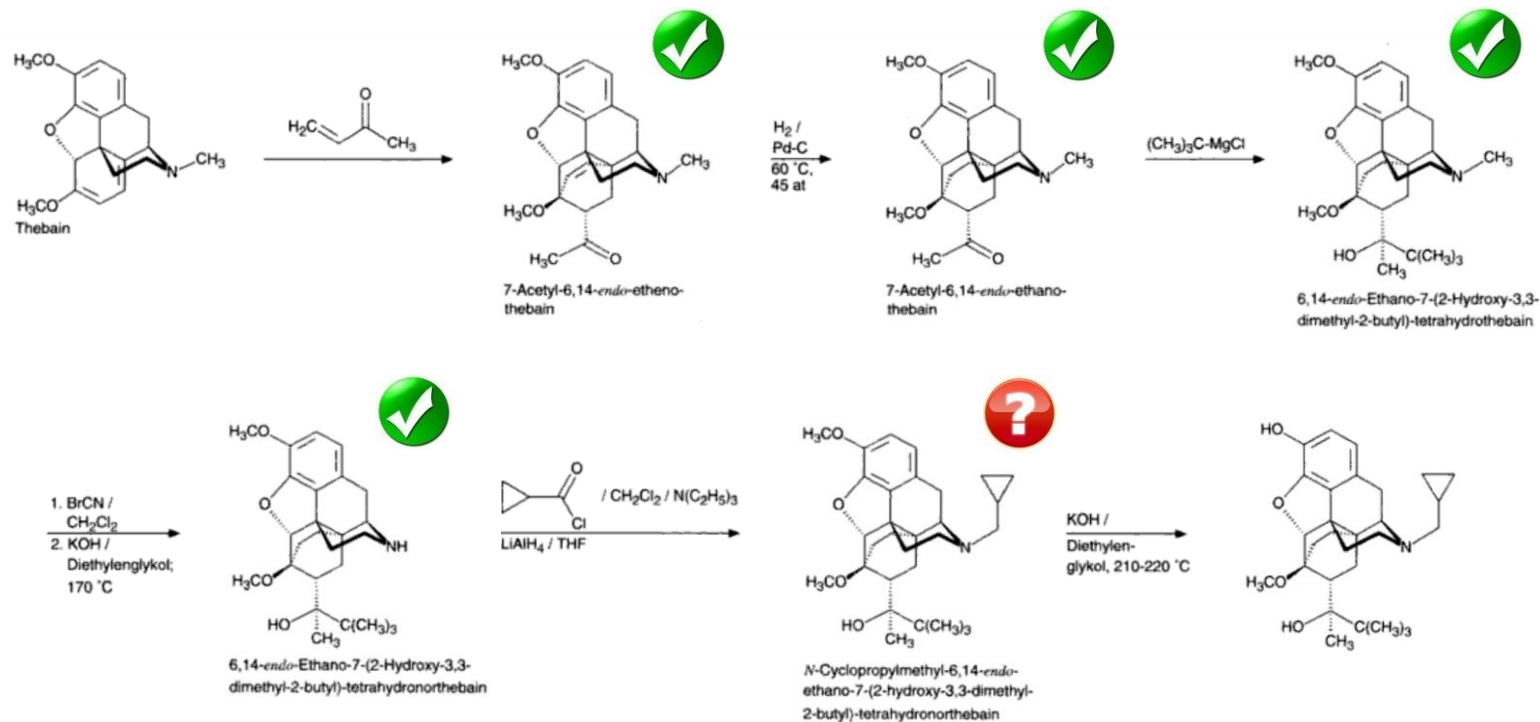
characterization of unknown compound by NMR:

structure elucidation based on 1D and 2D NMR spectra



synthesizing natural product

# Synthesis of Buprenorphine

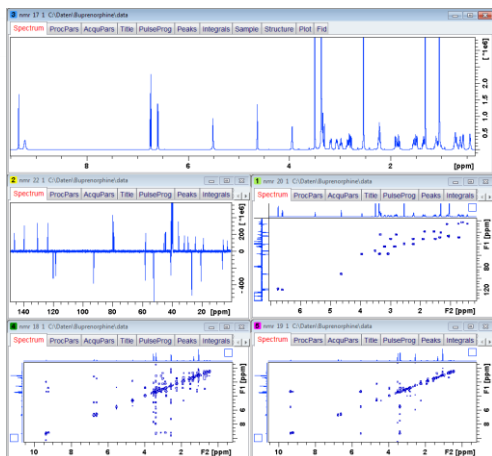


⇒ verifying synthesis steps by NMR

# CMC-se – Assisted Structure Verification



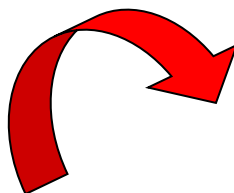
Acquired Data



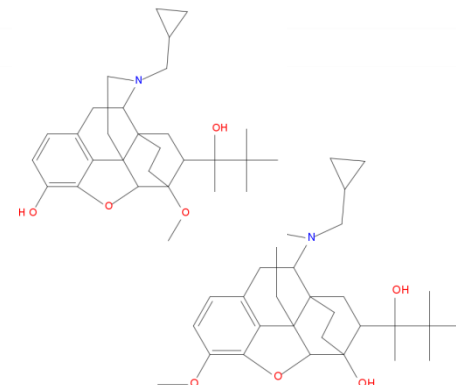
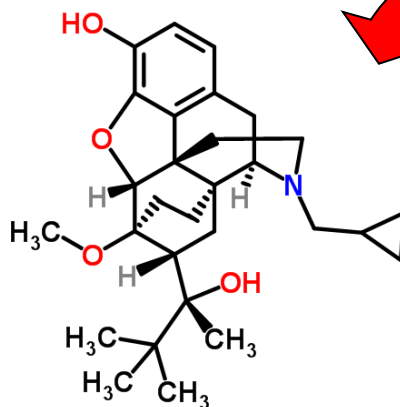
Populates Table

The screenshot shows the CMC-se software interface with a table of NMR data. The table has columns for Name, DB, Exp, NMR, Func, Group, and a grid of data points. The data points are color-coded (green, red, blue) to represent different types of correlations or assignments.

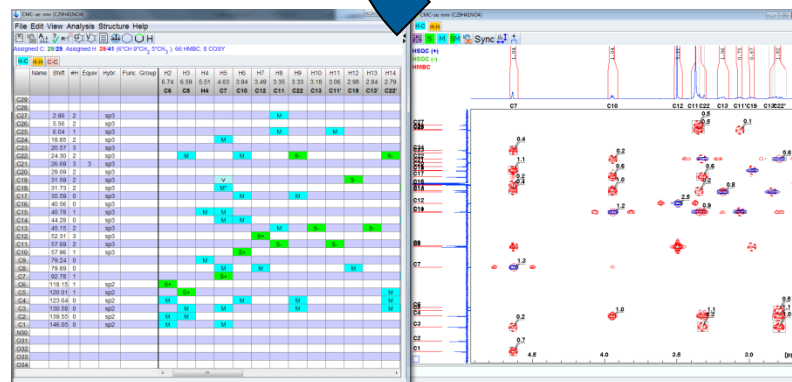
Table Refinement

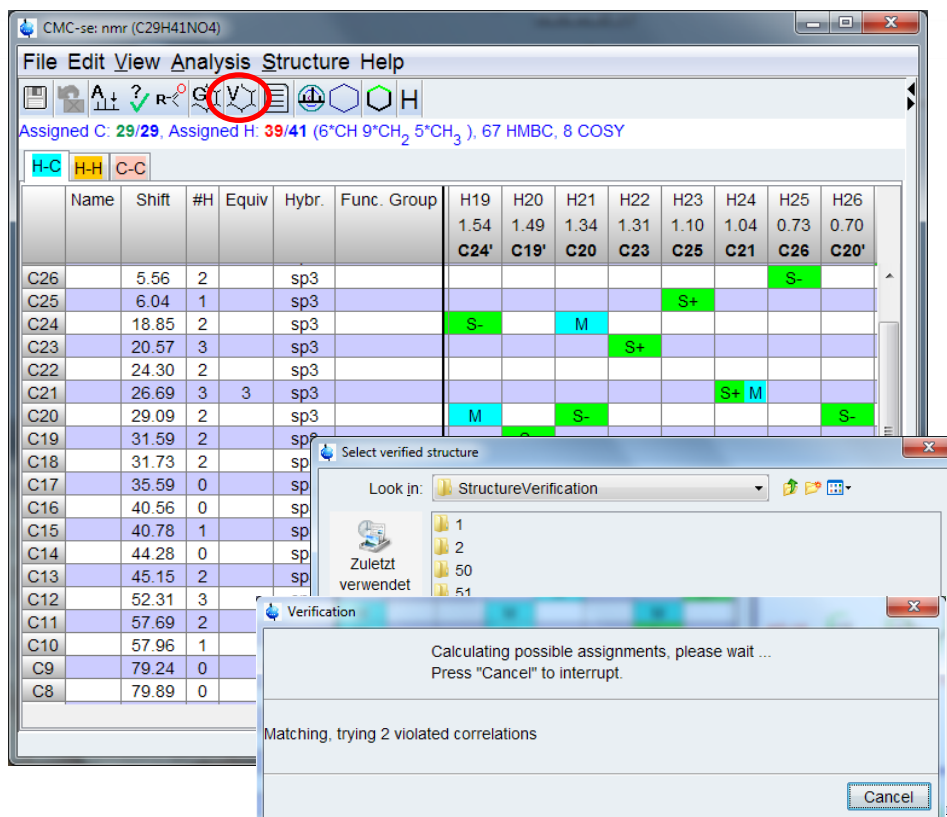


Structure Verification



Structure Proposals





CMC-se: nmr (C29H41NO4)

File Edit View Analysis Structure Help

Assigned C: 29/29, Assigned H: 39/41 (6<sup>°</sup>CH 9<sup>°</sup>CH<sub>2</sub> 5<sup>°</sup>CH<sub>3</sub>), 67 HMBC, 8 COSY

	Name	Shift	#H	Equiv	Hybr.	Func. Group	H19 1.54 C24'	H20 1.49 C19'	H21 1.34 C20	H22 1.31 C23	H23 1.10 C25	H24 1.04 C21	H25 0.73 C26	H26 0.70 C20'
C26		5.56	2		sp3							S-		
C25		6.04	1		sp3					S+				
C24		18.85	2		sp3		S-		M					
C23		20.57	3		sp3				S+					
C22		24.30	2		sp3									
C21		26.69	3	3	sp3						S+	M		
C20		29.09	2		sp3		M		S-				S-	
C19		31.59	2		sp3									
C18		31.73	2		sp3									
C17		35.59	0		sp3									
C16		40.56	0		sp3									
C15		40.78	1		sp3									
C14		44.28	0		sp3									
C13		45.15	2		sp3									
C12		52.31	3		sp3									
C11		57.69	2		sp3									
C10		57.96	1		sp3									
C9		79.24	0		sp3									
C8		79.89	0		sp3									

Select verified structure

Look in: StructureVerification

1  
2  
50  
51

Zuletzt verwendet

Verification

Calculating possible assignments, please wait ...  
Press "Cancel" to interrupt.

Matching, trying 2 violated correlations

Cancel

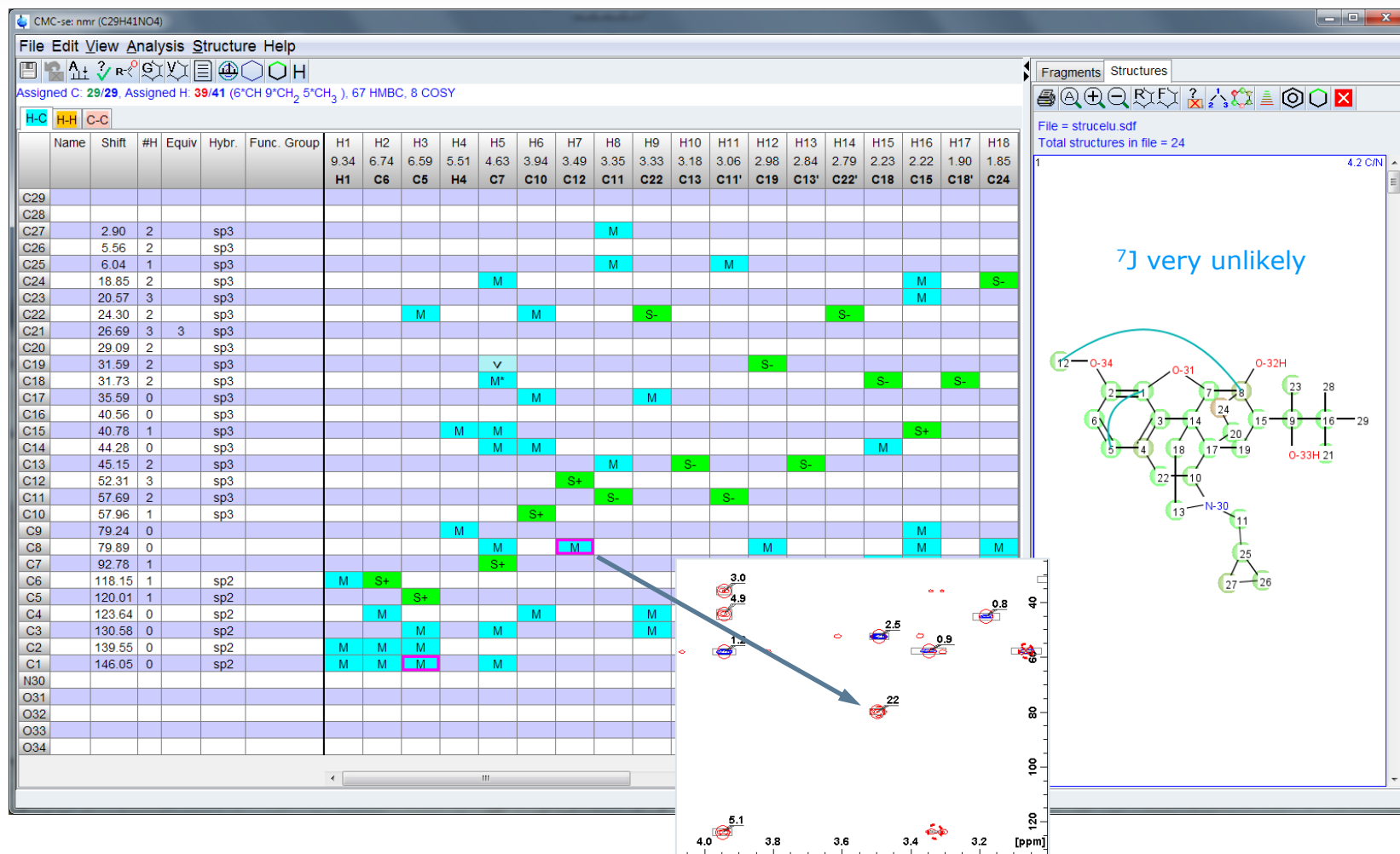
## Verify structure proposals

- provide projector .sdf file
- perform automated verification
- indicate selection of long verification paths
- perform automated spectra analysis
- manually inspect and refine correlation table

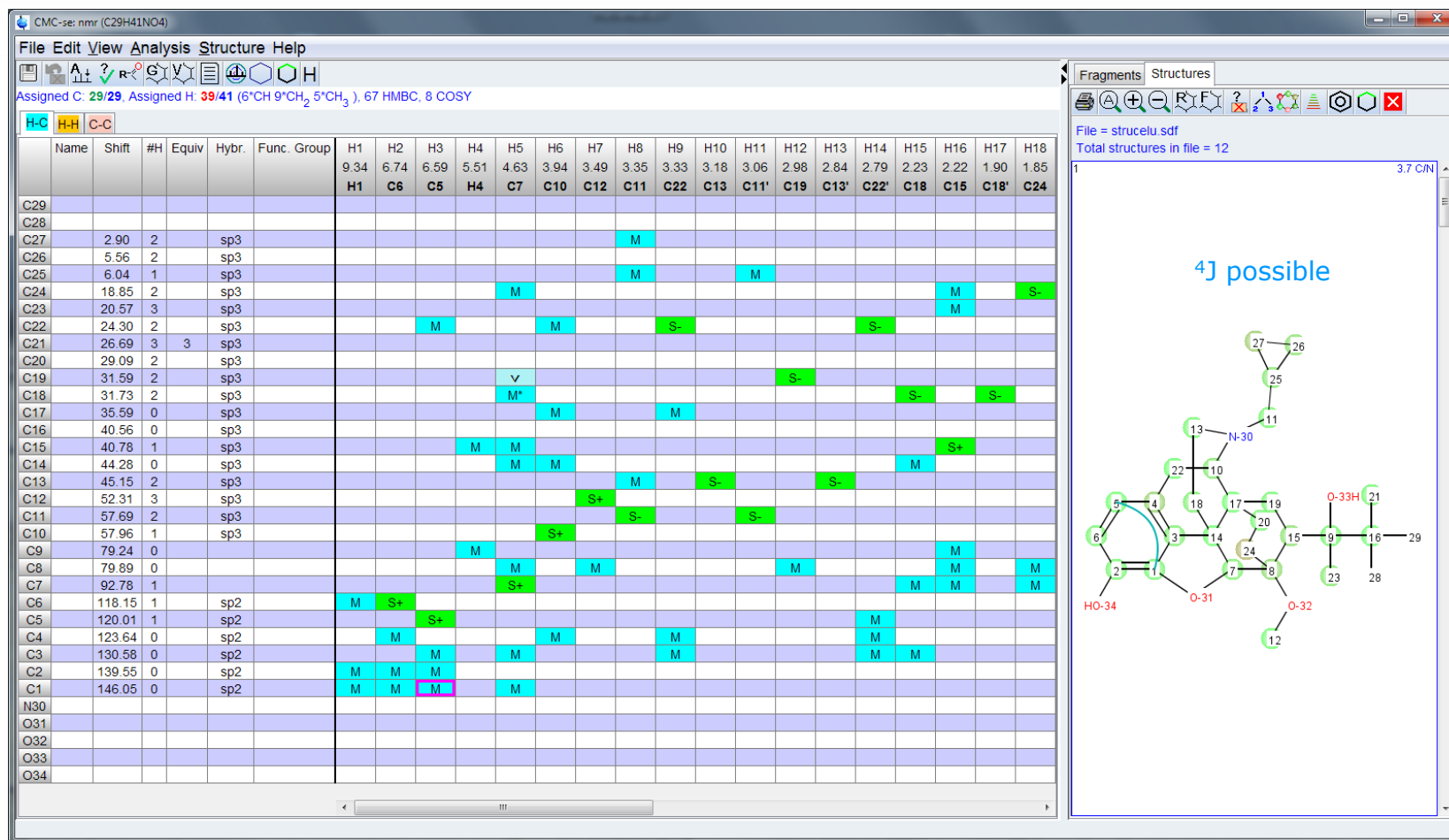
Verify structure proposals based on the correlation table



# CMC-se – Assisted Structure Verification



# CMC-se – Assisted Structure Verification



# CMC-se – Assisted Structure Verification

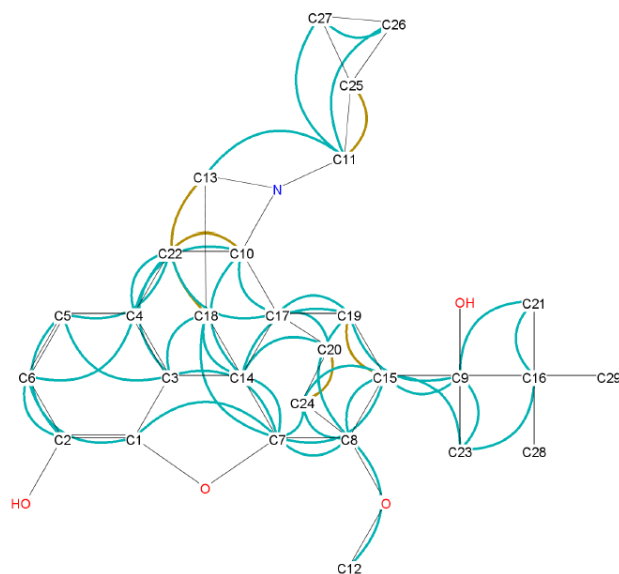


## CMC-se Report

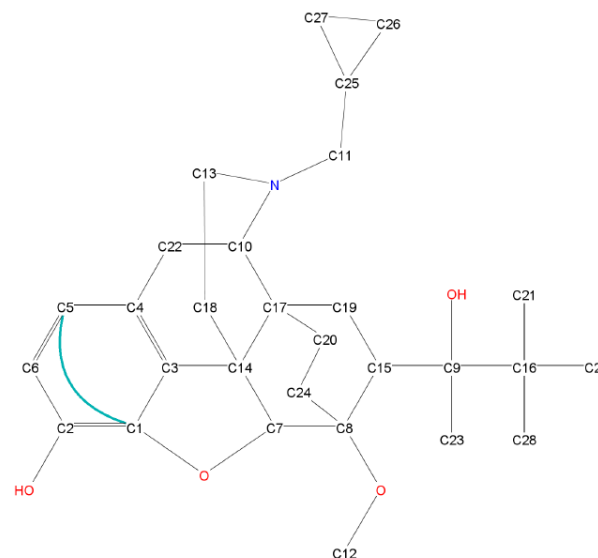
*nmr (C:\Daten\Buprenorphine\data\nmr\strucelu)*

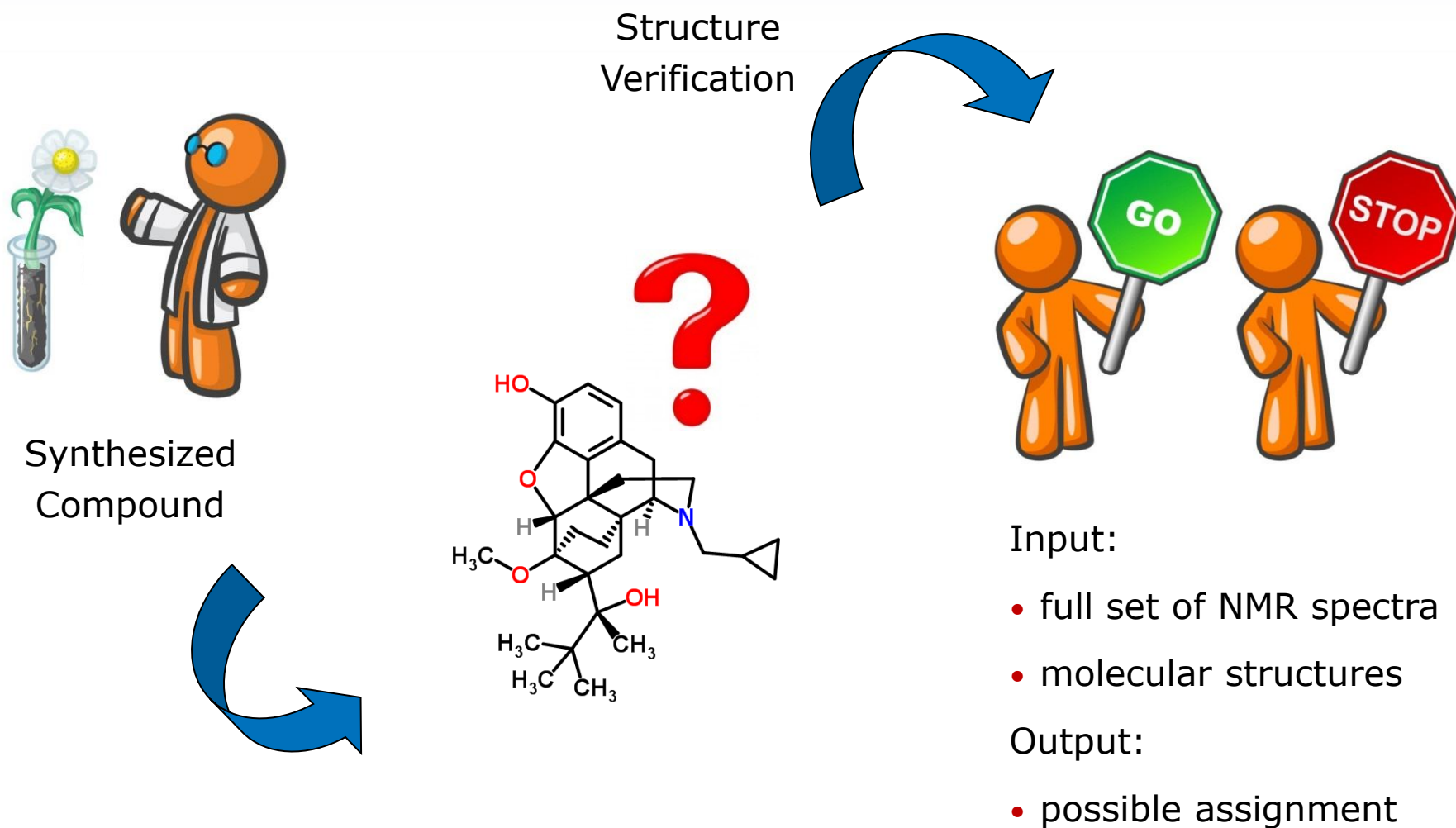


### Explained Correlations



### Incorrect Correlations





## mass spectrometry:

- ADC digitizer technology for accurate mass determination
- software package SmartFormula:
  - analysis of accurate mass and TIP for MS and MS/MS

## NMR spectroscopy:

- small volume probes: increase mass sensitivity
- cryoProbes: decrease noise and thus increase signal to noise
- software package CMCse:
  - assisted structure verification and elucidation
  - automated analysis of 1D and 2D NMR spectra

# Acknowledgment



Wim Vermeulen



Lukas Oberer



Application & Development Team

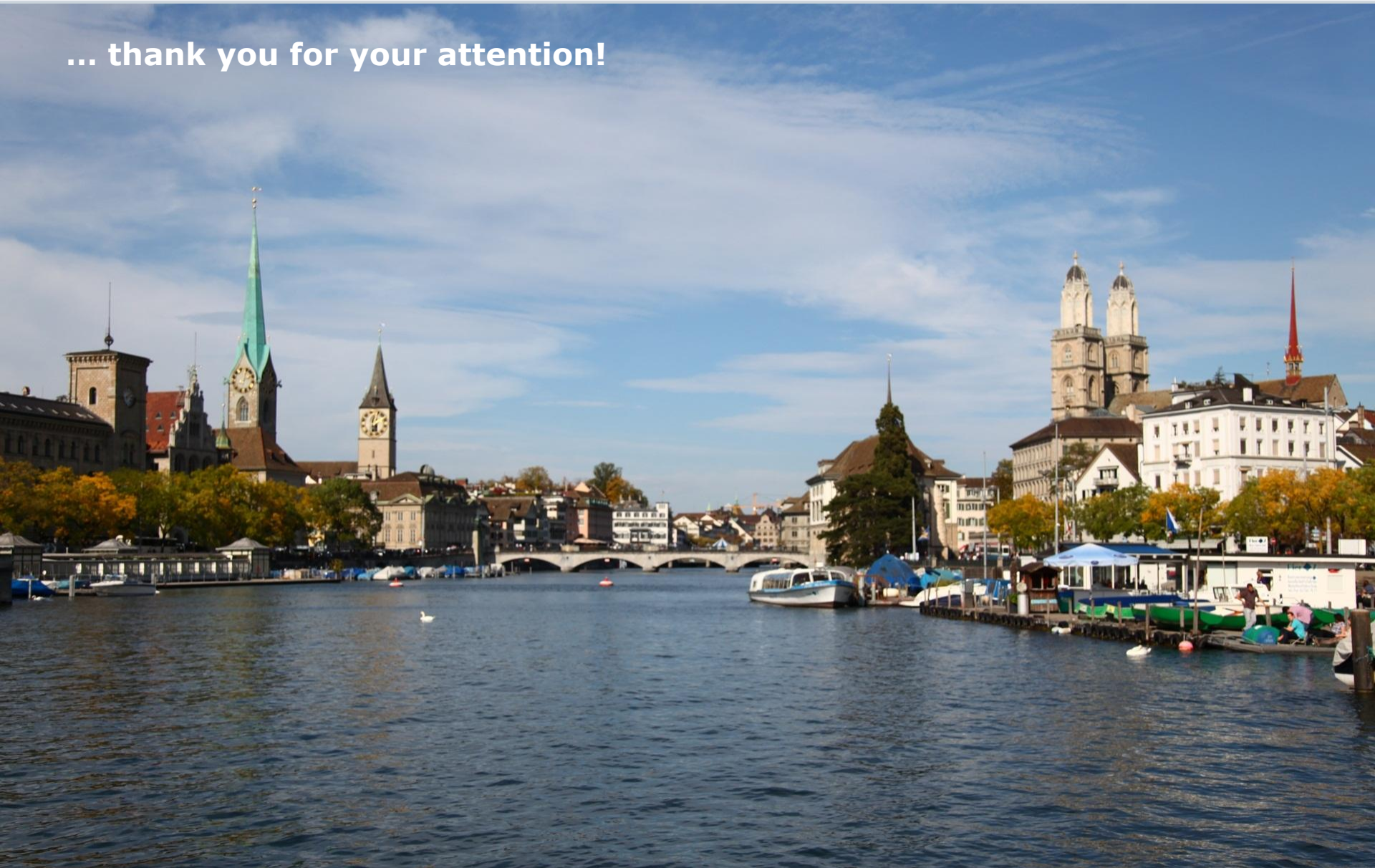


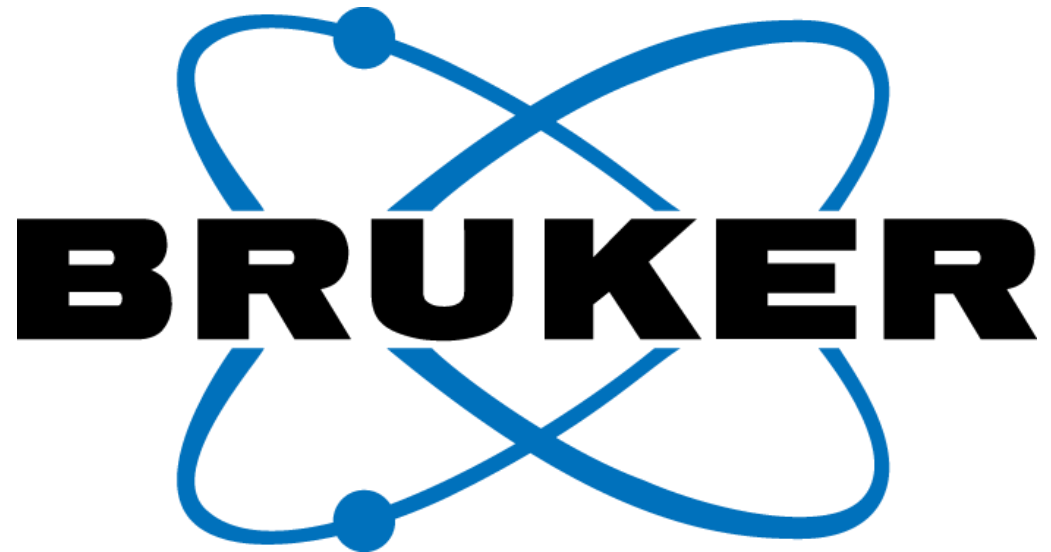


# Acknowledgment



... thank you for your attention!





Innovation with Integrity