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 **SP**ectrum **E**valuation of **R**egular polysaccharides

Polysaccharide structure:

Components	”type”	*Hex or HexNAc*
	relative configuration	*Glc or Man*
	absolute configuration	*D- or L-*
	ringsize	p *or* f
Linkages	position	$\rightarrow$*4) or* $\rightarrow$*6)*
	stereochemistry	α*- or* β*-*
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)

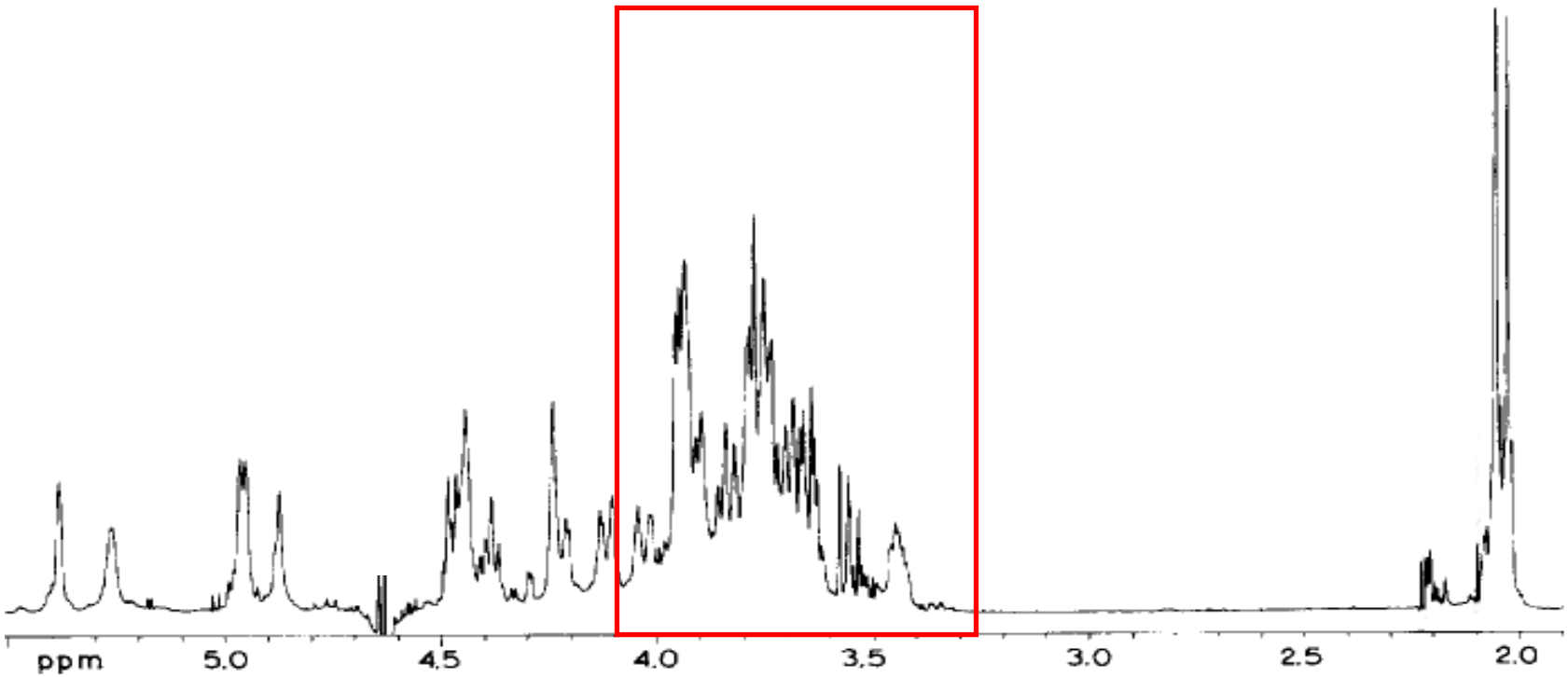
- contains many similar residues

- is very large

Typical ^{1}H spectrum

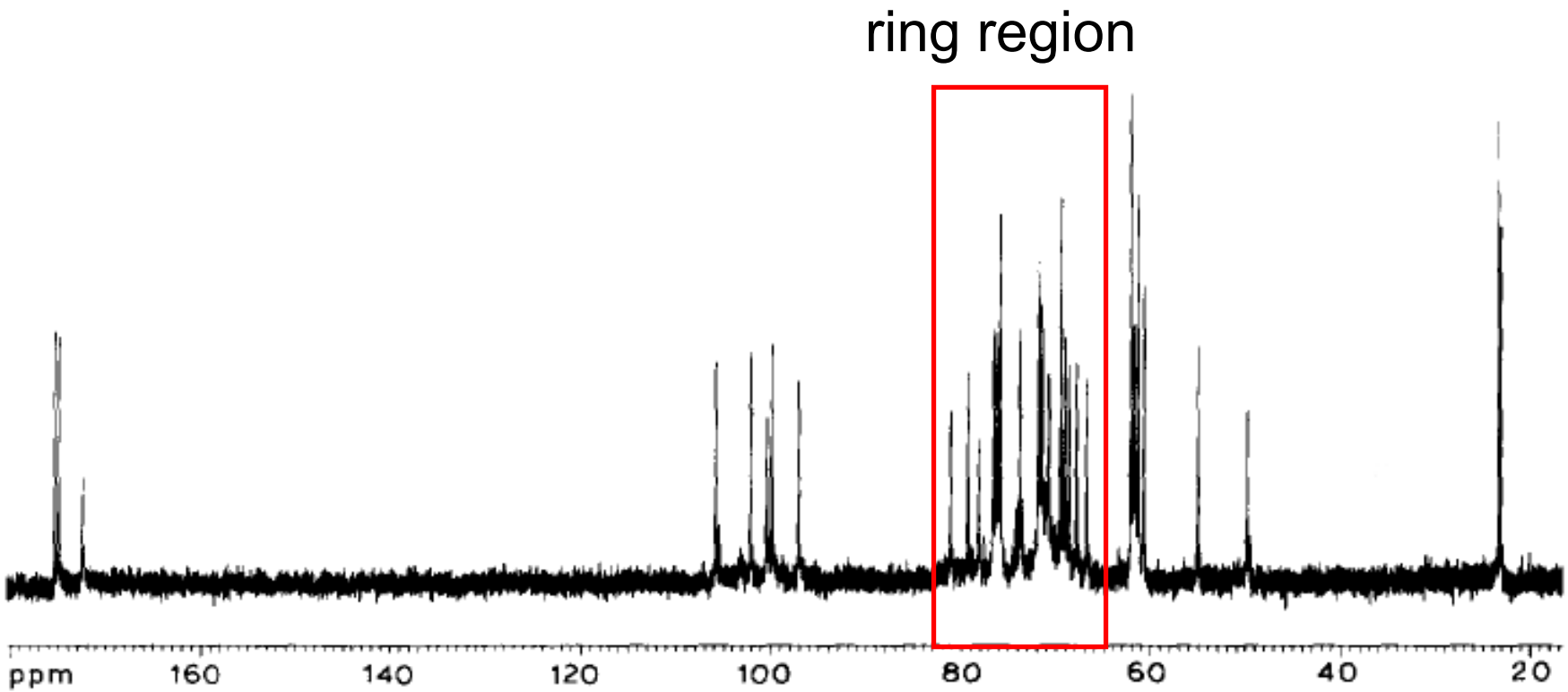
O-antigen of *E. coli* O113

reporter groups ring region reporter groups



Typical ^{13}C spectrum

O-antigen of *E. coli* O113



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!

Current approaches:

comparison with a database (SugaBase)

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

comparison with simulated NMR spectra (CASPER)

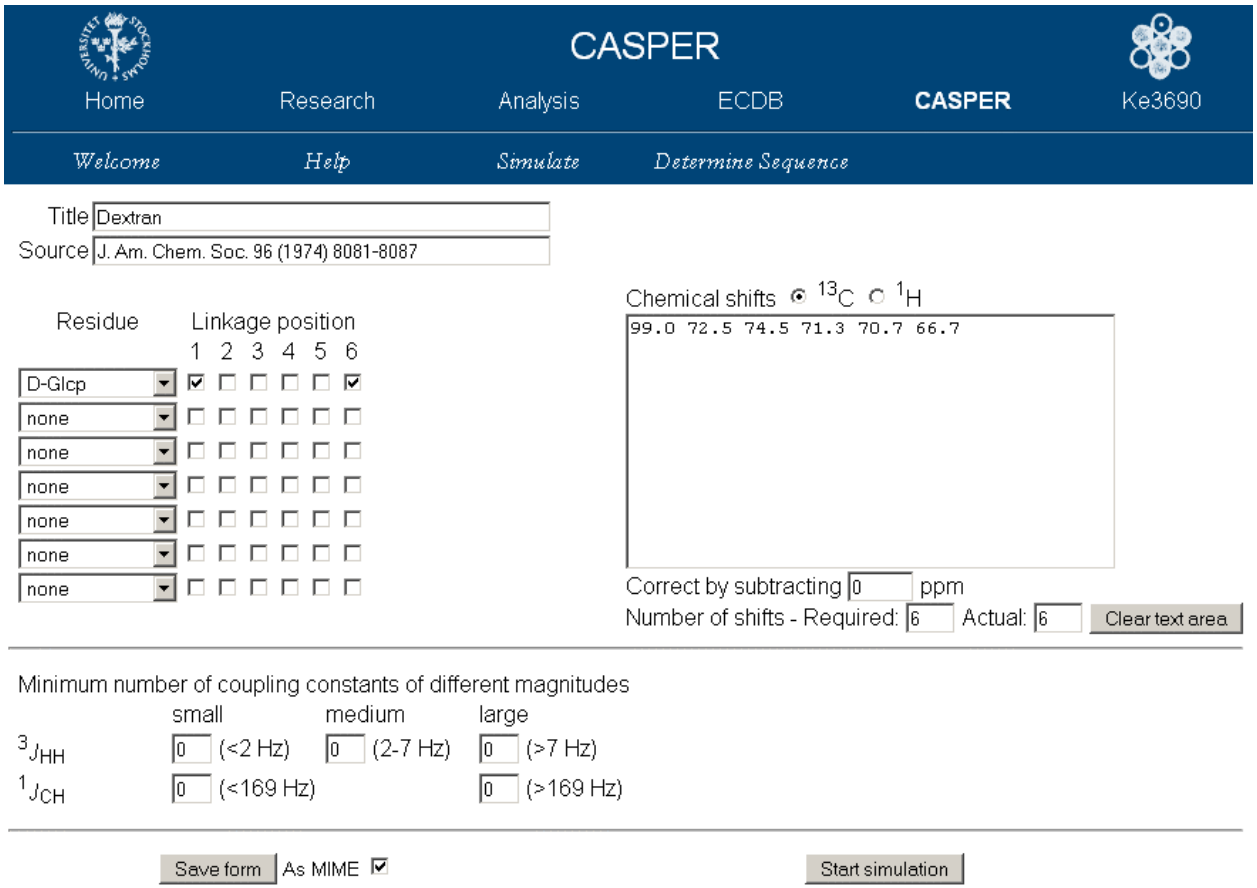
requires information about the components and linkages to limit the number of possible structures.

Artificial Neural Networks (ANN)

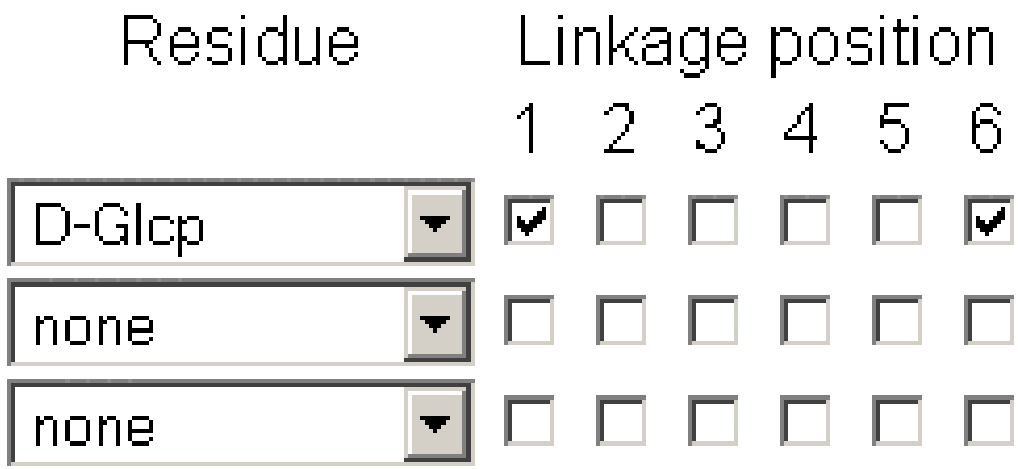
current application are limited to a single class of compounds.

CASPER

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

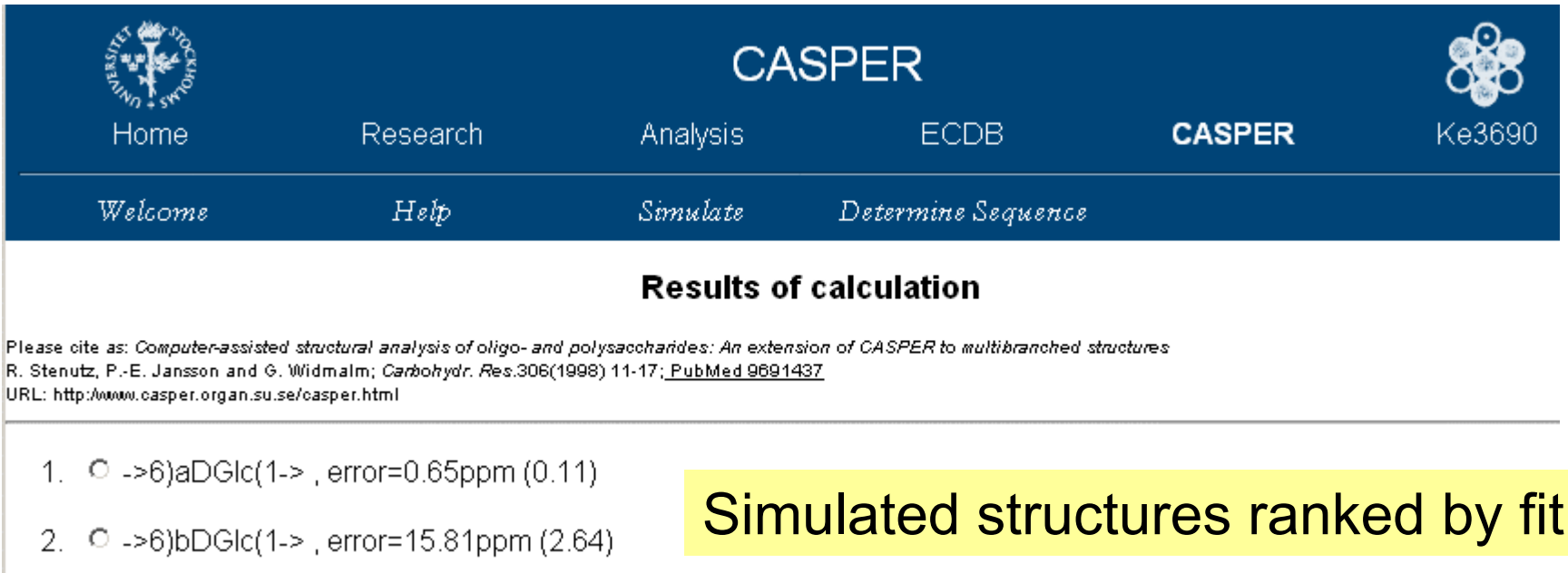


CASPER



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

CASPER



Simulated structures ranked by fit

CASPER

Simulated structure

->6)aDGlc^{i}(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 ^{13}C 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)aDGlc^{i}(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

CASPER

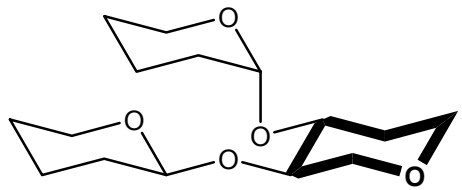
Experimental	Simulated	Exp-Sim	Assignment
99.00	98.90	0.10	aDGlc^{i} - 1
74.50	74.32	0.18	aDGlc^{i} - 3
72.50	72.39	0.11	aDGlc^{i} - 2
71.30	71.28	0.02	aDGlc^{i} - 5
70.70	70.66	0.04	aDGlc^{i} - 4
66.70	66.90	-0.20	aDGlc^{i} - 6

Simulation of NMR

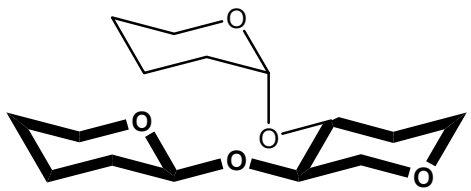
Computer **A**ssisted **SP**ectrum **E**valuation of **R**egular polysaccharides

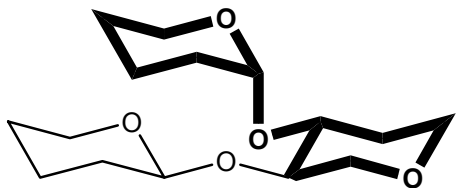
Chemical shift calculation

1) Start with monosaccharide

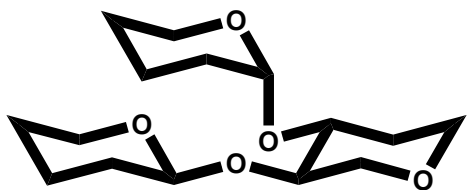


2) Add glycosylation shifts

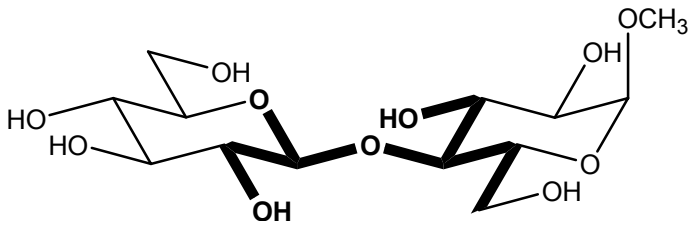




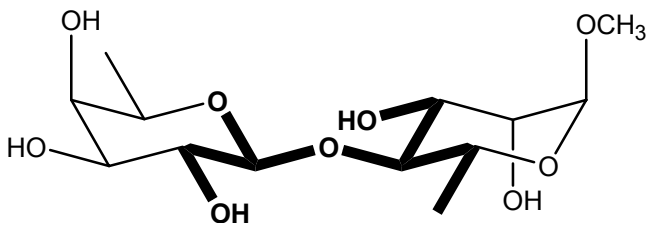
3) Add steric corrections



Glycosylation shifts



βDGlc(1→	6.45	-1.07	-0.18	-0.20	0.12	-0.19
→4)αDGlcOMe	-0.27	-0.25	-1.51	9.20	-1.36	-0.60



βDFuc(1→	6.85	-1.05	-0.13	-0.24	0.18	-0.23
→4)αDRhaOMe	-0.30	-0.66	-1.23	10.19	-1.35	-0.07

Line notation

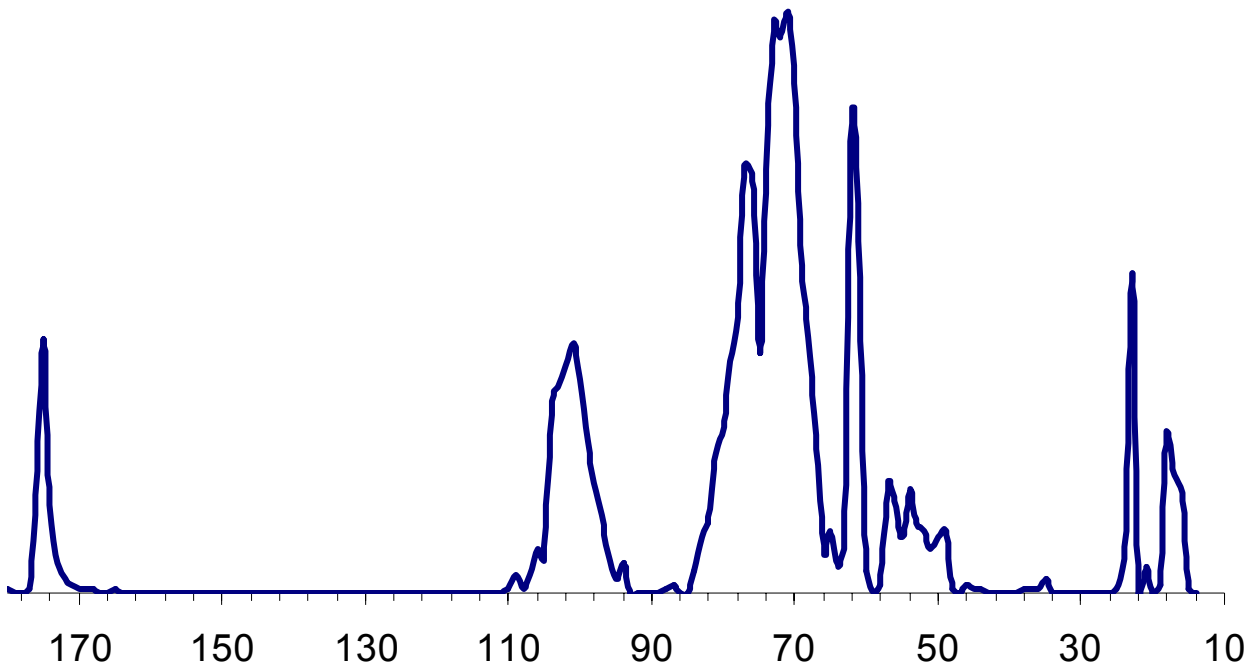
Lewis A	bDGal(-3)[aLFuc(-4)]bDGlcNAc
Dextran	-6)aDGlc(-

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

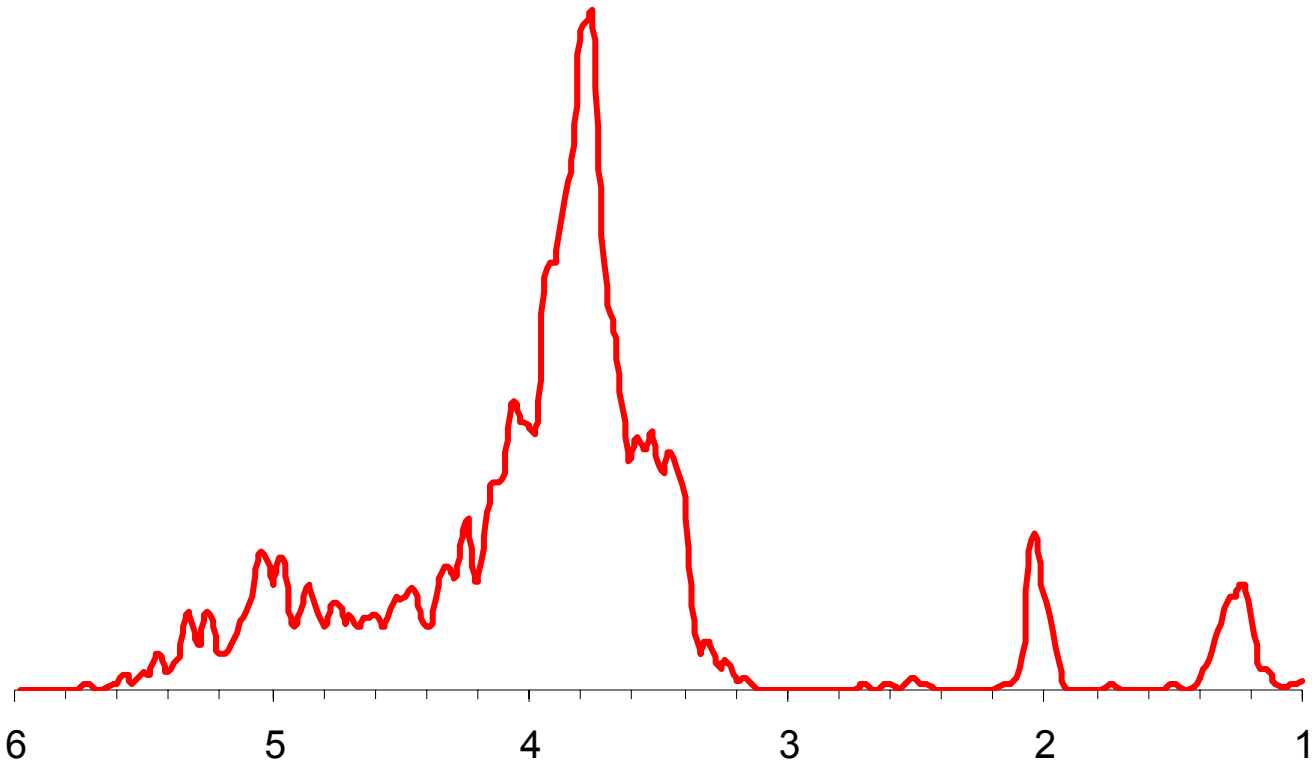
E. coli O-antigens

Serogroup	O6
Serotype	O6:K54:H10
Strain	A12b (see CCUG 11311)
ECDB#	50
Structure	->4)aDGalNAc(1->3)bDMan(1->4)bDMan(1->3)aDGlcNAc(1-> | bDGlcNAc(1->2)
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

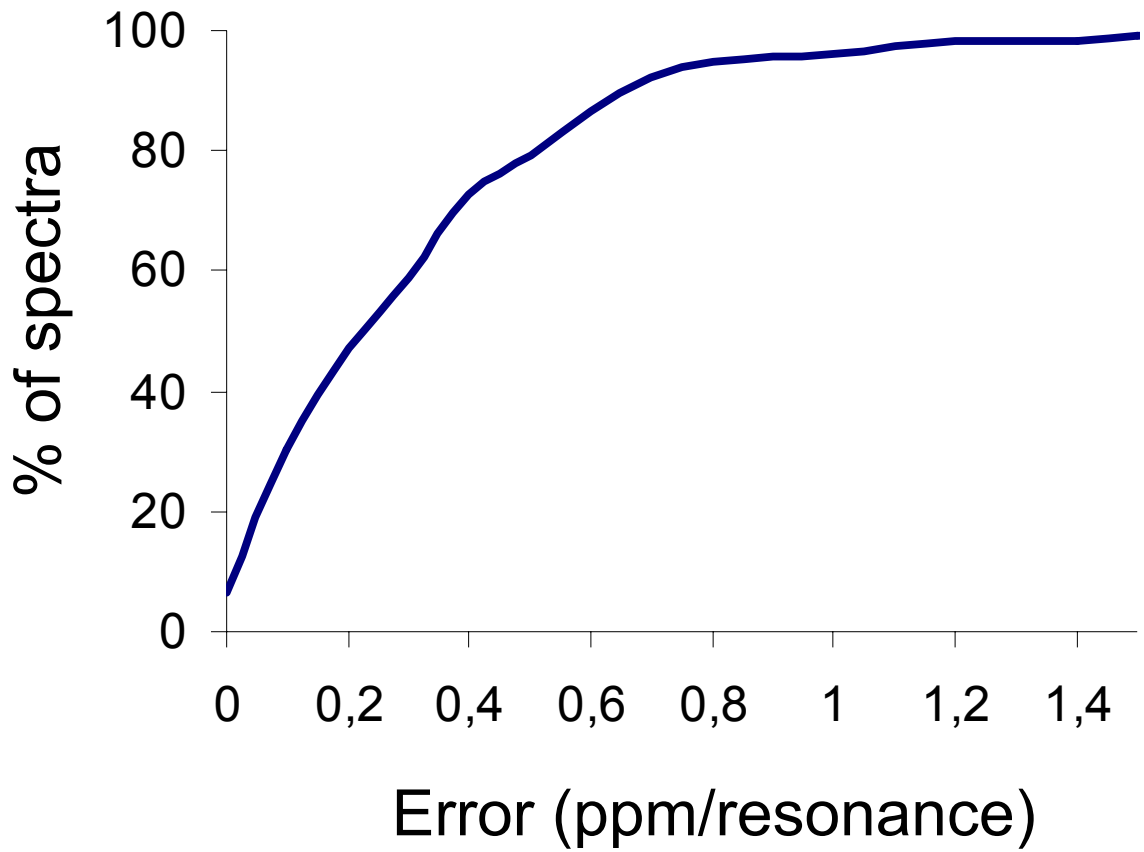
Sum of ^{13}C-spectra



Sum of ^{1}H-spectra



Prediction of published ^{13}C-NMR



Distribution of errors