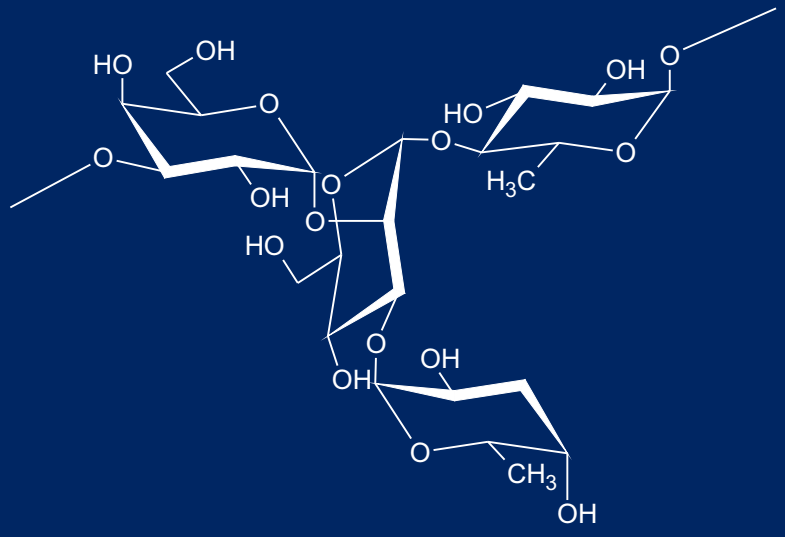
2D NMR in structural studies of natural compounds

model study:



(O-antigen of *Citrobacter freundii* O22)

Sugar analyzer:

Rha (2 eq, 6d)

Gal (4 eq)

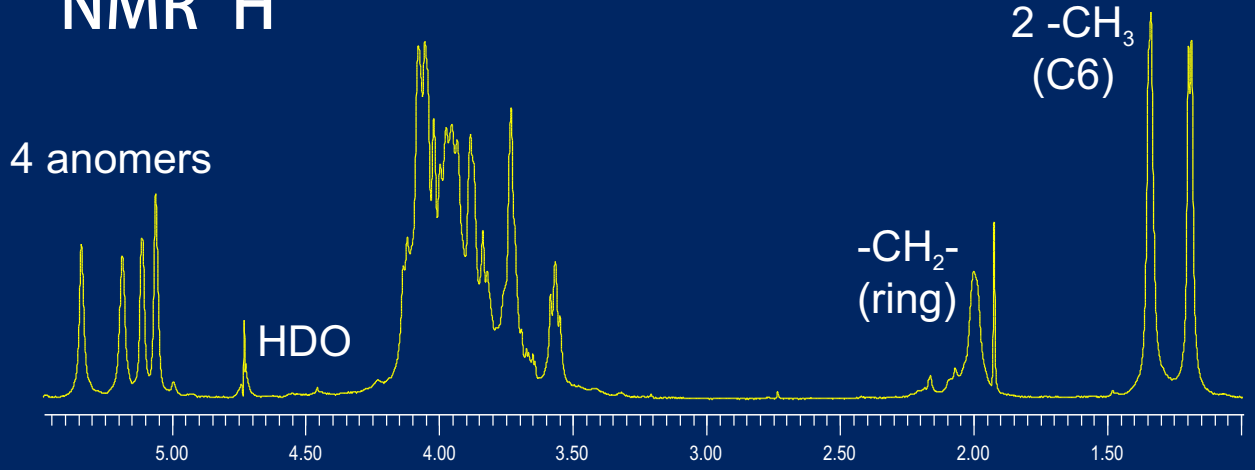
Man (3 eq)

Unknown

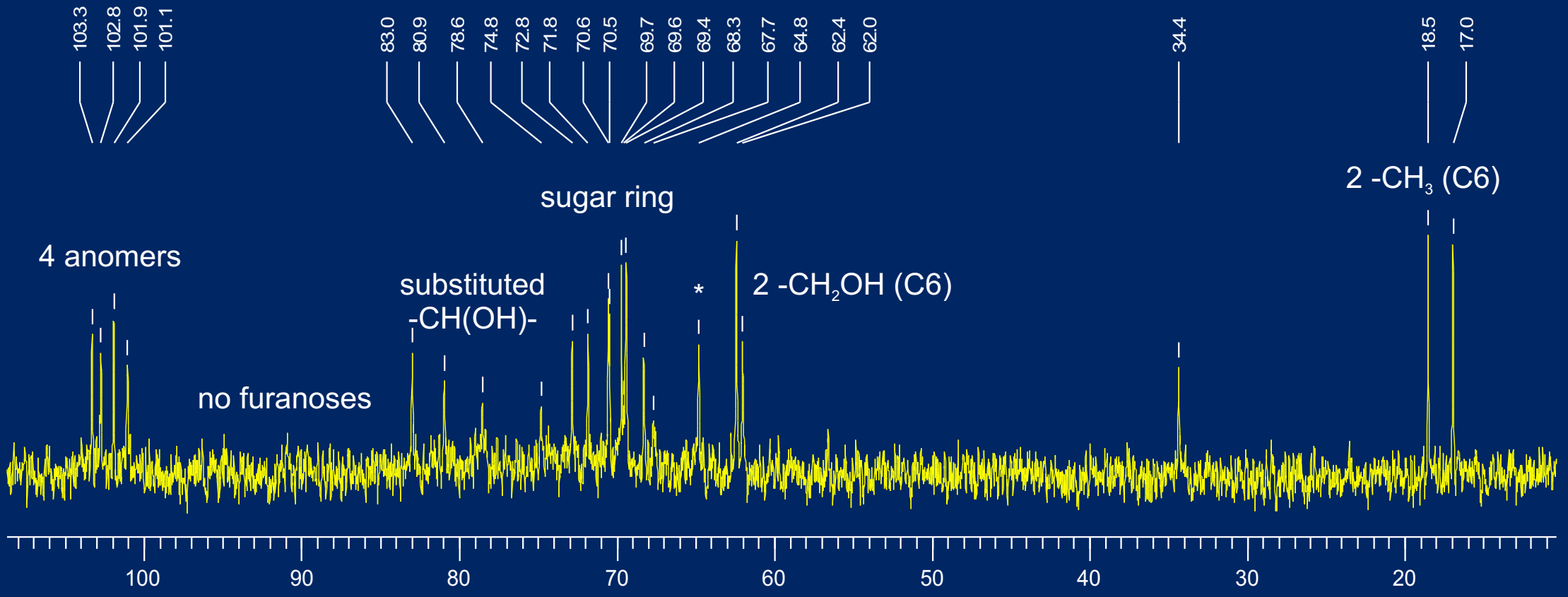
NMR ^{31}P

empty

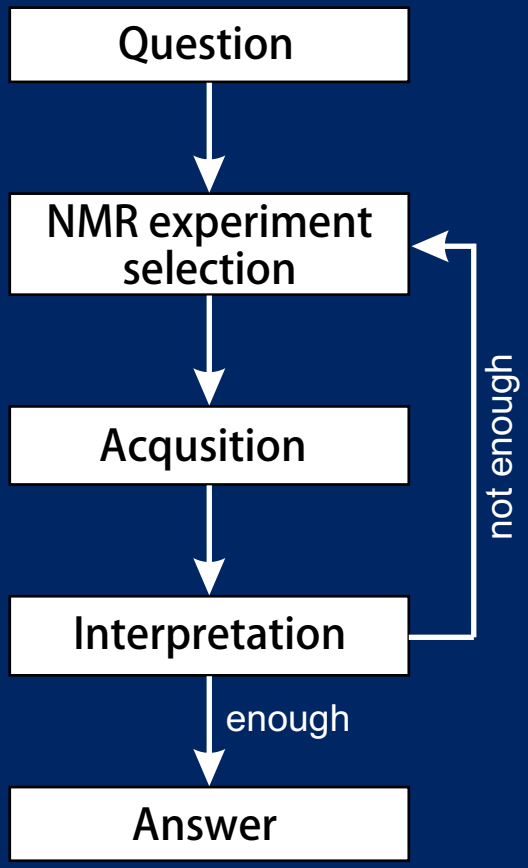
NMR ^{1}H



NMR ^{13}C BB



Experiments

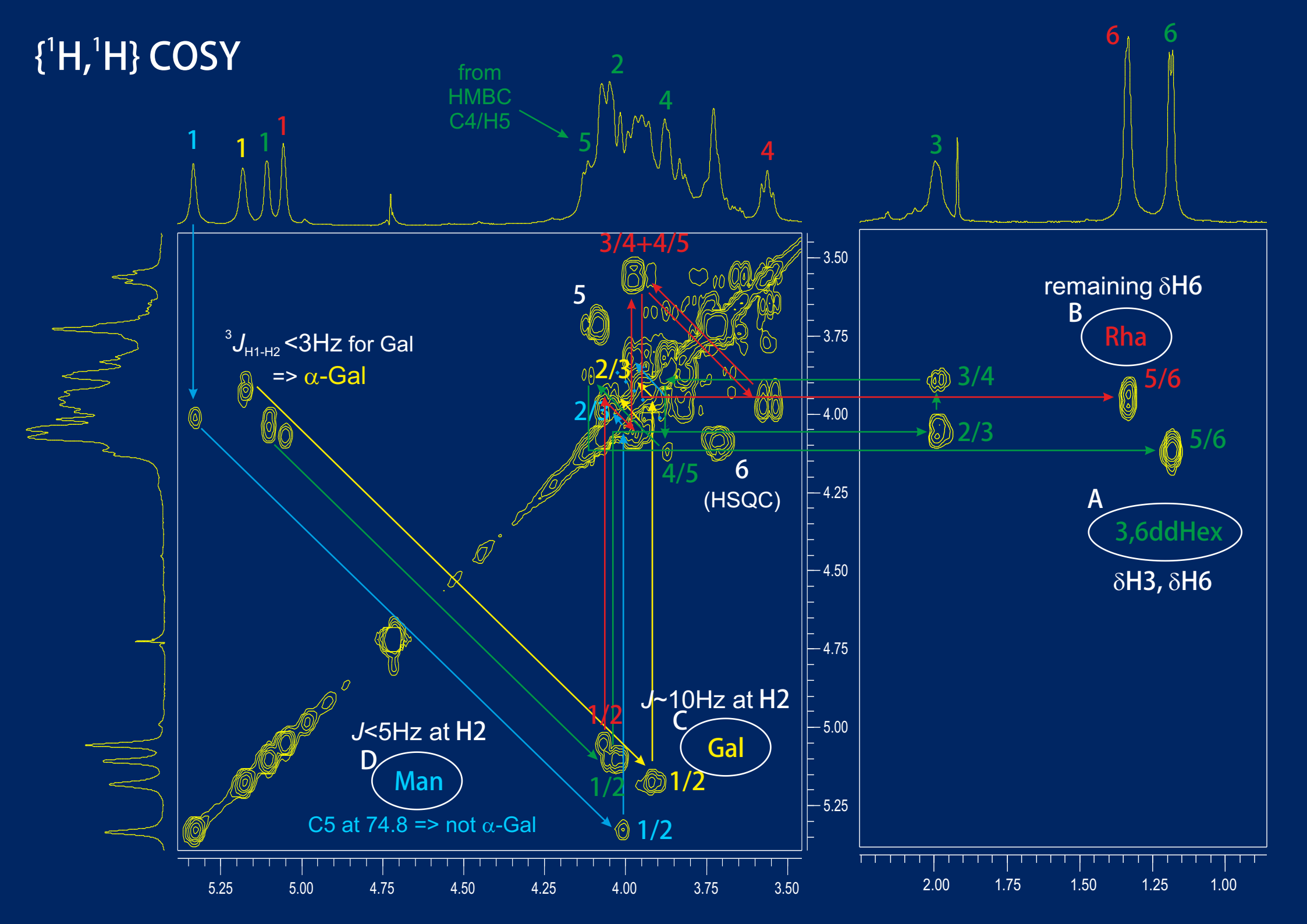


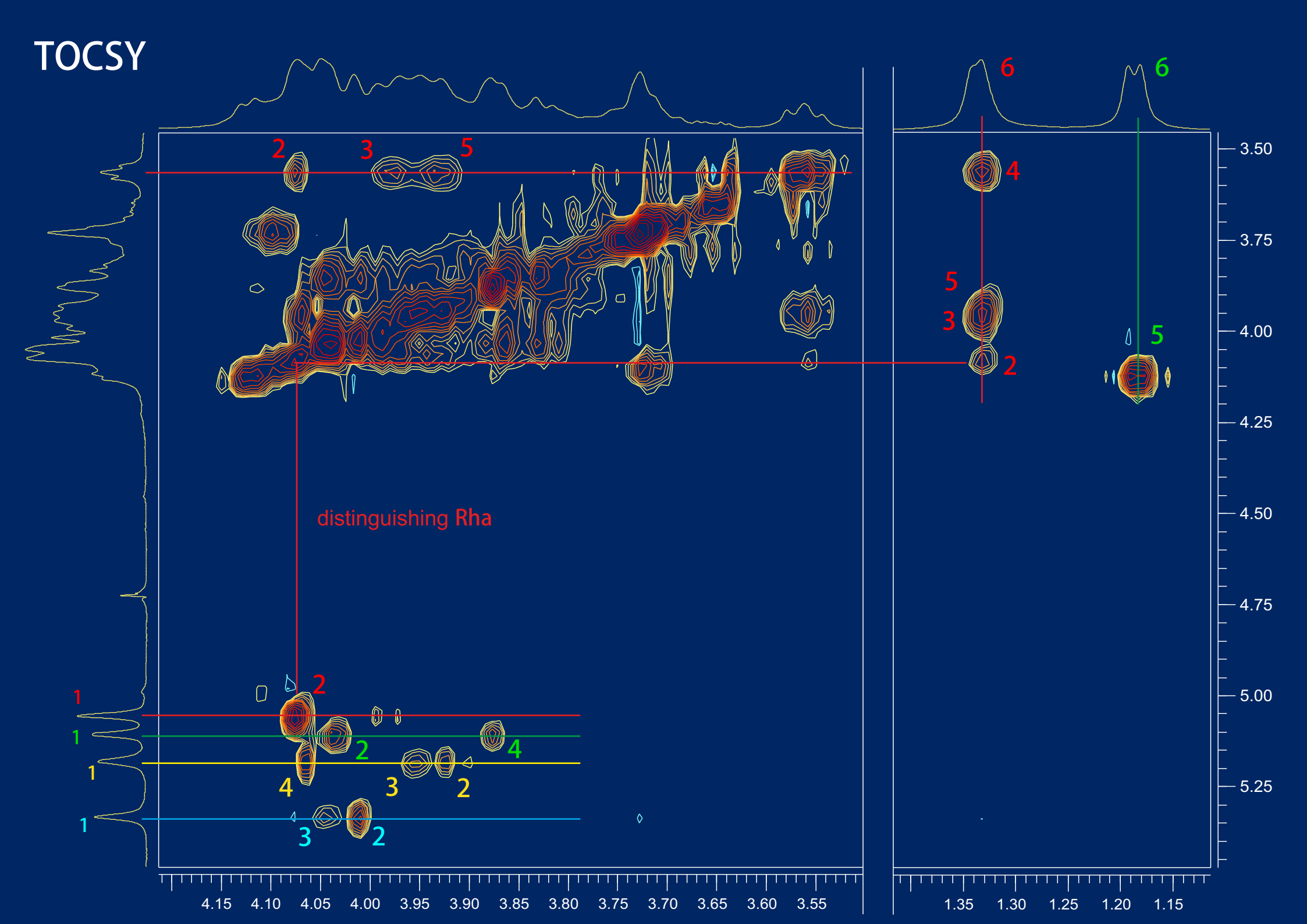
^{1}H 1D NMR	proton spectrum - *general information*
^{1}H HD diff	selective homonuclear decoupling - *revealing of signals of neighboring protons*
^{13}C, ^{31}P, ... BB	broad-band proton decoupled spectra - *additional information, "fingerprint", substitution positions*
^{13}C DEPT, APT, ^{15}N INEPT	edited selective polarization transfer - *revealing carbon and nitrogen protonation and sensitivity gain*
^{13}C Gated	undecoupled carbon spectrum - *heteronuclear coupling constants measurement*
^{1}H NOE diff	nuclear Overhauser effect measurement - *proton-proton spatial contacts*

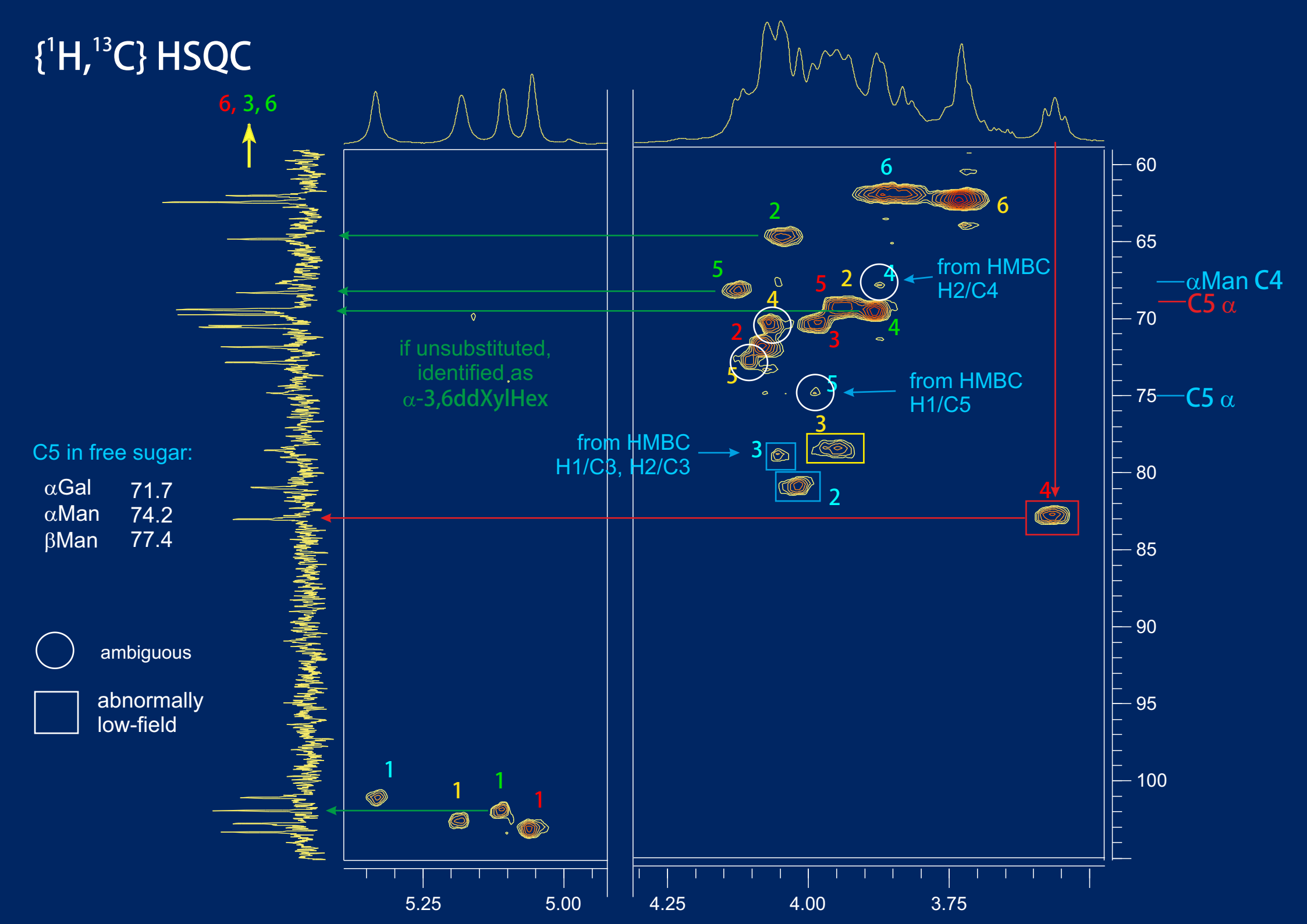
COSY, COSY-45	homonuclear spin correlation - *proton spectrum assignment*
COSY n-RCT	relayed coherence transfer in COSY - *proton spectrum assignment*
DQF COSY	double quantum filtered COSY - *assignment of proximal signals*
TOCSY	total homonuclear correlation - *distinguishing of proton spin systems*
NOESY, ROESY	homonuclear spatial correlation - *revealing of residue sequence and conformation studies*
DOSY	diffusion ordered spectroscopy - *separation of a spectrum into component subspectra*

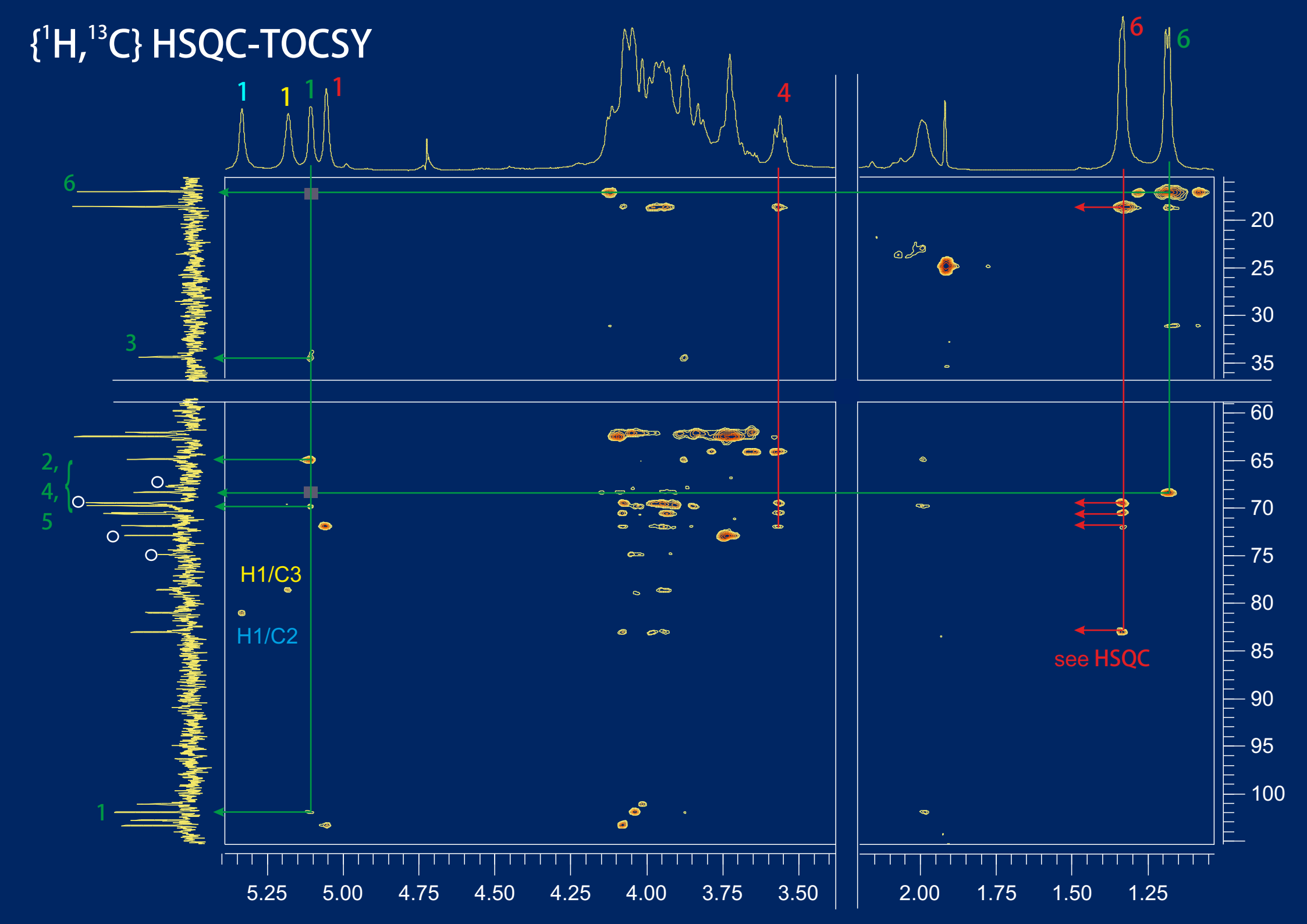
interpretation of spectra is NOT sequential

^{1}H, ^{13}C HSQC	proton-carbon spin correlation - *carbon spectrum assignment*
^{1}H, ^{31}P HSQC	proton-phosphorus spin correlation - *phosphate groups localization*
^{1}H, ^{13}C HMBC, ^{1}H, ^{15}N HMBC	multiple-bond heteronuclear spin correlation - *revealing modifier attachment pattern and residue sequence*
HSQC Relay	relayed heteronuclear spin correlation - *tracking neighboring carbons*
HSQC-TOCSY	total heteronuclear correlation - *distinguishing of residue spin systems*

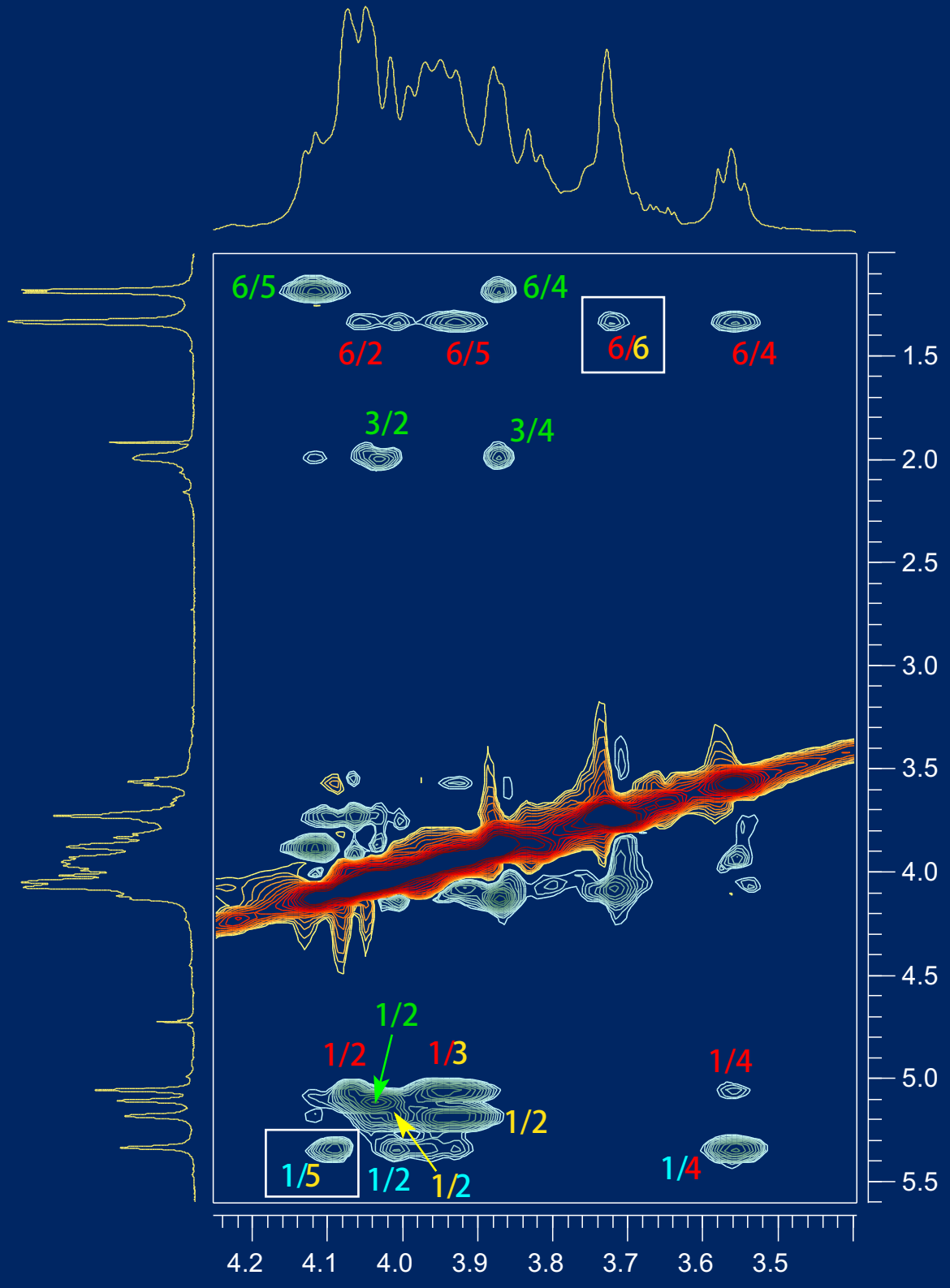

{1H,1H} COSY
from HMBC C4/H5
1
1
1
1
5
2
4
4
3
6
6
3/4+4/5
5
2/3
2/3
4/5
6
(HSQC)
3/4
2/3
5/6
5/6
remaining δH6
B
Rha
A
3,6ddHex
δH3, δH6
3JH1-H2 <3Hz for Gal
=> α-Gal
J~10Hz at H2
C
Gal
J<5Hz at H2
D
Man
C5 at 74.8 => not α-Gal
1/2
1/2
1/2
1/2
3.50
3.75
4.00
4.25
4.50
4.75
5.00
5.25
5.25
5.00
4.75
4.50
4.25
4.00
3.75
3.50
2.00
1.75
1.50
1.25
1.00


TOCSY
distinguishing Rha
4.15 4.10 4.05 4.00 3.95 3.90 3.85 3.80 3.75 3.70 3.65 3.60 3.55
1.35 1.30 1.25 1.20 1.15
3.50 3.75 4.00 4.25 4.50 4.75 5.00 5.25


{1H,13C} HSQC
6, 3, 6
if unsubstituted,
identified as
α-3,6ddXylHex
from HMBC
H2/C4
from HMBC
H1/C5
from HMBC
H1/C3, H2/C3
αMan C4
C5 α
C5 α
C5 in free sugar:
αGal 71.7
αMan 74.2
βMan 77.4
ambiguous
abnormally
low-field
5.25
5.00
4.25
4.00
3.75
60
65
70
75
80
85
90
95
100


{1H,13C} HSQC-TOCSY
1
1
1
1
4
6
6
6
3
2,
4,
5
1
H1/C3
H1/C2
see HSQC
20
25
30
35
60
65
70
75
80
85
90
95
100
5.25
5.00
4.75
4.50
4.25
4.00
3.75
3.50
2.00
1.75
1.50
1.25

ROESY



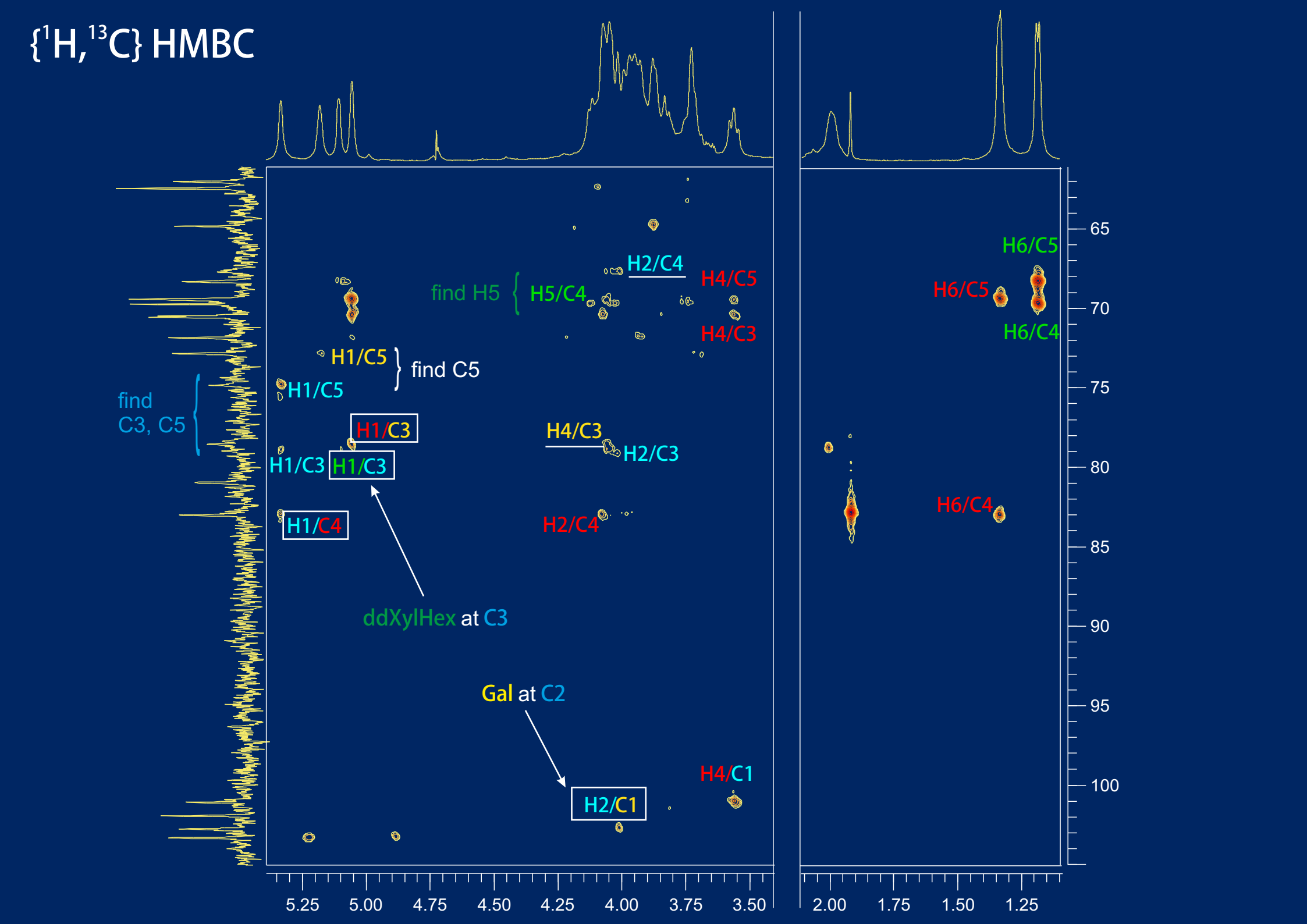
inter-residue
non-transglycosydic
contacts

by exclusion:
ddXylHex(1→?)Man

Gal(1→?)Man ? 2 3

Rha(1→3)Gal ? 3

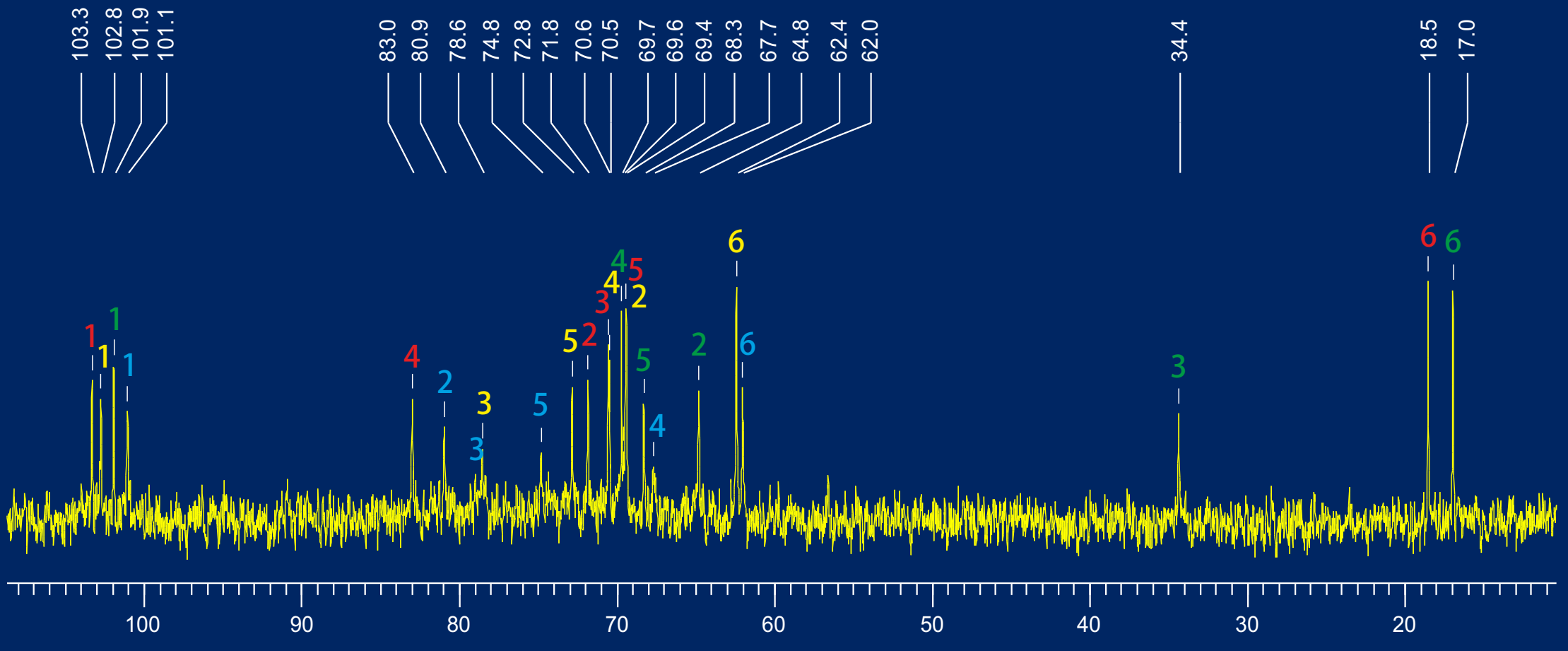
Man(1→4)Rha ? 4


{1H,13C} HMBC
H2/C4
H4/C5
find H5
H5/C4
H4/C3
H1/C5
find C5
H1/C5
find
C3, C5
H1/C3
H4/C3
H2/C3
H1/C3
H1/C3
H1/C4
H2/C4
ddXylHex at C3
Gal at C2
H4/C1
H2/C1
H6/C5
H6/C5
H6/C4
H6/C4
65
70
75
80
85
90
95
100
5.25
5.00
4.75
4.50
4.25
4.00
3.75
3.50
2.00
1.75
1.50
1.25

	C-1	C-2	C-3	C-4	C-5	C-6
α-Rhap	103.3	71.9	70.6	83.0	69.4	18.6
α-3,6ddXylp	102.0	64.8	34.4	69.7	68.3	17.0
α-Galp	102.8	69.4	78.6	70.5	72.9	62.5
α-Manp	101.1	80.9	79.0	67.7	74.8	62.0

	C-1	C-2	C-3	C-4	C-5	C-6
→4)α-Rhap	8.1	-0.2	-0.7	9.5	-0.1	0.6
α-3,6ddXylp	9.3	0.7	0.2	0.3	1.0	0.4
→3)α-Galp	9.3	-0.2	8.2	-0.1	1.2	0.1
→2,3)α-Manp	5.8	9.4	7.5	-0.5	0.6	-0.3

NMR ^{13}C BB



Absolute configurations

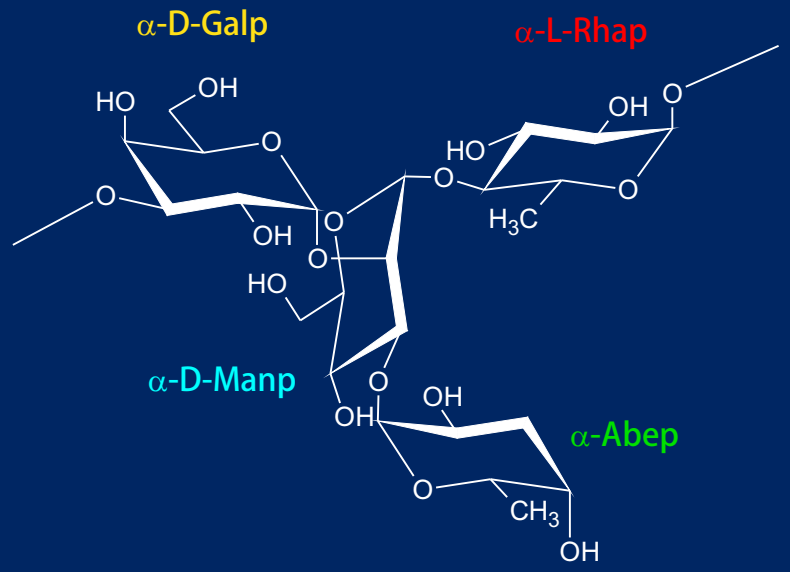
	C-1	C-2	C-3	C-4	C-5	C-6
→4)α-Rhap	8.1	-0.2	-0.7	9.5	-0.1	0.6
α-3,6ddXylp	9.3	0.7	0.2	0.3	1.0	0.4
→3)α-Galp	9.3	-0.2	8.2	-0.1	1.2	0.1
→2,3)α-Manp	5.8	9.4	7.5	-0.5	0.6	-0.3

residue pair	carbon	Theory		Experiment
		DD	DL	
Man→4αRha	C-4	7.6	9.2	9.5
Rha→3αGal	C-3	3.9	8.1	8.5
Gal→2αMan	C-2	9.7	6.0	9.4
Abe→3αMan	C-3	7.4	5.5	7.5

Gal is D (from enzymatic oxidation) =>
Rha is L =>
Man is D =>
3,6ddXyl is D (=Abequose)

Elucidated repeating unit

α-Abep(1
↓
3
→3)α-D-Galp(1→2)α-D-Manp(1→4)α-L-Rhap(1→



Details

The ^{13}C spectrum of polysaccharide demonstrated a regular structure. It contained signals for four sugar residues, including those for four anomeric carbons at 103.3, 102.8, 101.9 and 101.1, two unsubstituted CH_2OH groups, 15 oxygen-bearing sugar-ring carbons in the region 64-82, one CCH_2C group at 34.4 and two CH_3 groups at 18.6 and 17.0. Accordingly, the ^1H NMR spectrum contained signals for four anomeric protons at 5.34, 5.18, 5.10 and 5.06, signals of sugar-ring protons in region 3.5-4.2, one signal of a $C\text{-}CH_2\text{-}C$ group at 2.00 and signals of two CH_3 groups at 1.34 and 1.19. As judged by the absence of signals within ∂ 82-88 region, all sugar residues are in pyranose form [1].

The sugar analysis of the polysaccharide revealed residues of Rha, Man and Gal residues in the ratio ~1:1:1 and showed the presence of one more sugar, which has not been identified.

The ^1H and ^{13}C spectra of the polysaccharide were assigned using {^1H, ^1H} COSY, TOCSY, ROESY, {^1H,^{13}C} HSQC, {^1H,^{13}C} HSQC-TOCSY and {^1H,^{13}C} HMBC experiments (Tables 1 and 2). The spin system of Rha*p* was distinguished basing on TOCSY spectrum that showed correlations of Rha*p* H-6 (1.34) with all the other protons of Rha*p*. The Rha*p* ^1H signal assignment was completed by the COSY spectrum, which contained all correlations between neighboring protons in this residue. The assignment of Rha*p* ^{13}C signals was performed using the data of {^1H,^{13}C} HSQC experiment and confirmed by {^1H,^{13}C} HSQC-TOCSY, which revealed the correlations of Rha*p* H-2 with all carbons of the residue. The chemical shift for Rha*p* C-5 (69.4) indicated that this residue was in a-anomeric configuration [2]. The significant downfield displacement of Rha*p* C-4 signal (from 73.5 to 83.0) determined the substitution position as C-4.

The signal at 2.00 in the proton spectrum indicated the presence of a sugar deoxygenated at one of ring carbons. The further investigation (see below) identified this residue as 3,6-dideoxy-sugar (*3d6dHex*). The signals for protons from H-2 to H-4 were assigned using the COSY spectrum, which showed all correlations between neighboring protons in this residue. However, as judged by {^1H,^{13}C} HSQC experiment, the signal for *3d6dHex* H-4 at 3.88 possessed a complete overlap with the signals of protons corresponding to ^{13}C signals at 62.0 and 67.7. Due to this, data of COSY experiment could not be used for unambiguous assignment of the signal for *3d6dHex* H-5.

The ^{13}C signals from C-1 to C-4 were assigned using the data of {^1H,^{13}C} HSQC experiment. Due to the strong overlap of *3d6dHex* H-2 (with the signals of protons linked to carbons with chemical shifts 79.0 and 74.8) and H-4 signals (see above) in ^1H NMR spectrum at 4.04 and 3.88, respectively, assignment of signals for C-2 and C-4 was not clearly determined from {^1H,^{13}C} HSQC data and required data of {^1H,^{13}C} HSQC-TOCSY experiment, which contained H-1/C-2 and H-1/C-4 correlations at 5.11/64.8 and 5.11/69.7, respectively. The signal for *3d6dHex* H-5 was assigned basing on *intra*-residue C-4/H-5 correlation in HMBC spectrum at 69.7/4.12, and the H-5/H-6 correlation in COSY spectrum at 4.12/1.19 allowed to assign the signal for *3d6dHex* H-6. The signals for C-5 and C-6 were assigned using the data of {^1H,^{13}C} HSQC experiment.

The spin system of *3d6dHex* was distinguished basing on upfield chemical shifts of H-3 (2.00) and H-6 (1.19). The ^{13}C chemical shifts of *3d6dHex* appeared to be characteristic for the terminal 3,6-dideoxy-a-xylohexopyranoside [3].

The ^1H and ^{13}C NMR signals for *Hex-II* residue (Gal*p* or Man*p,* accordingly to the sugar analysis, H-1 at 5.34) were assigned using the data of COSY, {^1H,^{13}C} HSQC, and {^1H,^{13}C} HMBC experiments. The signal for *Hex-II* H-2 was assigned basing on H-1/H-2 correlation at 4.04/5.11 in COSY spectrum. As there were no H-2/H-3 and H-3/H-4 correlations observed in COSY, the signals for H-3 and H-5 were assigned using the data of {^1H,^{13}C} HSQC and {^1H,^{13}C} HMBC experiments. Particularly, the HMBC spectrum demonstrated the *intra*-residue H-1/C-5, H-1/C-3 and H-2/C-3 correlations at 5.34/74.8, 5.34/79.0 and 4.02/79.0, respectively, and allowed the assignment of C-3 and C-5 signals. The H-1/H-2 coupling constant <2 Hz revealed from ^1H NMR spectrum showed that *Hex-II* could not be the b-Gal*p* residue. The chemical shift for *Hex-II* C-5 was 74.8, while chemical shifts for C-5 of unsubstituted a-Gal*p*, a-Man*p* and a-Man*p* are 71.7, 74.2 and 77.4, respectively [2]. These data indicated that *Hex-II* had the a-*manno*-configuration, as the b-glycosilation effects of substitution at Gal*p* C-4 or C-6 or Man*p* do not exceed 1.8, and thus could not explain the observed chemical shift in all the other cases [4].

The signals for Man*p* H-3, H-4, H-5 and C-6 were assigned basing on correlations with corresponding atoms in the {^1H,^{13}C} HSQC spectrum. H-6a and H-6b signal assignment followed from H-5/H-6a and H-5/H-6b correlations in the COSY spectrum observed at 3.98/3.87 and 3.98/3.82, respectively. The assignment of Man*p* C-4 was done using the characteristic chemical shift (67.7) of Man*p* C-4 [2], and H-2/C-4 cross-peak at 67.7/4.02 in the {^1H,^{13}C} HMBC spectrum. The significant downfield displacement of the signals for Man*p* C-2 (from 71.5 to 83.0) and C-3 (from 71.5 to 79.0), as compared to the signals in unsubstituted residue [2], revealed that this residue was bisubstituted at C-2 and C-3.

The TOCSY spectrum showed the correlations of H-1 of the remaining Gal*p* with protons from Gal*p* H-2 to H-4. Together with COSY data, this allowed the assignment of Gal*p* H-2, H-3 and H-4. Due to the absence of H-4/H-5 correlation in COSY, further assignment required the data of {^1H,^{13}C} HSQC and HMBC experiments. Particularly, HMBC spectrum showed the H-1/C-5 correlation at 5.18/72.9 and the HSQC spectrum showed the C-5/H-5 correlation at 72.9/4.10. The signals of H-6a and H-6b were assigned basing on H-5/H6a and H-5/H6b correlations in COSY. The H-1/H-2 coupling constant <2Hz demonstrated the a-configuration of the residue. The Gal*p* ^{13}C signals except C-5 were assigned using the data of {^1H,^{13}C} HSQC experiment. The significant downfield displacement of Gal*p* C-3 signal (from 70.4 in the unsubstituted residue to 78.6) demonstrated the substitution at C-3.

The HMBC spectrum showed the following *inter*-residue cross-peaks: Rha*p* H-1 / Gal*p* C-3 at 5.06/78.6; *3d6dHex* H-1 / Man C-3 at 5.11/79.0; Gal*p* C-1 / Man*p* H-2 at 102.8/4.02; Man*p* H-1 / Rha*p* C-4 at 5.34/83.0. This confirmed the substitution pattern and revealed the sequence of residues.

The absolute configuration of Gal*p* was determined as D using the enzymatic oxidation method as described [5]. Taking the configuration of this residue as reference, glycosilation effect analysis was performed using the BIOPSEL database [4]. The comparison of the observed -glycosilation effects with the theoretical ones showed the following interrelations between possible absolute configurations of residues:

Gal*p*(12)Man*p* D,D-pair (observed effect: +9.5, theoretical: DD +9.7, DL +6.0); Rha*p*(13)Gal*p* D,L-pair (observed effect: +8.2, theoretical: DD +3.9, DL +8.1); Man*p*(14)Rha*p* D,L-pair (observed effect: +9.5, theoretical: DD +7.6, DL +9.2); *3d6dHex*(13)Man*p* D,D-pair (observed effect: +7.5, theoretical: DD +7.4, DL +5.5).

These data revealed the absolute configuration of all other residues but Gal*p*: D for Man, L for Rha and D for 3,6-deoxy-xylohexopyranoside. 3,6-deoxy-D-xylohexopyranose is called “abequose”. Thus the structure of repeating unit was elucidated as following [slide 11].

Tabular data

^{1}H NMR data (ppm) for the O-specific polysaccharide of *Citrobacter* PCM 1555

	H-1	H-2	H-3	H-4	H-5	H-6a	H-6b
-Abe(1-	5.11	4.04	2.00	3.88	4.12	1.19	
2,3)-D-Man*p*(1	5.34	4.02	4.05	3.87	~3.98	3.87	3.82
4)-L-Rha*p*(1	5.06	4.07	~3.98	3.56	3.94	1.34	
3)-D-Gal*p*(1	5.18	3.92	~3.95	4.07	4.10	3.75	3.69

^{13}C NMR data (ppm) for the O-specific polysaccharide of *Citrobacter* PCM 1555

	C-1	C-2	C-3	C-4	C-5	C-6
-Abe(1-	102.0	64.8	34.4	69.7	68.3	17.0
2,3)-D-Man*p*(1	101.1	81.0	79.0	67.7	74.8	62.0
4)-L-Rha*p*(1	103.3	71.9	70.6	83.0	69.4	18.6
3)-D-Gal*p*(1	102.8	69.4	78.6	70.5	72.9	62.5

Referenced from “details”:

1. Bock K, Redersen C. *Adv Carbohydr Chem Biochem,* 1983, **41**, 27-66.
2. Lipkind GM, Shaskov AS, Knirel YA, Vinogradov EV, Kochetkov NK. *Carbohydr Res,* 1988, **175**, 59-75.
3. Bundle DR, Josephson S. *Can J Chem,* 1978, **56**, 2686-2690.
4. Toukach FV, Shashkov AS. *Carbohydr Res,* 2001, **335**(2), 101-114.
5. Roth H, Segal S, Bertoli D. *Anal Biochem,* 1965, **10**, 35-52.

THIS WORK:

Katzenellenbogen E, Kocharova NA, Toukach FV, Górska S, Korzeniowska-Kowal A, Bogulska M, Gamian A, Knirel YA
“Structure of an abequose-containing O-polysaccharide from *Citrobacter freundii* O22 strain PCM 1555", *Carbohydr Res*, 2009, **344**(13), 1724-1728.