

On-line tools for the interpretation of NMR and MS-spectra

Roland Stenutz

*Department of Organic Chemistry,
Stockholm University
SE-106 91 Stockholm, Sweden*

roland@organ.su.se

On-line tools for the interpretation of NMR and MS-spectra



Overview

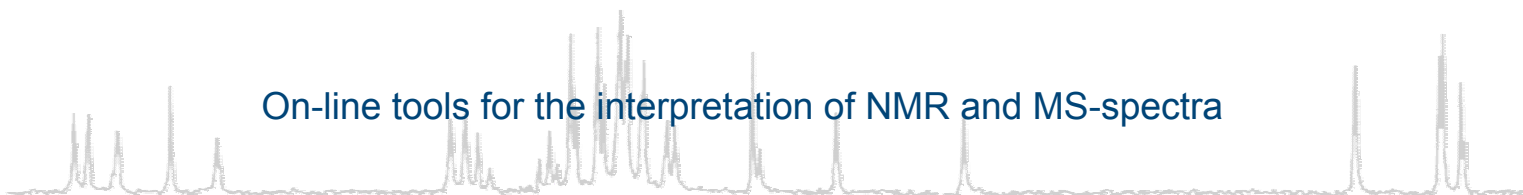
What is the problem?

How do you ask the right question?

What answer do you get?



On-line tools for the interpretation of NMR and MS-spectra



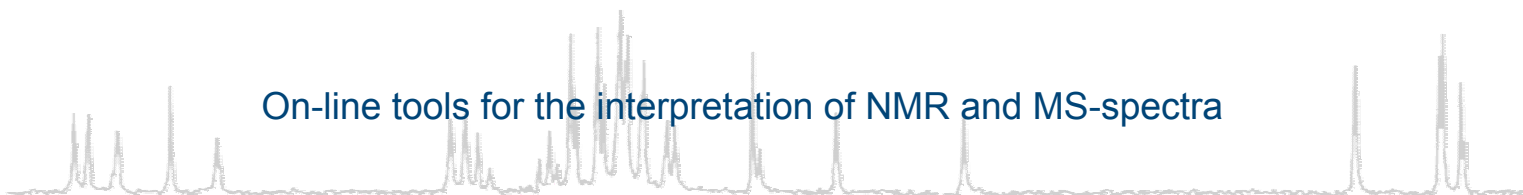
The problem

You have a spectrum of a (pure?) compound and want to know its structure...

...because other information about a compound can often be found only once its structure (identity) is known.



On-line tools for the interpretation of NMR and MS-spectra



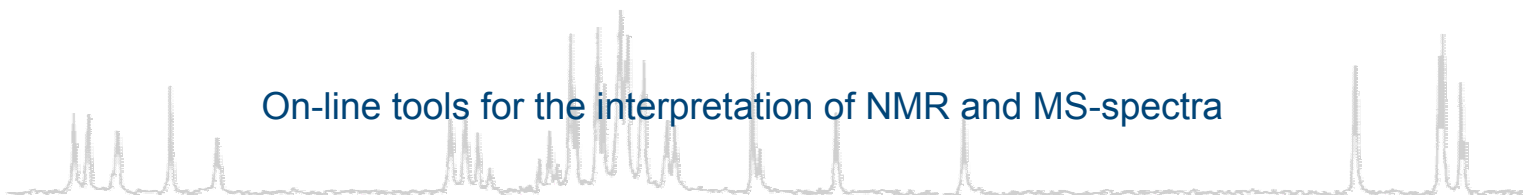
The question

What compounds have similar spectra?

Try to narrow down the problem using additional information – source, physical and chemical properties etc.



On-line tools for the interpretation of NMR and MS-spectra



The answer

You may

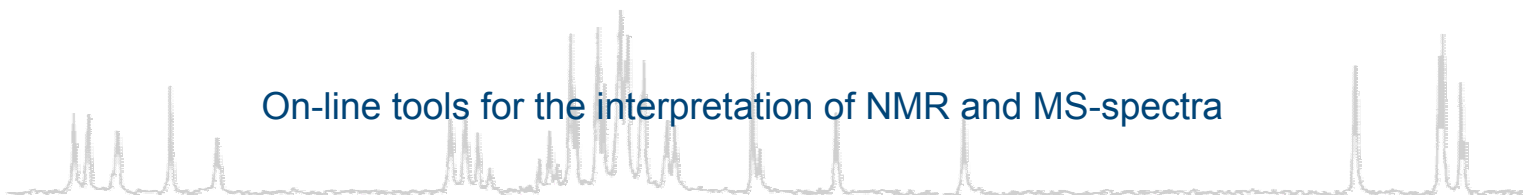
...identify the compound

...identify the class of compounds it belongs to
(partial structure)

...or just find out that it is unusual and worth
investigating



On-line tools for the interpretation of NMR and MS-spectra



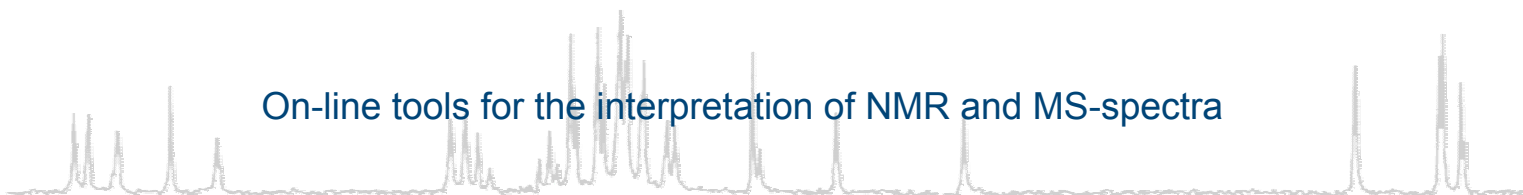
The answer

How well you succeed depends on...

...the amount and quality of the data

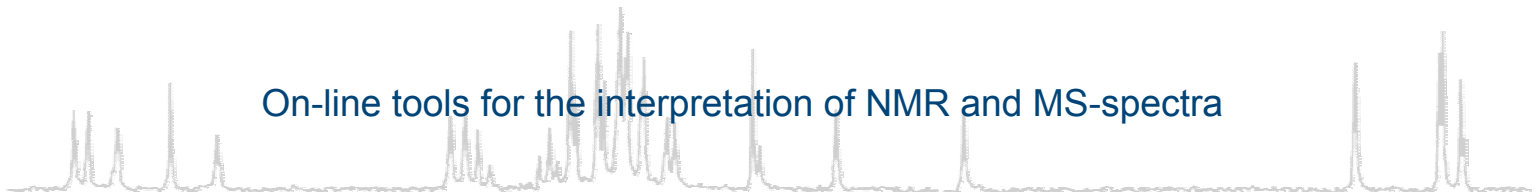
...the type of data (MS vs. NMR)

...the additional restraints you may impose on the answers



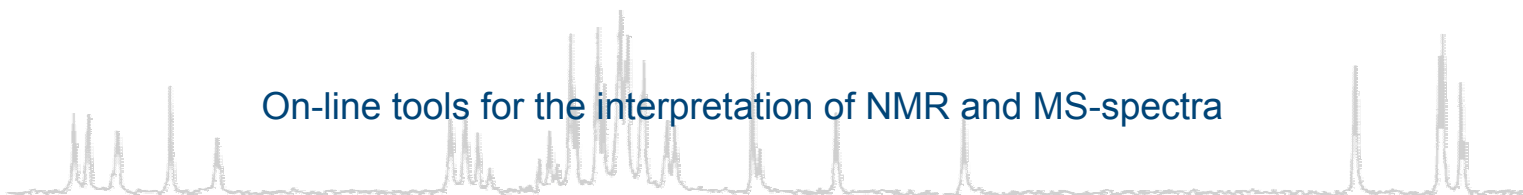
Polysaccharide structure:

Components	"type"	<i>Hex or HexNAc</i>
	relative configuration	<i>Glc or Man</i>
	absolute configuration	<i>D- or L-</i>
	ring size	<i>-p or -f</i>
Linkages	position	$\rightarrow 4)$ or $\rightarrow 6)$
	stereochemistry	$\alpha-$ or $\beta-$
Sequence		$\rightarrow 4)Glc(\rightarrow 4)Gal(\rightarrow$ or $\rightarrow 4)Gal(\rightarrow 4)Glc(\rightarrow$



NMR can be used to perform all of the steps in a structure determination
- *except the determination of the absolute configuration*

MS is insensitive to stereochemistry
- *e.g. can't distinguish between Glc & Man*



Part 1: NMR

Structure of bacterial polysaccharides

^{13}C -NMR

Polymers of repeating units

→ CASPER

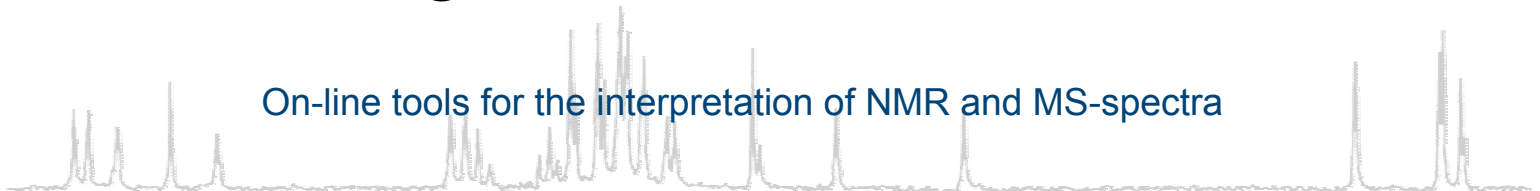
Structure of oligosaccharides from GP

^1H -NMR

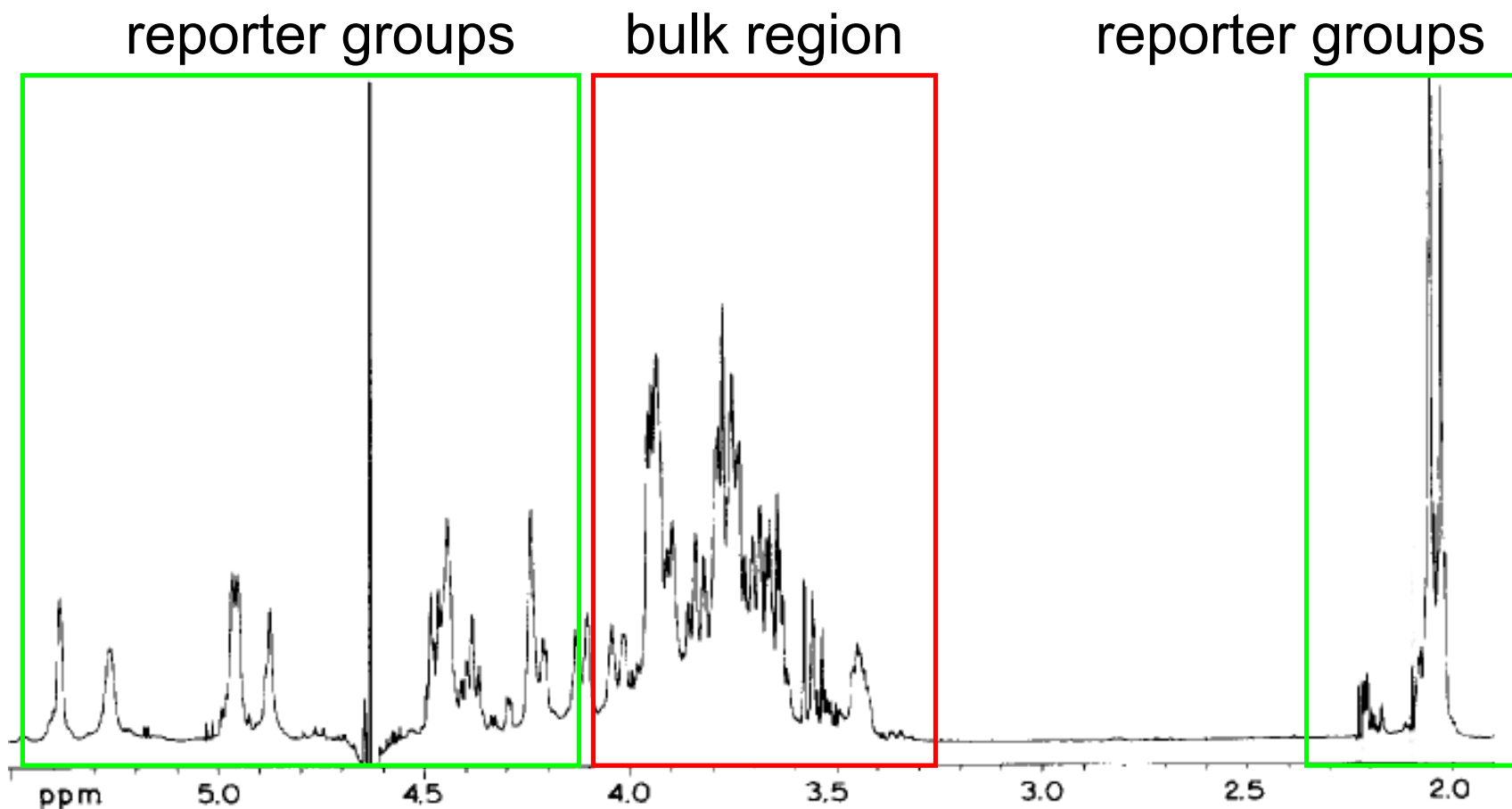
→ SugaBase



On-line tools for the interpretation of NMR and MS-spectra

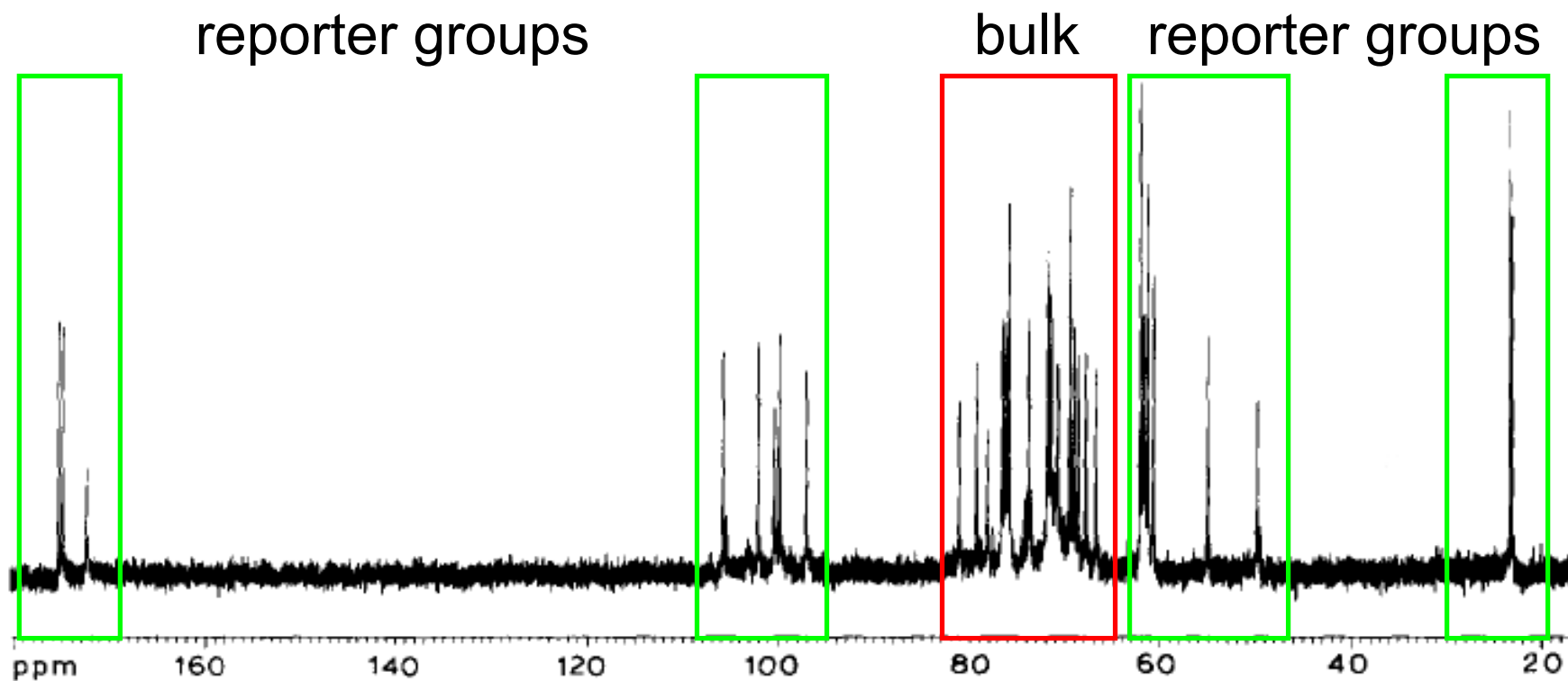


Typical ^1H spectrum (EC O113)



On-line tools for the interpretation of NMR and MS-spectra

Typical ^{13}C spectrum (EC O113)

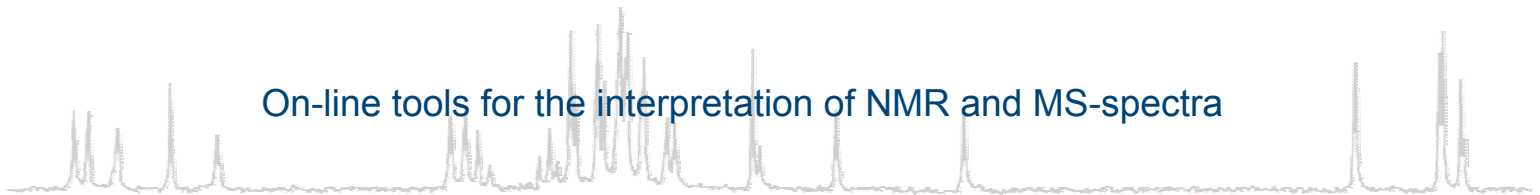


On-line tools for the interpretation of NMR and MS-spectra

Every polysaccharide has a unique 1D-NMR-spectrum -

i.e. all of the information about the structure is contained in a 1D-spectrum.

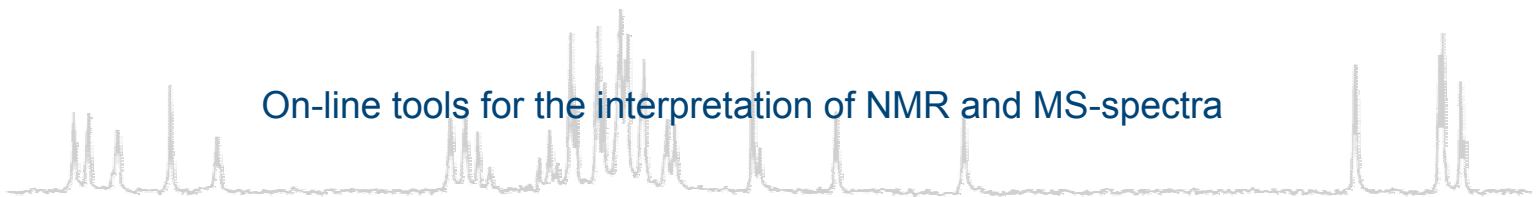
Most of the NMR-experiments are performed to assign the resonances and do not provide additional information about the structure.



Methods for the interpretation of
1D-spectra can save much time
and effort!



On-line tools for the interpretation of NMR and MS-spectra



Current approaches:

comparison with a database (SugaBase)

simple and accurate but limited to known structures or sub-structures.

comparison with simulated NMR spectra (CASPER)

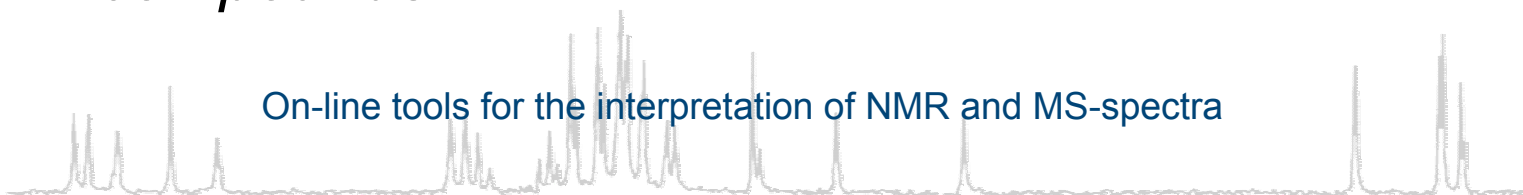
requires information about the components and linkages to limit the number of possible structures.

Artificial Neural Networks (ANN)

current application are limited to a single class of compounds.



On-line tools for the interpretation of NMR and MS-spectra



SugaBase

(www.boc.chem.uu.nl/sugabase/databases.html)

Carbon Chemical Shifts

Tolerance Limit: ppm.

Match Percentage: %.

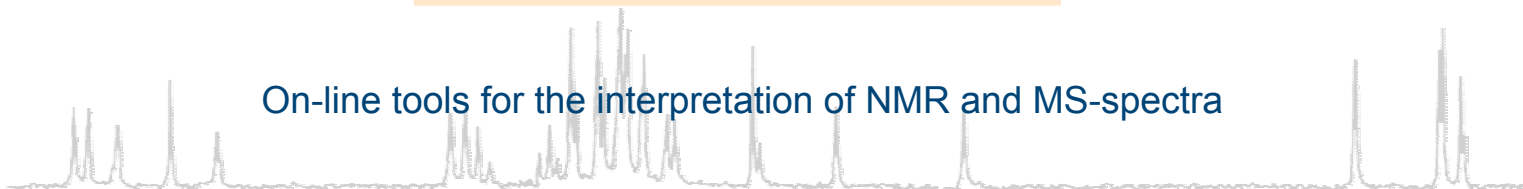
List of Carbon Chemical Shifts:

99.86
70.92
83.65
69.88

Carbon chemical shifts are relative to:

- ♦ *dioxane* = 67.40 ppm.
- ♦ *Methanol* = 49.72 ppm
- ♦ *Acetone* = 31.08 ppm
- ♦ *TMS* = 0.87 ppm (external)
- ♦ *DSS* = -1.84 ppm
- ♦ *TSP* = -1.97 ppm

On-line tools for the interpretation of NMR and MS-spectra



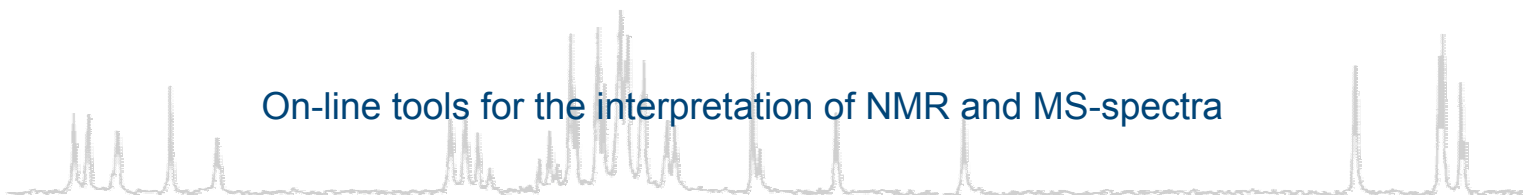
SugaBase

C#: P-0201-A00404
CC: CCSD:A00404
MHz 75
Temp 333
Solv D2O
Original Reference: Acetone
Reference Value : 31.45
Correction Applied: -0.37

CCSD-data + NMR

Residue	Linkage	Carbon	PPM	J	Hz	Note
a-D-Glcp		C-1	99.7			
		C-2	70.8			
		C-3	83.5			
		C-4	69.7			
		C-5	72.2			
		C-6	61.3			
b-L-Fucp	3	C-1	104.0			
		C-2	71.8			
		C-3	73.7			
		C-4	72.1			
		C-5	71.8			
		C-6	16.1			

On-line tools for the interpretation of NMR and MS-spectra



SweetDB – *glycosciences.de*



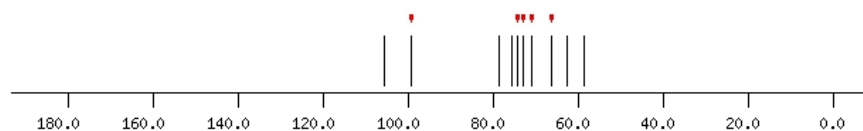
LinucsID: [Explore](#)

SweetDB [Bibliography](#) [Structure](#) [NMR](#) [MS](#) [Tools](#) [Related DBs](#) [About](#)

Searched for **peaks (spectra view)**. Results: **1 - 10 of 126**

[Structures](#)

[Details](#)

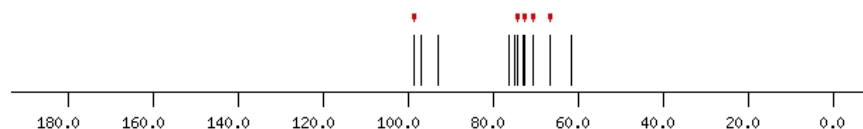


Spectrum matches query with 58.33%

[Explore](#)

[Structures](#)

[Details](#)



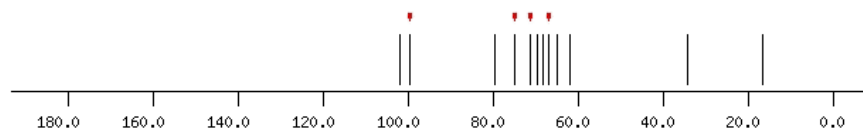
Spectrum matches query with 53.33%

[Explore](#)

[3D Co-ordinates](#)

[Structures](#)

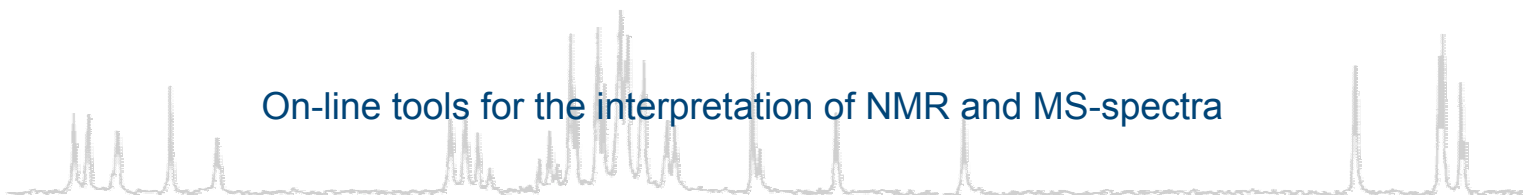
[Details](#)



Spectrum matches query with 53.33%



[Explore](#)

On-line tools for the interpretation of NMR and MS-spectra



CASPER

(<http://www.casper.organ.su.se/casper>)

CASPER

HomeResearchAnalysisECDBCASPERKe3690

WelcomeHelpSimulateDetermine Sequence

Title

Source

Residue	Linkage position					
	1	2	3	4	5	6
<input type="text" value="D-Glcp"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
<input type="text" value="none"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="text" value="none"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="text" value="none"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="text" value="none"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="text" value="none"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="text" value="none"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Chemical shifts ☒ ^{13}C ☐ ^1H

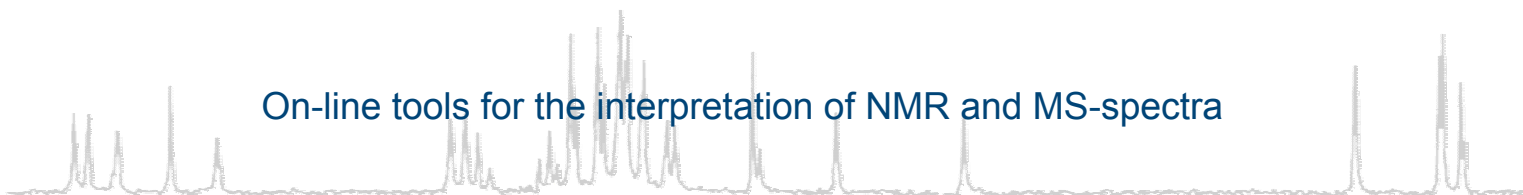
Correct by subtracting ppm
Number of shifts - Required: Actual:

Minimum number of coupling constants of different magnitudes

	small	medium	large
$^3J_{\text{HH}}$	<input type="text" value="0"/> (<2 Hz)	<input type="text" value="0"/> (2-7 Hz)	<input type="text" value="0"/> (>7 Hz)
$^1J_{\text{CH}}$	<input type="text" value="0"/> (<169 Hz)	<input type="text" value="0"/> (>169 Hz)	

As MIME ☒

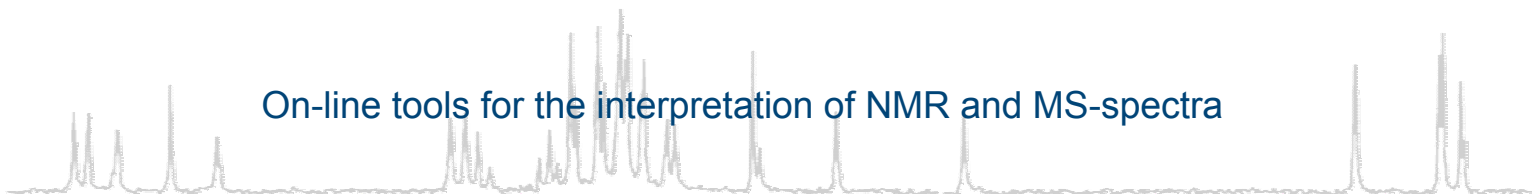
On-line tools for the interpretation of NMR and MS-spectra



CASPER

Residue	Linkage position					
	1	2	3	4	5	6
D-Glcp ▼	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
none ▼	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
none ▼	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Data from methylation analysis is used to limit the number of structures generated.



CASPER



Home

Research

Analysis

ECDB

CASPER



Ke3690

Welcome

Help

Simulate

Determine Sequence

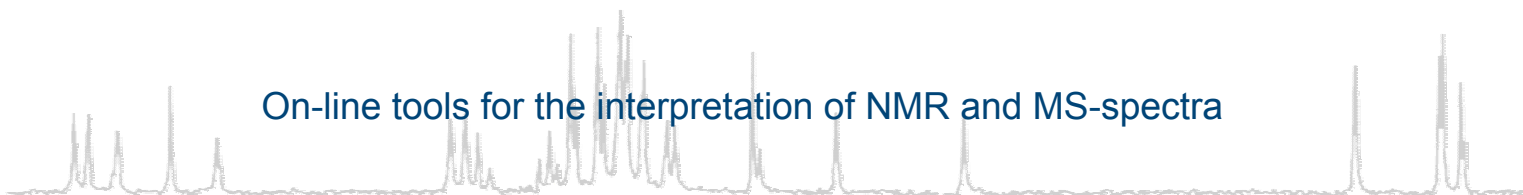
Results of calculation

Please cite as: *Computer-assisted structural analysis of oligo- and polysaccharides: An extension of CASPER to multibranched structures*
R. Stenutz, P.-E. Jansson and G. Widmalm; *Carbohydr. Res.* 306(1998) 11-17; [PubMed 9691437](#)
URL: <http://www.casper.org.se/casper.html>

1. ☐ ->6)aDGlc(1-> , error=0.65ppm (0.11)
2. ☐ ->6)bDGlc(1-> , error=15.81ppm (2.64)

Simulated structures ranked by fit

On-line tools for the interpretation of NMR and MS-spectra



CASPER

Simulated structure

->6)aDGlcⁱ(1->

->6)aDGlc ⁱ (1->	98.90	72.39	74.32	70.66	71.28	66.90	
	4.97	3.59	3.74	3.54	3.90	3.98	3.77

Assignment of ¹³C resonances

Experimental	Simulated	Exp-Sim	Assignment
99.00	98.90	0.10	aDGlc ⁱ - 1
74.50	74.32	0.18	aDGlc ⁱ - 3
72.50	72.39	0.11	aDGlc ⁱ - 2
71.30	71.28	0.02	aDGlc ⁱ - 5
70.70	70.66	0.04	aDGlc ⁱ - 4
66.70	66.90	-0.20	aDGlc ⁱ - 6

Error=0.65 ppm (0.11/shift), Systematic error=0.04 ppm, RMS error=0.13 ppm

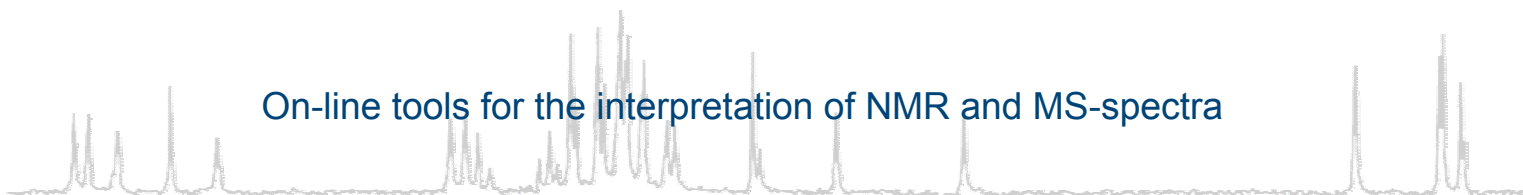
Experimental structure

->6)aDGlcⁱ(1->

->6)aDGlc ⁱ (1->	99.00	72.50	74.50	70.70	71.30	66.70	
	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.

JCAMP-format

On-line tools for the interpretation of NMR and MS-spectra

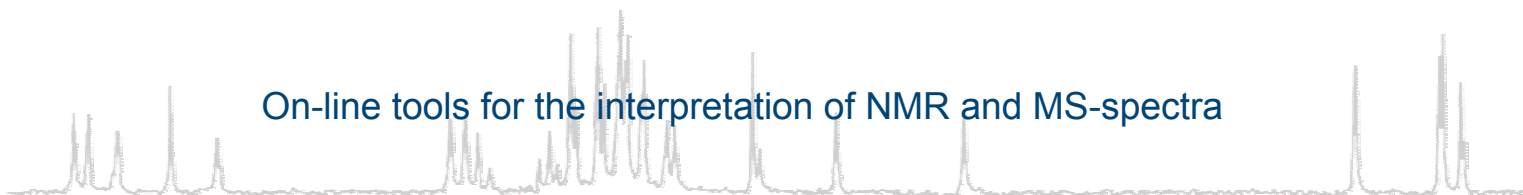


CASPER

Experimental	Simulated	Exp-Sim	Assignment
99.00	98.90	0.10	aDGlc ⁱ - 1
74.50	74.32	0.18	aDGlc ⁱ - 3
72.50	72.39	0.11	aDGlc ⁱ - 2
71.30	71.28	0.02	aDGlc ⁱ - 5
70.70	70.66	0.04	aDGlc ⁱ - 4
66.70	66.90	-0.20	aDGlc ⁱ - 6



On-line tools for the interpretation of NMR and MS-spectra



Simulation of NMR

Title
Source

	Residue	Linkage	'Reducing' end
1)	<input type="text" value="a"/> <input type="text" value="D-Glcp"/>	<input type="text" value="(->6)"/>	<input type="text" value="self"/>
2)	<input type="text" value="a"/> <input type="text" value="none"/>	<input type="text" value="(->2)"/>	<input type="text" value="residue 1"/>
3)	<input type="text" value="a"/> <input type="text" value="none"/>	<input type="text" value="(->2)"/>	<input type="text" value="residue 1"/>
4)	<input type="text" value="a"/> <input type="text" value="none"/>	<input type="text" value="(->2)"/>	<input type="text" value="residue 1"/>
5)	<input type="text" value="a"/> <input type="text" value="none"/>	<input type="text" value="(->2)"/>	<input type="text" value="residue 1"/>
6)	<input type="text" value="a"/> <input type="text" value="none"/>	<input type="text" value="(->2)"/>	<input type="text" value="residue 1"/>
7)	<input type="text" value="a"/> <input type="text" value="none"/>	<input type="text" value="(->2)"/>	<input type="text" value="residue 1"/>

¹³C-Chemical shifts

99.0 72.5 74.5 71.3 70.7 66.7

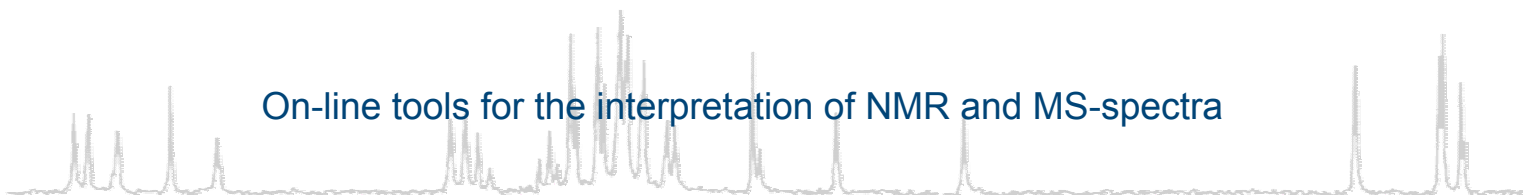
Correct by subtracting ppm

¹H-Chemical shifts

Correct by subtracting ppm

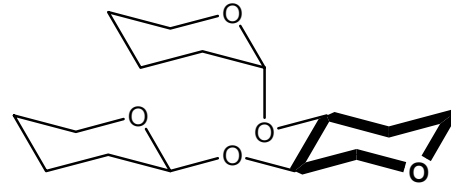
As MIME ☐

On-line tools for the interpretation of NMR and MS-spectra



Chemical shift calculation

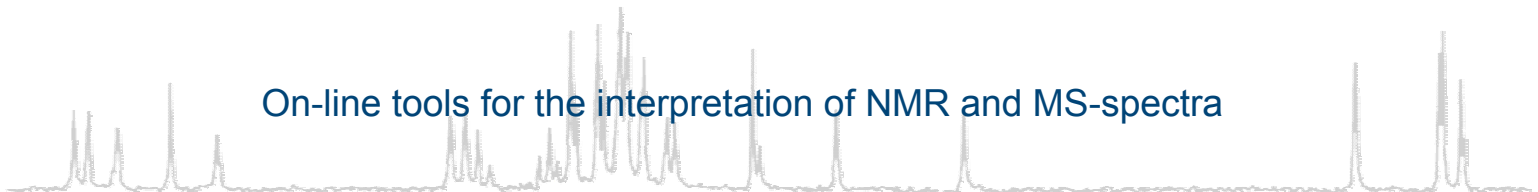
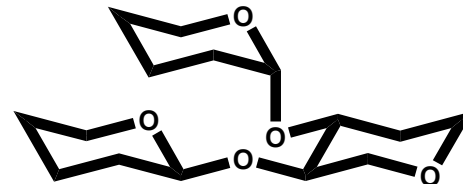
1) Start with monosaccharide



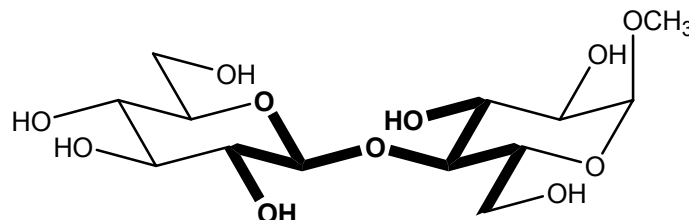
2) Add glycosylation shifts



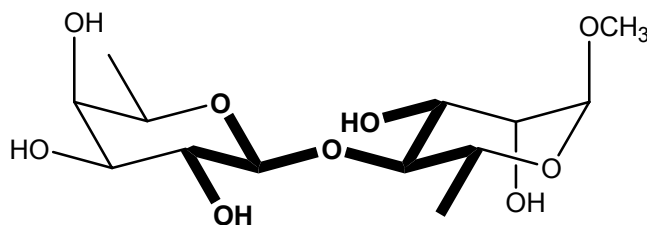
3) Add steric corrections



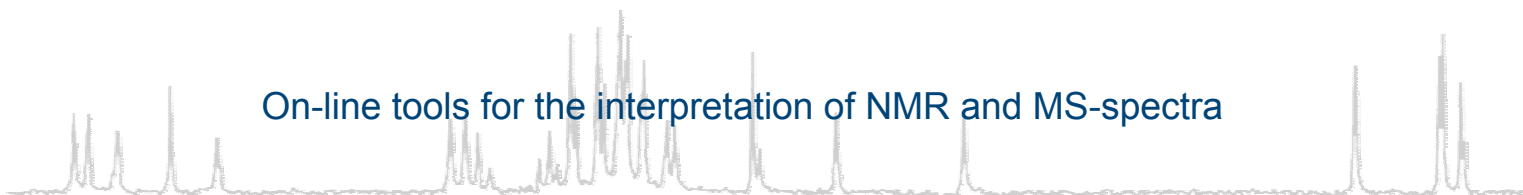
Glycosylation shifts



β DGlc (1→	6.45	-1.07	-0.18	-0.20	0.12	-0.19
→4) α DGlcOMe	-0.27	-0.25	-1.51	9.20	-1.36	-0.60



β DFuc (1→	6.85	-1.05	-0.13	-0.24	0.18	-0.23
→4) α DRhaOMe	-0.30	-0.66	-1.23	10.19	-1.35	-0.07

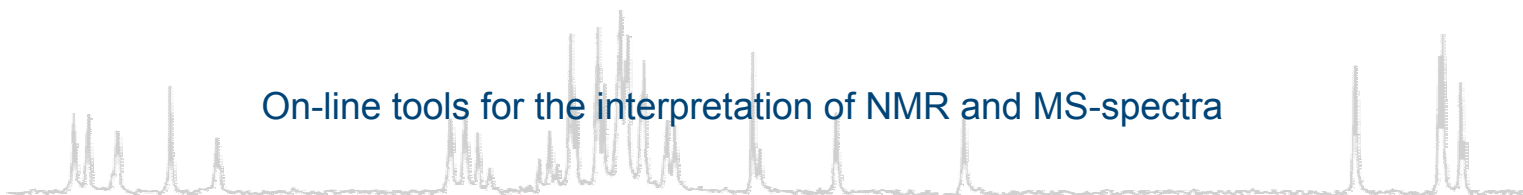


Caveats


It is assumed that the conformation in the PS is the same as in the disaccharide or trisaccharide fragments


In order to have transferable glycosylation shifts it is also assumed that the monosaccharides are rigid.

Extending the set of disaccharide and trisaccharide fragments used in the calculations may reduce these problems - but they are inherent to the approach.



ECDB – *E. coli* O-antigens



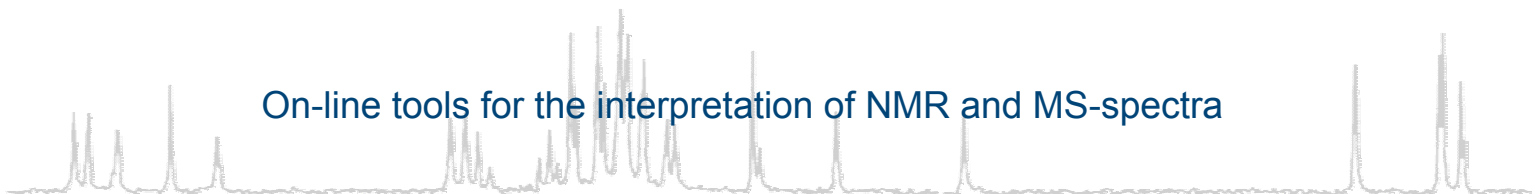
E. coli O-antigen Data Base



HomeResearchAnalysis**ECDB**CASPERKe3690

MainComponentsListSearch

ID	Serogroup	Serotype	Strain
85% <u>19</u>	O101	O101:K-:H33	[CCUG11402]
->6) aDGlcNAc (1->4) aDGalNAc (1->			
84% <u>49</u>			73-1
aDGlc (1->3) ->4) aDMan (1->2) aDMan (1->2) bDMan (1->3) aDGalNAc (1->			
82% <u>158</u>	O158		

On-line tools for the interpretation of NMR and MS-spectra



 <i>E. coli</i> O-antigen Data Base 	
Home	Research
Analysis	ECDB
CASPER	Ke3690
<i>Main</i>	<i>Components</i>
<i>List</i>	<i>Search</i>
Serogroup	O101
Serotype	O101:K:-H33
Strain	<u>CCUG 11402</u>
ECDB#	19
Structure	->6) aDGlcNAc (1->4) aDGalNAc (1->
CarbBank	
LinucsID	
Structure code	P2
Sugar components	DGlcNAc, DGalNAc
Non-sugar components	
Comments	Structure [1]
Identical to	
Cross-reacts with	
¹ H-NMR	
Conditions	Solvent: D2O; Temp.: 40C
Reference	TSP=0.00ppm
Comment	
¹³ C-NMR	175.4 175.2 99.2 98.0 78.3 72.5 72.0 71.3 70.7 68.1 66.2 61.1 54.9 50.9 22.9 22.7
Conditions	Solvent: D2O; Temp.: 40C
Reference	dioxan=67.4ppm
Comment	¹³ C-NMR [1] and unpublished
References	[1] Carbohydr. Res. 297 (1997) 297-299 [PubMed 9060191]

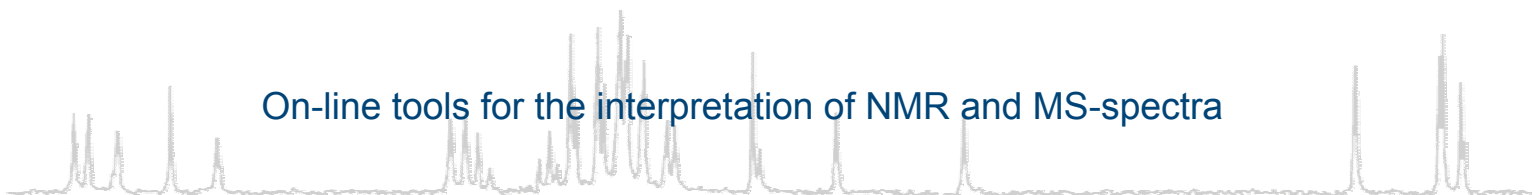
NMR

structure

bibliography



On-line tools for the interpretation of NMR and MS-spectra



SDBS –small molecules

www.aist.go.jp/RIODB/SDBS/menu-e.html

Welcome to SDBS

Integrated Spectral Data Base System for Organic Compounds

[Japanese](#)

Last updated : March 25, 2004

[National Institute of Advanced Industrial Science and Technology](#)
[Tsukuba, Ibaraki, Japan](#)

NMR: T.Saito, K.Hayamizu, M.Yanagisawa and O.Yamamoto
MS: N.Wasada
ESR: K.Someno
IR: S.Kinugasa, K.Tanabe and T.Tamura
Raman: K.Tanabe and J.Hiraishi

[Introduction](#)

[How to use](#)

[Search Compounds / Search NMR & MS / Display Spectra](#)

Access to this database is free of charge, however we would request users not to download more than 50 spectra and/or compound structure files in one day. All accesses are logged. If, for some specific purpose, more spectra are required the user should consult us, and explain the intended usage for the data. We also request that when the database data is used in a publication or presentation, a proper acknowledgement is given such as:
SDBSWeb: <http://www.aist.go.jp/RIODB/SDBS/> (access date)

On-line tools for the interpretation of NMR and MS-spectra

