F.V. Toukach

Methods of 1D and 2D NMR in structural elucidation of natural glycopolymers

Choice of the demo material

In genomics and proteomics primary structure elucidation is more automatized

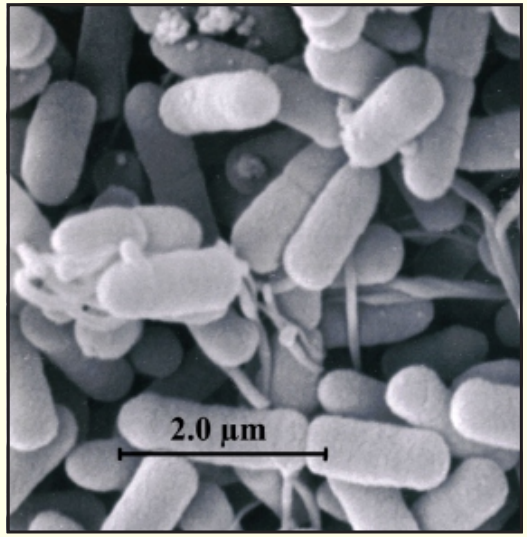
In glycomics:

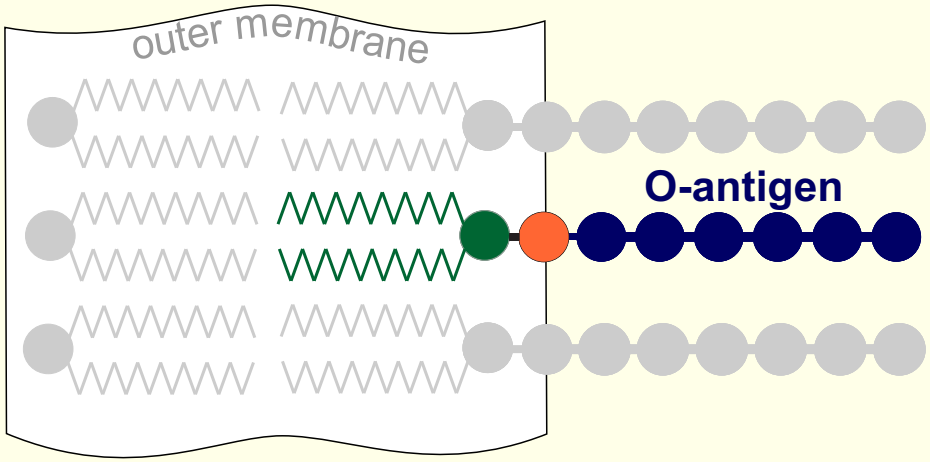
- greater chemical variativity *(especially for bacterial carbohydrates)*
- structure elucidation is not algorhythmic
- tertiary structure (=>biological properties) is more dependent on the *primary* structure

Polysaccharides of the Gram-negative bacterial outer membrane are antigenic

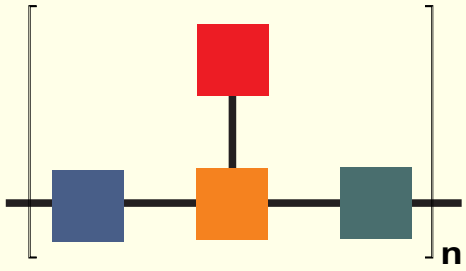
- knowledge of structure => serological classification, vaccines

Edwadrsiella tarda (strain 1153)
(Gram-negative enterobacteria of marine animals and reptiles, sometimes causes gastroenteritis)



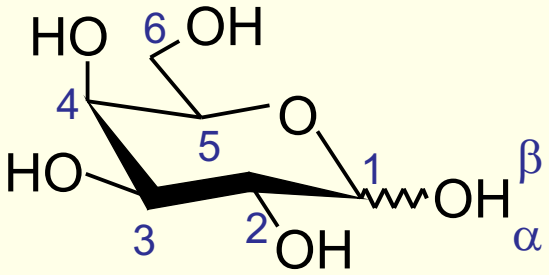


Carbohydrate structure



COMPLETE STRUCTURE

- structure of all residues,
  including non-sugars (aminoacids, aliphatic acids, etc.)
- substitution positions
- sequence of residues
- stoichiometry of residues
- phosphate and sulphate linkers
- number of repeating units and “frame positioning”

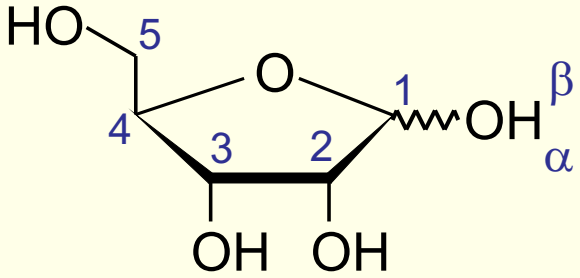


D-Gal

aldo-monosaccharide in pyranose form (example)

RESIDUE STRUCTURE

- carbon skeleton size (5-9)
- cycle size (pyranose, furanose, linear)
- spatial orientation of all -OH (ax/eq)
- anomeric configuration (α/β)
- absolute configuration (D/L)

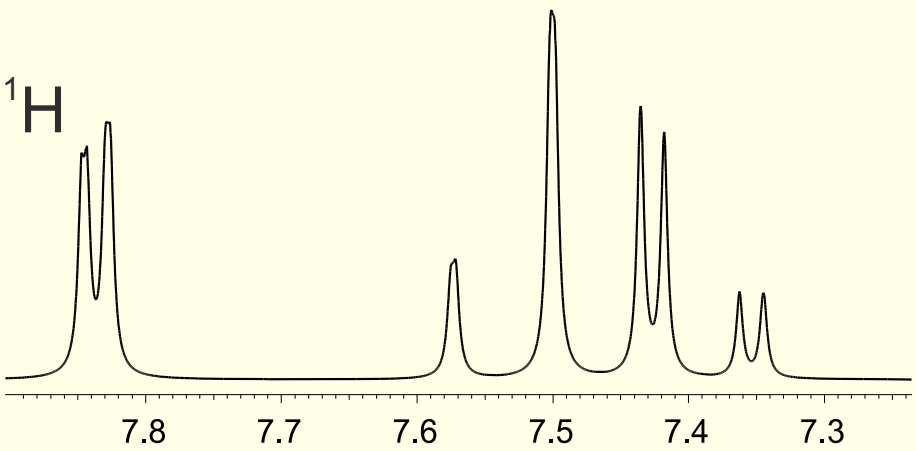


D-Rib

aldo-monosaccharide in furanose form (example)

- -H instead of -OH (deoxy-)
- -Nh_2 instead of -OH (amino-)
- -COOH instead of -CH_2OH (uronic acids)
- other functional groups

1D NMR spectrum

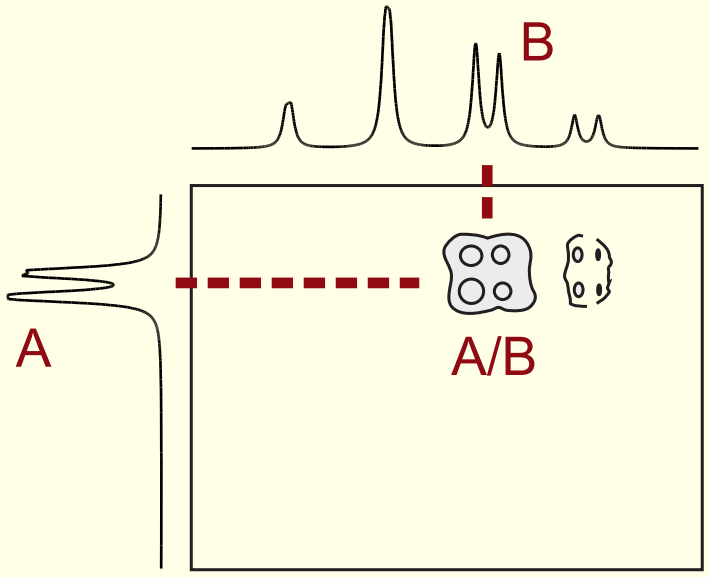


signal position
- distribution of electronic density

signal shape
- number of type of neighboring atoms

signal square
- number of equivalent atoms

2D NMR spectrum (correlation)



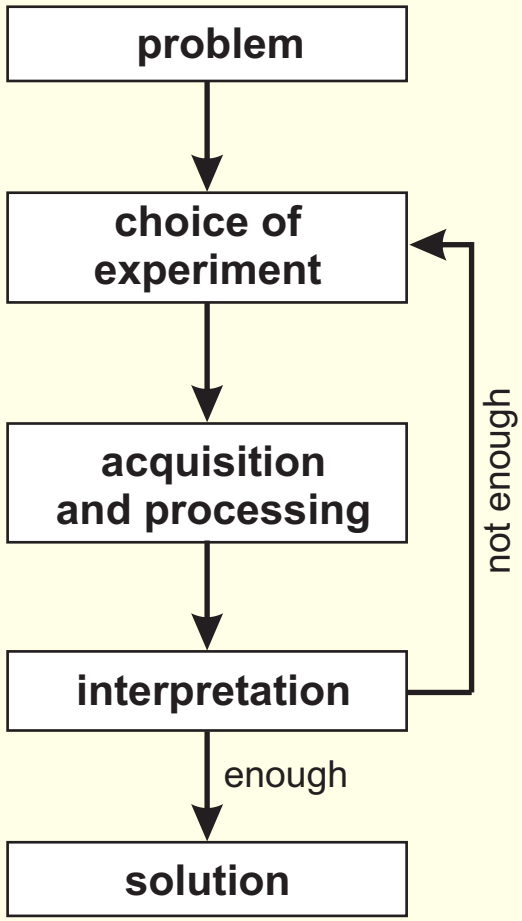
cross-peaks

- signal interactions

(spin coupling is via chemical bonds,
NOE is via space)

Non-correlational NMR experiments

1D 2D

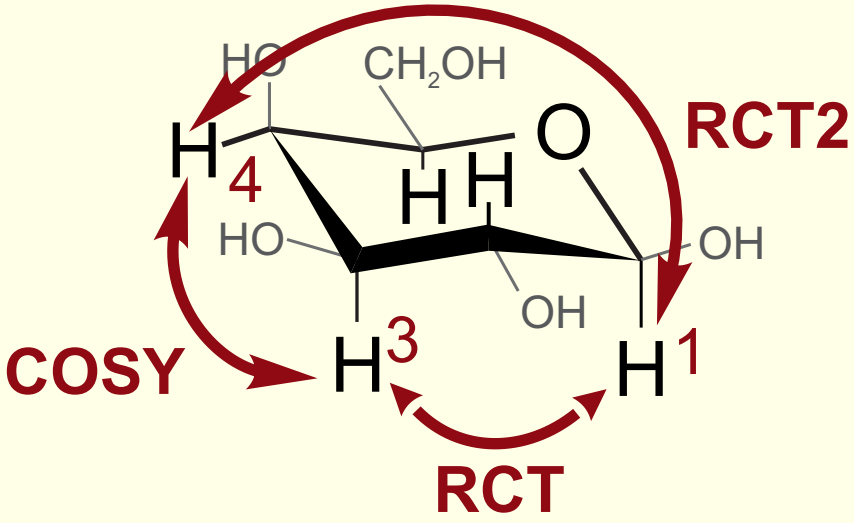


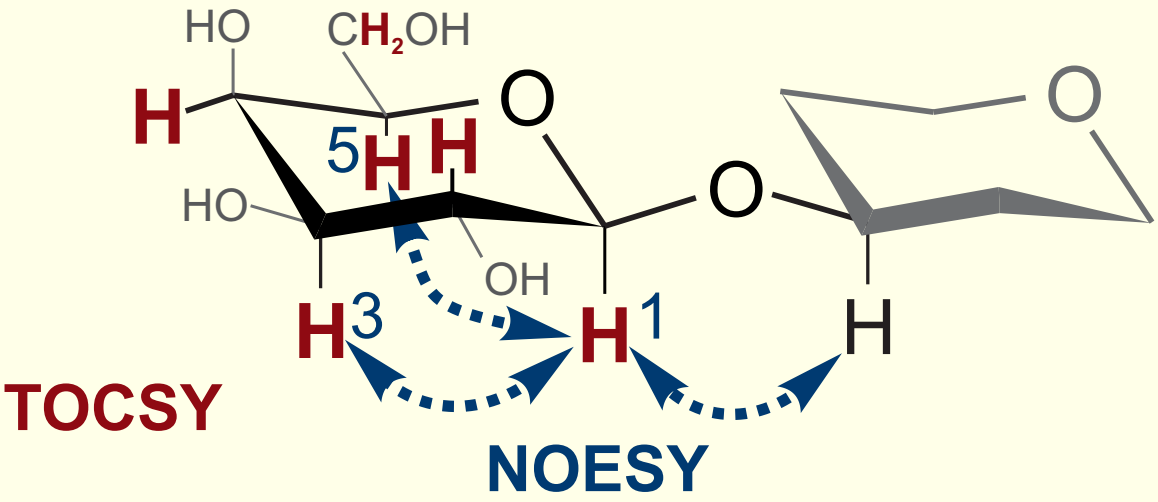
^{1}H	1D proton spectrum - *measurement of homonuclear spin coupling constants, general information, residue identification, basis for carbon spectrum assignment*
^{13}C BB	1D carbon spectrum with broad band proton decoupling - *detailed information, residue identification, substitution positions*
^{31}P BB, ^{15}N BB	other 1D spectra with broad band proton decoupling - *additional information*
APT, DEPT	edited carbon spectrum - *assignment of CH_2-groups*
^{13}C Gated, ^{31}P Gated	1D spectra without broad band decoupling - *measurement of homonuclear spin coupling constants, elucidation of anomeric configuration, conformation studies*
HH J-res	proton signals swept on multiplicity - *measurement of homonuclear spin coupling constants, general information, residue identification*
DOSY	proton spectrum swept on molecule correlation time - *separation of spectra into subspectra of components or of molecular parts that differ in mobility*

spectra interpretation is not algorhythmistic

Homonuclear correlations

1D 2D

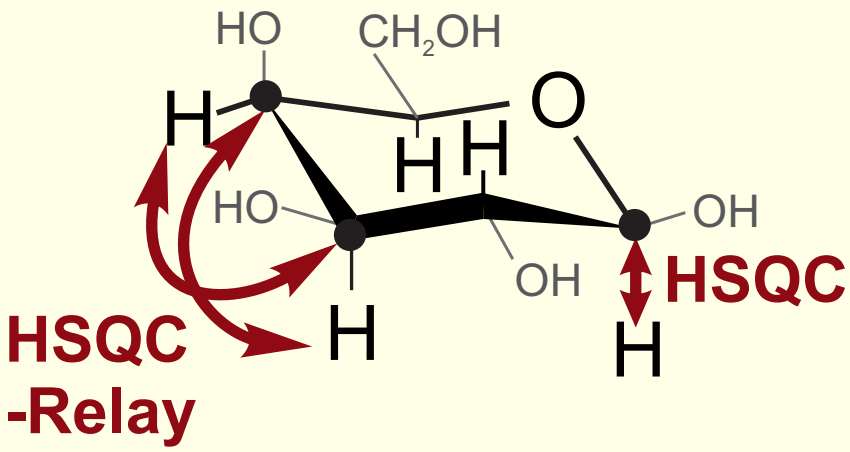


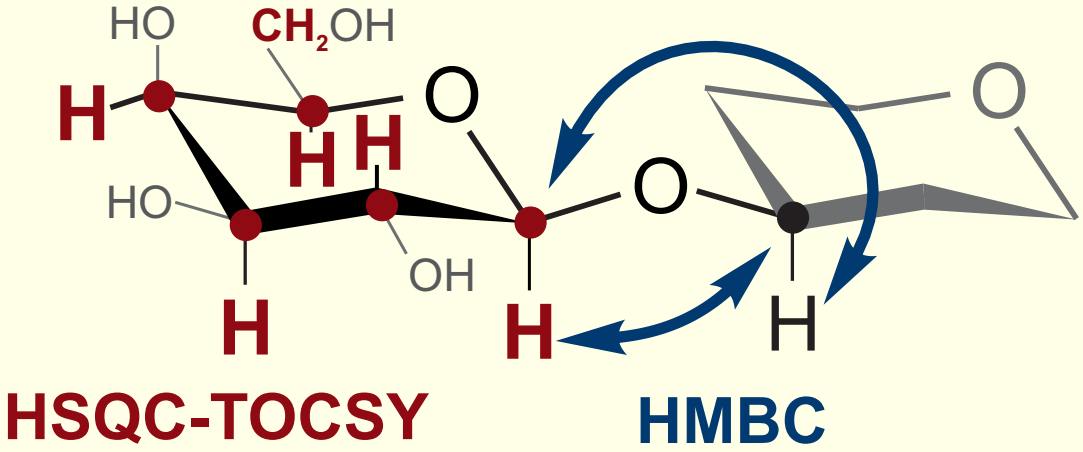


COSY	spin correlation between vicinal protons - *proton spectrum assignment*
RCT, RCT2	spin correlation with polarization transfer along vicinal proton pairs - *proton spectrum assignment*
DQF COSY	COSY without a diagonal line - *assignment of proximal signals*
^{1}H HD dif	differential selective decoupling - *H2 line shape analysis*
TOCSY (HOHAHA)	spin correlation with all protons in the spin system - *spin system distinguishing*
1D TOCSY	TOCSY for a single signal - *extraction of a residue spin system*
NOESY, ROESY	spatial correlation - *determination of the residue sequence, conformational studies*
^{1}H NOE dif	differential selective NOE measurement - *spatial contact studies*

Heteronuclear correlations

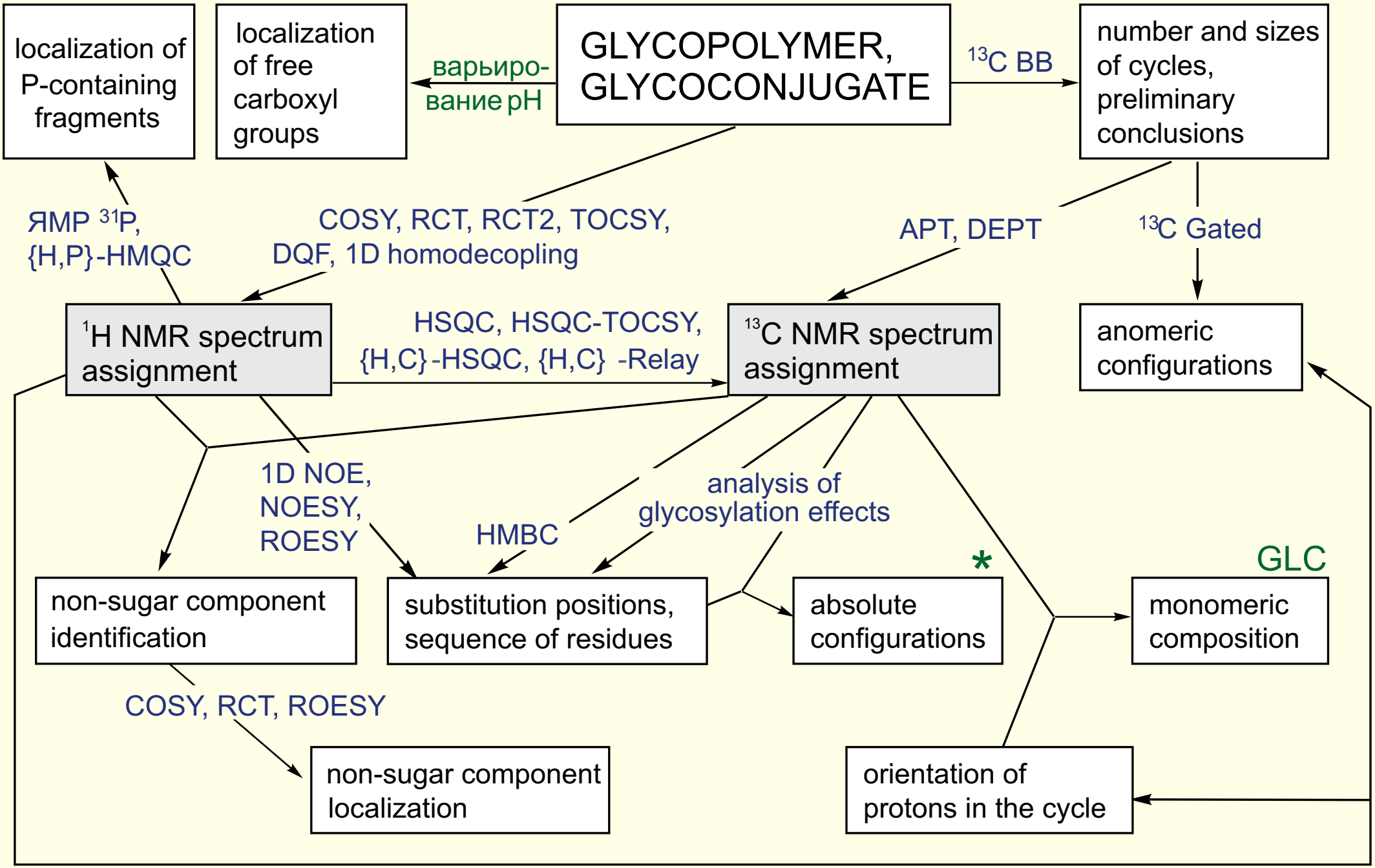
1D 2D





{^{1}H,^{13}C} HSQC	direct proton-carbon spin correlation - *carbon spectrum assignment*
{^{1}H,^{31}P} HSQC	proton-phosphorus spin correlation - *localization of phospate groups*
{^{1}H,^{13}C} HMBC	long-range proton-carbon spin correlation - *determination of the residue sequence*
{^{1}H,X} 1D HMBC	HMBC for a single signal - *assignment of protons around a certain heteroatom*
HSQC-Relay	carbon-carbon correlation via spin coupling of their vicinal protons - *поиск соседних углеродных атомов*
HSQC-TOCSY	correlation of carbons with all protons in the spin system (and vice versa) - assignment of *C5 using H6, and similar problems*
{^{1}H,X} 1D NOE	measurement of heteronuclear NOE - *conformational studies*

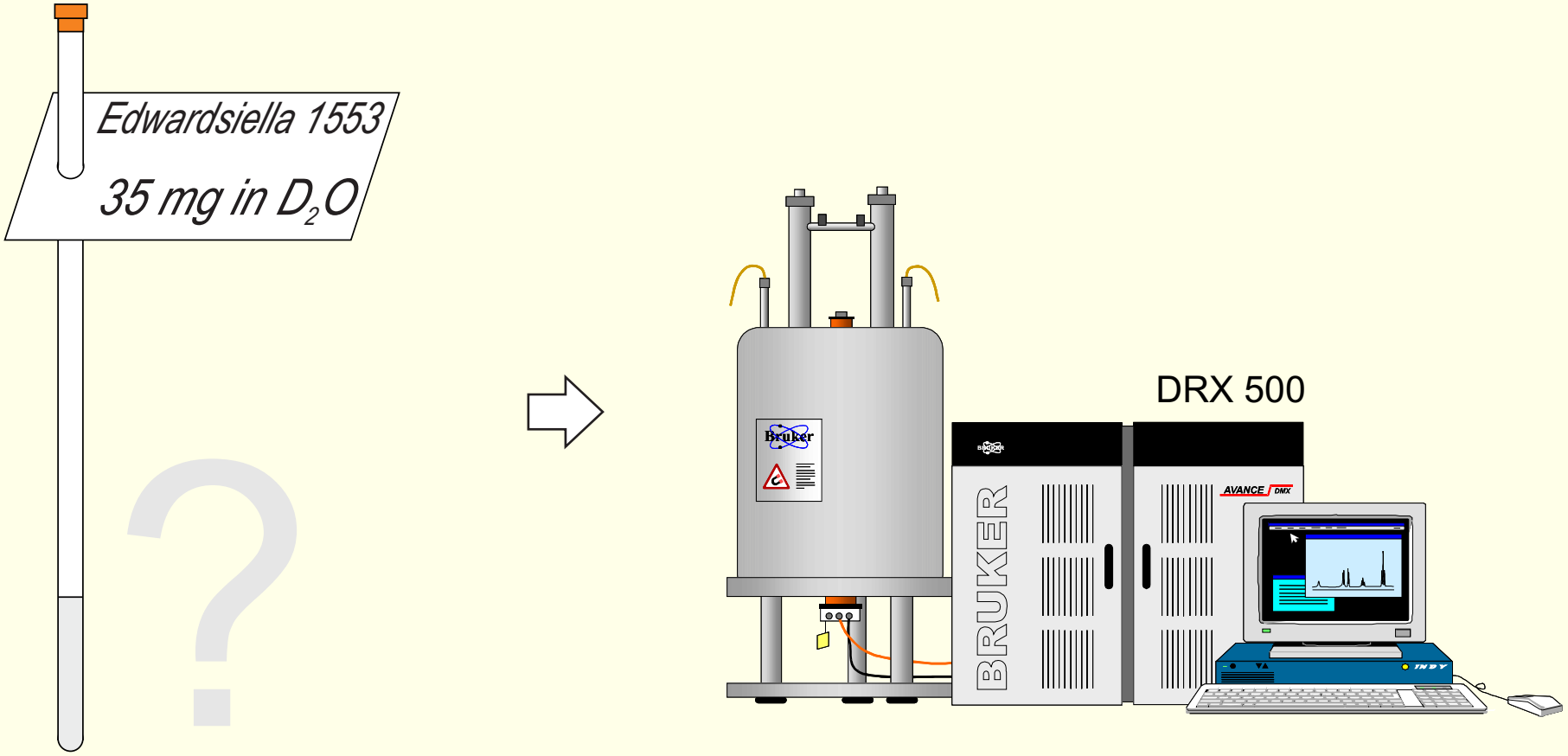
Approximate research schema



Source data

O-antigenic polysaccharide was extracted from the LPS of *Edwardsiella* 1153 bacteial cell wall
The ^{13}C NMR spectrum displayed regularity after the sample de-O-acetylation.

GLC (Sugar analyzer): at least **GlcN, Gal, GalA** are present in the unknown proportion

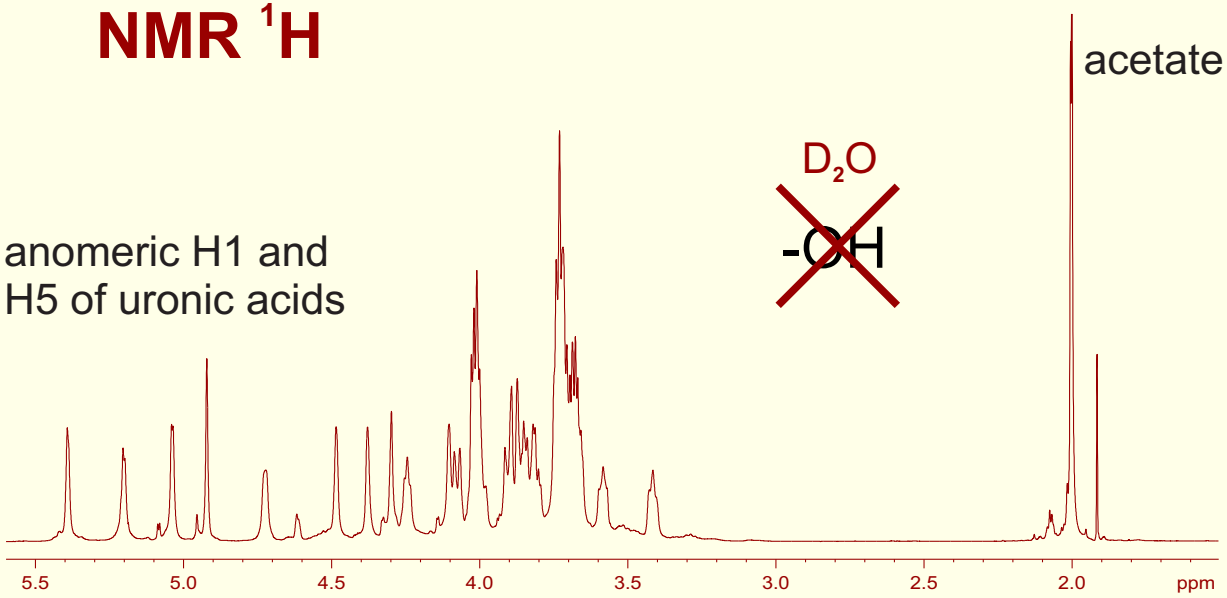


NMR ^{31}P

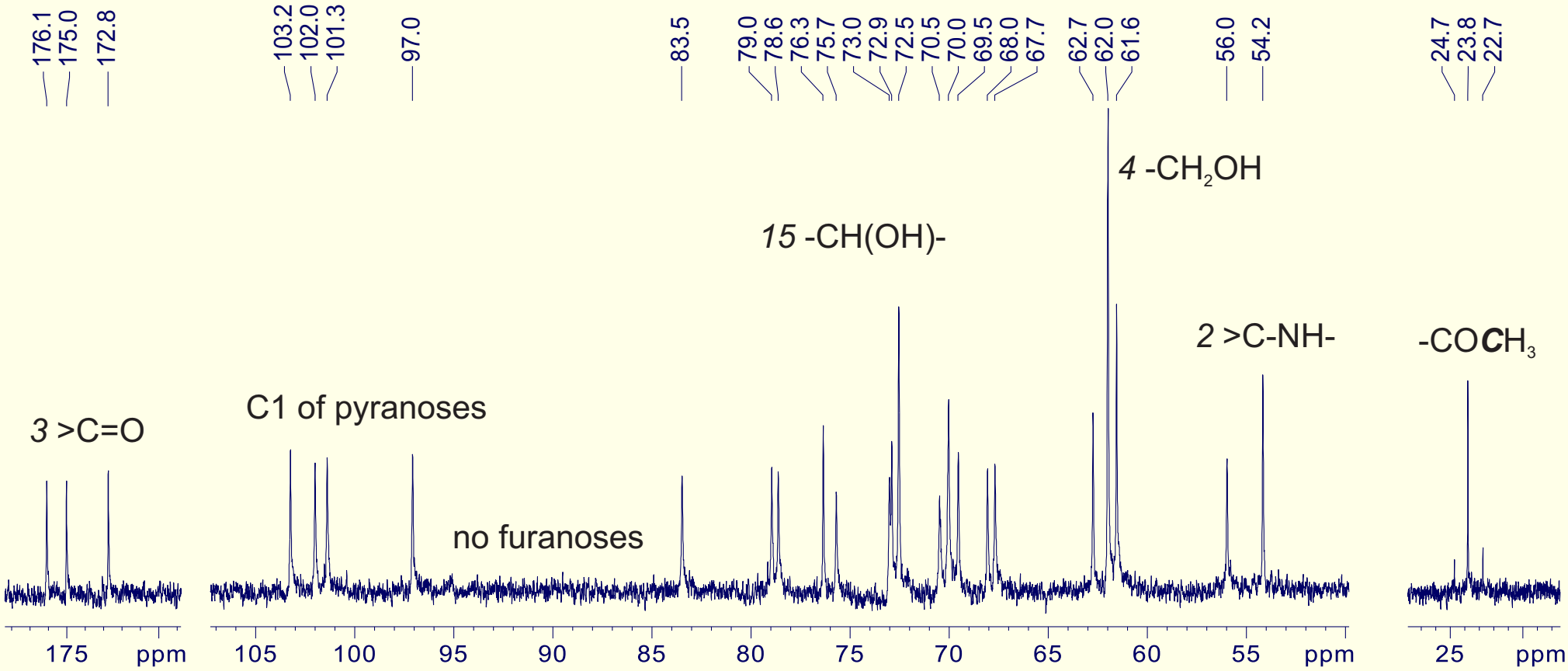
no signals

http://www.glyco.ac.ru/bcsdb3/

NMR ^{1}H



NMR ^{13}C BB



{^{1}H,^{1}H} COSY

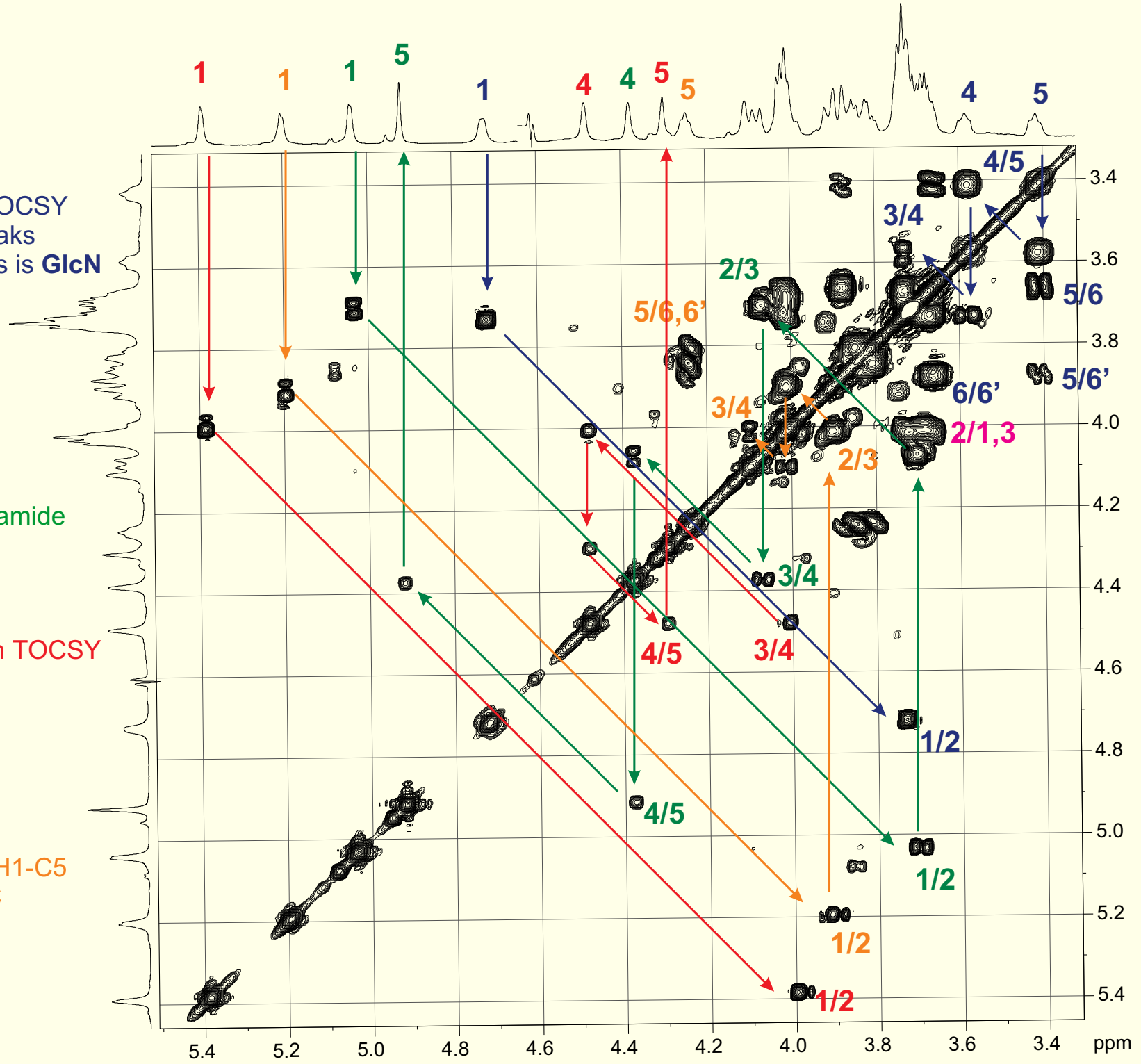
correlates protons with
vicinal proton neighbours

1-2 and **3-4-5-6,6'** are linked in TOCSY
5 assigned due to *three* cross-peaks
line shape of **4** => 3,4,5 ax => this is **GlcN**
(confirmation: C2 in HSQC)

1-2-3-4-5 are linked in COSY
(confirmation: TOCSY)
C6/H5 correlation in HMBC
and position of **5** => this is **GalA**-amide

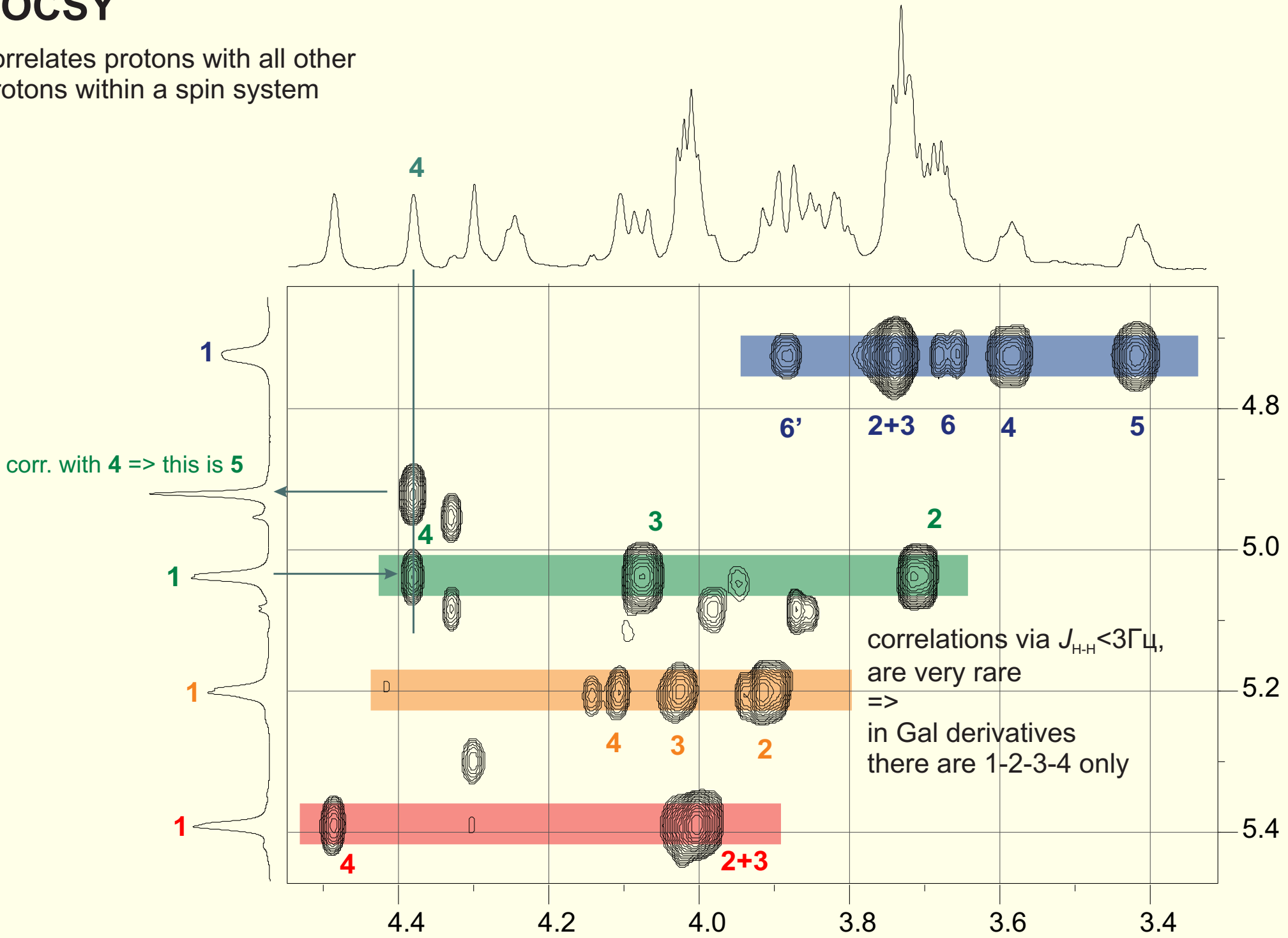
1-2 and **3-4-5** are linked via 1-4 in TOCSY
(confirmation: HSQC-TOCSY)
C6/H5 correlation in HMBC
=> this is **GalA**

1-2-3-4 and **5-6,6'** are linked via H1-C5
in HMBC and via H5/C5 in HSQC
(confirmation: **1**/**6** in ROESY
while **1**→**4** is proven)
+by exclusion => this is **Gal**



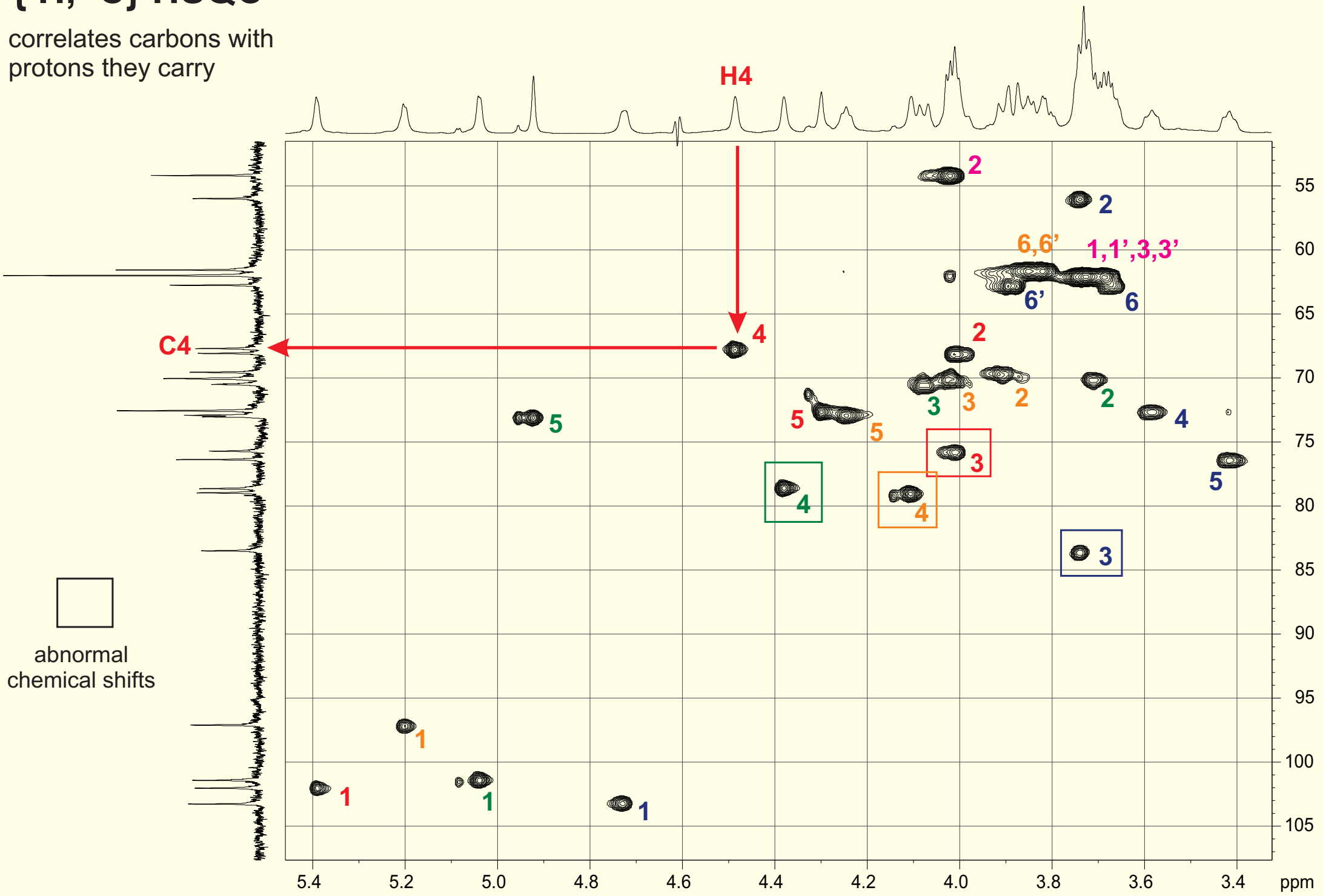
TOCSY

correlates protons with all other protons within a spin system



{^{1}H,^{13}C} HSQC

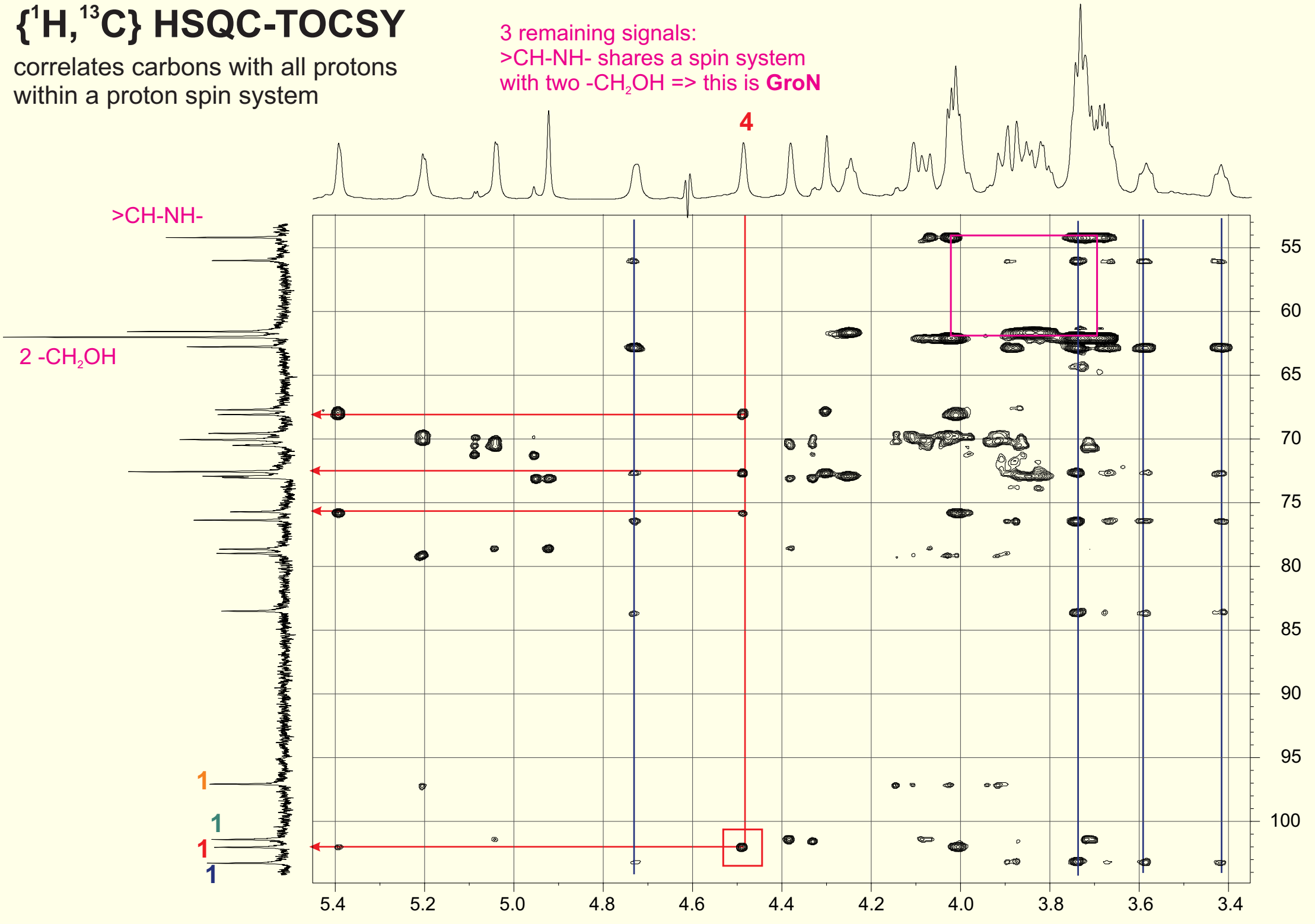
correlates carbons with protons they carry



{^{1}H,^{13}C} HSQC-TOCSY

correlates carbons with all protons within a proton spin system

3 remaining signals:
>CH-NH- shares a spin system
with two -CH_2OH => this is **GroN**



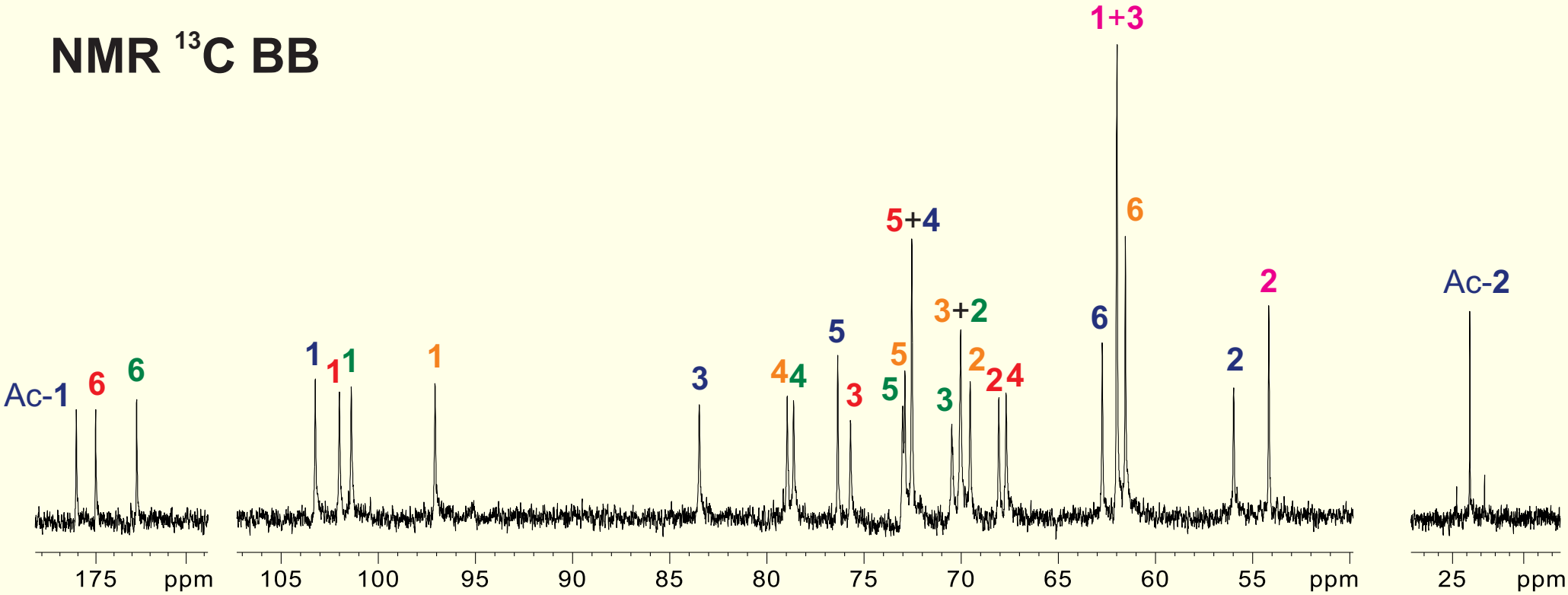
Substitution positions and anomeric configurations

	C1	C2	C3	C4	C5	C6
→3)-Gal*p*A^{I}-(1→	102.0 ***+7.2***	68.1	**75.8** ***+5.4***	67.7 *-4.0*	72.6	175.0
→3)-Glc*p*N-(1→	103.3 ***+7.1***	56.0 *-2.0*	**83.5** ***+8.4***	72.6 *+1.4*	76.4	62.8
Ac-(1→2)	176.1	23.8				
→4)-Gal*p*A^{II}-(1→	101.4 ***+6.6***	70.1	70.5	**78.7** ***+7.0***	73.0 *+0.8*	172.8 *-2.2*
GroN-(2→6)	62.0	54.2	62.0			
→4)-Gal*p*-(1→	97.1 ***+3.6***	69.6	70.1	**79.0** ***+8.4***	72.9 *+1.2*	61.6

GlcN
δC5 >76 => β-anomer
(confirmation: H1 line shape)

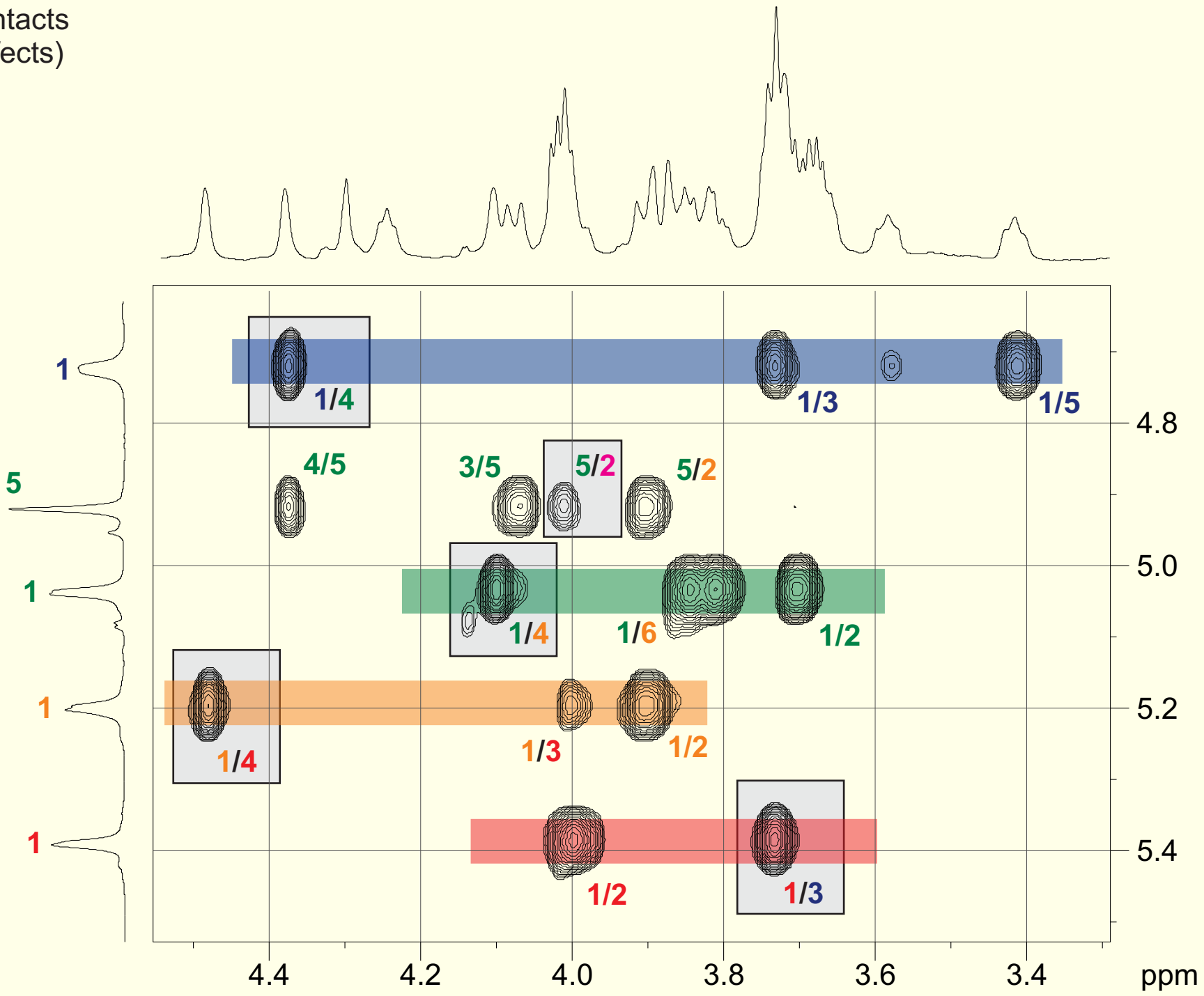
other:
δC5 <73 => α-anomers
(confirmations: H1 line shapes)

NMR ^{13}C BB



ROESY

shows proton spatial contacts
(Nuclear Overhauser Effects)



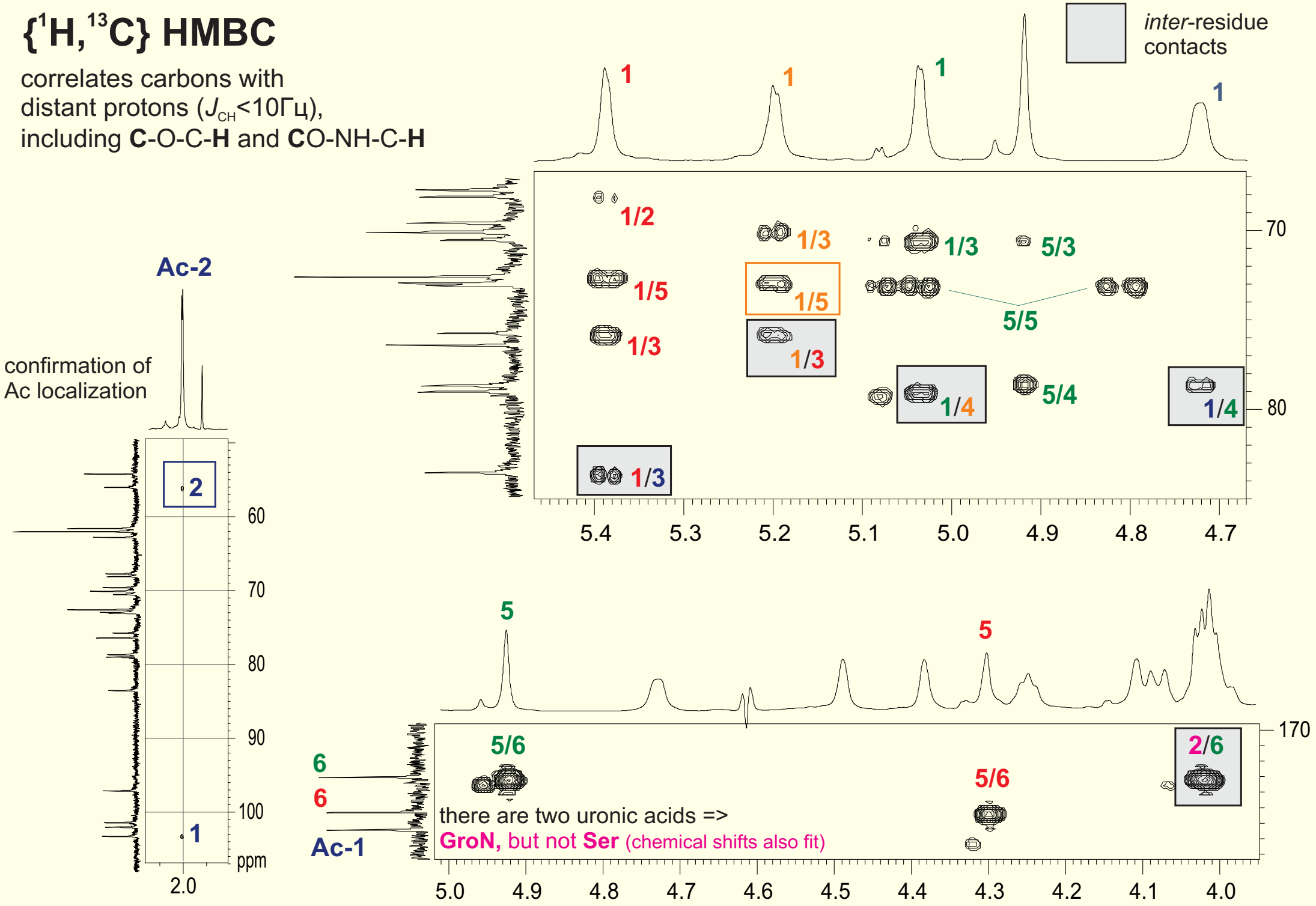
choice is based on
substitution positions
(HSQC);

confirmation: HMBC

inter-residue
contacts

{^{1}H,^{13}C} HMBC

correlates carbons with distant protons (J_{CH}<10Гц), including **C**-O-C-**H** and **C**O-NH-C-**H**



Absolute configurations

substitution effects are sensitive to absolute configurations and are deposited in the databases

	C1	C2	C3	C4	C5	C6
→3)-α-Gal*p*A^{I}-(1→	102.0 *+7.2*	68.1	**75.8** *+5.4*	67.7 *-4.0*	72.6	175.0
→3)-β-Glc*p*N-(1→	103.3 *+7.1*	56.0 *-2.0*	**83.5** *+8.4*	72.6 ***+1.4***	76.4	62.8
Ac-(1→2)	176.1	23.8				
→4)-α-Gal*p*A^{II}-(1→	101.4 *+6.6*	70.1	70.5	**78.7** *+7.0*	73.0 *+0.8*	172.8
GroN-(2→6)	62.0	54.2	62.0			
→4)-α-Gal*p*-(1→	97.1 ***+3.6***	69.6	70.1	**79.0** *+8.4*	72.9 *+1.2*	61.6

both GalA are D

(from the values of optical rotation of their S-butylglycosides)

residue pair	atom	database		experiment
		DD	DL	
Galα1→3GalA^{I}	C-1	**3.3**	8.3	3.6
GalA^{I}β1→3**GlcN**	C-4	**0.7**	-1.3	1.4

all chiral residues have D-configuration

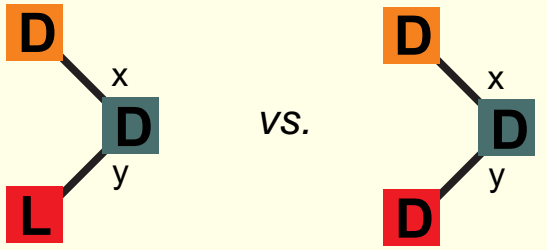
Sources of ^{13}C chemical shift data

BCSDB:
experimental NMR data

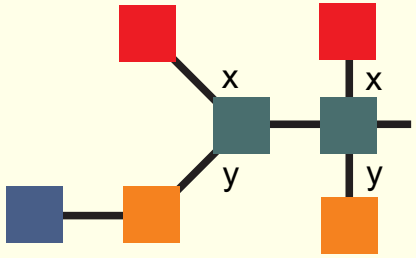
BCSDB/BIOPSEL:
^{13}C chemical shifts of monomers,
dimers (linear fragments),
and trimers (branching points).
substitution effects.
http://www.glyco.ac.ru/bcsdb3/nmr.html

	C1	C2	C3	C4	C5	C6
→3)-α-D-Gal*p*A^{I}-(1→	102.0 *+7.2*	68.1	**75.8** *+5.4*	67.7 *-4.0*	72.6	175.0
→3)-β-D-Glc*p*N-(1→	103.3 *+7.1*	56.0 *-2.0*	**83.5** *+8.4*	72.6 ***+1.4***	76.4	62.8
Ac-(1→2)	176.1	23.8				
→4)-α-D-Gal*p*A^{II}-(1→	101.4 *+6.6*	70.1	70.5	**78.7** *+7.0*	73.0 *+0.8*	172.8
GroN-(2→6)	62.0	54.2	62.0			
→4)-α-D-Gal*p*-(1→	97.1 ***+3.6***	69.6	70.1	**79.0** *+8.4*	72.9 *+1.2*	61.6

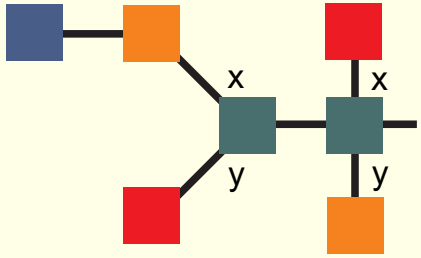
Example cases when incremental calculation is demanded:



vs.



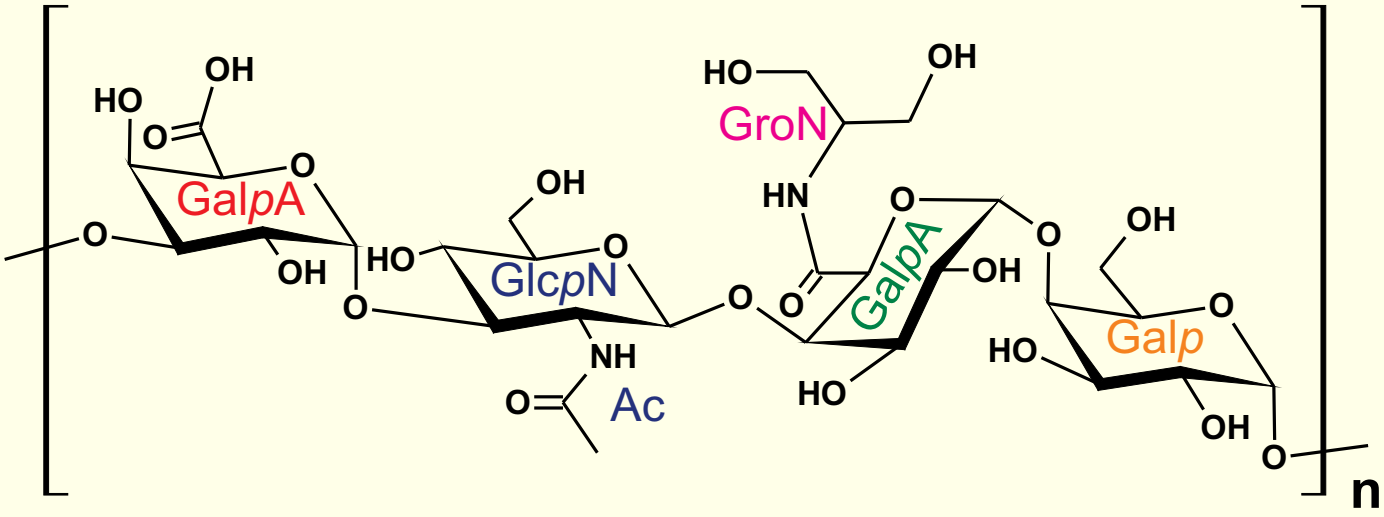
vs.



Structure of the chemical repeating unit of *Edwardsiella* 1153

GroN-(2
\6
→3)-α-D-Gal*p*A^{I}-(1→3)-β-D-Glc*p*NAc-(1→4)-α-D-Gal*p*A^{II}-(1→4)-α-D-Gal*p*-(1→

HSQC => in the native polysaccharide Gal is O-acetylated at position 2 (45% of units) or 3 (other 45% of units)



***30** hours of work:*

extraction and purification	4
de-O-acetylation	2
monomeric composition	1
optical rotation	1
NMR spectra acquisition	8
interpretation	6
documentation and presentation	3
O-acetate localization	4+1