

CASPER

A tool for the interpretation of
NMR spectra

Roland Stenutz

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

roland@organ.su.se



Computer **A**ssisted **S**pectrum **E**valuation of **R**egular polysaccharides

A faint, horizontal NMR spectrum is visible at the bottom of the slide, showing several distinct peaks of varying heights.

Polysaccharide structure:

Components	"type"	<i>Hex or HexNAc</i>
	relative configuration	<i>Glc or Man</i>
	absolute configuration	<i>D- or L-</i>
	ringsize	<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*



The interpretation of NMR spectra becomes difficult if the polysaccharide -

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



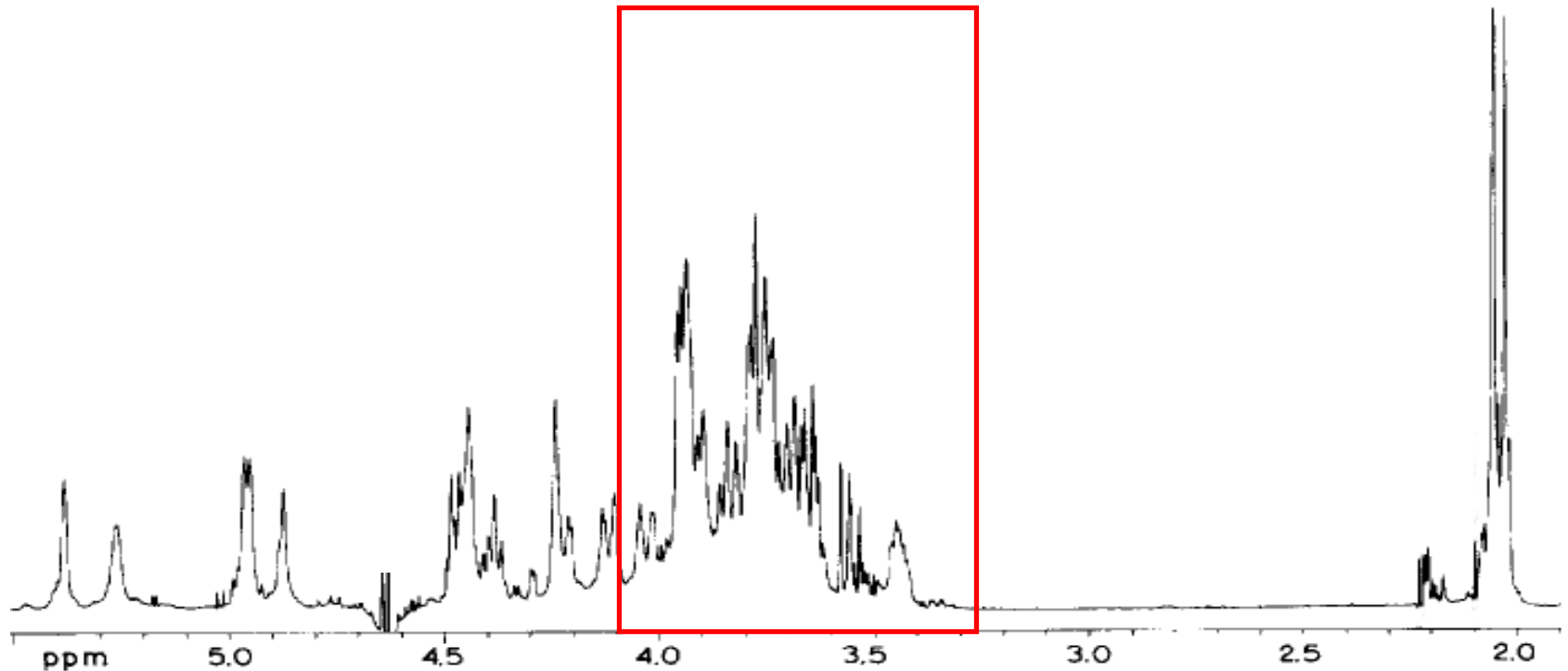
Typical ^1H spectrum

O-antigen of *E. coli* O113

reporter groups

ring region

reporter groups



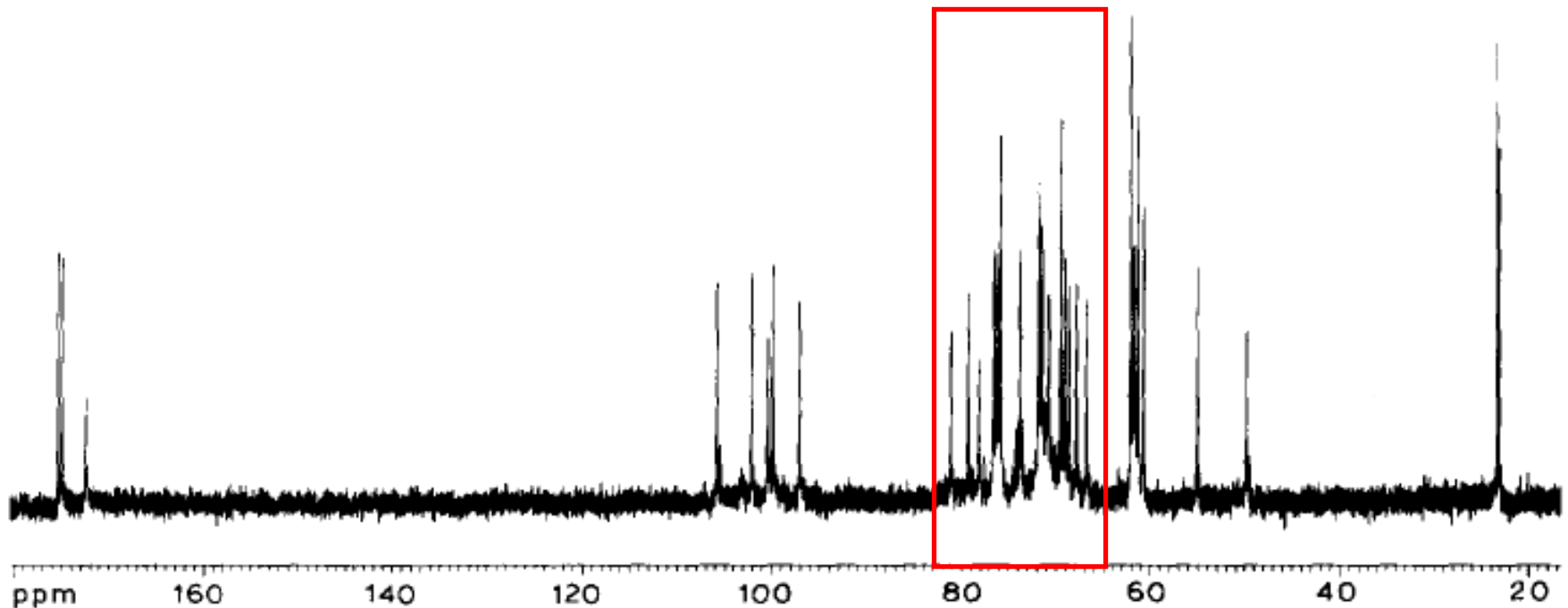
Computer Assisted **S**pectrum Evaluation of **R**egular polysaccharides



Typical ^{13}C spectrum

O-antigen of *E. coli* O113

ring region



Computer Assisted **S**pectrum Evaluation of **R**egular polysaccharides



Polysaccharides have unique 1D-spectra
- *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!



Computer **A**ssisted **S**pectrum **E**valuation of **R**egular polysaccharides



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.




CASPER

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



CASPER



Ke3690

HomeResearchAnalysisECDBCASPER

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 ☒

Computer Assisted **S**pectrum Evaluation of Regular polysaccharides



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



Computer Assisted **S**pectrum Evaluation of Regular polysaccharides



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

Computer Assisted **S**pectrum Evaluation of Regular polysaccharides



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



Computer Assisted **S**pectrum Evaluation of **R**egular polysaccharides



Simulation of NMR

Title
Source

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

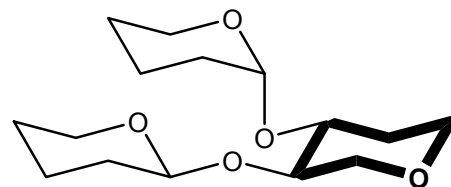
As MIME ☐

Computer Assisted **S**pectrum Evaluation of Regular polysaccharides



Chemical shift calculation

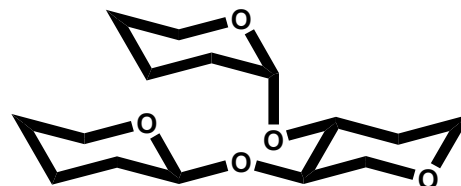
1) Start with monosaccharide



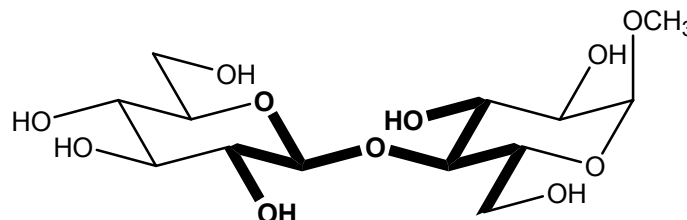
2) Add glycosylation shifts



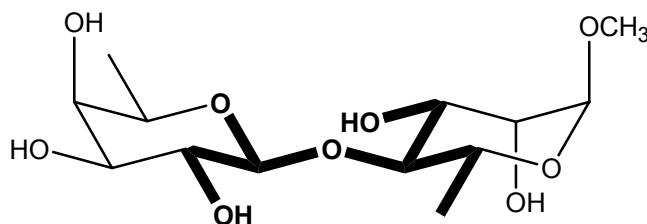
3) Add steric corrections



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

Computer Assisted **S**pectrum **E**valuation of **R**egular polysaccharides

Line notation

Lewis A bDGal(-3)[aLFuc(-4)]bDGlcNAc

Dextran -6)aDGlc(-

<http://www.casper.organ.su.se/casper/cgi-bin/ln.cgi?structure>



Computer Assisted **S**Pectrum Evaluation of **R**egular polysaccharides



E. coli O-antigens

Serogroup	O6
Serotype	O6:K54:H10
Strain	A12b (see CCUG 11311)
ECDB#	50

Structure	$\begin{array}{c} \rightarrow 4) \alpha \text{DGalNAc} (1 \rightarrow 3) \beta \text{DMan} (1 \rightarrow 4) \beta \text{DMan} (1 \rightarrow 3) \alpha \text{DGlcNAc} (1 \rightarrow \\ \\ \beta \text{DGlcNAc} (1 \rightarrow 2) \end{array}$
-----------	---

CarbBank	34570
LinucsID	9686
Structure code	P4.1
Sugar components	2 DMan, 2 DGlcNAc, DGalNAc
Non-sugar components	
Comments	Structure [1]

Identical to	
Cross-reacts with	Escherichia coli O6:K2:H1, O6:K13:H1 [1]

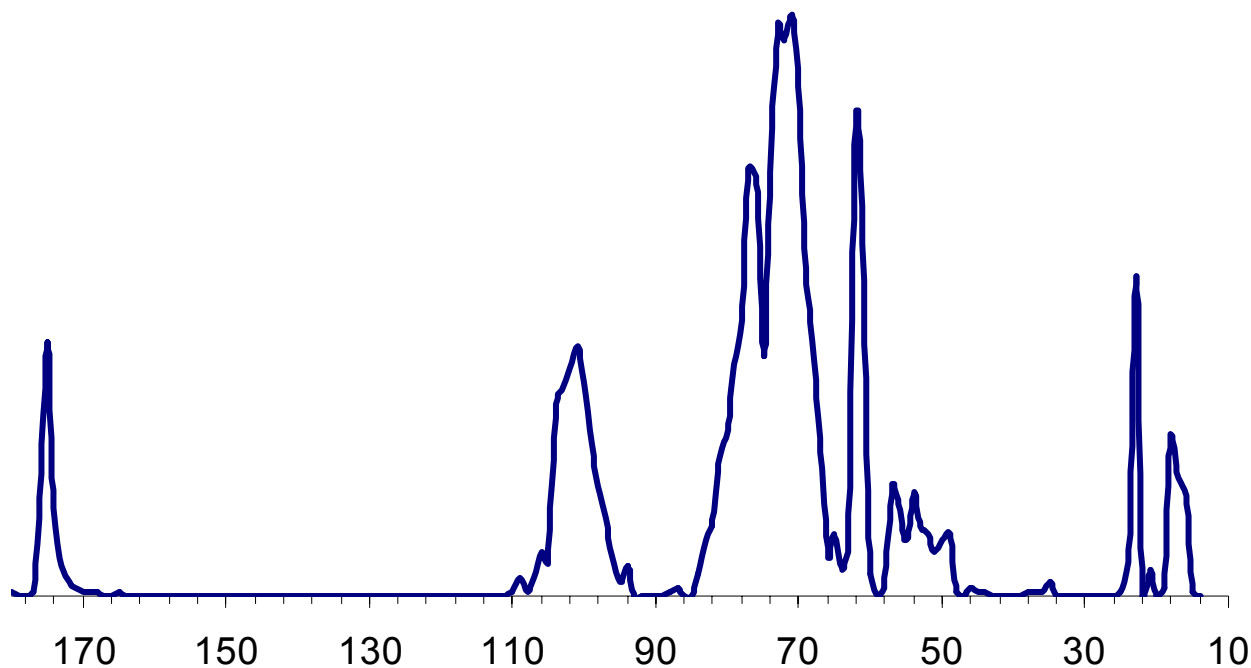
¹ H-NMR	4.81 4.01 3.84 3.75 3.57 3.87 3.73 4.92 4.14 4.01 3.69 4.18 3.84 3.78 5.24 4.29 4.09 4.09 4.31 3.77 3.77 4.73 4.39 3.77 3.67 3.46 3.97 3.63 4.78 3.73 3.59 3.47 3.39 3.89 3.75
--------------------	--



Computer Assisted **S**pectrum Evaluation of Regular polysaccharides



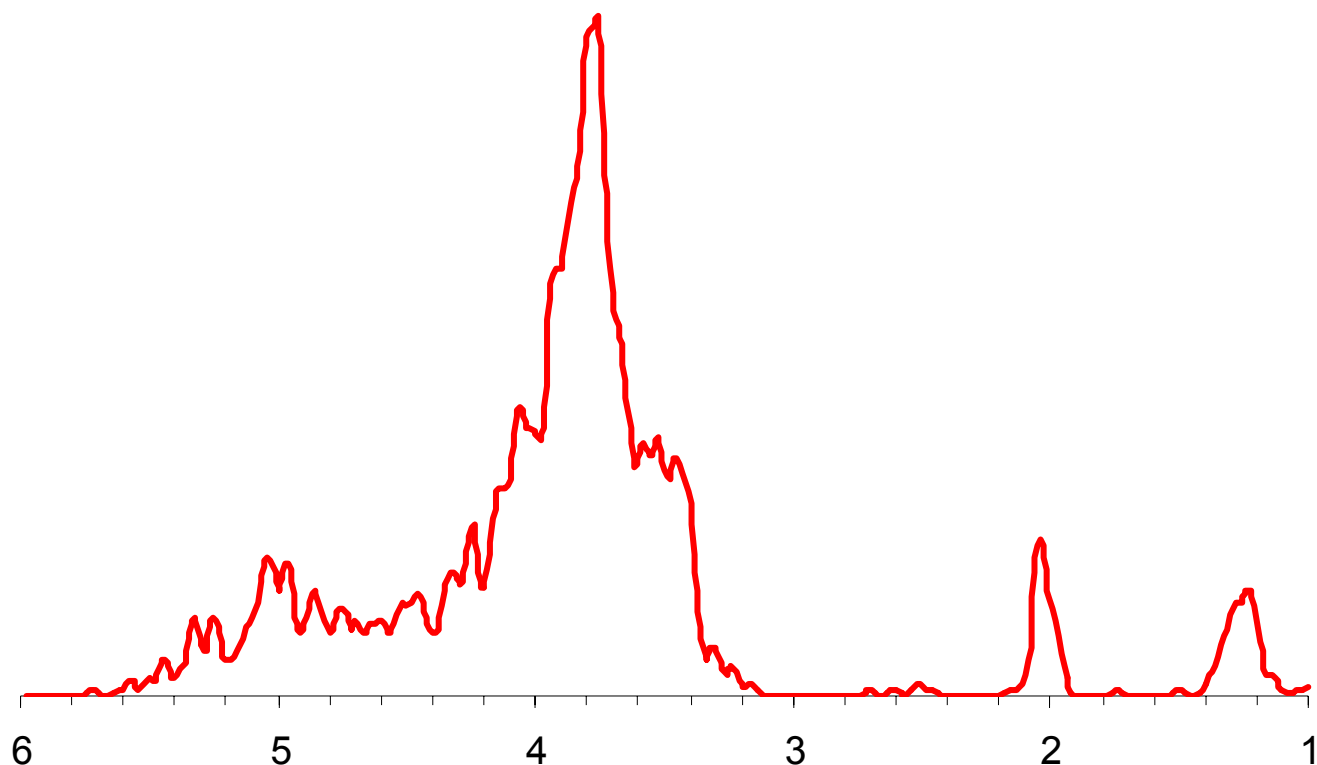
Sum of ^{13}C -spectra



Computer Assisted **S**pectrum Evaluation of **R**egular polysaccharides

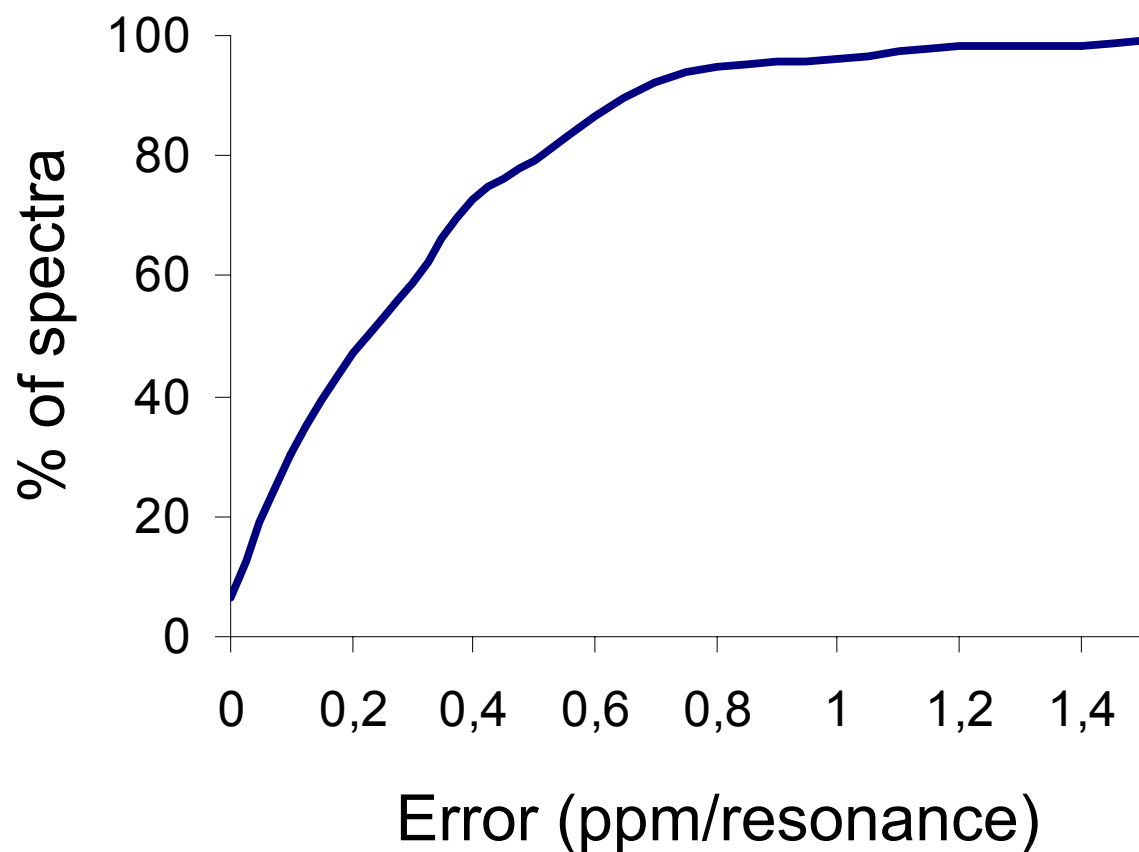


Sum of ^1H -spectra



Computer Assisted **S**pectrum Evaluation of **R**egular polysaccharides

Prediction of published ^{13}C -NMR



Computer Assisted **S**pectrum **E**valuation of **R**egular polysaccharides



Distribution of errors

