



Ф.В. Тоукач

**Спектроскопия ЯМР  
в исследовании природных углеводов**

<http://toukach.ru/rus/nmrglyco.htm>

# Методы ЯМР

ЯМР - основной метод работы с первичной структурой в растворе

## **Инновационные методы**

(пространственное строение белков, нерегулярные полимеры, ...)

отдельный специалист для каждого исследования

## **Стандартные методы**

(первичная структура и конформация, <5kDa)

- достаточная квалификация ЯМР-сервисов
- небогатые образцы
- доступные приборы

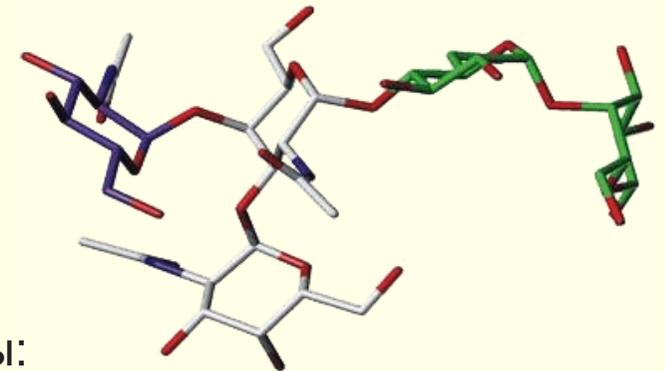
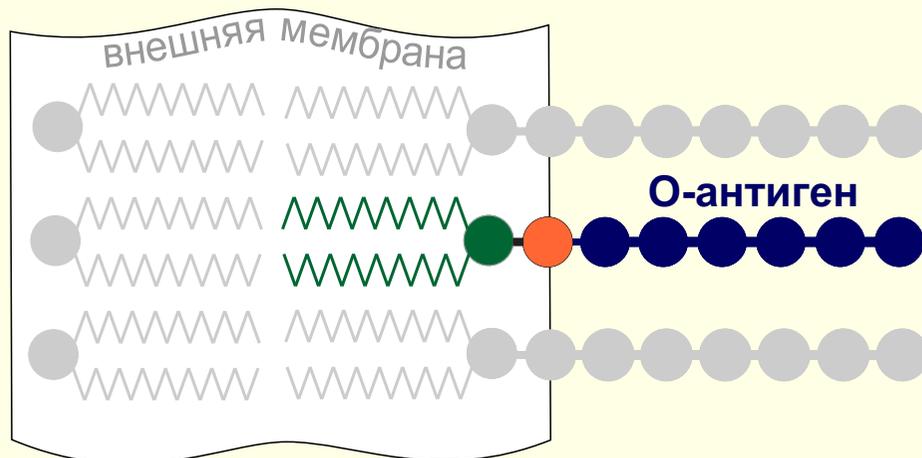
# Биомакромолекулы: ДНК/РНК, белки, полисахариды, гликоконъюгаты

В геномике и протеомике установление первичной структуры уже автоматизировано

В гликомике:

- больше химическая вариативность (*особенно для бактериальных углеводов*)
- установление строения не алгоритмизировано
- *первичная* структура в большей степени определяет третичную (биологические свойства)

Сахариды внешней мембраны бактерий являются антигенами ⇒

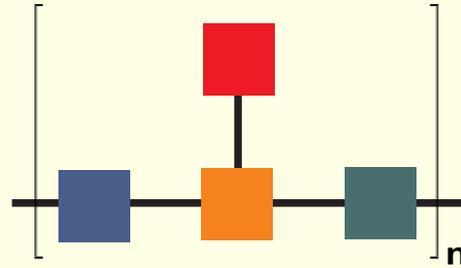


знание структуры:

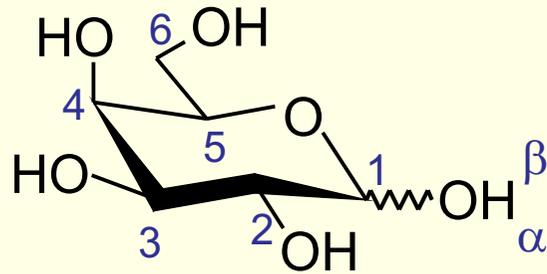
- молекулярное объяснение иммунного ответа
- разработка вакцин и иммуностимуляторов
- серотипирование

# Строение углеводов

## ПОЛНАЯ СТРУКТУРА

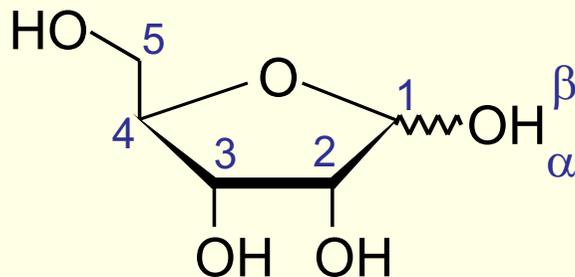


- строение всех остатков, включая неуглеводные (аминокислоты, алифатические кислоты и пр.)
- положения замещения
- последовательность соединения остатков
- стехиометрия остатков
- фосфатные и сульфатные группы
- число звеньев и “положение рамки”



### D-Gal

альдо-моносахарид  
в пиранозной форме (пример)



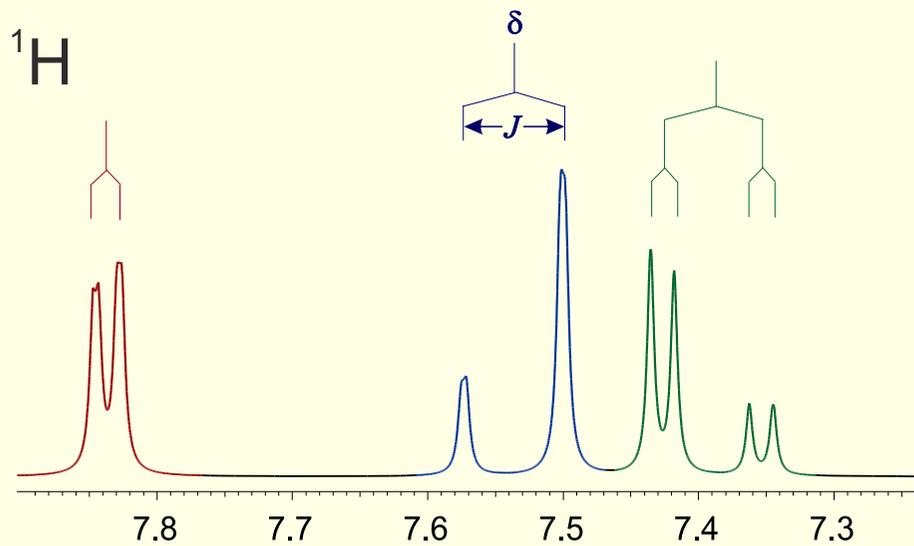
### D-Rib

альдо-моносахарид  
в фуранозной форме (пример)

## СТРОЕНИЕ ОСТАТКА

- размер углеродного скелета (5-9)
- размер цикла (пираноза, фураноза, линейный)
- пространственные положения всех -ОН (ax/eq)
- аномерная конфигурация ( $\alpha/\beta$ )
- абсолютные конфигурации (D/L)
- -H вместо -ОН (деокси-)
- -NH<sub>2</sub> вместо -ОН (амино-)
- -COOH вместо -CH<sub>2</sub>ОН (уроновые кислоты)
- другие функциональные группы

# Одномерный спектр ЯМР



**положение сигналов**

- распределение электронной плотности

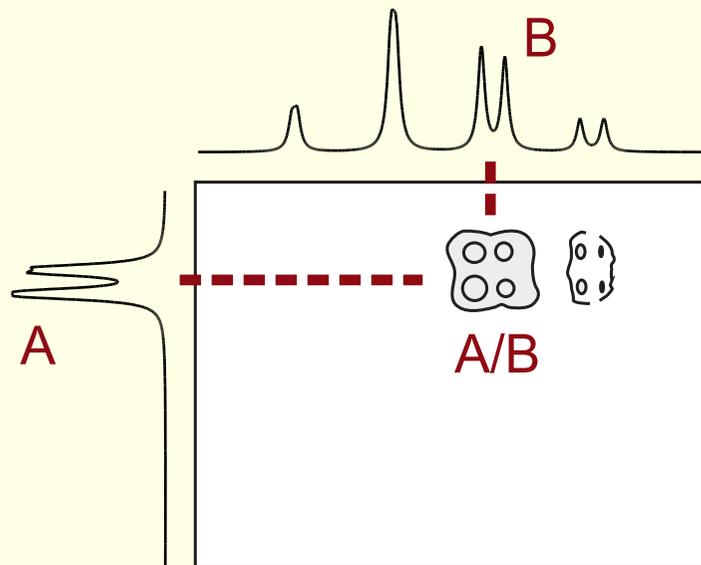
**форма сигналов**

- число и тип соседних атомов

**площадь сигналов**

- количество эквивалентных атомов

# Двумерный корреляционный спектр

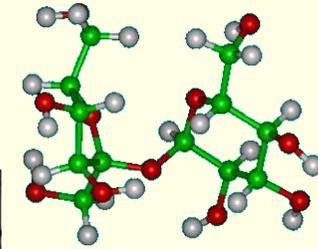
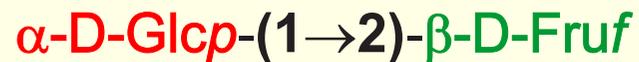
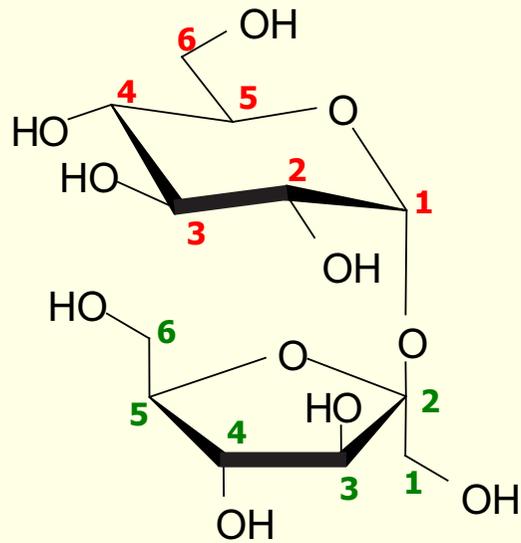


**кросс-пики**

- взаимодействие между сигналами

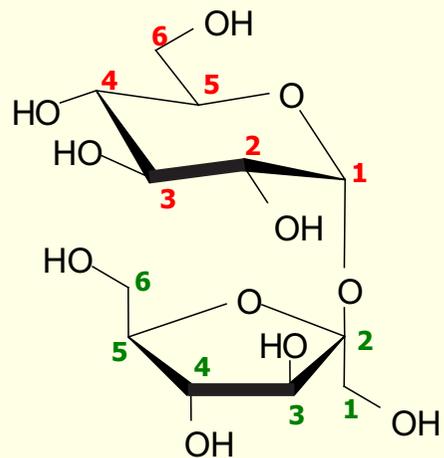
(спиновое - через связи,  
NOE - через пространство)

# Модельный объект



Дисахарид  $\alpha$ -D-Glcp-(1 $\rightarrow$ 2)- $\beta$ -D-Fruf (сахароза):

- содержит два (разных) остатка: глюкозу и фруктозу
- доступна в любой концентрации
- простые спектры (сигналы не перекрываются)



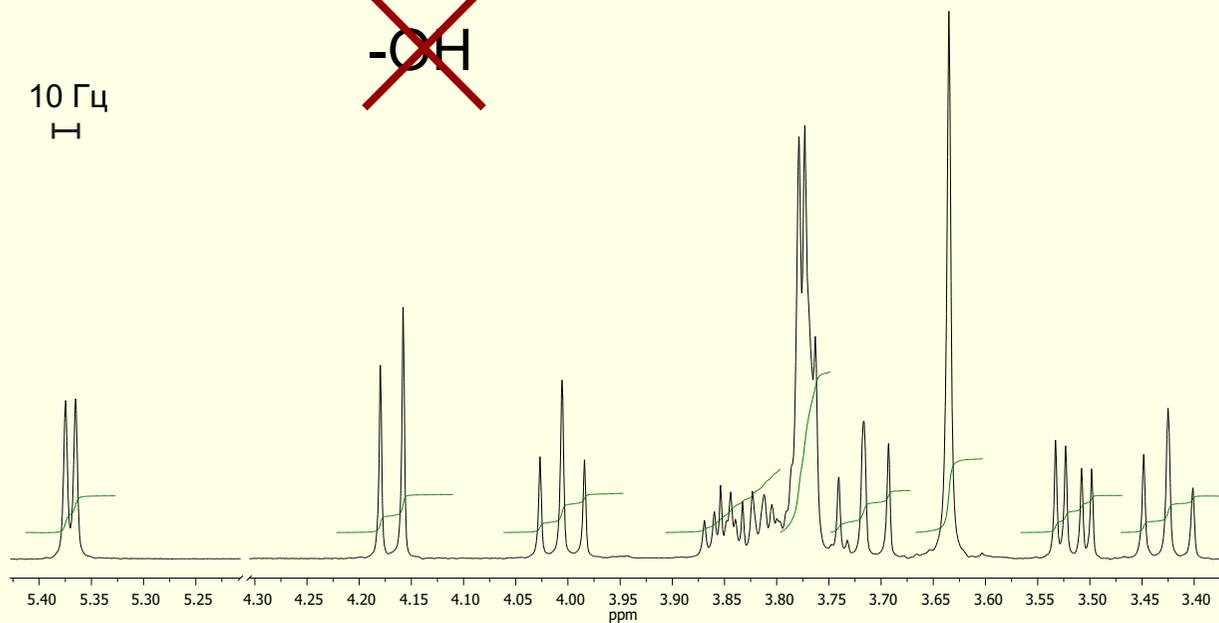
$\alpha$ -D-Glcp-(1 $\rightarrow$ 2)- $\beta$ -D-Fruf  
(сахароза)

**NMR  $^1\text{H}$**

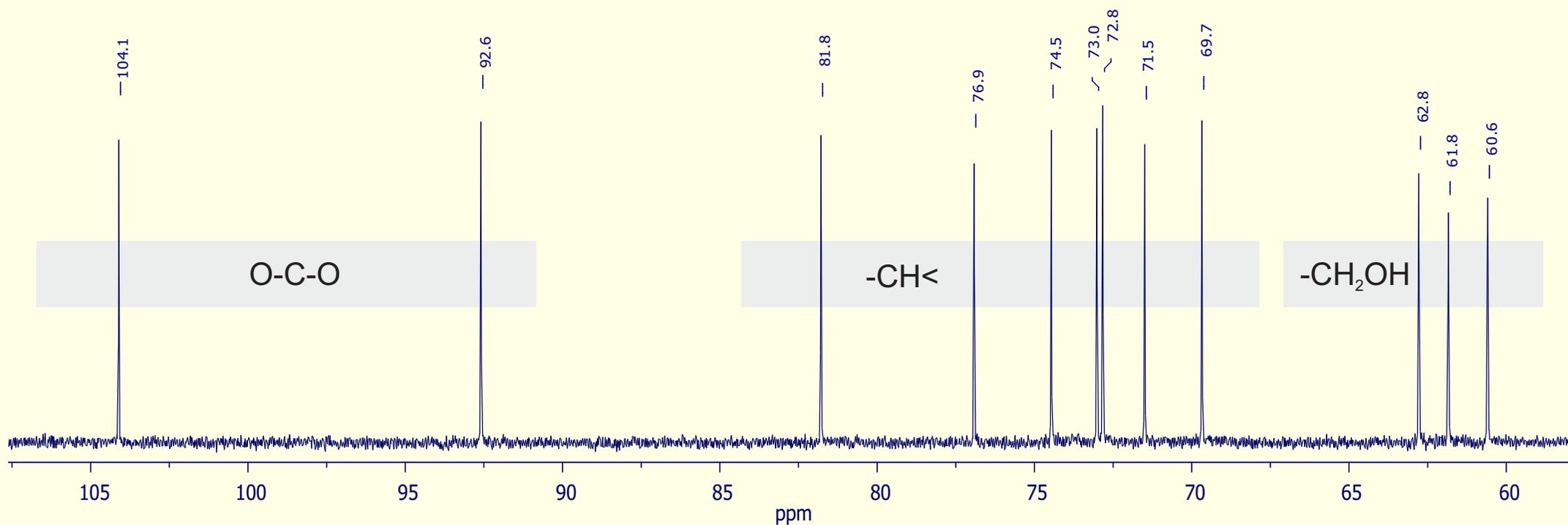


**1D 400MHz,  $\text{D}_2\text{O}$**

10 Гц  
H

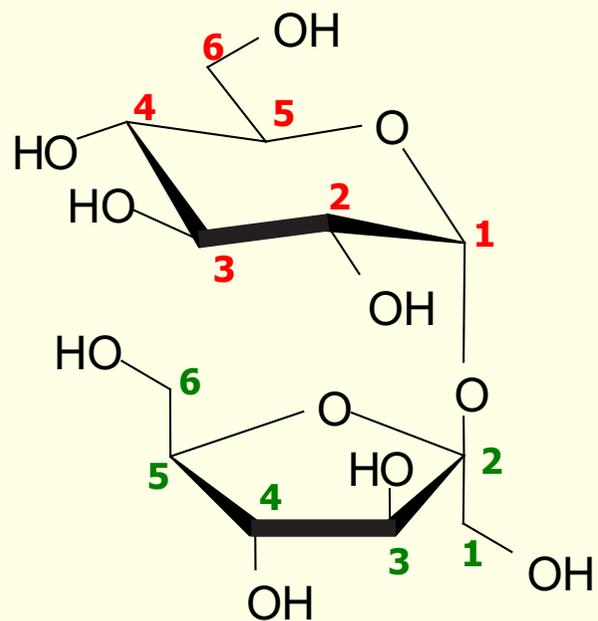


**NMR  $^{13}\text{C}$  BB**

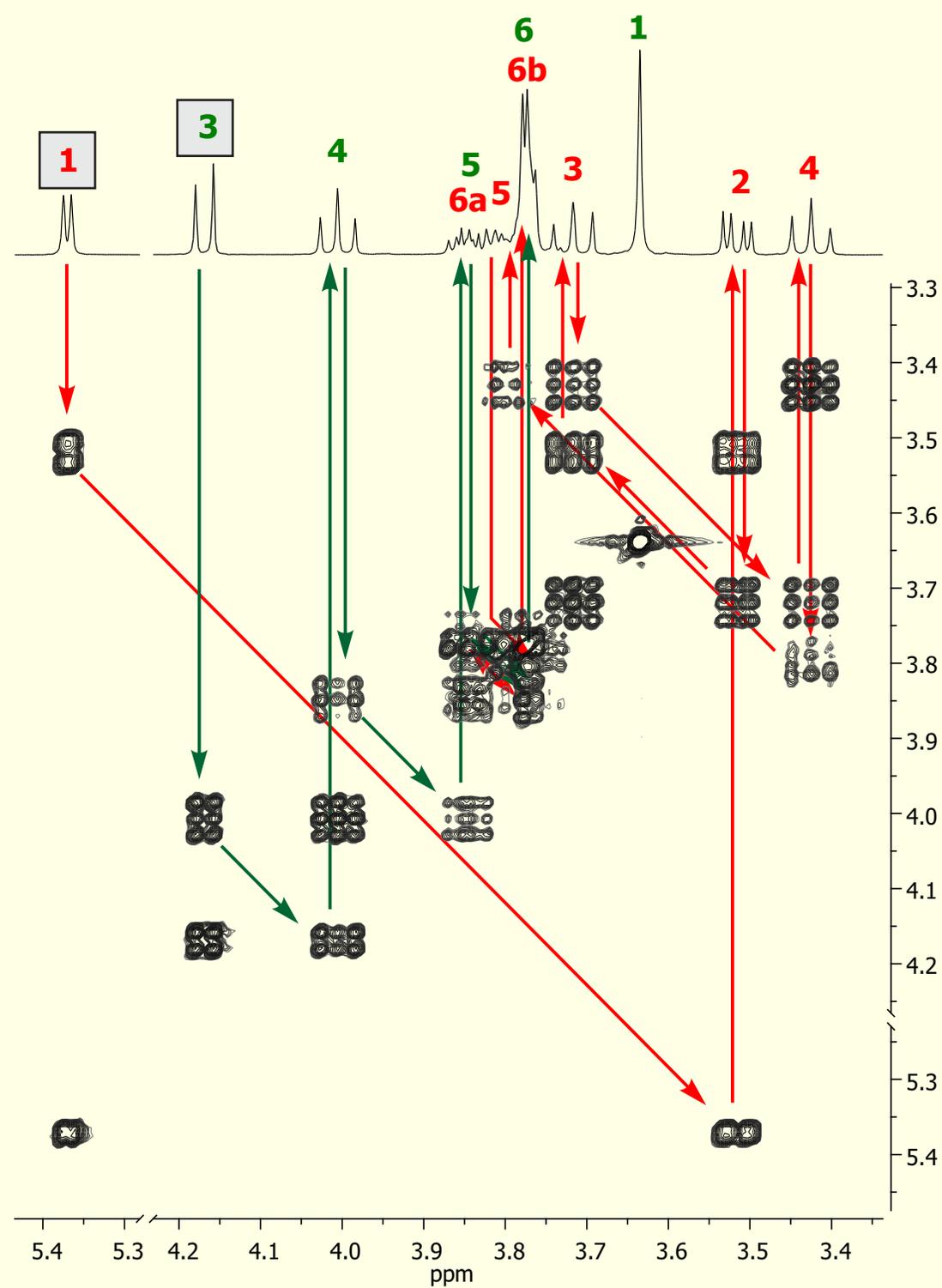
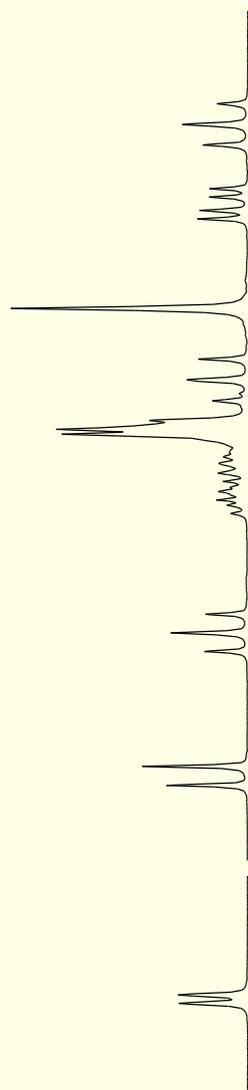


# {<sup>1</sup>H,<sup>1</sup>H} COSY

показывает корреляции протонов  
с вицинальными протонами-соседями

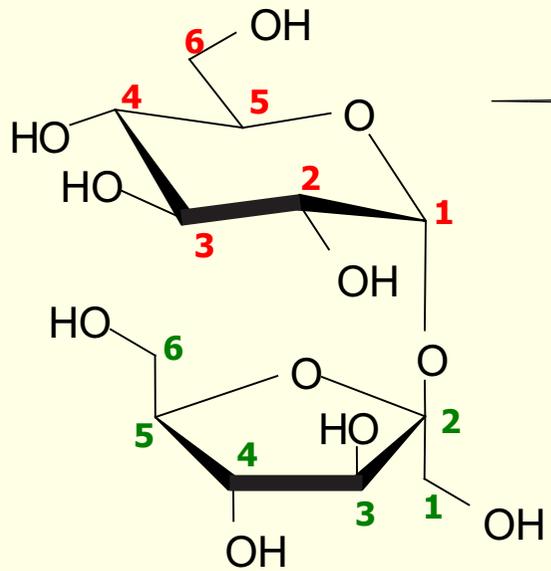


$\alpha$ -D-Glcp-(1→2)-β-D-Fruf  
(сахароза)

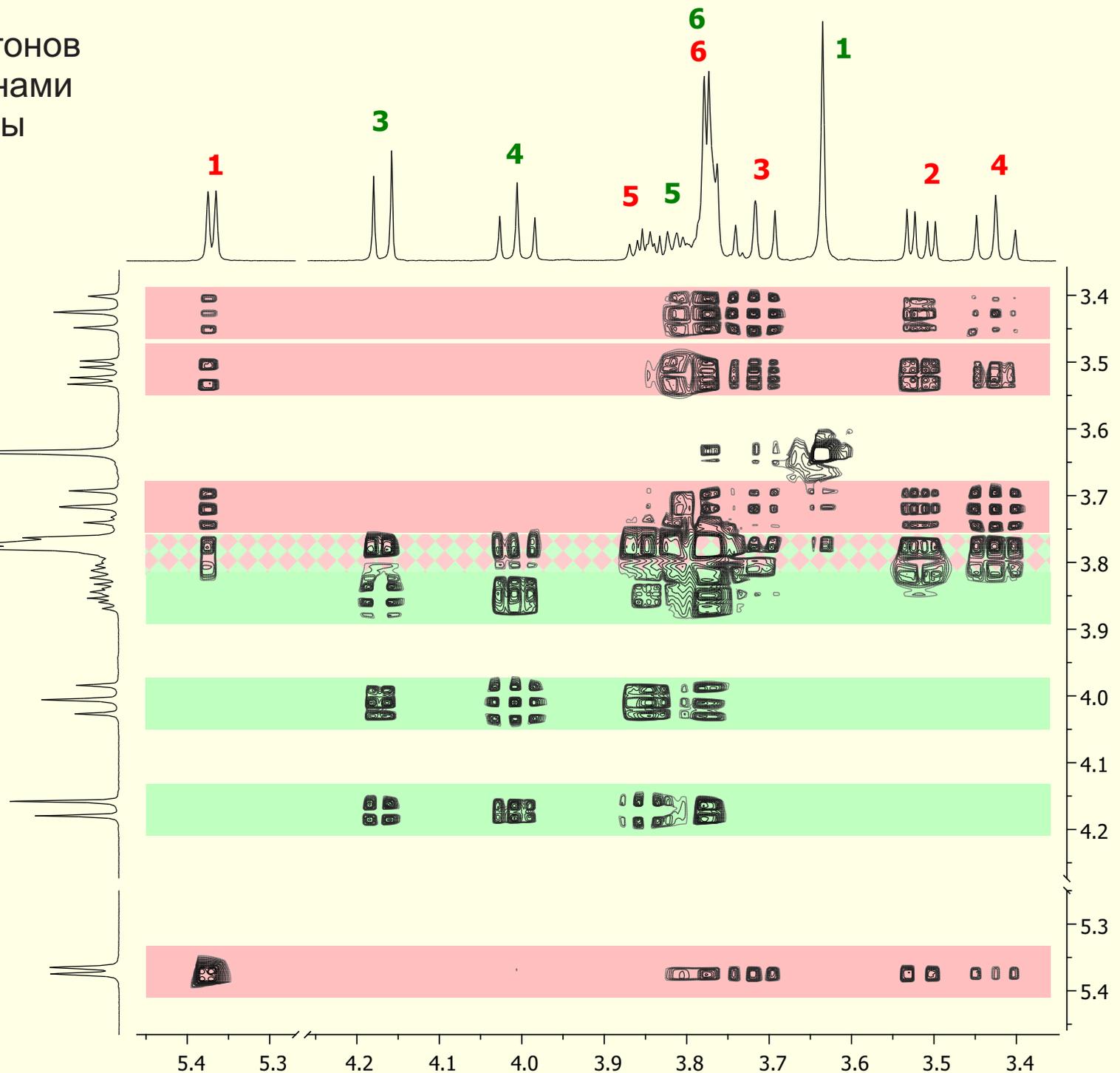


# TOCSY

показывает корреляции протонов со всеми остальными протонами в пределах спиновой системы

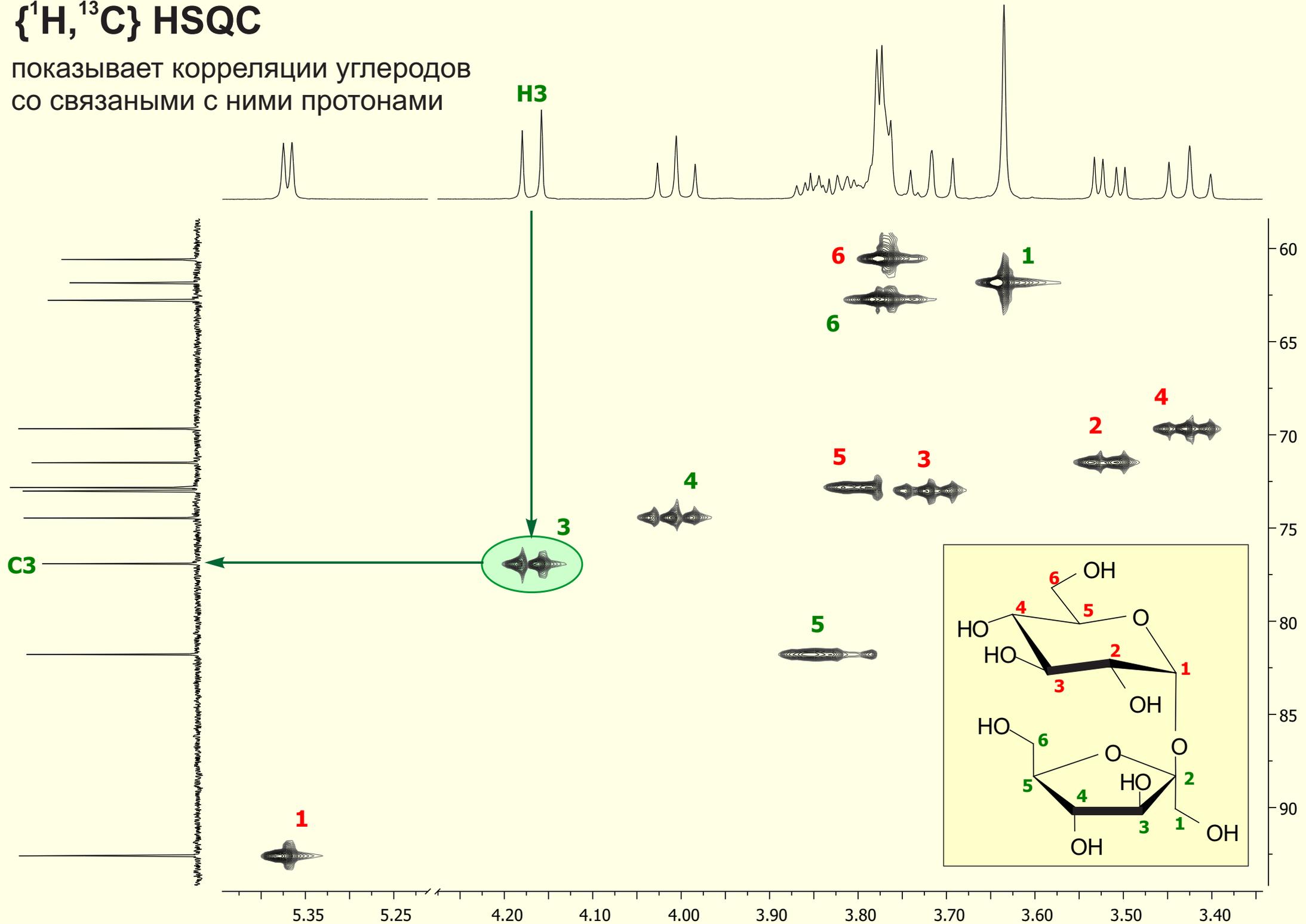


$\alpha$ -D-Glcp-(1 $\rightarrow$ 2)- $\beta$ -D-Fruf  
(сахароза)



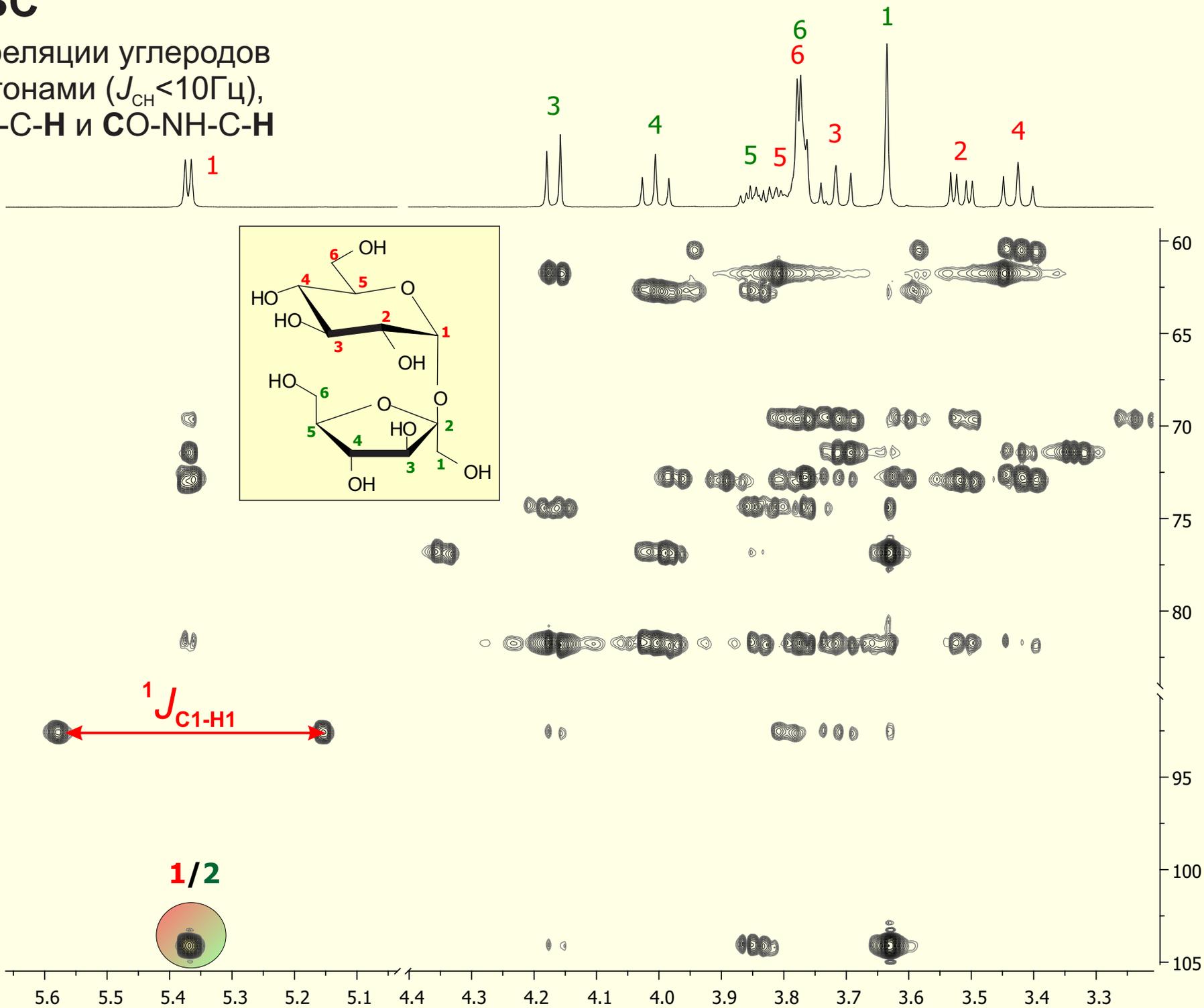
# {<sup>1</sup>H, <sup>13</sup>C} HSQC

показывает корреляции углеродов со связанными с ними протонами



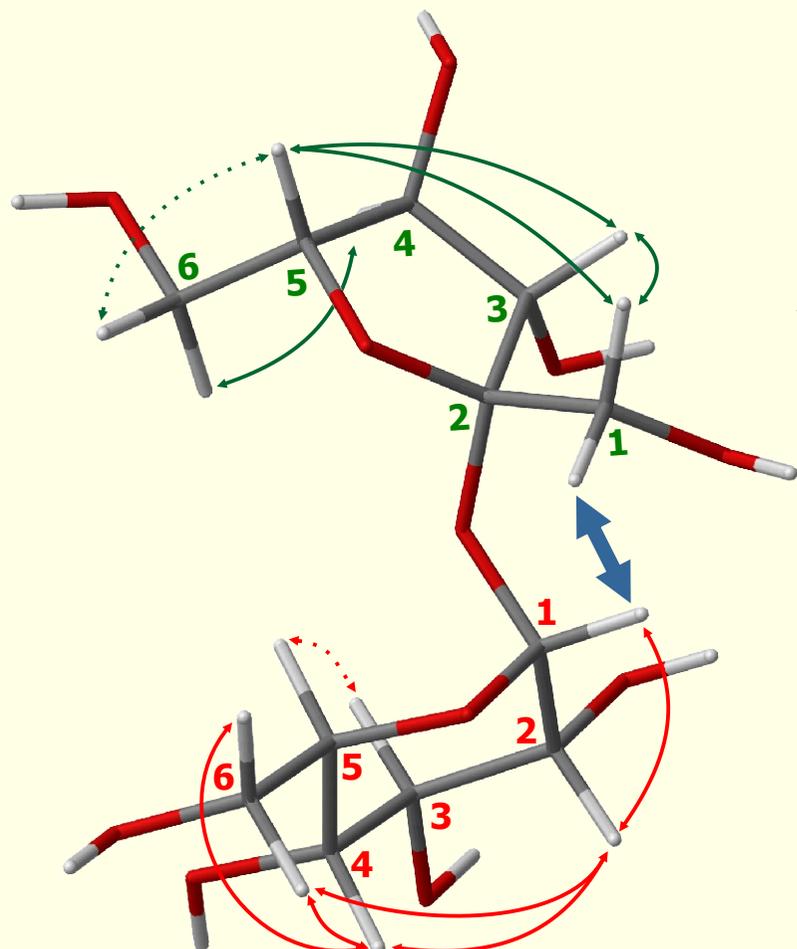
# {<sup>1</sup>H, <sup>13</sup>C} НМВС

показывает корреляции углеродов с дальними протонами ( $J_{CH} < 10\text{Гц}$ ), в том числе **C-O-C-H** и **CO-NH-C-H**

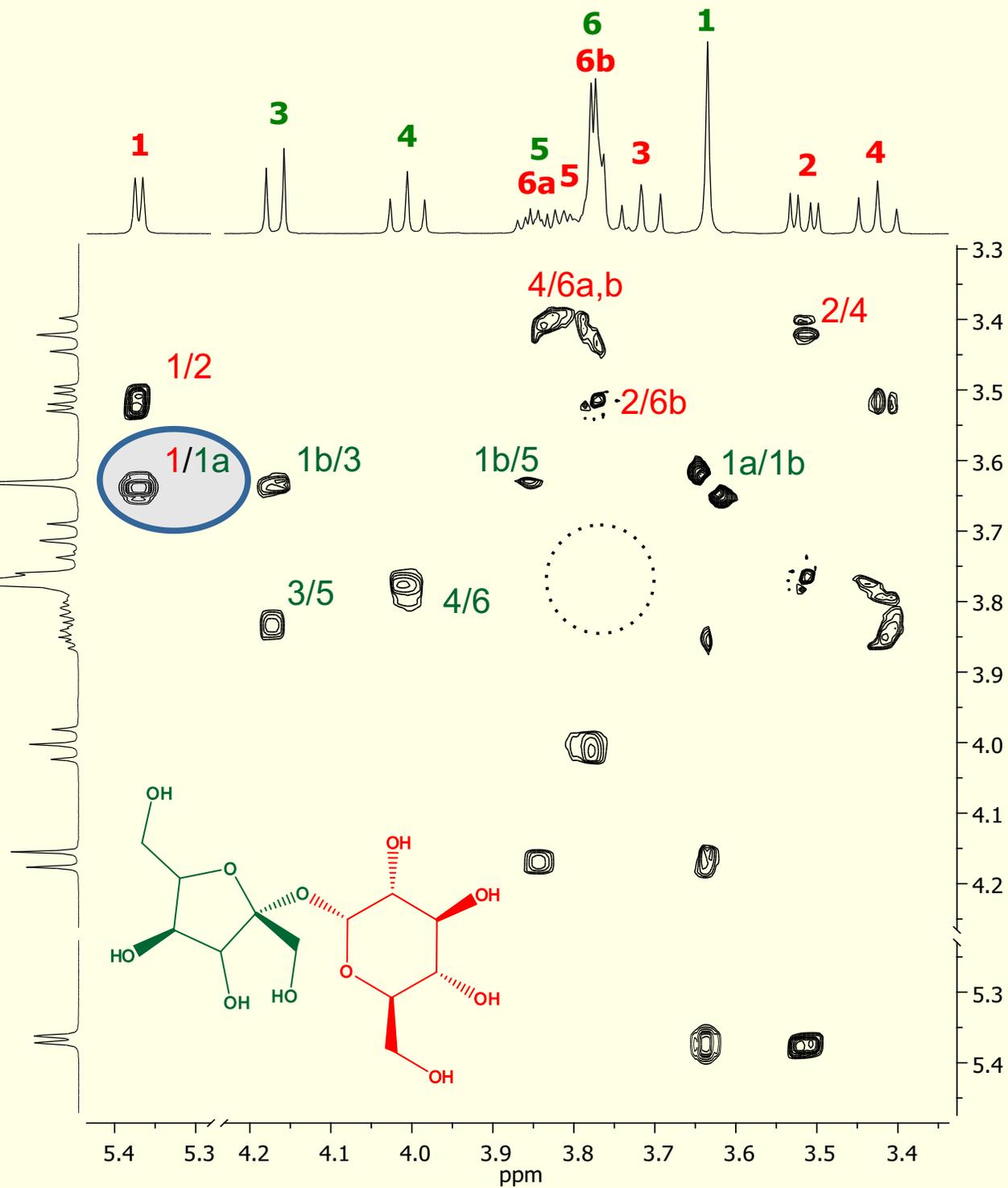


# NOESY

показывает пространственные контакты протонов (ядерные эффекты Оверхаузера)



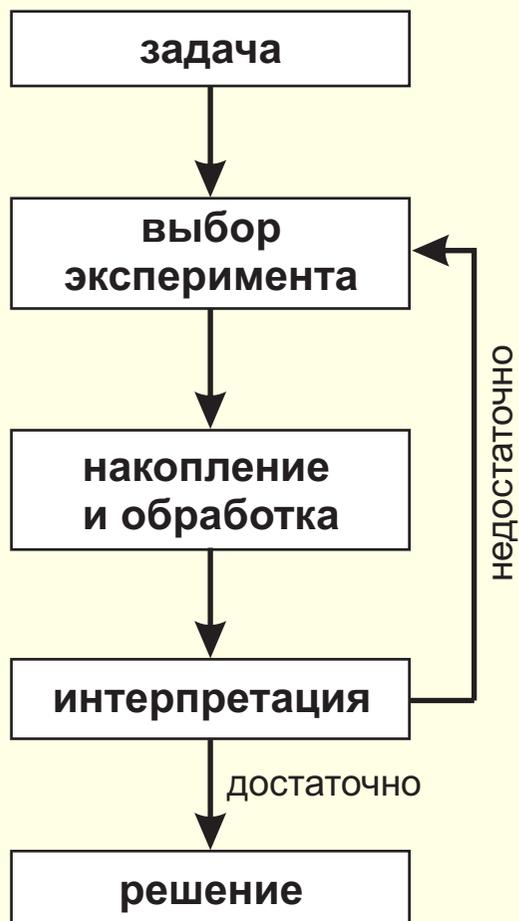
$\alpha$ -D-Glcp-(1 $\rightarrow$ 2)- $\beta$ -D-Fruf  
(сахароза)



# Некорреляционные эксперименты ЯМР

1D

2D



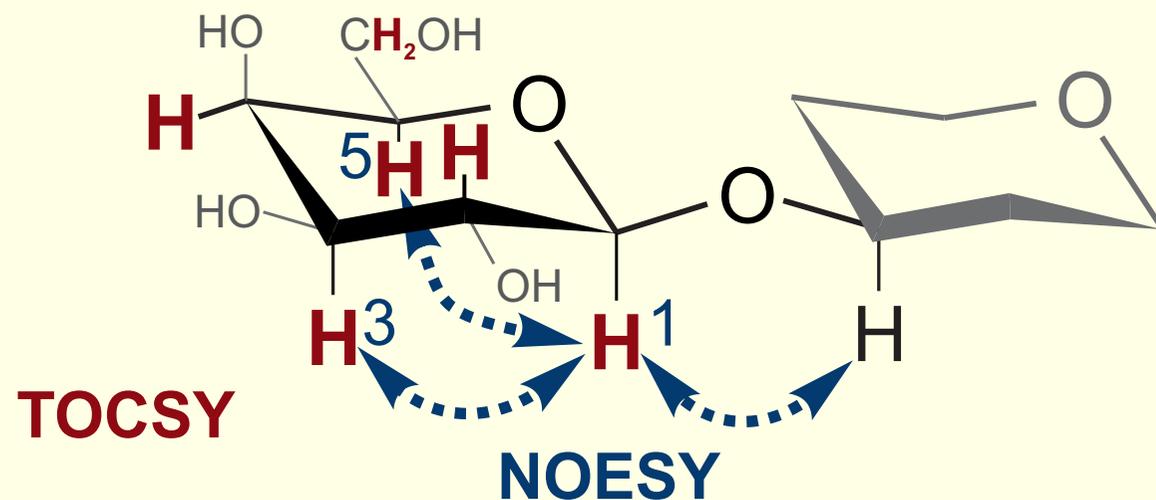
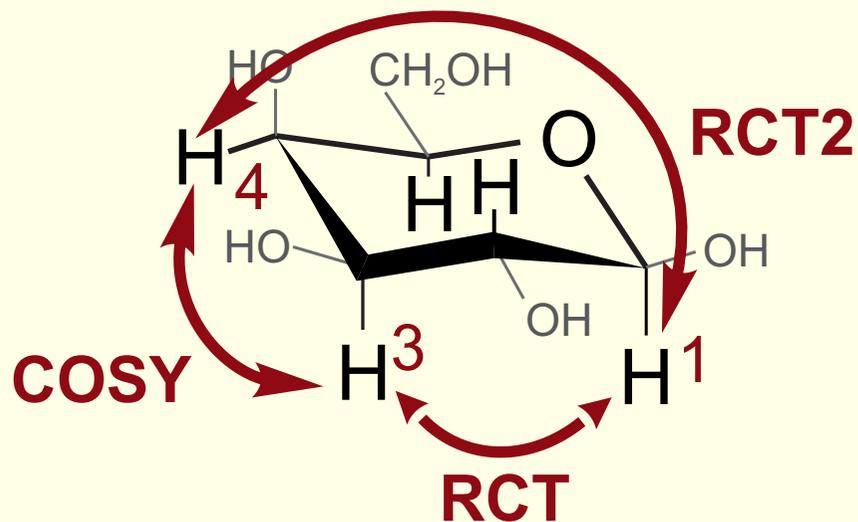
<b><math>^1\text{H}</math></b>	одномерный протонный спектр - <i>измерение гомоядерных КССВ, общая информация, идентификация остатков, базис для отнесения углеродного спектра</i>
<b><math>^{13}\text{C}</math> ВВ</b>	одномерный углеродный спектр с широкополосной развязкой от протонов - <i>детальная информация, идентификация остатков, позиции замещения</i>
<b><math>^{31}\text{P}</math> ВВ, <math>^{15}\text{N}</math> ВВ</b>	одномерные спектры с широкополосной развязкой от протонов - <i>дополнительная информация</i>
<b>APT, DEPT</b>	отредактированный углеродный спектр - <i>выявление <math>\text{CH}_2</math>-групп</i>
<b><math>^{13}\text{C}</math> Gated, <math>^{31}\text{P}</math> Gated</b>	одномерные спектры без развязки от протонов - <i>измерение гетероядерных КССВ, выяснение аномальной конфигурации, изучение конформации</i>
<b>HN J-res</b>	развертка протонного спектра по мультиплетности - <i>измерение гомоядерных КССВ, общая информация, идентификация остатков,</i>
<b>DOSY</b>	развертка протонного спектра по времени корреляции молекулы - <i>разделение спектра на подспектры компонентов смеси или подспектры фрагментов с разной подвижностью</i>



интерпретация спектров НЕ является последовательным алгоритмом

# Гомоядерные корреляции

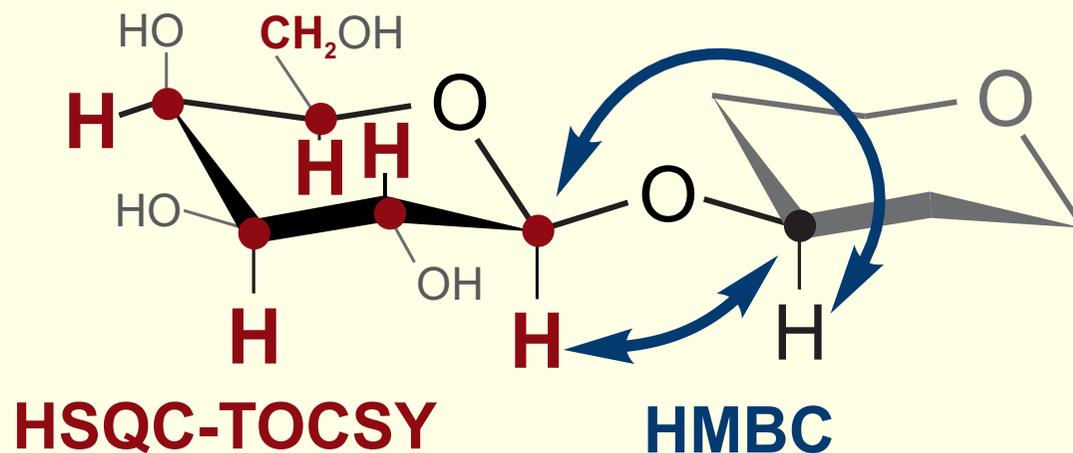
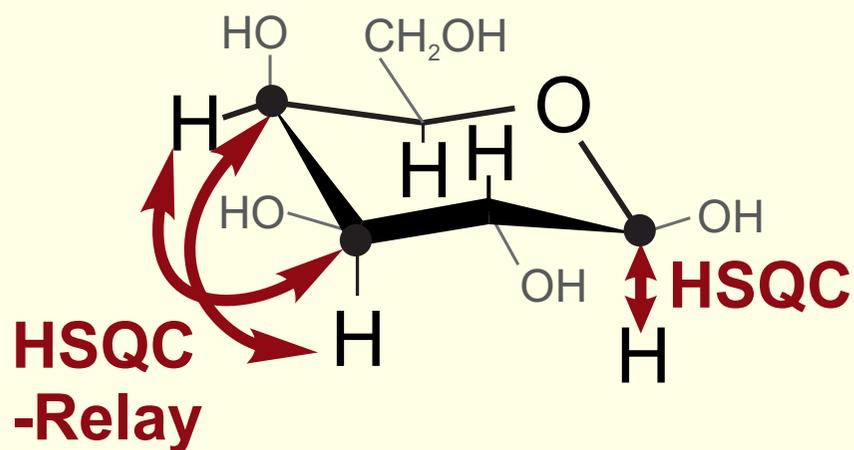
1D 2D



<b>COSY</b>	спиновая корреляция между вицинальными протонами - <i>отнесение протонного спектра</i>
<b>RCT, RCT2</b>	спиновые корреляции по цепочке вицинальных КССВ - <i>отнесение протонного спектра</i>
<b>DQF COSY</b>	COSY без диагональной линии - <i>отнесение близколежащих сигналов</i>
<b><sup>1</sup>H HD dif</b>	разностный селективный двойной резонанс - <i>выявление формы линии H2</i>
<b>TOCSY (нонана)</b>	спиновая корреляция со всеми протонами спиновой системы - <i>разделение спиновых систем</i>
<b>1D TOCSY</b>	TOCSY для одного сигнала - <i>выявление спиновой системы остатка</i>
<b>NOESY, ROESY</b>	пространственная корреляция - <i>выяснение последовательности остатков, конформации</i>
<b><sup>1</sup>H NOE dif</b>	разностное селективное измерение ЯЭО - <i>изучение пространственных контактов</i>

# Гетероядерные корреляции

1D 2D



$\{^1\text{H}, ^{13}\text{C}\}$ HSQC	прямая протон-углеродная спиновая корреляция - <i>отнесение углеродного спектра</i>
$\{^1\text{H}, ^{31}\text{P}\}$ HSQC	протон-фосфорная спиновая корреляция - <i>локализация фосфатных групп</i>
$\{^1\text{H}, ^{13}\text{C}\}$ HMBC	дальняя протон-углеродная корреляция - <i>выяснение последовательности остатков</i>
$\{^1\text{H}, \text{X}\}$ 1D HMBC	HMBC для одного сигнала - <i>выявление протонов в окружении гетероядра</i>
HSQC-Relay	углерод-углеродная корреляция через вицинальные протонные КССВ - <i>поиск соседних углеродных атомов</i>
HSQC-TOCSY	корреляция углеродов со всеми протонами спиновой системы (и наоборот) - <i>поиск C5 по H6 и аналогичные задачи</i>
$\{^1\text{H}, \text{X}\}$ 1D NOE	измерение гетероядерного ЯЭО - <i>исследование конформации</i>

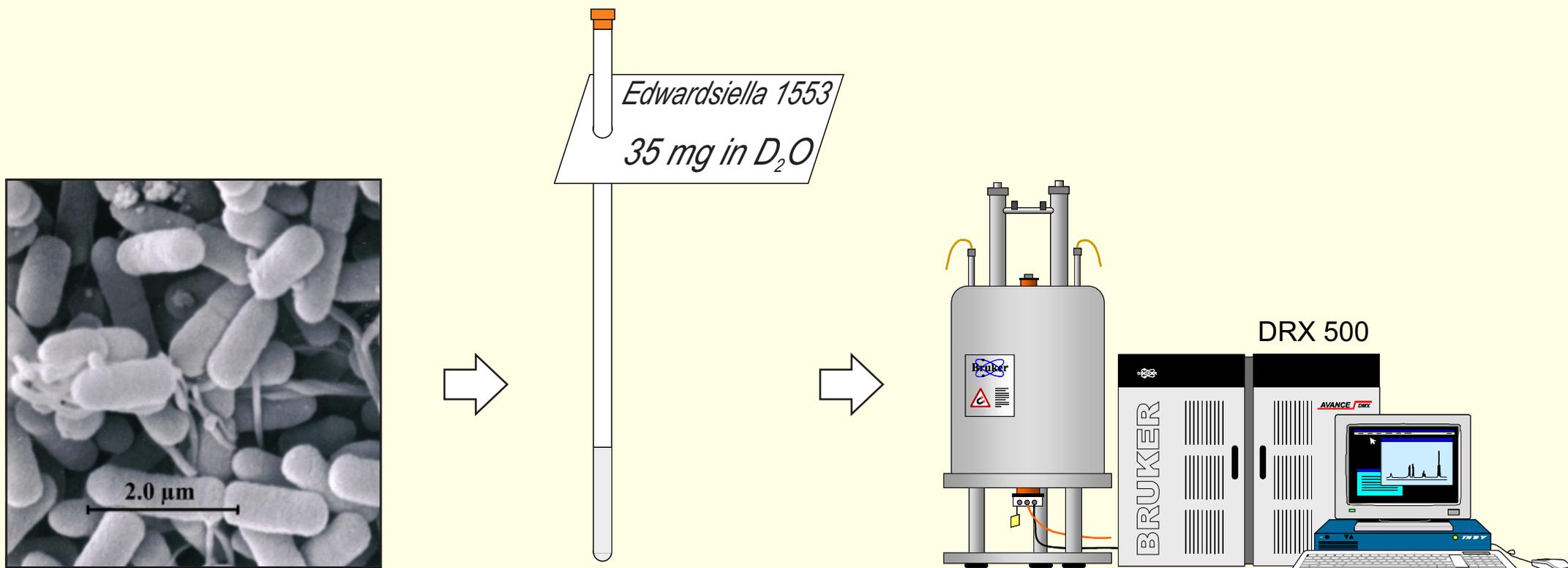
# Реальное исследование

## *Edwardsiella tarda* (штамм 1153)

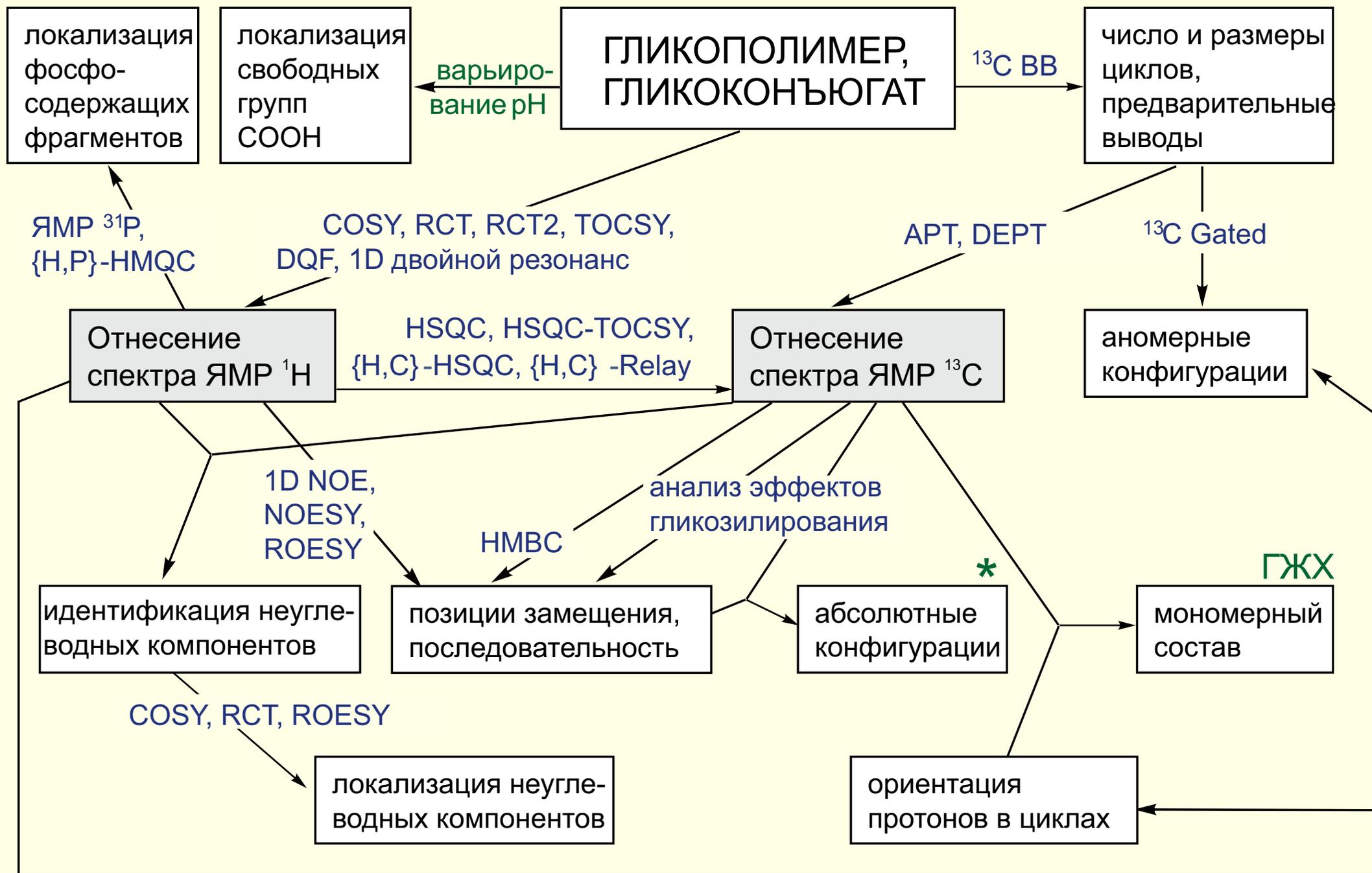
(энтеробактерия морских животных и рептилий + причина гастроэнтерита)

Из клеточной стенки *Edwardsiella* 1153 выделен O-антигенный полисахарид.  
Спектр  $^{13}\text{C}$  приобрел признаки регулярности после де-O-ацетилирования образца.

ГЖХ (Sugar analyzer): присутствуют как минимум **GlcN**, **Gal**, **GalA** в неизвестных пропорциях



# Эксперименты ⇒ информация



# NMR <sup>31</sup>P

нет сигналов



<http://csdb.glycoscience.ru>

# NMR <sup>13</sup>C ВВ

176.1  
175.0  
172.8

103.2  
102.0  
101.3  
97.0

83.5

79.0

78.6

76.3

75.7

73.0

72.9

72.5

70.5

70.0

69.5

68.0

67.7

62.7

62.0

61.6

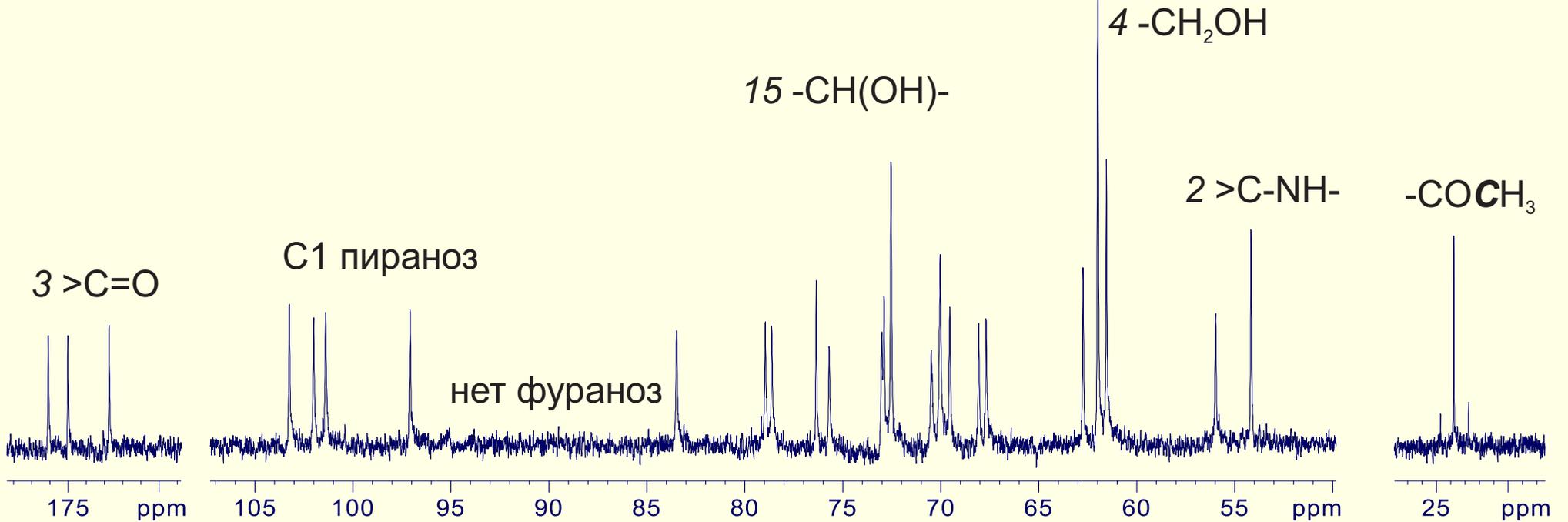
56.0

54.2

24.7

23.8

22.7

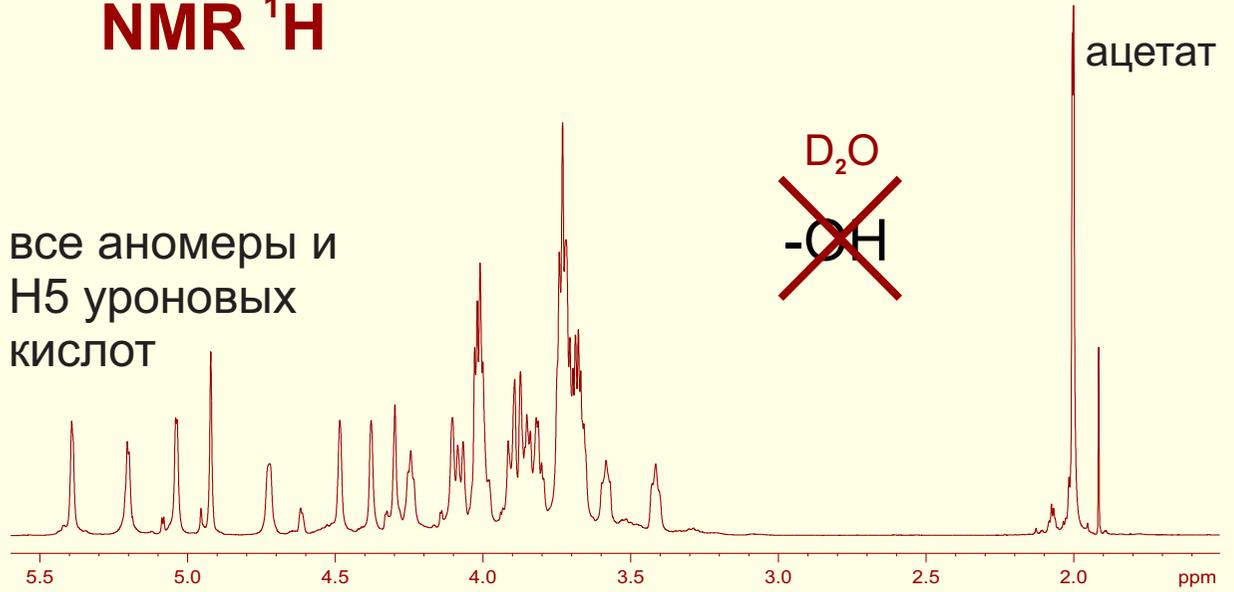


# NMR <sup>1</sup>H

все аномеры и  
H5 уроновых  
кислот



ацетат



# {<sup>1</sup>H,<sup>1</sup>H} COSY

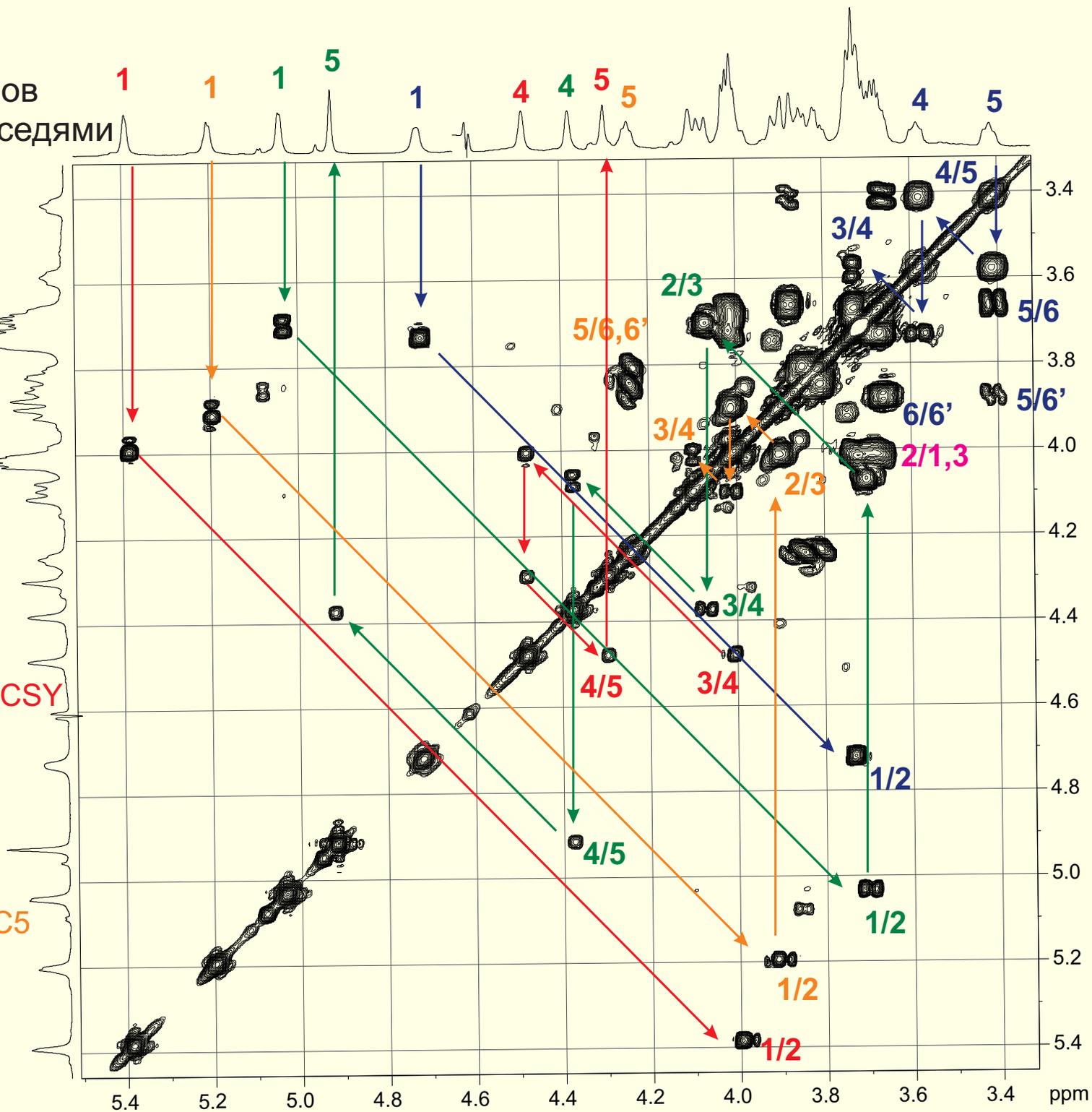
показывает корреляции протонов  
с вицинальными протонами-соседями

1-2 и 3-4-5-6,6' связаны в TOCSY  
5 выявлен, т.к. три кросс-пика  
форма 4 => 3,4,5 ах => это **GlcN**  
(подтверждение: C2 в HSQC)

1-2-3-4-5 связаны в COSY  
(подтверждение: TOCSY)  
корреляция C6/H5 в HMBC  
и положение 5 => это **GalA-амид**

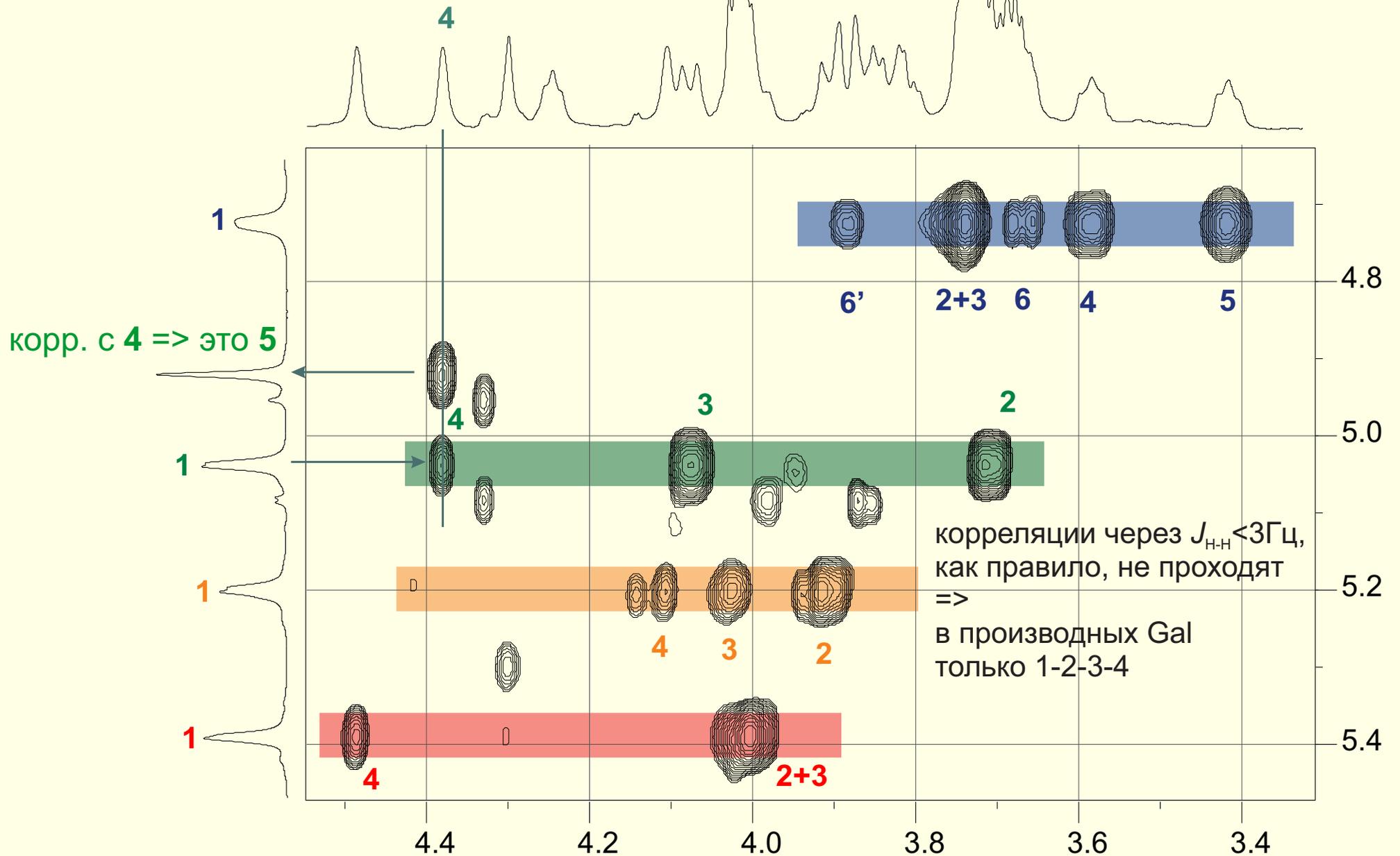
1-2 и 3-4-5 связаны через 1-4 в TOCSY  
(подтверждение: HSQC-TOCSY)  
корреляция C6/H5 в HMBC  
=> это **GalA**

1-2-3-4 и 5-6,6' связаны через H1-C5  
в HMBC и через H5/C5 в HSQC  
(подтверждение: 1/6 в ROESY  
при известной связи 1->4)  
+методом исключения => это **Gal**



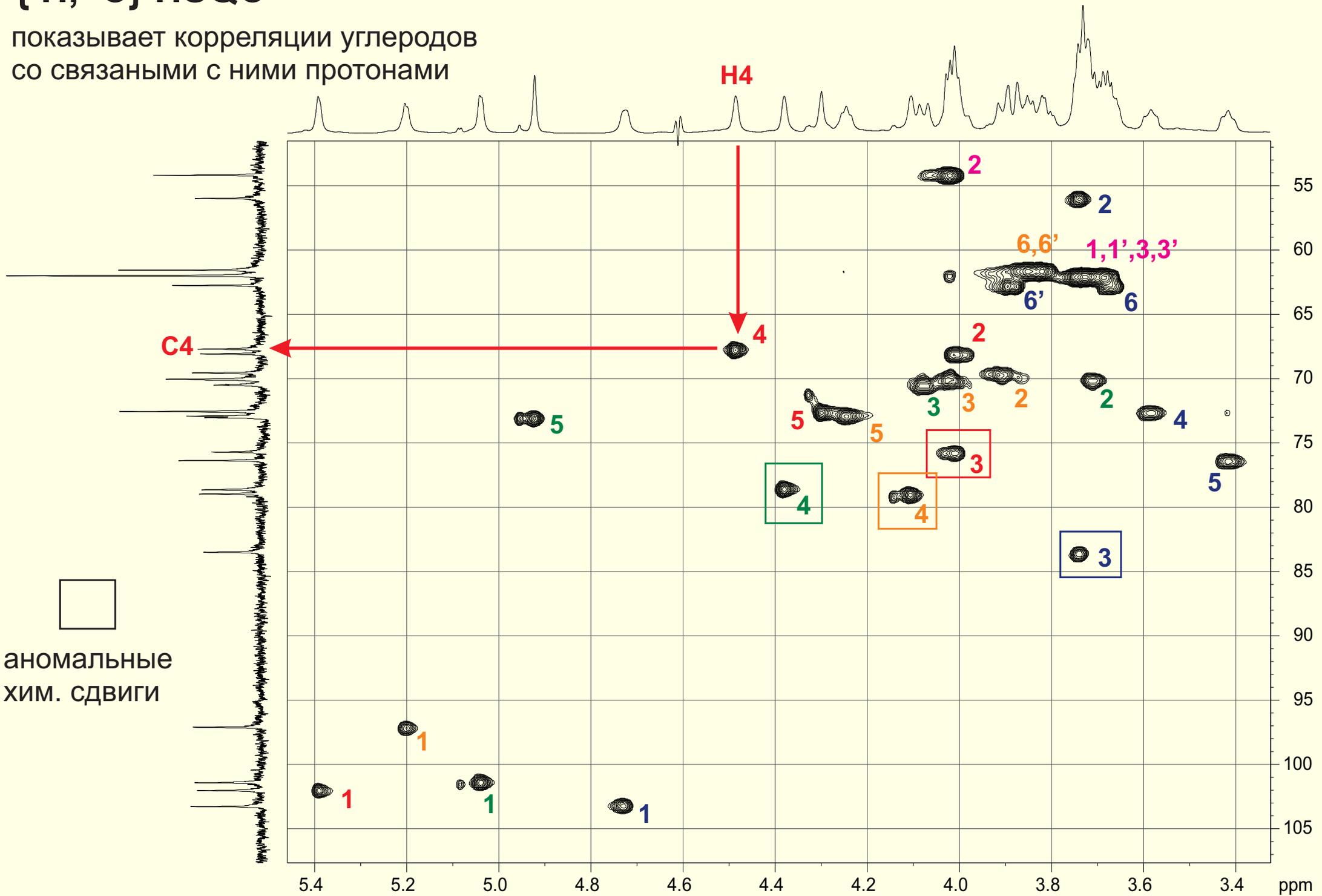
# TOCSY

показывает корреляции протонов со всеми остальными протонами в пределах спиновой системы



# $\{^1\text{H}, ^{13}\text{C}\}$ HSQC

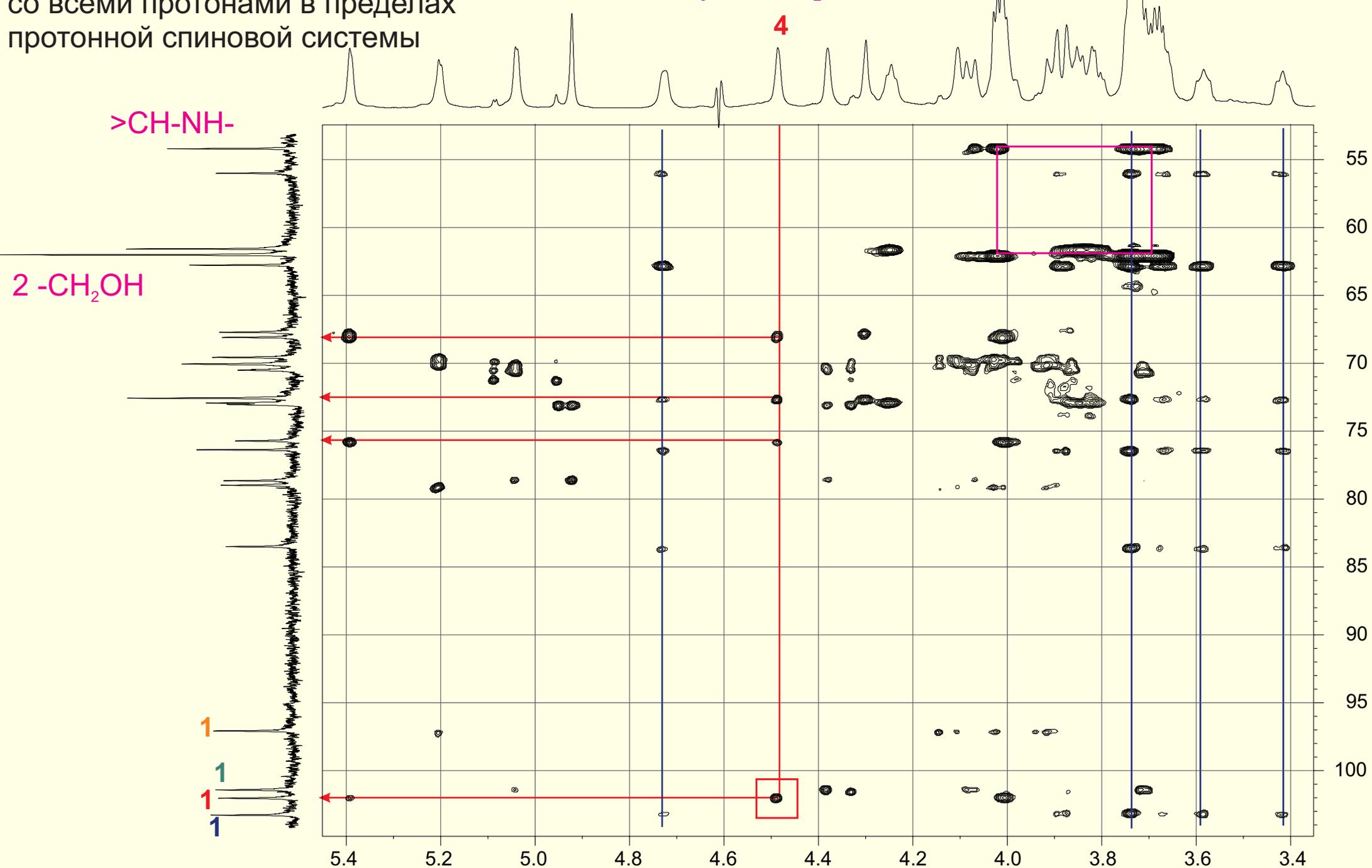
показывает корреляции углеродов  
со связанными с ними протонами



# {<sup>1</sup>H, <sup>13</sup>C} HSQC-TOCSY

показывает корреляции углеродов со всеми протонами в пределах протонной спиновой системы

3 оставшихся сигнала:  
>CH-NH- находится в одной спиновой системе с двумя -CH<sub>2</sub>OH => **GroN**



# ROESY

показывает пространственные контакты протонов (ядерные эффекты Оверхаузера)

**GlcN(1→4)GalA**

**GalA(1→3)GlcN**

**Gal(1→3 или 4)GalA**

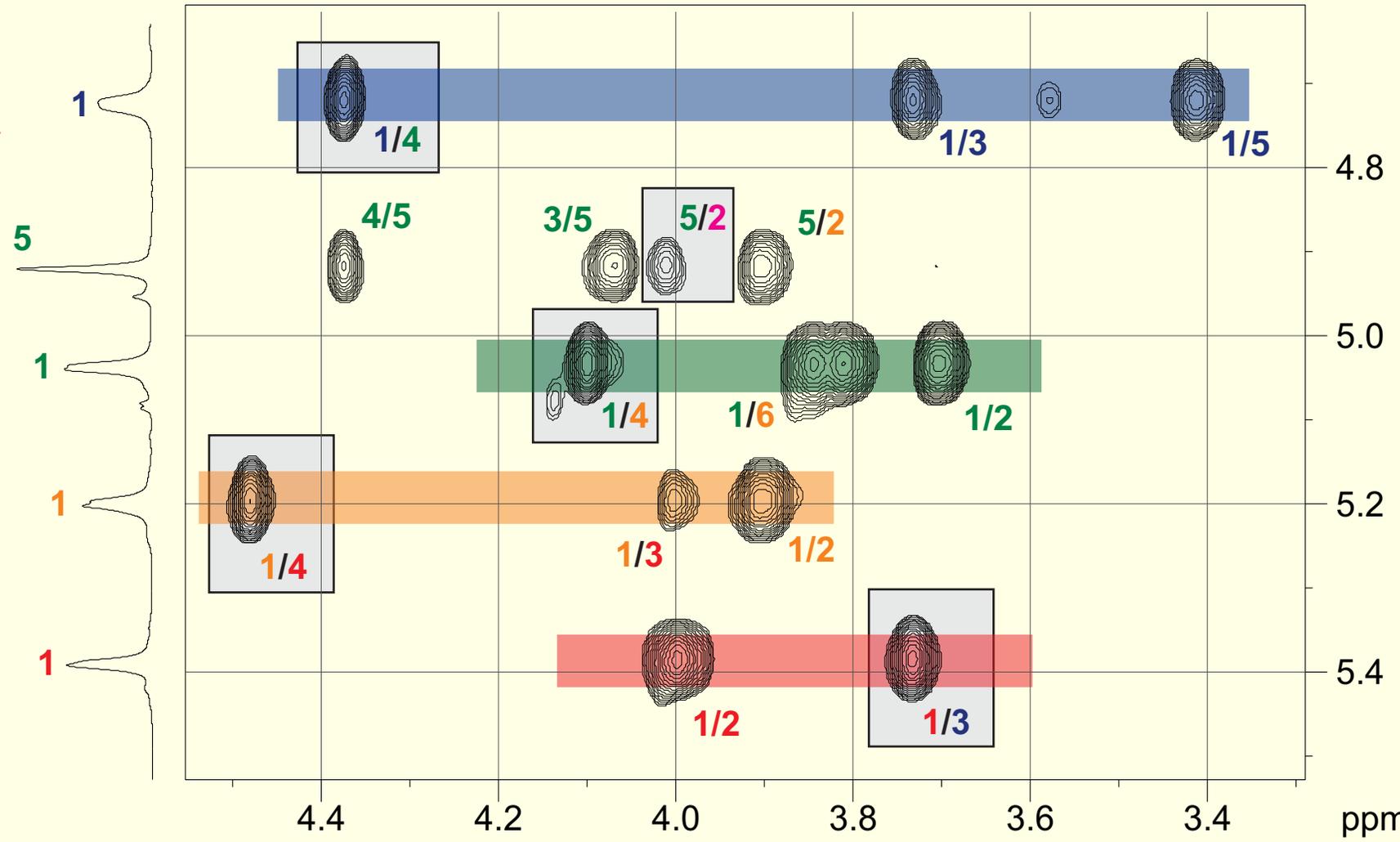
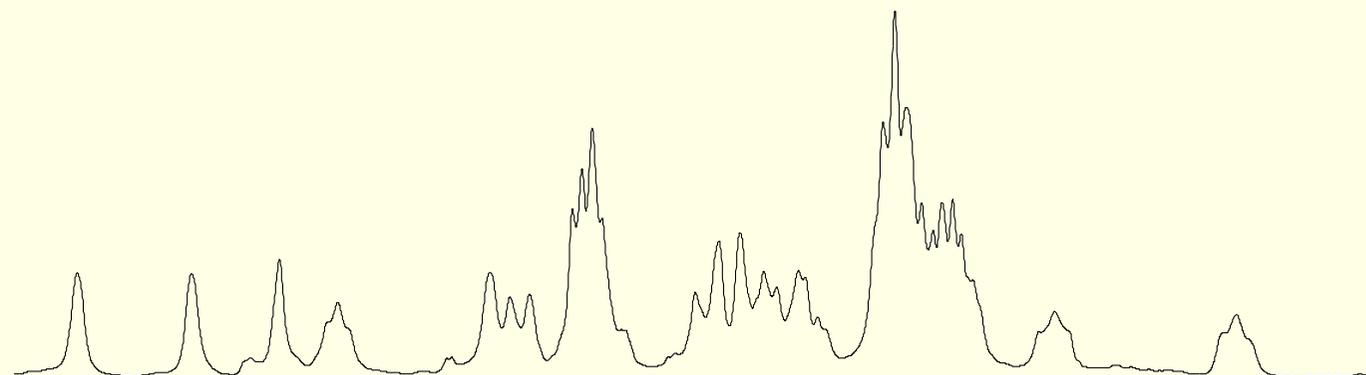
**GalA(1→4 или 6)Gal**

выбор на основании  
позиций замещения  
(HSQC);

подтверждение: HMBC



трансгликозидные  
контакты



# Позиции замещения и аномерные конфигурации

	C1	C2	C3	C4	C5	C6
→3)-GalpA <sup>I</sup> -(1→	102.0 <b>+7.2</b>	68.1	<b>75.8</b> <b>+5.4</b>	67.7 -4.0	72.6	175.0
→3)-GlcN-(1→	103.3 <b>+7.1</b>	56.0 -2.0	<b>83.5</b> <b>+8.4</b>	72.6 +1.4	76.4	62.8
Ac-(1→2)	176.1	23.8				
→4)-GalpA <sup>II</sup> -(1→	101.4 <b>+6.6</b>	70.1	70.5	<b>78.7</b> <b>+7.0</b>	73.0 +0.8	<b>172.8</b> -2.2
GroN-(2→6)	62.0	54.2	62.0			
→4)-Galp-(1→	97.1 <b>+3.6</b>	69.6	70.1	<b>79.0</b> <b>+8.4</b>	72.9 +1.2	61.6

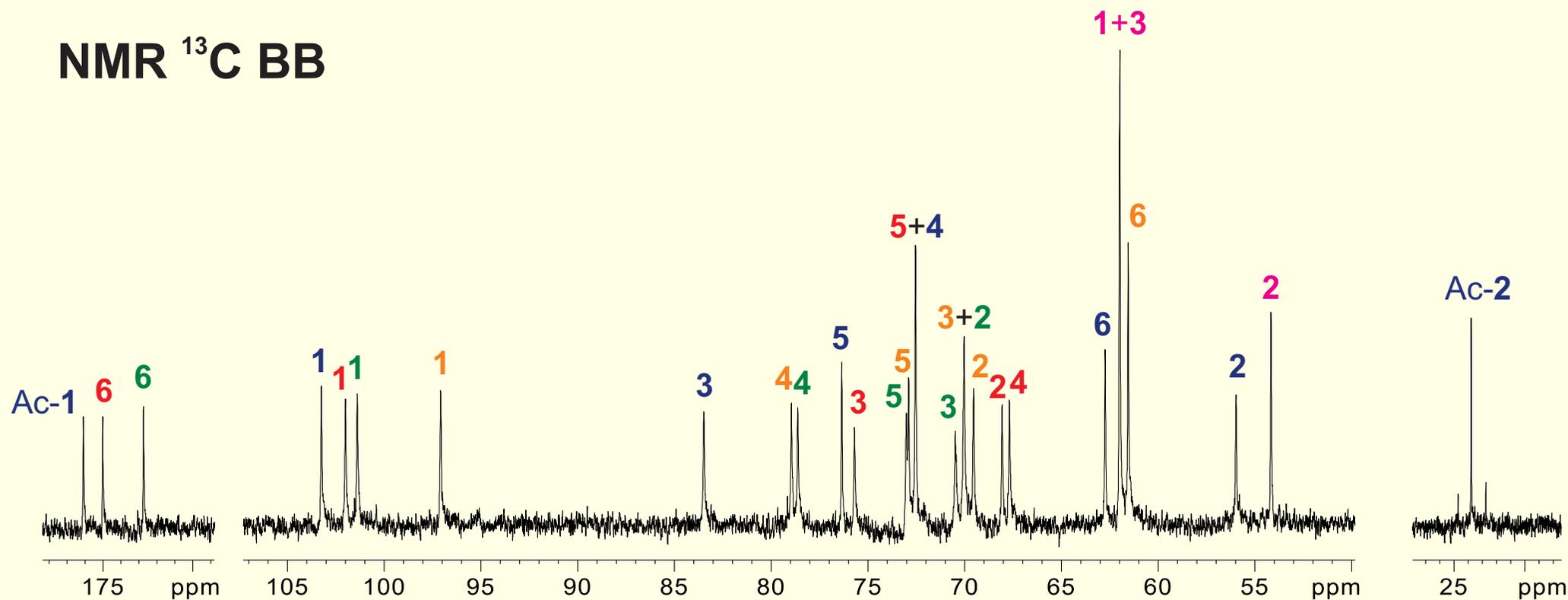
GlcN

$\delta C5 > 76 \Rightarrow \beta$ -аномер  
(подтверждение: форма H1)

остальные:

$\delta C5 < 73 \Rightarrow \alpha$ -аномеры  
(подтверждения: форма H1)

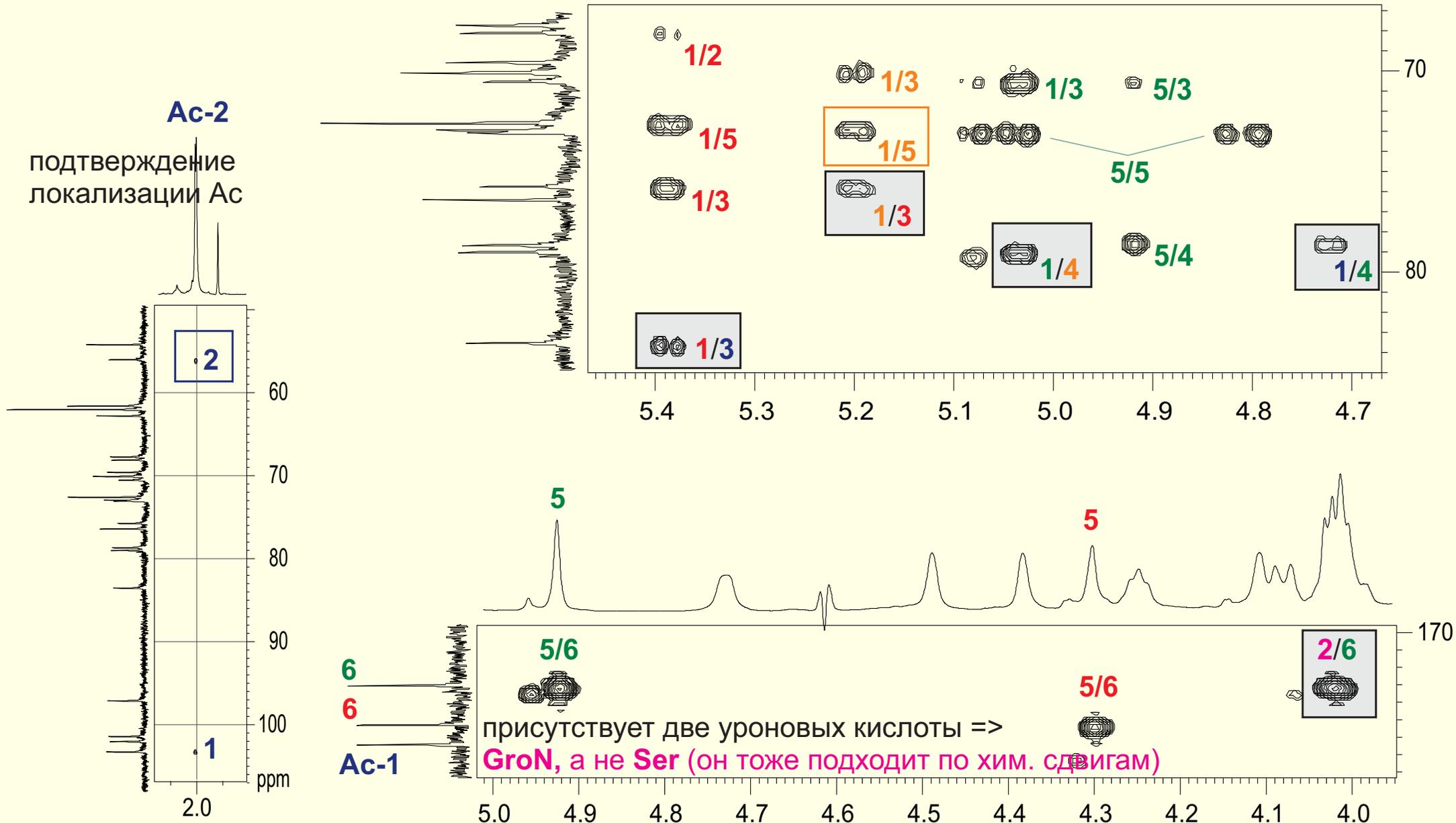
## NMR <sup>13</sup>C BB



# {<sup>1</sup>H, <sup>13</sup>C} НМВС

показывает корреляции углеродов с дальними протонами ( $J_{CH} < 10$  Гц), в том числе **C-O-C-H** и **CO-NH-C-H**

□ трансликозидные контакты



# Абсолютные конфигурации

эффекты замещения чувствительны к абсолютным конфигурациям и сведены в базы данных

	C1	C2	C3	C4	C5	C6
→3)-α-GalpA <sup>I</sup> -(1→	102.0 +7.2	68.1	<b>75.8</b> +5.4	67.7 -4.0	72.6	175.0
→3)-β-GlcpN-(1→	103.3 +7.1	56.0 -2.0	<b>83.5</b> +8.4	72.6 <b>+1.4</b>	76.4	62.8
Ac-(1→2)	176.1	23.8				
→4)-α-GalpA <sup>II</sup> -(1→	101.4 +6.6	70.1	70.5	<b>78.7</b> +7.0	73.0 +0.8	172.8
GroN-(2→6)	62.0	54.2	62.0			
→4)-α-Galp-(1→	97.1 <b>+3.6</b>	69.6	70.1	<b>79.0</b> +8.4	72.9 +1.2	61.6

обе GalA - D

(из величины оптического вращения их S-бутилгликозидов)

пара остатков	атом	теория		эксперимент
		DD	DL	
Gal <sup>α</sup> 1→3GalA <sup>I</sup>	C-1	<b>3.3</b>	8.3	3.6
GalA <sup>I</sup> β1→3GlcN	C-4	<b>0.7</b>	-1.3	1.4



все оптически-активные остатки имеют D-конфигурацию

# Откуда брать хим. сдвиги $^{13}\text{C}$ ?



## CSDB:

экспериментальные данные ЯМР



## GODDESS / GRASS:

химические сдвиги мономеров, димеров (линейные участки), и тримеров (точки разветвления).

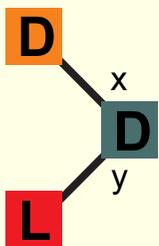
**эффекты замещения**

<http://csdb.glycoscience.ru/help/nmr.html>

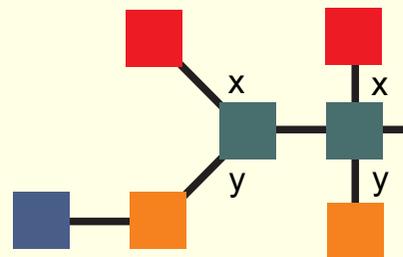
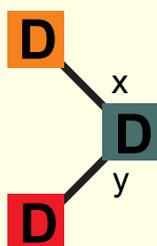


	C1	C2	C3	C4	C5	C6
$\rightarrow 3)\text{-}\alpha\text{-D-GalpA}^1\text{-(1}\rightarrow$	102.0 +7.2	68.1	<b>75.8</b> +5.4	67.7 -4.0	72.6	175.0
$\rightarrow 3)\text{-}\beta\text{-D-GlcpN-(1}\rightarrow$	103.3 +7.1	56.0 -2.0	<b>83.5</b> +8.4	72.6 <b>+1.4</b>	76.4	62.8
Ac-(1 $\rightarrow$ 2)	176.1	23.8				
$\rightarrow 4)\text{-}\alpha\text{-D-GalpA}^1\text{-(1}\rightarrow$	101.4 +6.6	70.1	70.5	<b>78.7</b> +7.0	73.0 +0.8	172.8
GroN-(2 $\rightarrow$ 6)	62.0	54.2	62.0			
$\rightarrow 4)\text{-}\alpha\text{-D-Galp-(1}\rightarrow$	97.1 <b>+3.6</b>	69.6	70.1	<b>79.0</b> +8.4	72.9 +1.2	61.6

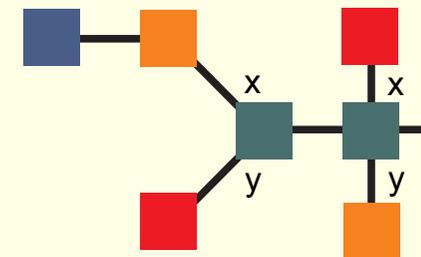
Примеры необходимости расчета по инкрементной схеме:



vs.



vs.



# Методы ЯМР-моделирования

## Статистические

*химические сдвиги  $^{13}\text{C}$  и  $^1\text{H}$*

- ☺ Существует база данных (CSDB)
- ☺ Прослеживается до публикаций
- ☹ Медленно (~ 1 мин)

параметризовано для углеводов:  
GODDESS

## Эмпирические

*химические сдвиги, форма сигналов*

- ☺ Очень быстро (~ 0.01 сек)
- ☹ Требуется модель
- ☹ Требуются специальные БД

параметризовано для углеводов:  
GODDESS, BIOPSEL, CASPER

## Квантовомеханические

*геометрия + все параметры ЯМР*

- ☺ Не зависят от баз данных
- ☹ Низкая точность (>3 м.д.)
- ☹ Очень медленно (~ 1 неделя)

не параметрические:  
GAUSSIAN, PRIRODA

## Прочие

*(нейронные сети, регрессия, ММ-QM)*

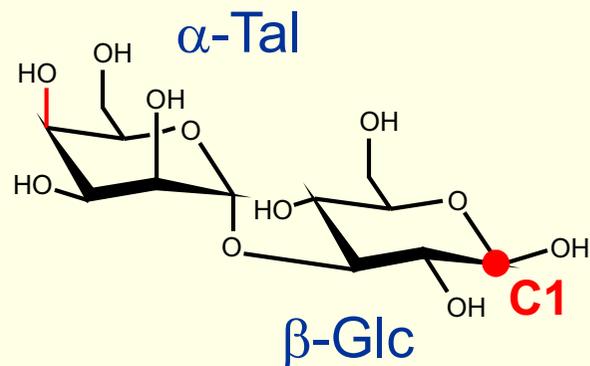
- ☹ Нет преимуществ в глико-ЯМР

НЕ параметризовано для углеводов

# Статистическая модель ЯМР

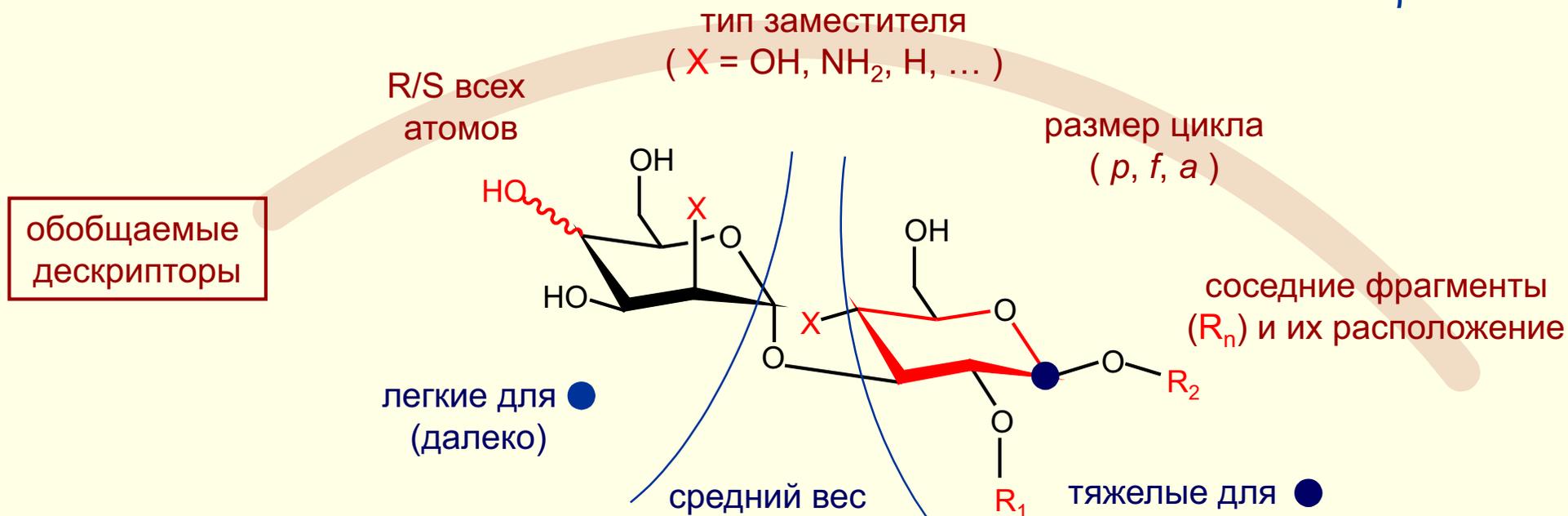
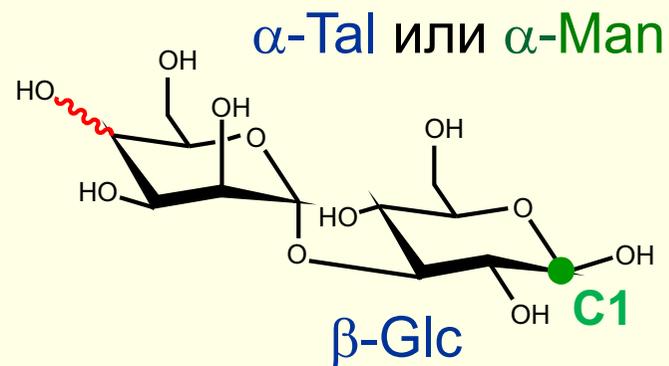
редкий фрагмент структуры

обобщенные фрагменты, в т. ч. распространенные



пример обобщения

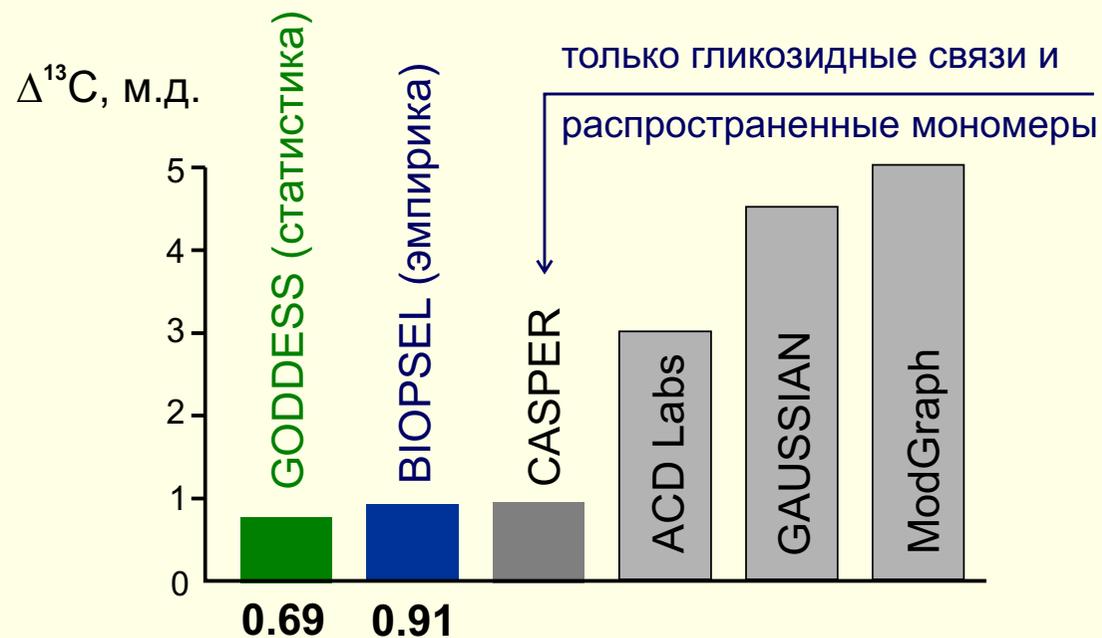
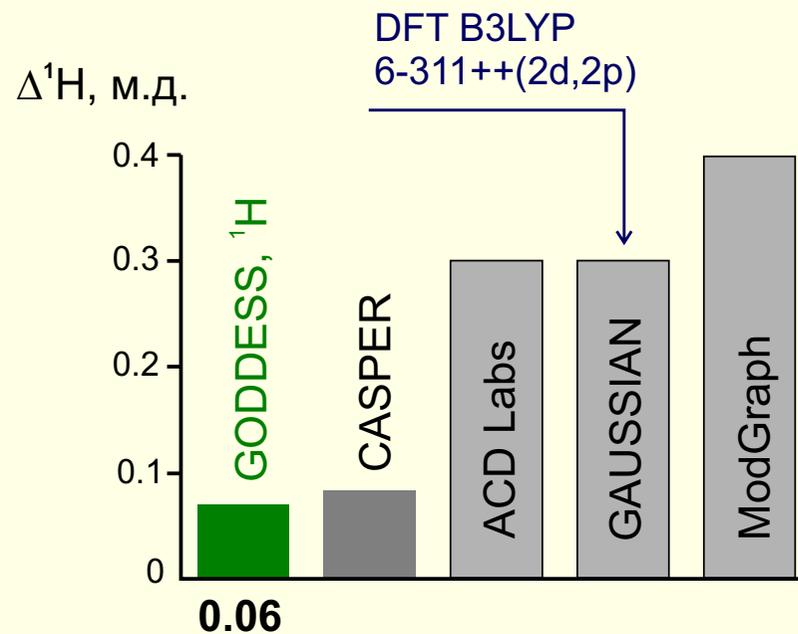
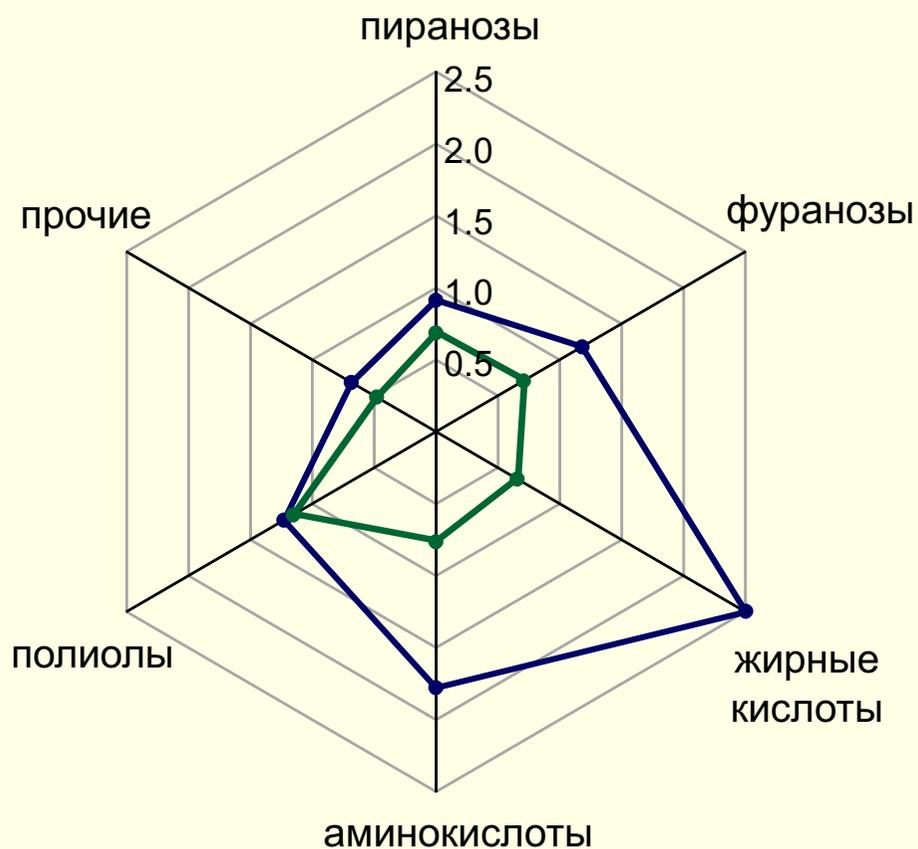
(конфигурация C4 донора)



- каждый олигомерный фрагмент, содержащий нужный атом, обобщается от малых весов к большим, пока содержащие его структуры не будут найдены в базе
- ХС усредняются с учетом выбросов, их достоверность оценивается по качеству данных и общему «весу» обобщений

# Программы предсказания спектров

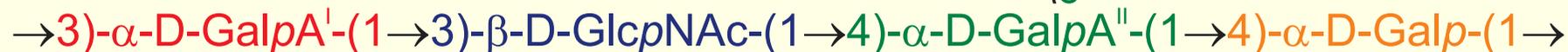
сравнение модель-эксперимент ( $\Delta^{13}\text{C}$ , м.д.)  
на выборке ~32000 XC  
из всех классов биогликанов



# Предсказание спектров в CSDB

GroN-(2

\6



Поддерживается:

**<sup>1</sup>D:** <sup>13</sup>C ВВ, цветовой код,  
таблицы отнесения <sup>1</sup>H и <sup>13</sup>C

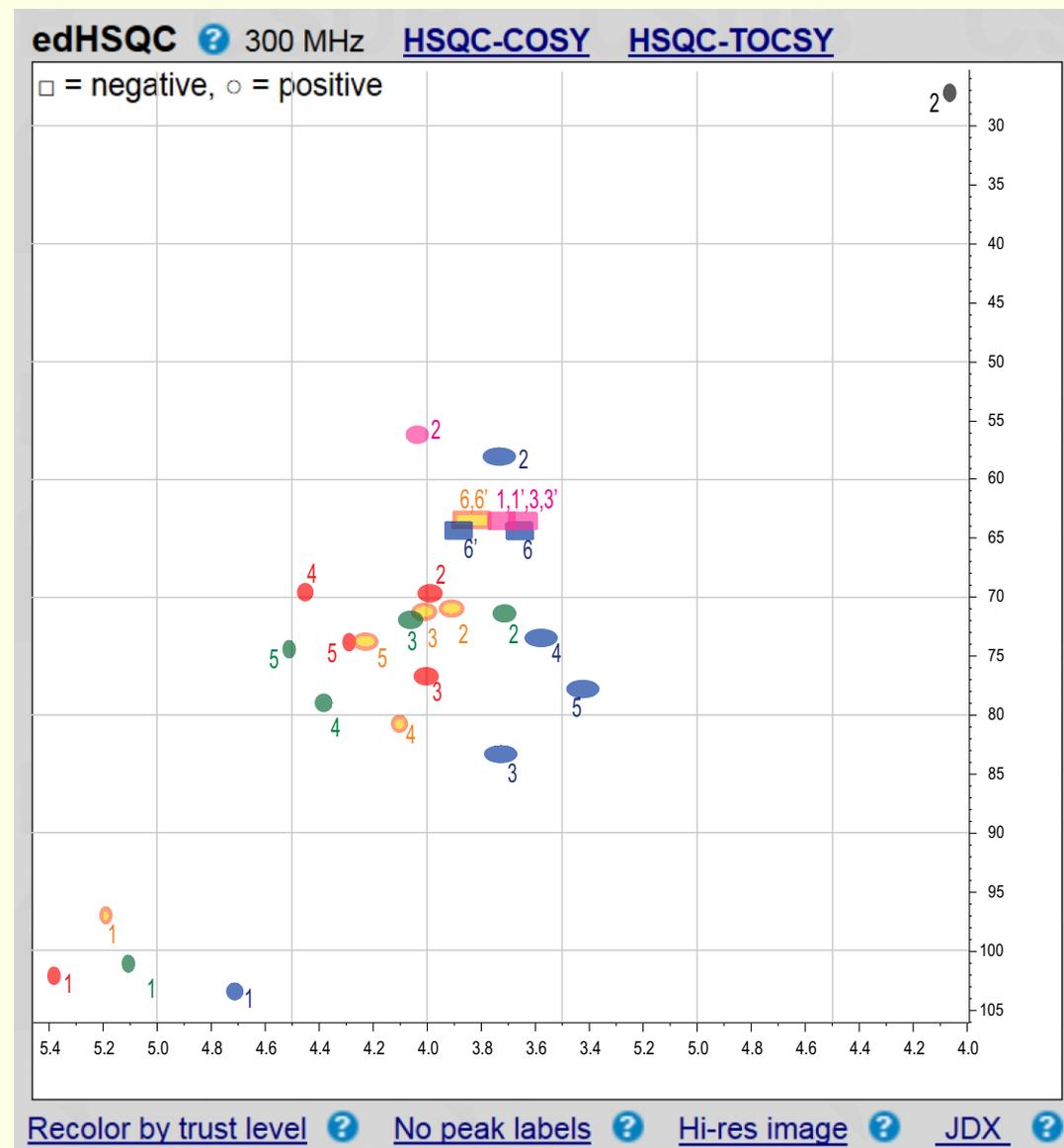
**<sup>1</sup>H / <sup>1</sup>H:** COSY, TOCSY,  
DQF COSY, COSY RCT

**<sup>1</sup>H / <sup>13</sup>C:** edHSQC, HMBC,  
HSQC-COSY, HSQC-TOCSY

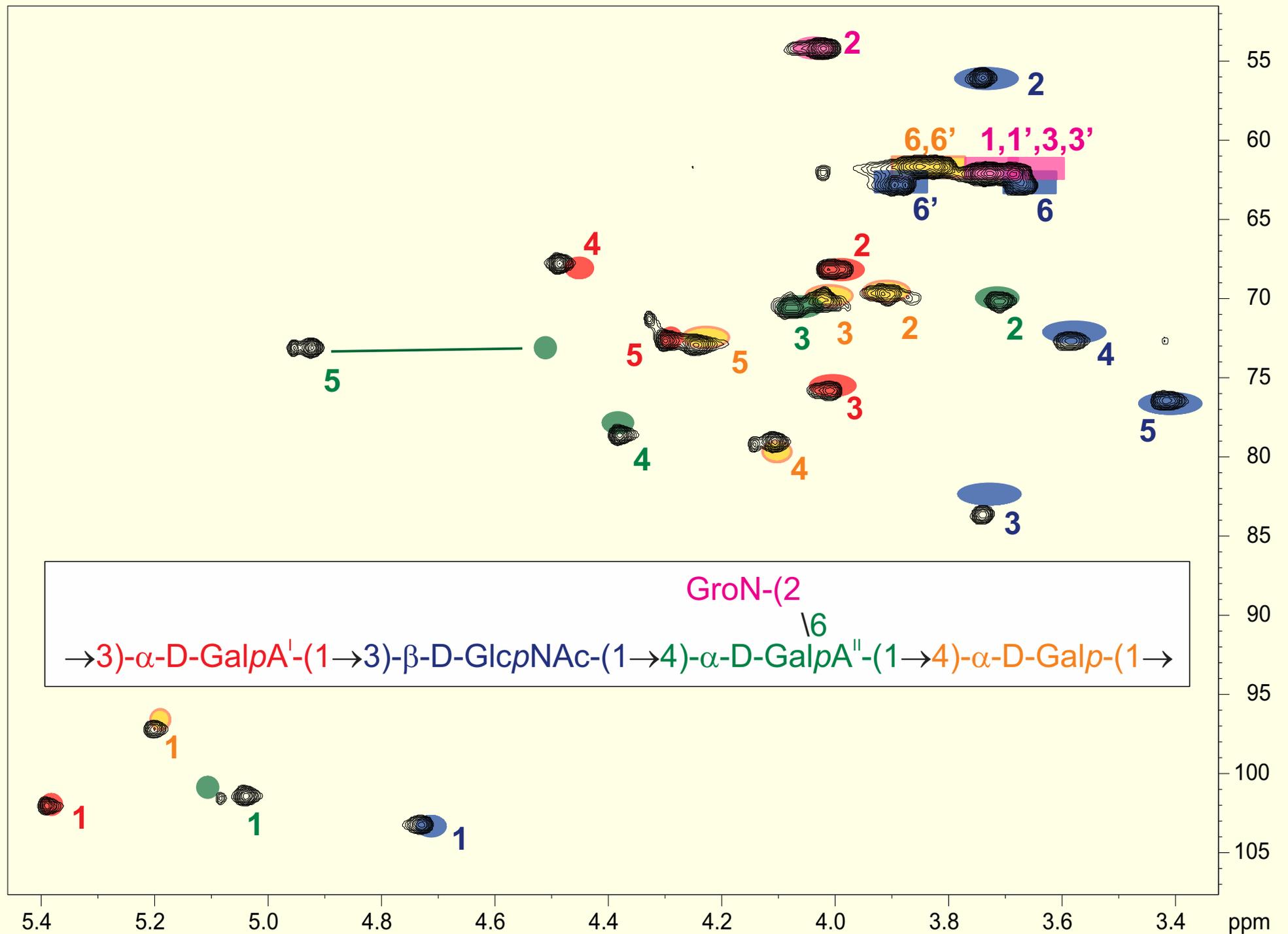
эмпирическая оценка KCCB  
+ частота спектрометра  
=> ширина кросс-пигов

Вывод:

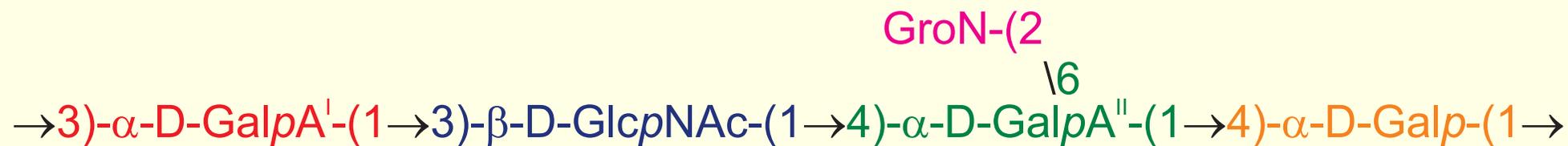
таблицы, спектры, «живая» работа в  
браузере, экспорт в CSV и Jcamp-DX



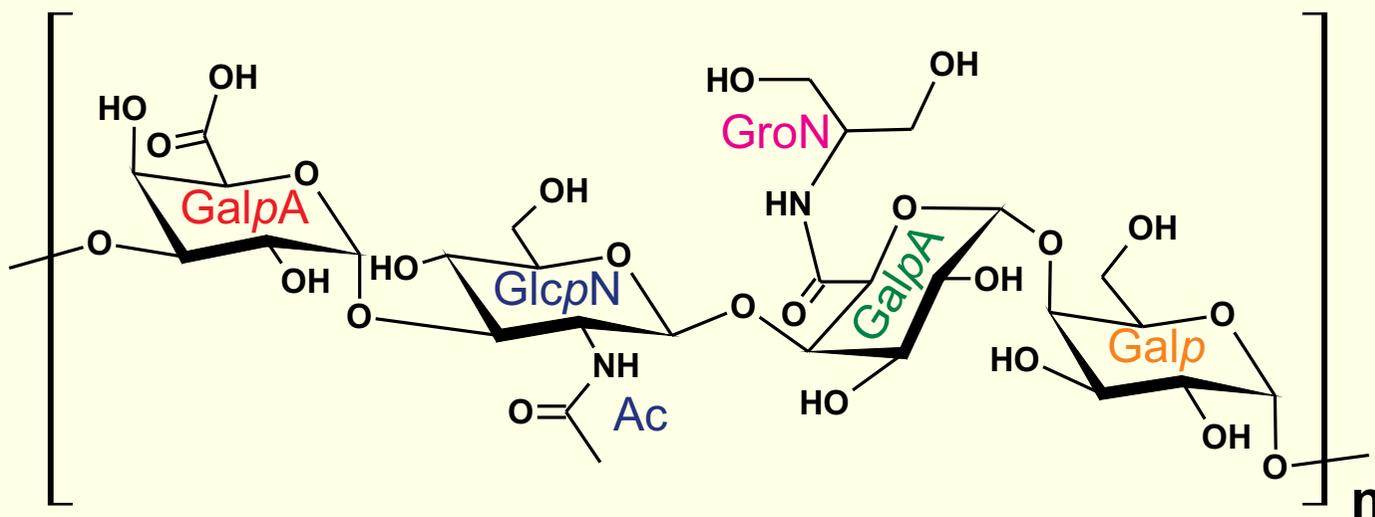
# Эксперимент и симуляция



# Строение повторяющегося звена *Edwardsiella* 1153



HSQC => в нативном полисахариде остаток Gal O-ацетилирован  
в положения 2 (45% звеньев) либо 3 (другие 45% звеньев)



установлено за **30** часов:

выделение и очистка	4
деацетилирование	2
анализ мономерного состава	1
оптические методы	1
накопление ЯМР спектров	8
интерпретация	6
оформление	3
локализация O-ацетатов	4+1

# Возможности метода

