

Structure Determination of Carbohydrates

Internet tools for the interpretation
of NMR spectra

Roland Stenutz

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

roland@organ.su.se

Internet tools for the interpretation of NMR spectra
Roland Stenutz, Organic Chemistry, Stockholm University



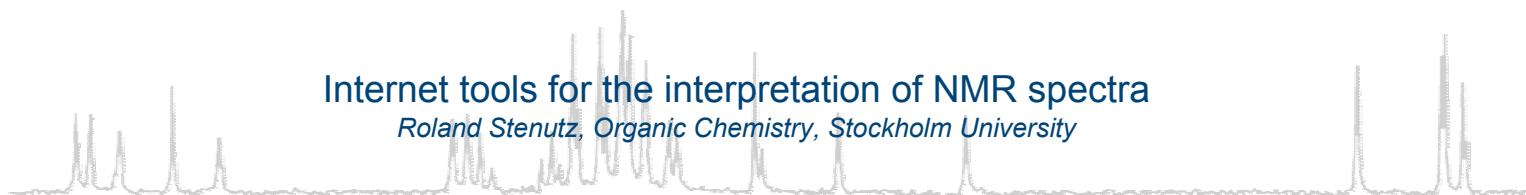
Polysaccharide structure:

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
		$\rightarrow 4)Gal(\rightarrow 4)Glc(\rightarrow$



Internet tools for the interpretation of NMR spectra

Roland Stenutz, Organic Chemistry, Stockholm University



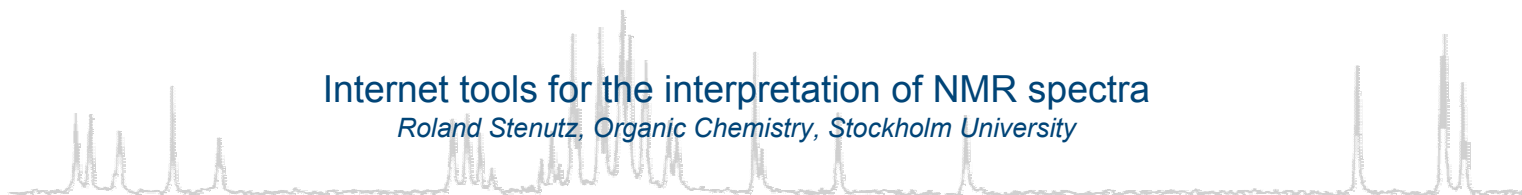
NMR can be used to perform all of the steps in a structure determination

- *except the determination of the absolute configuration*



Internet tools for the interpretation of NMR spectra

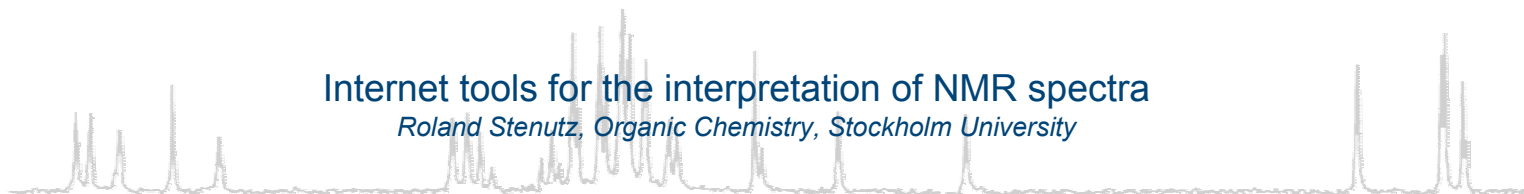
Roland Stenutz, Organic Chemistry, Stockholm University



Many oligo- and polysaccharides contain at least one sugar that occurs naturally only in one configuration.

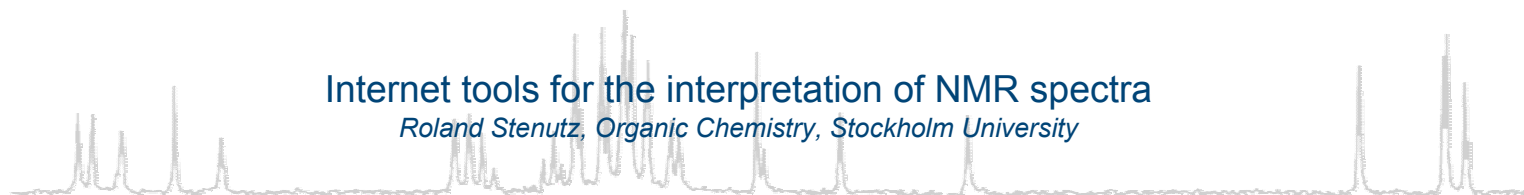
e.g. D-Glc, D-GlcA, D-GlcNAc, D-GalNAc

The relative configuration of the other residues can be determined by NMR.



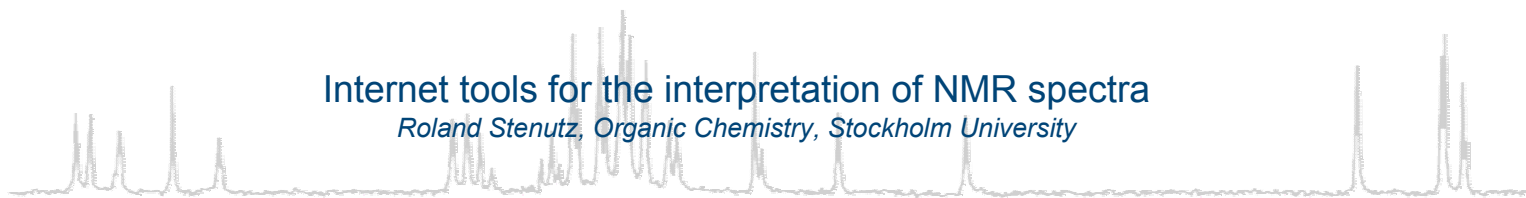
The interpretation of NMR spectra becomes difficult if the polysaccharide -

- *is irregular (or a mixture)*
- *contains many similar residues*
- *is very large*



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.



NMR-experiments

Assignment of ^1H -resonances

$^1\text{H}, ^1\text{H}$ -COSY, $^1\text{H}, ^1\text{H}$ -TOCSY

Assignment of ^{13}C -resonances

$^1\text{H}, ^{13}\text{C}$ -HETCOR or HSQC

Sequence determination

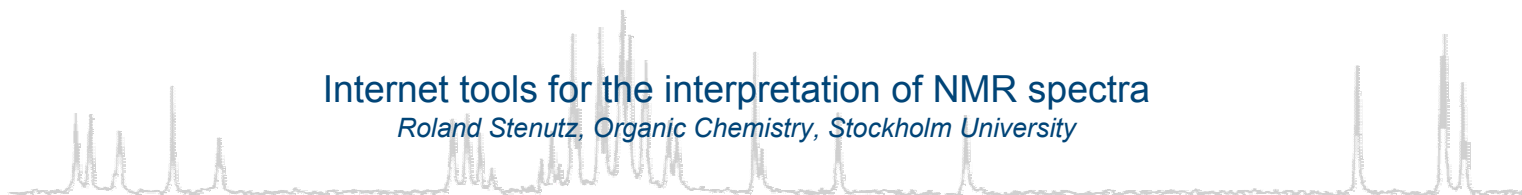
NOESY (Short $^1\text{H}, ^1\text{H}$ distances)

HMBC ($^3J_{\text{COCH}}$)



Internet tools for the interpretation of NMR spectra

Roland Stenutz, Organic Chemistry, Stockholm University

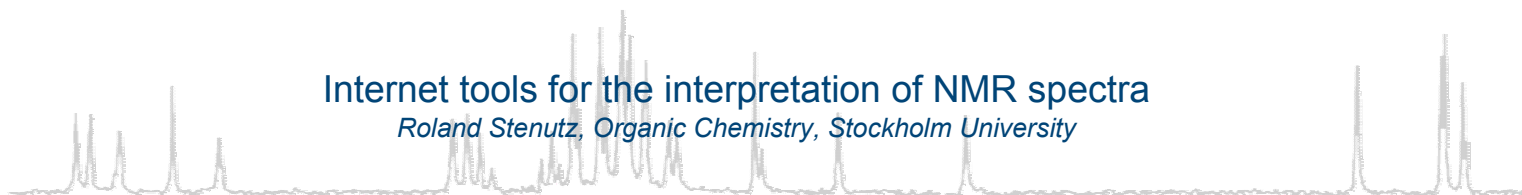


Methods for the interpretation of
1D-spectra can save much time
and effort!



Internet tools for the interpretation of NMR spectra

Roland Stenutz, Organic Chemistry, Stockholm University



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.

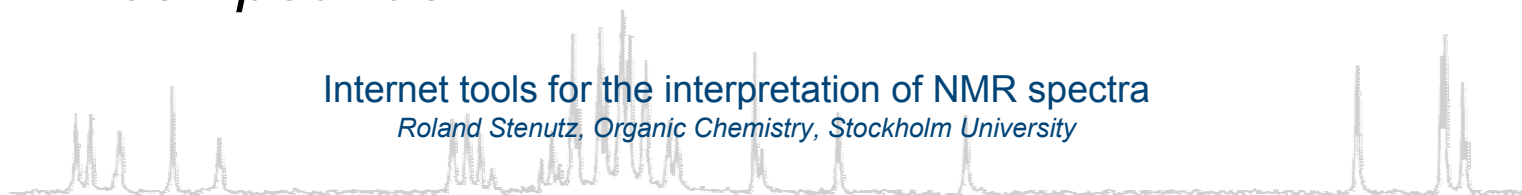
Artificial Neural Networks (ANN)

current application are limited to a single class of compounds.



Internet tools for the interpretation of NMR spectra

Roland Stenutz, Organic Chemistry, Stockholm University



SugaBase

(<http://www.boc.chem.uu.nl/sugabase/databases.html>)

Carbohydrate Structure

Help: [\[Search level\]](#) [\[Strict linkages\]](#) [\[Nomenclature\]](#)

Search level:

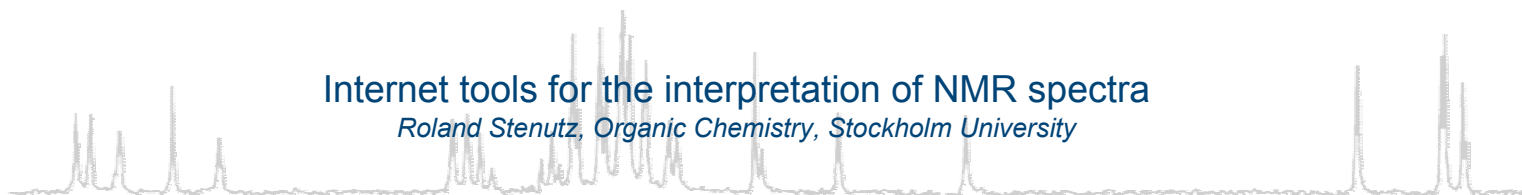
Strict linkage check: ☐

Carbohydrate Structure:



Internet tools for the interpretation of NMR spectra

Roland Stenutz, Organic Chemistry, Stockholm University



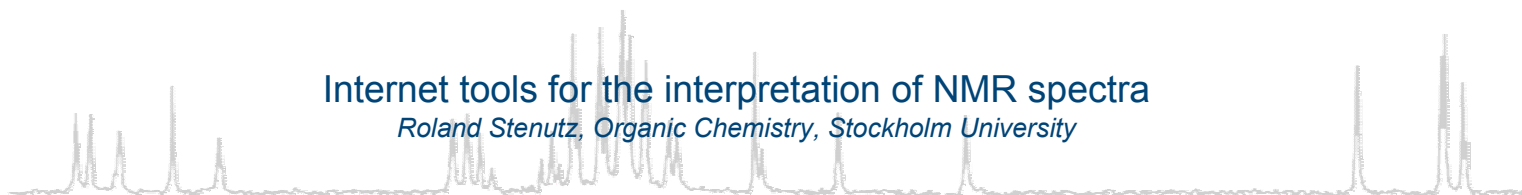
SugaBase

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

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			

Internet tools for the interpretation of NMR spectra

Roland Stenutz, Organic Chemistry, Stockholm University



Sweet-DB

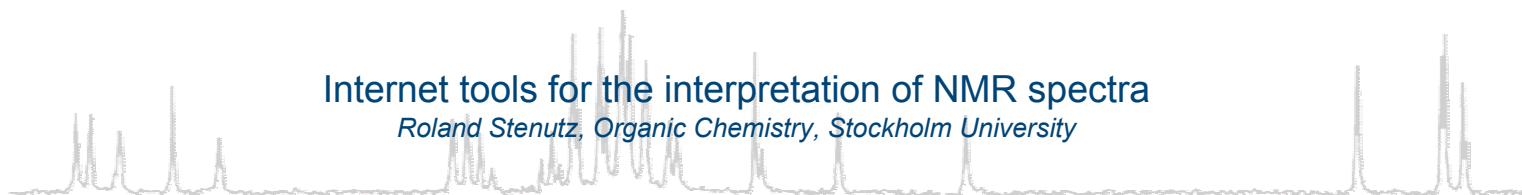


<http://glycosciences.de> hosts several tools and databases for the carbohydrate chemist

Sweet-DB is an interface to CarbBank/SugaBase



Internet tools for the interpretation of NMR spectra
Roland Stenutz, Organic Chemistry, Stockholm University



Search with NMR-data

NMR Information / Advanced Search

Peaks :

99.7	70.8	83.5	69.7	72.2	61.3
104.0	71.8	73.7	72.1	71.8	16.1

☐ Proton ☒ Carbon

Tolerance : PPM

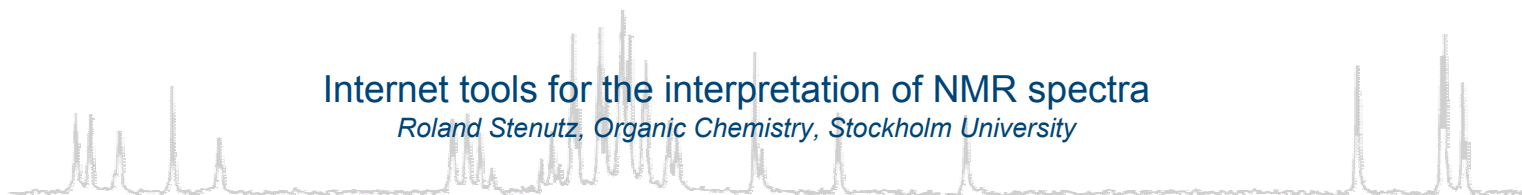
[Example 1: Proton Search](#)

[Example 2: Carbon Search](#)



Internet tools for the interpretation of NMR spectra

Roland Stenutz, Organic Chemistry, Stockholm University

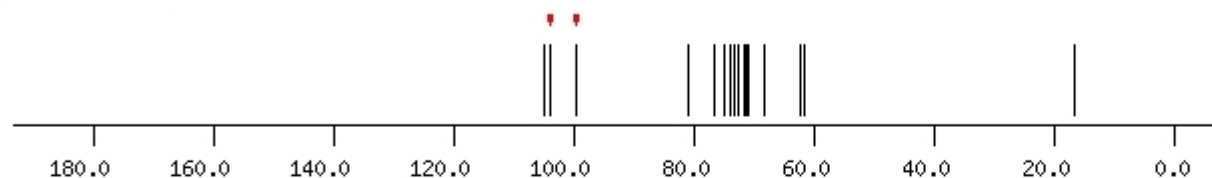


Search results

Searched for **peaks (spectra view)**. Results: **1 - 10** of **174**

Structures

Details



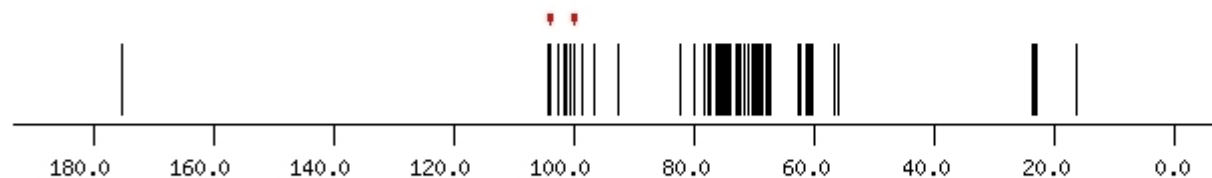
Spectrum matches query with 95.00%

Explore

3D Co-ordinates

Structures

Details



Spectrum matches query with 94.00%

Explore

3D Co-ordinates

Structures

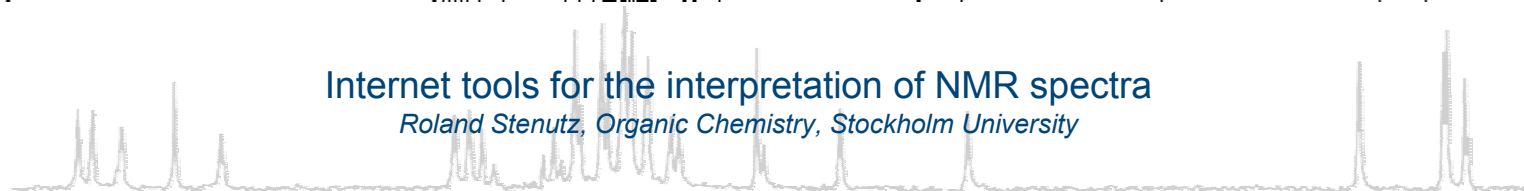
Details



Spectrum matches query with 92.00%

Internet tools for the interpretation of NMR spectra

Roland Stenutz, Organic Chemistry, Stockholm University



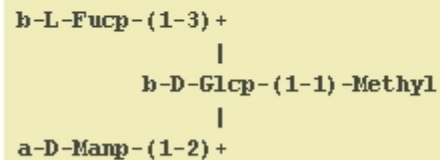
Search results

Searched for **nmr information**. Results: 1 - 10 of 174

Spectra

Details

Structure matches query with 95.00%



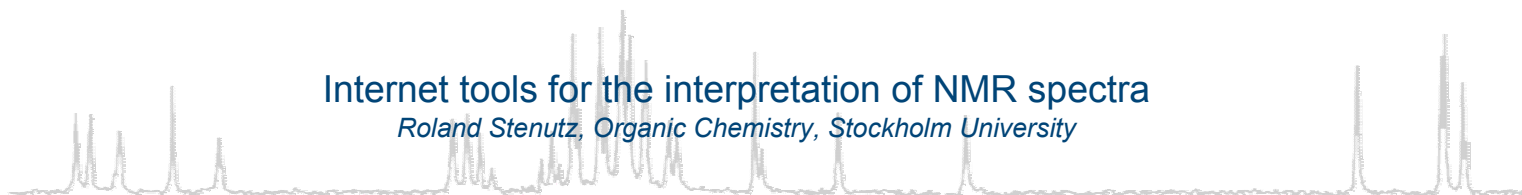
Explore

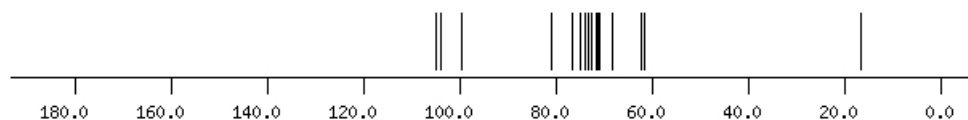
3D Co-ordinates



Internet tools for the interpretation of NMR spectra

Roland Stenutz, Organic Chemistry, Stockholm University

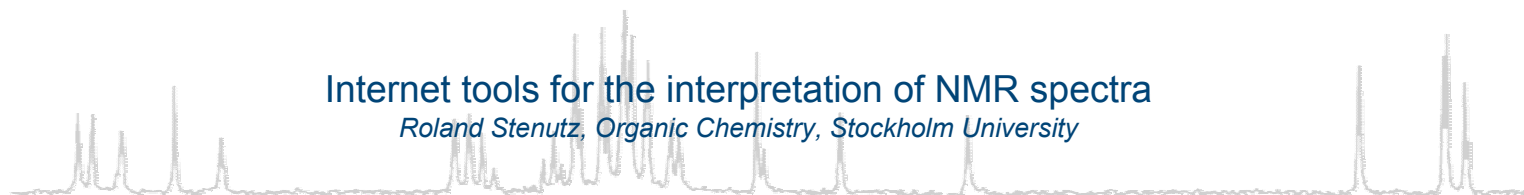




MHZ	75						
Temperature	333						
Solvent	D2O						
Residue	Linkage	Carbon	PPM	JFrom	JTo	Hz	Note
b-D-Glcp		C-1	105			0	
b-D-Glcp		C-2	74.800003			0	
b-D-Glcp		C-3	80.900002			0	
b-D-Glcp		C-4	71.400002			0	
b-D-Glcp		C-5	76.5			0	
b-D-Glcp		C-6	61.599998			0	
b-L-Fucp	3	C-1	104			0	
b-L-Fucp	3	C-2	72.400002			0	
b-L-Fucp	3	C-3	74			0	
b-L-Fucp	3	C-4	72.599998			0	
b-L-Fucp	3	C-5	71.400002			0	
b-L-Fucp	3	C-6	16.6			0	
a-D-Manp	2	C-1	99.599998			0	
a-D-Manp	2	C-2	71.199997			0	
a-D-Manp	2	C-3	71			0	
a-D-Manp	2	C-4	68.199997			0	
a-D-Manp	2	C-5	73.099998			0	
a-D-Manp	2	C-6	62.099998			0	

Internet tools for the interpretation of NMR spectra

Roland Stenutz, Organic Chemistry, Stockholm University



CASPER

(<http://www.casper.org.se>)



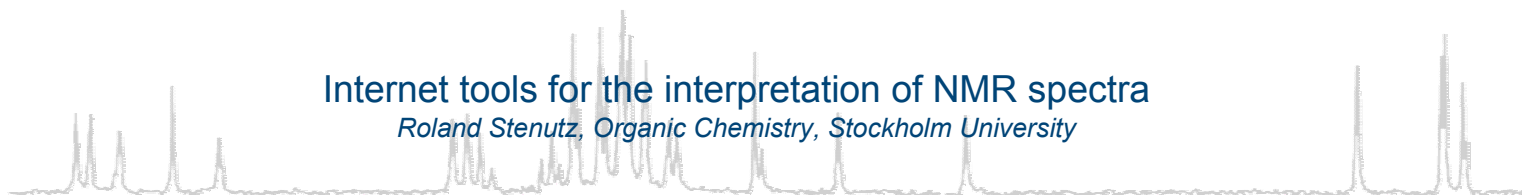
[Main page](#) | [Help](#) | [Simulate spectrum](#) | [Sequence determination](#)
[Stockholm University](#) | [Swedish University of Agricultural Sciences](#) | [The Karolinska Institute](#)

- [Help using CASPER](#)
- [Sequence determination](#)
Determine the sequence of a poly- or oligo-saccharide from NMR chemical shifts and the results of methylation and sugar analysis.
[Example 1 - a \(1->6\) glucan](#)
[Example 2 - a Shigella LPS](#)
- [Structure simulation](#)
Build a structure, simulate its ^{13}C - and ^1H -NMR spectra, and, optionally assign experimental spectra.
[Example 1 - Assignment of the \$^{13}\text{C}\$ -spectrum of methyl \$\beta\$ -cellobioside.](#)
- [Literature](#) List of references.
- [Guide to methylation analysis](#)



Internet tools for the interpretation of NMR spectra

Roland Stenutz, Organic Chemistry, Stockholm University



CASPER

Title

Source

	Residue	Linkage	'Reducing' end
1)	<input type="text" value="a"/> D-GlcpOMe	<input type="text" value="not linked"/>	<input type="text" value="none"/>
2)	<input type="text" value="b"/> L-Fucp	<input type="text" value="(1->3)"/>	<input type="text" value="residue 1"/>
3)	<input type="text" value="a"/> none	<input type="text" value="(1->2)"/>	<input type="text" value="residue 1"/>
4)	<input type="text" value="a"/> none	<input type="text" value="(1->2)"/>	<input type="text" value="residue 1"/>
5)	<input type="text" value="a"/> none	<input type="text" value="(1->2)"/>	<input type="text" value="residue 1"/>
6)	<input type="text" value="a"/> none	<input type="text" value="(1->2)"/>	<input type="text" value="residue 1"/>

¹³C-Chemical shifts

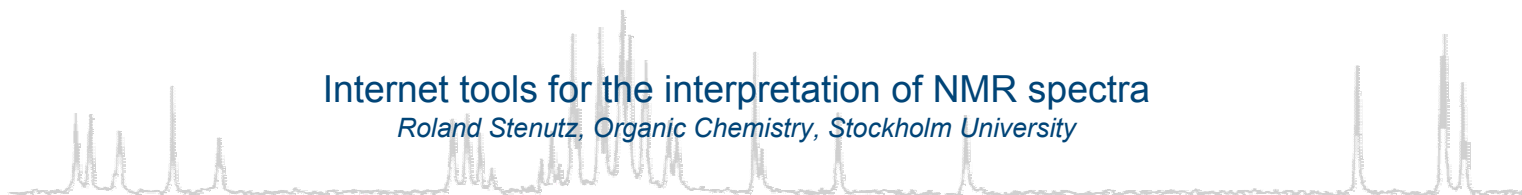
Load shifts from

Correct by subtracting ppm



Internet tools for the interpretation of NMR spectra

Roland Stenutz, Organic Chemistry, Stockholm University



CASPER

Calculated shifts

Results of calculation - Microsoft Internet Explorer

Address: <http://www.casper.org.se/casper/cgi-bin/build.cgi>

b-L-Fucp-(1-3)-a-D-Glcp-(1-1)-Methyl

Simulated structure

bLFuc(1->3)aDGlcOMe	99.86	70.92	83.65	69.88	72.31	61.51	55.82
->3)aDGlcOMe	4.86	3.65	3.85	3.61	3.67	3.77	3.87
bLFuc(1->	104.06	71.91	73.81	72.14	71.88	16.16	
	4.60	3.57	3.68	3.76	3.82	1.29	

Assignment of ¹³C resonances

Experimental	Simulated	Exp-Sim	Assignment
104.00	104.06	-0.06	bLFuc - 1
99.70	99.86	-0.16	aDGlcOMe - 1
83.50	83.65	-0.15	aDGlcOMe - 3
73.70	73.81	-0.11	bLFuc - 3
72.20	72.31	-0.11	aDGlcOMe - 5
72.10	72.14	-0.04	bLFuc - 4
71.80	71.91	-0.11	bLFuc - 2
71.80	71.88	-0.08	bLFuc - 5
70.80	70.92	-0.12	aDGlcOMe - 2
69.70	69.88	-0.18	aDGlcOMe - 4
61.30	61.51	-0.21	aDGlcOMe - 6
16.10	16.16	-0.06	bLFuc - 6

Error=1.39 ppm (0.12/shift), Systematic error=-0.12 ppm, RMS error=0.13 ppm

Experimental structure

bLFuc(1->3)aDGlcOMe	99.70	70.80	83.50	69.70	72.20	61.30	n.d.
->3)aDGlcOMe	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
bLFuc(1->	104.00	71.80	73.70	72.10	71.80	16.10	
	n.d.	n.d.	n.d.	n.d.	n.d.		

When using CASPER in your research please cite as:

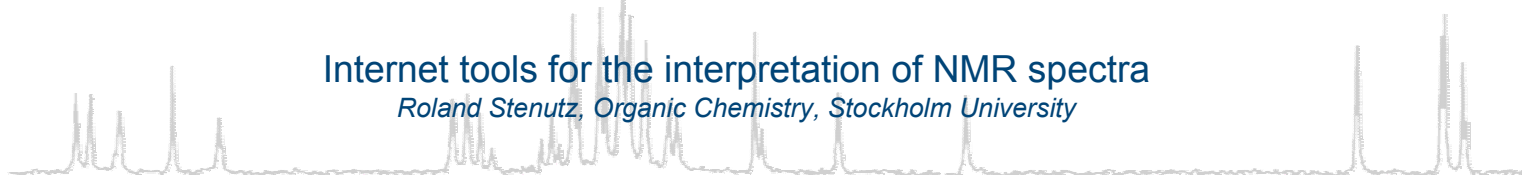
Assignment

Experimental shifts



Internet tools for the interpretation of NMR spectra

Roland Stenutz, Organic Chemistry, Stockholm University



CASPER

b-L-Fucp-(1-3)-a-D-Glcp-(1-1)-Methyl

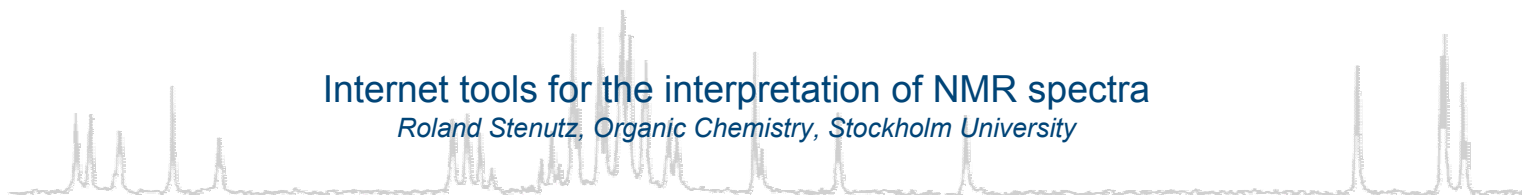
Simulated structure

bLFuc(1->3)aDGlcOMe

->3)aDGlcOMe	99.86	70.92	83.65	69.88	72.31	61.51	55.82	
	4.86	3.65	3.85	3.61	3.67	3.77	3.87	3.44
bLFuc(1->	104.06	71.91	73.81	72.14	71.88	16.16		
	4.60	3.57	3.68	3.76	3.82	1.29		

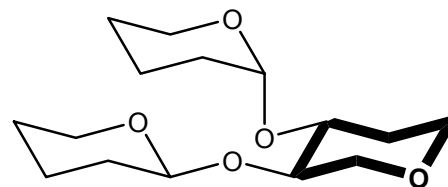
Internet tools for the interpretation of NMR spectra

Roland Stenutz, Organic Chemistry, Stockholm University

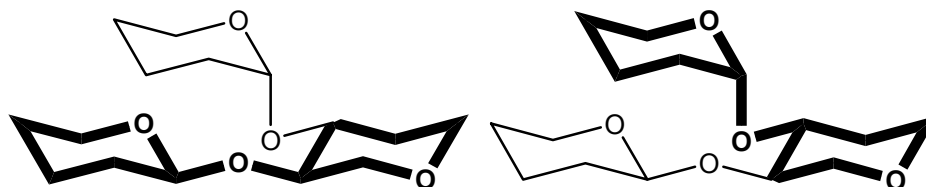


Chemical shift calculation

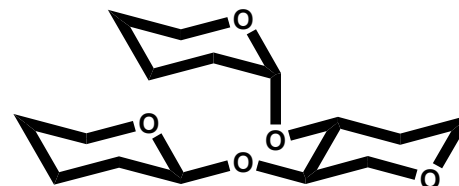
1) Start with monosaccharide



2) Add glycosylation shifts

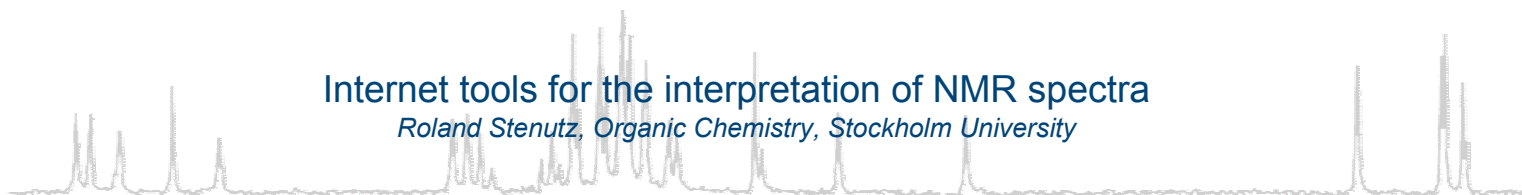


3) Add steric corrections

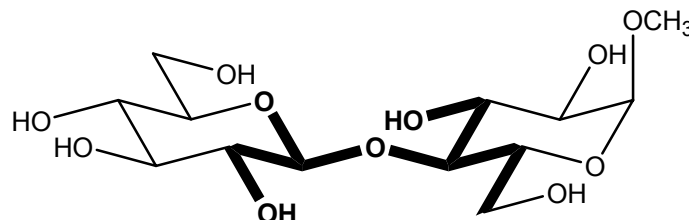


Internet tools for the interpretation of NMR spectra

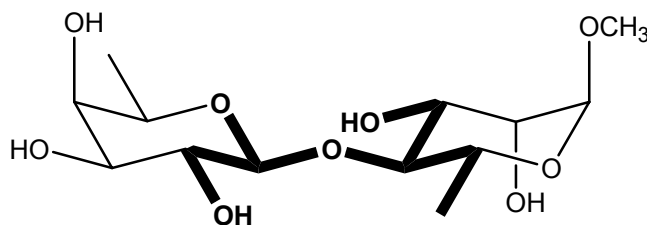
Roland Stenutz, Organic Chemistry, Stockholm University



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

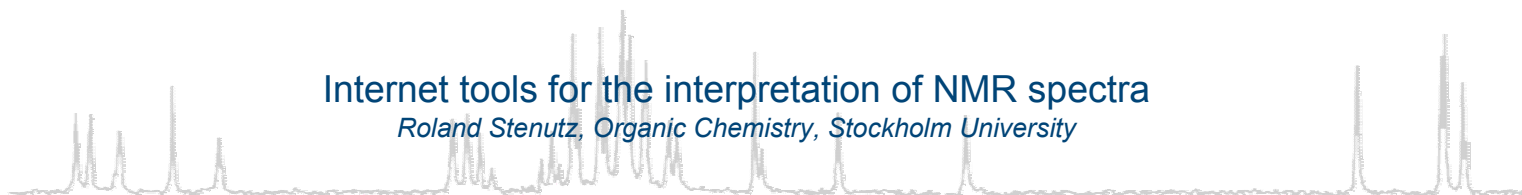


β DFuc (1→	6.85	-1.05	-0.13	-0.24	0.18	-0.23
→4) α DRhaOMe	-0.30	-0.66	-1.23	10.19	-1.35	-0.07



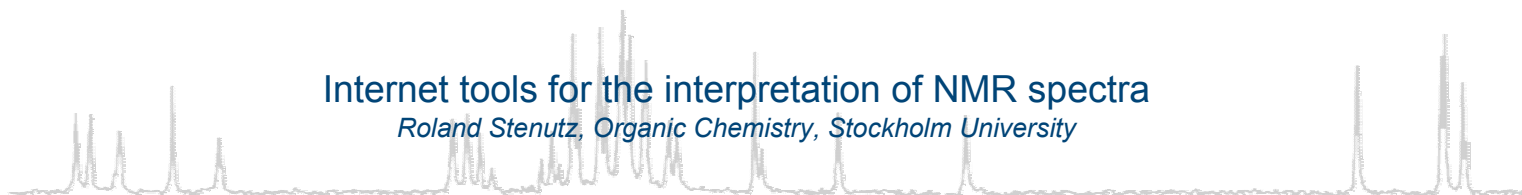
Internet tools for the interpretation of NMR spectra

Roland Stenutz, Organic Chemistry, Stockholm University



Sequence determination

- 1) Start with all possible components
- 2) Generate all topologies and sequences
- 3) Calculate spectra and compare with experimental data



CASPER

Components
&
linkage positions

Chemical shifts

Additional restrictions

The screenshot shows the CASPER web interface. A red circle highlights the 'Residue' dropdown menu (set to 'D-Glcp') and the 'Linkage position' table. A red arrow points from the text 'Components & linkage positions' to this circle. Another red circle highlights the 'Chemical shifts' text area, which contains the values '99.0 72.5 74.5 71.3 70.7 66.7'. A red arrow points from the text 'Chemical shifts' to this circle. A third red circle highlights the 'Additional restrictions' section, specifically the coupling constant settings for $^3J_{HH}$ and $^3J_{OH}$. A red arrow points from the text 'Additional restrictions' to this circle.

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"/>
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"/>
none		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Chemical shifts: 99.0 72.5 74.5 71.3 70.7 66.7

Load shifts from: Bläddra...

Correct by subtracting: ppm

Number of shifts - Required: Actual: Clear text area

Minimum number of coupling constants of different magnitudes

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

Save form Start simulation

CASPER

<u>Residue</u>	<u>Linkage position</u>					
	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"/>
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"/>
none ▼	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

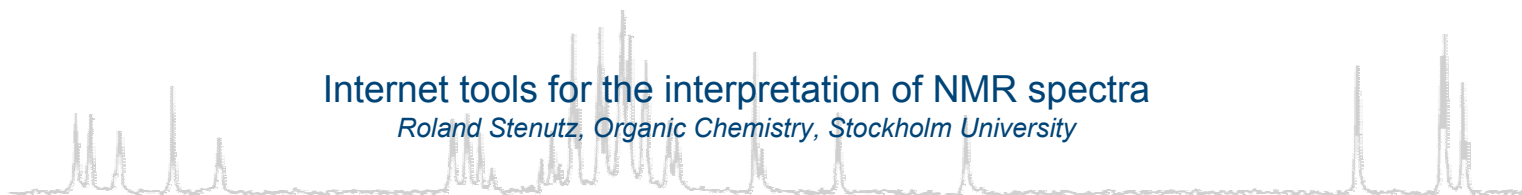
Chemical shifts ☒ α C ☐ α H

Source of chemical shifts ☒ text area ☐ file

99.0 72.5 74.5 71.3 70.7 66.7

Internet tools for the interpretation of NMR spectra

Roland Stenutz, Organic Chemistry, Stockholm University



CASPER

Best fitting structures

1. [Structure 1](#), error=0.65 (0.11)
->6)aDGlc(1->
2. [Structure 2](#), error=15.81 (2.64)
->6)bDGlc(1->

Structure 1

->6)aDGlc(1->

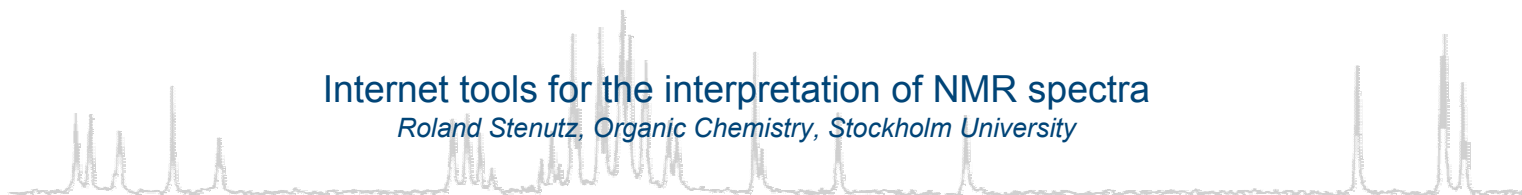
Simulated structure

->6)aDGlc(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	



Internet tools for the interpretation of NMR spectra

Roland Stenutz, Organic Chemistry, Stockholm University



Main weakness of CASPER -

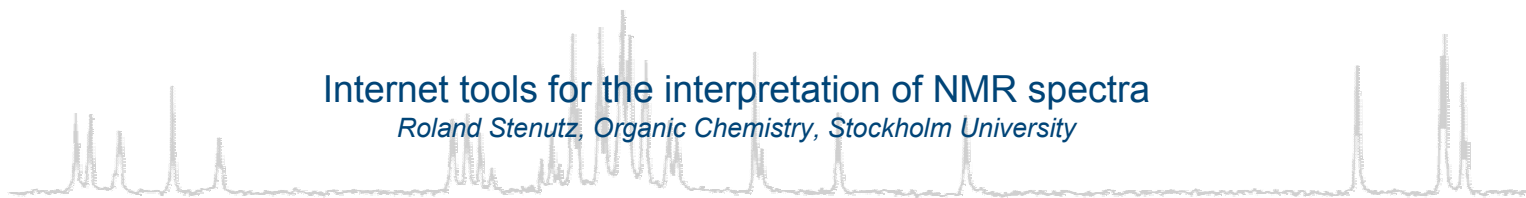
- requires data from chemical analyses to limit the number of structures that are simulated.

An Opportunity for ANN:s?



Internet tools for the interpretation of NMR spectra

Roland Stenutz, Organic Chemistry, Stockholm University

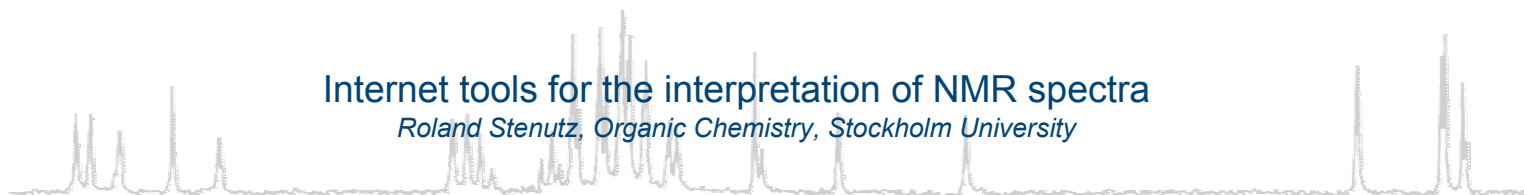


Caveats

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.



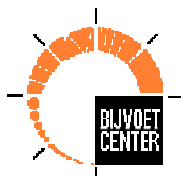
JRP-2 members



Deutsches Krebsforschungszentrum



Stockholm University



Bijvoet Center for Biomolecular Research



Forschungszentrum Borstel



Swedish University of Agricultural Sciences



Internet tools for the interpretation of NMR spectra

Roland Stenutz, Organic Chemistry, Stockholm University

