



УДК 547.458.022 : 543.422.25

АНАЛИЗ СПЕКТРОВ ^{13}C -ЯМР ОЛИГОСАХАРИДОВ
С ПОМОЩЬЮ МИКРОКОМПЬЮТЕРА

Жубаев В. И.

*Институт органической химии им. Н. Д. Зелинского
Академии наук СССР, Москва*

Описаны принципы работы программы ANMROL, предназначенной для установления того, однозначно ли спектр ^{13}C -ЯМР олиго- или полисахарида подтверждает его структуру. Программа написана на языке Бейсик и реализована на микрокомпьютере «Искра-226». Обсуждаются результаты применения программы для анализа спектров ряда линейных и разветвленных олигосахаридов.

Спектроскопия ^{13}C -ЯМР широко используется в структурном анализе олиго- и полисахаридов. Накоплен обширный материал по спектрам ^{13}C -ЯМР углеводов [1—3], и сопоставление спектра вновь полученного олиго- или полисахарида с литературными данными обычно используется для его идентификации или установления строения. Как недавно показано [4—6], расчет спектров линейных полисахаридов по аддитивной схеме с помощью компьютера может быть весьма эффективным подходом к установлению строения этих полимеров.

При работе по синтезу олиго- и полисахаридов часто приходится сталкиваться с решением следующего вопроса: достаточно ли спектра ^{13}C -ЯМР для однозначного установления строения продукта реакции гликозилирования, или для этой цели должны быть привлечены и другие методы, например метилирование, поляриметрия, периодатное окисление и т. п. Иными словами, вопрос состоит в том, соответствует ли спектр ^{13}C -ЯМР только одному из возможных типов связи между гликозилирующим и гликозилируемым компонентами реакции, или может найтись такой изомер продукта, структуре которого этот спектр не противоречит. Для ответа на этот вопрос необходимо тщательное сопоставление экспериментального спектра ЯМР со спектрами соответствующих модельных соединений и литературными данными о возможных пределах изменения α - и β -эффектов гликозилирования для каждого типа гликозидной связи. Проводимые при этом расчеты довольно трудоемки, и для автоматизации этого процесса нами разработана программа ANMROL (Analysis of NMR spectra of oligosaccharides). Программа написана на языке Бейсик и реализована на микрокомпьютере «Искра-226». В настоящей статье описаны принципы работы этой программы, ее возможности и ограничения, а также обсуждены результаты ее использования для анализа спектров ряда линейных и разветвленных олигосахаридов. Копия программы и подробная инструкция по ее использованию могут быть получены от автора по запросу.

ОБЩИЕ СВЕДЕНИЯ О ПРОГРАММЕ

Программа ANMROL применима для анализа спектров ^{13}C -ЯМР дисахаридов $S_2 \rightarrow S_1$ и их производных, содержащих это дисахаридное звено, в том числе олиго- и полисахаридов. В последнем случае дополнительные моносахаридные остатки, входящие в состав этих соединений, обозначаются сокращениями $S_3(i)$, где i — целое число. Настоящий вариант программы может быть использован для олигосахаридов, вплоть до гексасахаридов, и полисахаридов, содержащих до шести моносахаридных остатков в повторяющемся звене.

Типы структур олиго- и полисахаридов, исследованных с помощью программы ANMROL

Тип структуры	Использованные модельные соединения
(I): S2 → S1	S2; S1
(II): S2 → S1 → R	S2; S1 → R
(III): S2 → S1 → S3(1)	S2; S1 → S3(1)
(IV): S2 → [S3(1) →] S1	S2; S3(1) → S1
(V): S2 → [S3(1) →] S1 → R	S2; S3(1) → S1 → R
(VI): S2 → S1 → [S3(2) →] S3(1)	S2; S1 → [S3(2) →] S3(1)
(VII): S3(2) → S2 → S1 → S3(1)	S3(2) → S2; S1 → S3(1)
(VIII): → S2 → S1 → S3(1) →	R → S2; S1 → S3(1) → R
(IX): → S3(2) → S3(1) → [S2 →] S1 →	S2; → S3(2) → S3(1) → S1 →

В результате выполнения расчетов делается вывод о типе связи между остатками S2 и S1, входящими в состав данного соединения, дается предполагаемое отнесение сигналов в спектре и вычисляются величины эффектов гликозилирования. Если данному экспериментальному спектру могут соответствовать несколько изомеров дисахаридного звена, то такая информация сообщается для каждого случая. Для связей с участием остатков S3(i) принимается, что они идентичны аналогичным связям в использованных модельных соединениях.

В табл. 1 приведены примеры типов структур олиго- и полисахаридов, для анализа ^{13}C -ЯМР-спектров которых может быть использована программа ANMROL. Межмономерная связь, для которой проводится определение возможного типа, показана на формулах стрелкой с точками.

Программа ANMROL написана в предположении, что все моносахариды, входящие в состав исследуемых соединений, являются гексопиранозами. Она может быть использована без изменений и для случая пентопираноз, но если хотя бы один из моносахаридных остатков содержит более шести атомов углерода или находится в фуранозной форме, применение программы невозможно без ее модификации.

ПРИНЦИПЫ РАБОТЫ ПРОГРАММЫ И ПОДГОТОВКА ИСХОДНЫХ ДАННЫХ

Анализ спектра ^{13}C -ЯМР исследуемого соединения состоит в сопоставлении набора химических сдвигов сигналов экспериментального спектра $Y(j)$ (j — порядковый номер сигнала в спектре) и наборов данных $X(k, m)$ для модельных соединений, где k — номер моносахаридного остатка, m — номер атома углерода*.

Все сигналы, присутствующие в спектре исследуемого соединения, можно разбить на две группы: 1) «невозмущенные» сигналы, т. е. сигналы, положение которых не изменяется существенно при образовании связи S2 → S1; они относятся к атомам углерода, удаленным от атомов, участвующих в образовании гликозидной связи; 2) «возмущенные» сигналы, химические сдвиги которых заметно изменены по сравнению с моделями за счет α - и β -эффектов гликозилирования. С логической точки зрения задача определения типа межмономерной связи состоит в идентификации в спектре невозмущенных и возмущенных сигналов.

При анализе спектра олигосахаридов, содержащего N моносахаридных остатков и не имеющего свободной полуацетальной группы, или полисахарида с N моносахаридными остатками в повторяющемся звене программа требует введения набора $6N$ значений химических сдвигов $Y(j)$, расположенных в порядке уменьшения величины δ , т. е. от слабого поля к сильному.

* Термином «модельное соединение» обозначено соединение, содержащее остаток моносахарида, интерпретированный ^{13}C -ЯМР-спектр которого сравнивается со спектром исследуемого олиго- или полисахарида. Этот остаток моносахарида мы будем называть «моделью» для остатка соответствующего моносахарида в исследуемом соединении, а относящиеся к его атомам сигналы в спектре модельного соединения — «модельным набором данных».

Если в состав анализируемого соединения входят один или несколько остатков пентоз, то серию дополняют соответствующим количеством значений $Y(j) = 0$, так, чтобы общее количество сигналов было равно $6N$. При наличии остатков уроновых кислот слабopольные сигналы, отвечающие атомам углерода карбоксильных групп, не включают в серию, а вместо них вводят нулевые значения $Y(j)$.

В случае восстанавливающего олигосахарида со свободной полуацетальной группой экспериментальные данные химических сдвигов необходимо разбить на три серии. Основная серия $Y(j)$ включает в себя $6N - 6$ сигналов единичной интенсивности, относящихся к атомам углерода всех моносахаридных остатков, кроме остатка, находящегося на восстанавливаемом конце цепи. Две дополнительные серии $Y1(j)$ и $Y2(j)$ содержат по 6 сигналов меньшей интенсивности, они соответствуют двум аномерным формам восстанавливающего моносахаридного остатка. Сигналы основной серии вводятся в начале выполнения программы, а сигналы дополнительных серий используются только в том случае, когда они относятся к остатку S1; их вводят после идентификации сигналов остатков S3 (i) и S2.

Наиболее удобными модельными соединениями для сопоставления с экспериментальным спектром являются моно- или олигосахариды, образующиеся при ретросинтетическом расщеплении исследуемого соединения по связи S2 \rightarrow S1 (см. табл. 1). При синтетической работе спектры таких соединений обычно доступны и интерпретированы к моменту начала исследования более сложного соединения. Это требование не является, однако, обязательным, и необходимые модельные наборы данных $X(k, m)$ могут быть получены и другими путями.

Ввод данных для модельных соединений проводится на разных этапах расчета по запросу программы. При этом в случае пентоз и уроновых кислот в качестве сигнала для атома углерода C6 необходимо вводить нулевое значение, т. е. принимать $X(k, 6) = 0$.

Поскольку незамещенные моносахариды во многих случаях могут быть использованы в качестве модельных соединений для остатков S2 и S1, в библиотеку программы включены данные по спектрам α - и β -аномеров пяти наиболее распространенных гексоз и 6-дезоксигексоз — глюкозы, галактозы, маннозы, рамнозы и фукозы. Данные для первых четырех моносахаридов взяты из работы [5], для фукозы — из обзора [1]. Предусмотрена возможность дальнейшего расширения библиотеки. При работе с производными моносахаридов, данные для которых имеются в библиотеке, количество вводимых данных для модельных соединений значительно уменьшается.

Работа программы начинается с ввода наборов значений $Y(j)$ и $X(k, m)$, а также сокращений для S1 и S2. Дальнейшие вычисления включают в себя следующие этапы:

- 1) поиск в экспериментальном спектре невозмущенных сигналов, относящихся к остаткам S3 (i), и их вычитание из спектра;
- 2) определение конфигурации при C1 остатка S2, идентификация сигналов этого моносахарида и их вычитание из спектра;
- 3) выбор необходимого набора модельных данных для остатка S1;
- 4) идентификация возмущенных и невозмущенных сигналов остатка S1 и нахождение возможных вариантов типа связи между остатками S2 и S1.

Основное арифметическое действие, производимое при выполнении программы, — это вычисление разностей $D(j, k, m)$ между химическими сдвигами одного из сигналов в спектре исследуемого соединения и сигнала одного из атомов углерода в модельном соединении:

$$D(j, k, m) = Y(j) - X(k, m). \quad (1)$$

Невозмущенные сигналы идентифицируются в экспериментальном спектре как сигналы $Y(j)$, для которых при заданных k и m абсолютная величина разности $D(j, k, m)$ минимальна, т. е. принимается, что

$$\left. \begin{aligned} Y_0(k, m) = Y(j); \quad E(k, m) = D(j, k, m) \\ \text{при } |E(k, m)| = \min(|D(j, k, m)|) \end{aligned} \right\}, \quad (2)$$

где $Y_0(k, m)$ — химический сдвиг, приписываемый атому m остатка k в исследуемом спектре, $E(k, m)$ — разность химических сдвигов сигналов одних и тех же атомов углерода в спектре исследуемого и модельного соединения.

ИДЕНТИФИКАЦИЯ СИГНАЛОВ ОСТАТКОВ S3(i)

Моделями для остатков S3 (i) могут служить остатки соответствующих моносахаридов, замещенные по той же OH-группе и имеющие ту же конфигурацию при C1, что и остаток S3(i), в исследуемом соединении.

При образовании гликозидной связи S2 → S1 положение сигналов, относящихся к остаткам S3 (i), не должно заметно изменяться. Таким образом, эти сигналы относятся к невозмущенным, и их поиск производится в соответствии с условием (2). Программа запрашивает сокращение для остатка S3 (i) и модельный набор данных для него. После проведения поиска и вычисления величин $Y_0(k, m)$ и $E(k, m)$, где $k = i + 2$, сигналы $Y(j)$, соответствующие найденным значениям $Y_0(k, m)$, вычитаются из экспериментального спектра. Операции далее повторяются для следующего остатка S3 (i) до тех пор, пока сигналы, относящиеся к остаткам S3 (i), не будут идентифицированы. В итоге этого этапа работы в экспериментальном спектре остаются неидентифицированными 12 сигналов, соответствующие остаткам S2 и S1.

ИДЕНТИФИКАЦИЯ СИГНАЛОВ ОСТАТКА S2 И ОПРЕДЕЛЕНИЕ КОНФИГУРАЦИИ ПРИ C1 ОСТАТКА S2

Для решения этой задачи необходимы наборы модельных данных, соответствующих α - и β -аномеру остатка S2 — $X_l(2, m)$, где $l = 1$ для α -аномера и $l = 2$ для β -аномера. В тех случаях, когда остаток S2 в исследуемом соединении присутствует на невосстанавливаемом конце цепи (например, в структурах типа (I)–(VI) и (IX), см. табл. 1), моделью для него служит незамещенный моносахарид, при нахождении же остатка S2 внутри цепи (см. структуры типа (VII) и (VIII)) необходимо использовать в качестве моделей замещенные остатки S2, входящие в состав дисахаридного звена S3 (i) → S2.

Как известно, превращение моносахарида S2 в дисахарид S2 → S1 приводит к сильному изменению химического сдвига сигнала атома C1 остатка S2. При этом могут наблюдаться два случая: «нормальный» эффект гликозилирования, когда химический сдвиг рассматриваемого сигнала близок к химическому сдвигу аналогичного сигнала в метилгликозиде S2, и «аномальный» эффект гликозилирования, когда сигнал атома C1 остатка S2 находится в значительно более сильном поле, чем для метилгликозида. Последний случай характерен для циклогексилгликозидов и дисахаридных звеньев S2 → S1, в которых остаток гексозы S1 содержит в β -положении по отношению к гликозилируемой OH-группе экваториальный атом водорода (аксиальную OH-группу) при соблюдении определенных стереохимических требований (см. [7]).

В соответствии с этим программа производит преобразование модельных наборов данных для незамещенного остатка S2 в модельные наборы, отвечающие остатку моносахарида, входящему в состав дисахаридного звена при нормальном и аномальном эффекте гликозилирования. Это преобразование производится путем прибавления к исходным данным $X_l(2, m)$ инкрементов $Z_p^l(m)$, отражающих влияние образования гликозидной связи на химические сдвиги сигналов атомов углерода остатка S2 ($p = 2$ при нормальном эффекте гликозилирования, $p = 4$ — при аномальном, $k = 2$):

$$X_{l+p}(k, m) = X_l(k, m) + Z_p^l(m). \quad (3)$$

Включенные в программу значения инкрементов $Z_p^l(m)$ (табл. 2) получены из рассмотрения литературных данных по спектрам ^{13}C -ЯМР олигосахаридов [1–3, 5]. После выполнения преобразования по уравнению

Введенные в программу значения инкрементов для учета влияния замещения при С1 на химические сдвиги сигналов остатка моносахарида

m (номер атома С)	$Z_2^l(m)$ при нормальном эффекте гликозилирования *	
	$l=1$ (α -аномер)	$l=2$ (β -аномер)
1	7,3(7,0) **	7,4(6,7) **
2	-0,5	-1,0
3	0,4	0
5	0,5	0

* $Z_2^1(4) = Z_2^1(6) = 0$. Набор значений $Z_4^l(m)$ для аномального эффекта гликозилирования отличается только значением для $m=1$: $Z_4^1(1) = Z_4^2(1) = 4,5$.

** Значения без скобок — при экваториальном положении ОН-группы при С2, в скобках — при аксиальном положении.

(3) для замещенного по С1 остатка S2 имеются четыре модельных набора данных: величины $X_3(2, m)$ и $X_4(2, m)$ соответствуют сигналам атомов углерода α - и β -аномеров остатка S2 при нормальном эффекте гликозилирования, а величины $X_5(2, m)$ и $X_6(2, m)$ — аналогичным сигналам при аномальном эффекте гликозилирования. При использовании таких данных можно рассматривать сигналы, относящиеся к атомам остатка S2 в экспериментальном спектре, как невозмущенные. Их поиск в спектре производится в соответствии с условием (2).

Каждому из использованных наборов соответствует свой набор найденных величин химических сдвигов и разностей $E_l(2, m)$. Критерием для выбора правильного модельного набора служит величина суммы абсолютных значений этих разностей:¹

$$S_l = \sum_{m=1}^6 |E_l(k, m)|. \quad (4)$$

При $S_3 < S_5$ и $S_4 < S_6$ аномальные эффекты гликозилирования отсутствуют, и выбор конфигурации при С1 остатка S2 делается в пользу той конфигурации, набор модельных данных которой обеспечивает меньшее значение величин S_l , т. е. при $S_3 < S_4$ делается вывод об α -конфигурации, а при $S_3 > S_4$ — о β -конфигурации.

Если хотя бы одна из сумм S_5 или S_6 меньше, чем суммы, отвечающие нормальным эффектам гликозилирования, программа делает вывод о существовании в данном дисахаридном звене аномального эффекта гликозилирования. В этом случае для вывода о конфигурации при С1 остатка S2 используются закономерности, подробно обсужденные в работе [7]. В настоящем варианте программы учтен только наиболее часто встречающийся случай 1 \rightarrow 3-связи между остатками S2 и S1. В соответствии с этим при обнаружении аномального эффекта гликозилирования программа предлагает ответить на вопросы, одинакова ли абсолютная конфигурация S1 и S2 и содержит ли остаток S1 аксиальную ОН-группу при С2 или С4. В зависимости от ответов выбор конфигурации при С1 остатка S2 производится как показано в табл. 3. Если же аксиальные ОН-группы в остатке S1 отсутствуют или находятся при С1 или С3, то вывод о конфигурации при С1 остатка S2 при аномальном эффекте гликозилирования делается на основании тех же критериев, которые используются при нормальном эффекте.

Значения $Y_0(2, m)$, соответствующие выбранному модельному набору данных, вычитаются из экспериментального спектра, после чего в нем остается лишь шесть сигналов, отвечающих атомам углерода замещенного остатка S1.

Выбор конфигурации при С1 остатка S2 при аномальном эффекте гликозилирования и 1 → 3-связи

Положение аксепальной ОН-группы в остатке S1	Абсолютная конфигурация остатков S1 и S2	
	одинаковая	разная
2	β	α
4	α	β

ВЫБОР НЕОБХОДИМОГО НАБОРА МОДЕЛЬНЫХ ДАННЫХ ДЛЯ ОСТАТКА S1

Если остаток S1 в исследуемом соединении не находится в узле разветвления, то моделью для него может служить соответствующий незамещенный моносахарид. При этом в тех случаях, когда остаток S1 находится на восстанавливающем конце цепи и имеет свободную полуацетальную группу (структуры типа (I)), необходимо использовать два набора модельных данных, отвечающих α - и β -аномерам остатка S1 — $X_l(1, m)$, где $l = 1, 2$. В структурах типа (II), (III), (VII) и (VIII) достаточно данных по спектрам той аномерной формы, которая присутствует в исследуемом соединении.

Для остатка S1, входящего в состав узла разветвления (структуры (IV), (V), (IX)) в качестве модели должны быть использованы данные для остатка S1 в дисахаридном звене $S3(i) \rightarrow S1$.

Если остаток S1 в исследуемом соединении имеет свободную полуацетальную гидроксильную группу, то, как отмечалось выше, относящиеся к нему сигналы в экспериментальном спектре входят в состав двух дополнительных серий — $Y1(j)$ и $Y2(j)$. Эти серии используют в расчетах поочередно и возникает проблема правильного выбора модельного набора для последующих вычислений, т. е. определение того, соответствуют ли сигналы первой из введенных дополнительных серий $Y1(j)$ спектру α - или β -аномера остатка S1. Для решения этой задачи величины $X_l(1, m)$ преобразуются в набор $U_l(j)$, в котором значения химических сдвигов расположены в порядке убывания $\delta(j)$ — номер сигнала при таком расположении), и проводится вычисление сумм T_l по уравнению (5):

$$T_l = \sum_{j=1}^6 |Y1(j) - U_l(j)|. \quad (5)$$

При $T_1 < T_2$ делается вывод об α -конфигурации при С1 остатка S1 для рассматриваемой экспериментальной серии, в противном случае — о β -конфигурации. Соответствующая серия $X_l(1, m)$ применяется в дальнейшей работе.

После окончания работы по анализу серии $Y1(j)$ программа предлагает повторить анализ для другой дополнительной серии $Y2(j)$. В этом случае конфигурация при С1 остатка S1 автоматически выбирается противоположной найденной в первом случае.

Если сигналы остатка S1 входят в основную серию экспериментальных данных $Y(j)$, то программа спрашивает, можно ли использовать α - или β -метилгликозид S1 как модель для остатка S1 в исследуемом соединении. При положительном ответе значения химических сдвигов рассматриваемого моносахарида преобразуются в величины $X_3(1, m)$ и $X_4(1, m)$ по уравнению (3) ($k = 1, p = 2$) и соответствующий модельный набор используется в дальнейшей работе.

ИДЕНТИФИКАЦИЯ СИГНАЛОВ ОСТАТКА S1 И НАХОЖДЕНИЕ ВОЗМОЖНЫХ ВАРИАНТОВ ТИПОВ СВЯЗИ

Следующий этап работы программы — идентификация возмущенных и невозмущенных сигналов, относящихся к моносахаридному остатку S1. Решение этой задачи начинается с поиска сигналов, соответствующих

Стандартные значения границ эффектов гликозилирования, введенные в программу

Эффект	$F1$ (нижняя граница)	$F2$ (верхняя граница)
α -Эффект	3,3	10,5
β -Эффект		
При 1 \rightarrow 4-связи		
на С3	-2,8	1,2
на С5	-2,3	1,5
При 1 \rightarrow 3-связи		
на С2	-3,5	1,3
на С4	-4,3	1,5
При 1 \rightarrow 2-связи		
на С1	-3,5	1,0
на С3	-2,5	0,8
Дальние эффекты	-0,5	1,0

атомам С1 и С6. Эти сигналы достаточно сильно отличаются по своим химическим сдвигам от остальных сигналов, и даже в тех случаях, когда они возмущены за счет эффектов гликозилирования, величины $Y_0(1, 1)$ и $Y_0(1, 6)$ могут быть безошибочно найдены в соответствии с условием (2).

Определение величин $E(1, 6)$ позволяет проверить гипотезу о гликозилировании 6-оксигруппы в остатке S1. При $E(1, 6) > 2,5$ программа делает вывод о существовании 1 \rightarrow 6-связи в дисахариде S2 \rightarrow S1, отнесение сигналов к атомам С2, С3 и С4 остатка S1 проводится по условию (2), а единственный сигнал, оставшийся неидентифицированным, относится к сигналу атома С5.

Если величина $E(1, 6) \leq 2,5$, то связь между остатками S2 и S1 должна осуществляться через ОН-группу при С2, С3 или С4 остатка S1. При каждом из вариантов один из сигналов, относящихся к атомам С1—С5 остатка S1, должен быть возмущен за счет α -эффекта гликозилирования, два — за счет β -эффектов гликозилирования, а два сигнала — остаться невозмущенными.

Для идентификации возмущенных сигналов остатка S1 величины $D(j, k, m)$, найденные по уравнению (1), сопоставляются с данными о возможных пределах изменения величин α - и β -эффектов гликозилирования, т. е. принимается, что

$$Y_0(1, m) = Y(j); E(1, m) = D(j, 1, m) \quad (6)$$

при $F1 < E(1, m) < F2$, где $F1$ и $F2$ — соответственно нижний и верхний пределы изменения эффектов гликозилирования. Стандартные значения для границ эффектов, занесенные в библиотеку программы, приведены в табл. 4. Предусмотрена возможность изменения этих границ в режиме диалога, что позволяет использовать программу для определения типа связи в тех случаях, когда величины эффектов гликозилирования отличаются от обычных, например для случая разветвленных олигосахаридов.

Работа программы начинается с поиска тех сигналов, химические сдвиги которых могут быть объяснены возмущением сигналов модельного соединения за счет α -эффектов гликозилирования. Обычно при этом возникает от двух до шести вариантов возможной идентификации таких сигналов и соответственно возможного типа связи в дисахаридном звене.

Проверка того, реален ли тот или иной тип гликозидной связи, основывается на очевидных соображениях. Если m — номер атома углерода остатка S1, сигнал которого возмущен за счет α -эффекта гликозилирования, то сигналы атомов с номерами $m - 1$ и $m + 1$ должны быть возмущены за счет β -эффектов, а сигналы остальных атомов остатка S1 должны оставаться невозмущенными. При выполнении этих условий можно сделать вывод, что рассматриваемый тип связи между остатками S2 и S1 возможен, в противном случае необходимо отвергнуть эту гипотезу.

Для невозмущенных сигналов остатка S1 изменение химических сдвигов в результате гликозилирования не должно выходить за пределы изменения «дальних» эффектов (см. табл. 4), а для атомов углерода, сигналы которых возмущены за счет β -эффектов гликозилирования, — за границы изменения последних, которые несколько различны для разных типов связи.

Как правило, после перебора всех возможных вариантов оказывается, что возможен только один тип связи между остатками S2 и S1 (см. ниже). Если при заданных границах эффектов гликозилирования ни одна из гипотез о типе связи не является удовлетворительной, программа предлагает пользователю изменить эти границы, выбор необходимых изменений обычно достаточно очевиден из рассмотрения полученных результатов.

ВЫВОД РЕЗУЛЬТАТОВ АНАЛИЗА

По мере работы программы ANMROL результаты вычислений выводятся на дисплей; по желанию пользователя они могут быть напечатаны. Распечатка включает следующие данные:

- 1) химические сдвиги экспериментального спектра;
- 2) сокращения для остатков S3 (i) и таблицы значений $Y_0(k, m)$ и $E(k, m)$ для соответствующего моносахаридного остатка;
- 3) сообщение об аномальном эффекте гликозилирования, если таковой наблюдается;
- 4) сокращение для остатка S2 и конфигурация при S1 этого остатка;
- 5) таблица значений $Y_0(2, m)$ и $E(2, m)$, а также разностей между $Y_0(2, m)$ и значениями $X_l(2, m)$ для свободного моносахарида соответствующей конфигурации;
- 6) сокращение для остатка S1;
- 7) значения $X_l(1, m)$ и $Y_0(1, m)$ для сигналов экспериментального спектра, которые могут быть объяснены возмущением за счет α -эффекта гликозилирования, и число возможных вариантов приписывания сигналов остатка S1;
- 8) результаты анализа каждого из этих вариантов. Если рассматриваемый вариант внутренне непротиворечив, то печатается тип связи между остатками S2 и S1 и таблица значений $Y_0(1, m)$, $E(1, m)$ и разностей между $Y_0(1, m)$ и величиной $X_l(1, m)$ для свободного незамещенного моносахарида. В противном случае сообщается, что данный вариант невозможен по величинам дальних эффектов или β -эффектов, и приводятся величины $E(1, m)$, выходящие за пределы заданных границ эффектов гликозилирования;
- 9) число возможных вариантов типа связи между остатками S2 и S1.

При анализе спектра олигосахарида, содержащего остаток S1 со свободной полуацетальной группой, распечатка результатов приводится отдельно для каждой серии экспериментальных данных.

В итоге после завершения работы программы выдается информация о том, какие типы межмономерной связи S2 \rightarrow S1 возможны для исследуемого дисахарида или дисахаридного звена олиго- или полисахарида по данным спектра ^{13}C -ЯМР; наряду с этим проводится внутренне непротиворечивое отнесение сигналов спектра ^{13}C -ЯМР и вычисляются значения эффектов гликозилирования, характерные для данного дисахаридного звена.

ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ, ПОЛУЧЕННЫХ ПРИ РАБОТЕ С ПРОГРАММОЙ

Программа ANMROL в разных вариантах используется в нашей лаборатории в течение двух лет и показала свою полезность. Первые варианты программы были разработаны в связи с анализом спектров ^{13}C -ЯМР ди- и олигосахаридных фрагментов О-специфических полисахаридов сальмонеллы серогрупп С₂ и С₃ [8—10]. Далее, был проведен анализ спектров ряда олигосахаридов, ранее полученных в нашей лаборатории [11—14] и синтезированных в последнее время в связи с биохимическими и конформационными исследованиями [15, 16]. Структура олигосахаридов в исследуе-

мых сериях была такой, что выбор модельных соединений не вызывал затруднений.

К настоящему времени программа ANMROL использована для анализа спектров ^{13}C -ЯМР 30 олигосахаридов структурных типов (I)—(IV) (см. табл. 1). Однозначный ответ о типе связи между остатками S2 и S1 не был получен лишь в трех случаях; кроме того, в четырех из 10 исследованных дисахаридов однозначный выбор типа связи достигался лишь для одной из серий сигналов, отвечающих моносахаридному остатку S1. Использование программы ANMROL для анализа спектров полисахаридов будет описано отдельно.

Полученные результаты показывают, что в подавляющем большинстве случаев спектра ^{13}C -ЯМР достаточно для однозначного доказательства структуры синтетического олигосахаридов. Более того, использование накопленного в последнее время материала о связи величины α - и β -эффектов гликозирования со структурой остатка S1 и конфигурацией при C1 остатка S2 [5], который игнорируется в программе ANMROL, позволяет во всех сомнительных случаях осуществить однозначное определение типа гликозидной связи (см. табл. 5).

Так, для дисахаридов $D\text{-Man}(\beta 1 \rightarrow 3)D\text{-Gal}$ (случаи 1 и 2 в табл. 5) результаты расчета по программе ANMROL не позволяют исключить альтернативную структуру с $\beta 1 \rightarrow 2$ -типом гликозидной связи. Однако сопоставление величин кажущихся эффектов гликозирования для сигналов атомов C2 и C1 остатка β -галактозы, соответствующих такому типу связи, с уточненными значениями α - и β -эффектов гликозирования для этого случая [5] указывает на существование значительных различий и позволяет уверенно исключить альтернативную структуру. Точно так же альтернативные структуры с $\beta 1 \rightarrow 2$ - и $\beta 1 \rightarrow 4$ -типами связи, которые представляются возможными для α -аномера дисахаридов, могут быть отброшены на основании несоответствия кажущихся и уточненных величин β -эффекта на атоме C3 остатка галактозы, хотя в этом случае различия менее значительны.

Аналогичные результаты получены и для других дисахаридов, в которых программа ANMROL дает неоднозначный выбор типа связи для одной из аномерных форм (случаи 3—6); в каждом случае обнаруживается один или несколько α - или β -эффектов, для которых кажущиеся значения при альтернативных структурах не соответствуют уточненным величинам. Особо следует остановиться лишь на примере дисахаридов $L\text{-Rha}(\alpha 1 \rightarrow 3)D\text{-Gal}$ (случай 4); альтернативная его структура с $\alpha 1 \rightarrow 2$ -связью может быть отвергнута на основании небольшого, но характерного различия β -эффектов на атоме C1 остатка α -галактозы. Именно величина этого эффекта, который становится более заметным при β -конфигурации остатка галактозы, является определяющим фактором, позволяющим отвергнуть альтернативные структуры с 2,4-замещением остатка галактозы для разветвленного трисахаридов (случаи 7 и 8) или его метилгликозида (случай 9): можно уверенно предсказать, что наблюдаемые эффекты гликозирования на C1 остатка β -галактозы совместимы лишь с присутствием моносахаридных заместителей у C3 и C4.

Этот подход — использование уточненных значений величин α - и β -эффектов гликозирования для выбора между альтернативными структурами в тех редких случаях, когда программа ANMROL не дает однозначного определения типа связи, — является, по-видимому, более целесообразным, чем внесение в библиотеку программы уточненных сведений об эффектах гликозирования. Полученные результаты показывают, что во всех исследованных примерах спектр ^{13}C -ЯМР позволяет однозначно определить структуру олигосахаридов. Существуют, однако, и трудные случаи — например, выбор между $1 \rightarrow 3$ - и $1 \rightarrow 2$ -связями при $S1 = \alpha\text{-D-Gal}$ и конфигурации $\alpha\text{-L}$ или $\beta\text{-D}$ при C1 остатка S2, когда вывод о типе связи делается на основании небольшого различия величин лишь одного β -эффекта; в этих случаях желательное независимое подтверждение типа связи между моносахаридными компонентами. Не исключено, что при дальнейшей работе будут обнаружены и другие трудные случаи, а также случаи

Использование уточненных значений α - и β -эффектов гликозилирования для выбора типа связи S2 \rightarrow S1 при неоднозначном ответе программы ANMROL

№ п. п.	Соединение	Тип связи S2 \rightarrow S1		Характерные эффекты			
		истинный	альтернативный	k	m	E^* _{каж}	E^{2*} _{ист}
1	(I): S1 = D-Gal(α) S2 = D-Man	β 1-3	β 1-4 β 1-2	1	3	-1,7	0,3
				1	3	-1,7	-1,0
2	(I): S1 = D-Gal(β) S2 = D-Man	β 1-3	β 1-2	1	2	9,6	7,8
				1	1	-0,2	-1,4
3	(I): S1 = D-Gal(α) S2 = D-Glc	α 1-4	α 1-3	2	1	8,0	3,3
				1	3	9,5	4,9
				1	4	-0,6	-3,7
4	(I): S1 = D-Gal(α) S2 = L-Rha	α 1-3	α 1-4 α 1-2	2	1	8,4	5,6
				1	1	0,1	-0,5
5	(I): S1 = D-Glc(α) S2 = D-Man	α 1-3	α 1-4 α 1-2	1	4	10,1	7,9
				1	3	-2,3	0,3
				2	1	6,8	3,4
				1	2	8,3	4,0
				1	1	0,1	-1,8
				1	3	-2,3	-1,3
6	(I): S1 = D-Glc(β) S2 = D-Man	α 1-2	α 1-3	2	1	5,6	6,7
				1	3	3,7	6,8
				1	2	0,3	-1,2
7	(IV): S1 = D-Gal(α) S2 = L-Rha S3(1) = D-Glc(α 1-	α 1-3	α 1-2	1	2	6,3	8,6
				1	1	0,2	-0,5
8	(IV): S1 = D-Gal(β) S2 = L-Rha S3(1) = D-Glc(α 1-	α 1-3	α 1-2	1	1	0,2	-1,3
9	(V): S1 = D-Gal(β) S2 = L-Rha S3(1) = Glc(α 1- R = Me	α 1-3	α 1-2	1	1	0,2 ^{3*}	-1,3 ^{3*}

* Вычисленные кажущиеся значения $E(k, m)$ при выборе альтернативного типа связи.

^{2*} Уточненные значения $E(k, m)$, ожидаемые для альтернативного типа связи (по [5]).

^{3*} Эффекты замещения относительно метил- β -D-галактозида.

неоднозначной идентификации типа связи с помощью программы ANMROL даже при использовании уточненных значений эффектов гликозилирования.

Отнесение сигналов в спектре ^{13}C -ЯМР не является целью производимых вычислений, а используется лишь для того, чтобы идентифицировать невозмущенные сигналы спектра и сделать возможным последующий анализ изменений химического сдвига возмущенных сигналов. В целом, при поиске невозмущенных сигналов по условию (2) при достаточно большом различии величин $X(k, m)$ для разных атомов углерода правильность идентификации сигналов не вызывает сомнений. Однако в противном случае результат отнесения зависит от последовательности, в которой выполняются операции поиска, и соседние по номеру сигналы $Y(j)$ могут быть «перепутаны». Ошибка в идентификации сигналов может, далее, возникнуть в тех случаях, когда химический сдвиг одного из невозмущенных сигналов в экспериментальном спектре несколько изменен, например, за счет конформационных факторов. Наконец, может сложиться и такая ситуация, что один из сигналов, возмущенных за счет β -эффектов гликозилирования, при поиске по условию (2) будет ошибочно принят за невозмущенный, так как окажется ближайшим к одному из сигналов $X(k, m)$ модельного набора. Это при дальнейшей работе должно привести к неправильному значению для кажущегося β -эффекта гликозилирования.

Для оценки того, как часто встречаются такого рода ошибки отнесения и насколько существенно их значение, мы сравнили отнесения сигналов в спектрах ^{13}C -ЯМР для ряда метилгликозидов — производных дисахаридов и разветвленных трисахаридов, проведенные по данным гетероядерного двойного резонанса и с помощью программ ANMROl. Полная сводка обнаруженных различий приведена в примечаниях к таблице спектральных данных, включенной в работу [16].

Оказалось, что неточная идентификация сигналов — явление, встречающееся заметно чаще, чем неоднозначность определения типа связи: из 168 сигналов, входящих в состав рассматриваемой серии спектров, ошибочное отнесение было обнаружено в 26 случаях. Обычно имеет место перестановка соседних невозмущенных сигналов, что не оказывает никакого влияния на ход дальнейших вычислений. В шести случаях было показано ошибочное отнесение сигналов, возмущенных за счет β -эффектов, к невозмущенным сигналам. Однако, как правило, изменение величины кажущегося эффекта гликозилирования в результате этого невелико и лежит в пределах от 0,1 до 0,5 м. д.; лишь в одном случае оно было более значительным (1,2 м. д.). Таким образом, ошибки в отнесении сигналов, допускаемые программой ANMROl, не имеют существенного значения для определения типа связи; более того, найденные с помощью программы величины эффектов гликозилирования являются обычно достаточно хорошим приближением к истинным значениям.

Программа ANMROl использует значительно меньшее количество параметров, чем программы для анализа структуры линейных полисахаридов по данным ^{13}C -ЯМР-спектров [4—6], и вследствие этого результаты расчетов с ее помощью заметно меньше зависят от точности определения этих параметров. Это делает особенно плодотворным использование программы ANMROl в тех случаях, когда закономерности, определяющие зависимость величин эффектов гликозилирования от структуры, мало изучены, например для разветвленных олиго- и полисахаридов. Информации, получаемой с помощью рассматриваемой программы, вполне достаточно при синтетической работе, а легкость проведения вычислений с помощью микро-ЭВМ — ее несомненное достоинство.

ЛИТЕРАТУРА

1. Шашков А. С., Чижов О. С. // Биоорган. химия. 1976. Т. 2. № 4. С. 437—497.
2. Bock K., Pedersen C., Pedersen H. // Adv. in Carbohydr. Chem. and Biochem. 1984. V. 42. P. 193—226.
3. Bradbury J. H., Jenkins G. A. // Carbohydr. Res. 1984. V. 126. № 1. P. 125—156.
4. Липкинд Г. М., Шашков А. С., Кочетков Н. К. // Биоорган. химия. 1987. Т. 13. № 6. С. 833—841.
5. Lipkind G. M., Shashkov A. S., Knirel Y. A., Vinogradov E. V., Kochetkov N. K. // Carbohydr. Res. 1988. V. 175. № 1. P. 59—75.
6. Jansson P.-E., Kenne L., Wildwalm G. // Carbohydr. Res. 1987. V. 168. № 1. P. 67—77.
7. Kochetkov N. K., Chizhov O. S., Shashkov A. S. // Carbohydr. Res. 1984. V. 133. № 2. P. 173—185.
8. Торгов В. И., Паносян К. А., Шibaев В. Н. // Биоорган. химия. 1986. Т. 12. № 4. С. 559—561.
9. Торгов В. И., Паносян К. А., Шibaев В. Н. // Биоорган. химия. 1986. Т. 12. № 4. С. 562—564.
10. Torgov V. I., Panossian C. A., Shibaev V. N. // Carbohydr. Res. 1987. V. 161. № 1. P. 97—112.
11. Торгов В. И., Шibaев В. Н., Шашков А. С., Кочетков Н. К. // Биоорган. химия. 1980. Т. 6. № 12. С. 1860—1871.
12. Торгов В. И., Кудашова О. В., Шibaев В. Н., Кочетков Н. К. // Биоорган. химия. 1982. Т. 8. № 4. С. 114—119.
13. Торгов В. И., Шibaев В. Н., Кочетков Н. К. // Биоорган. химия. 1984. Т. 10. № 7. С. 946—953.
14. Торгов В. И., Паносян К. А., Смельянский А. Т., Шibaев В. Н. // Биоорган. химия. 1985. Т. 11. № 1. С. 83—90.
15. Торгов В. И., Дружинина Т. Н., Нечаев О. А., Шibaев В. Н. // Биоорган. химия. 1987. Т. 13. № 7. С. 947—957.
16. Нечаев О. А., Торгов В. И., Шibaев В. Н., Мамян С. С. // Биоорган. химия. 1988. Т. 14. № 3. С. 359—369.

Поступила в редакцию
27.1.1988

MICROCOMPUTER ANALYSIS OF ^{13}C NMR SPECTRA
OF OLIGOSACCHARIDES

SHIBAEV V. N.

*N. D. Zelinsky Institute of Organic Chemistry,
Academy of Sciences of the USSR, Moscow*

A microcomputer programme, ANMROL, has been developed to check whether structure of an oligo- or polysaccharide may be elucidated unambiguously from the ^{13}C NMR data. Principles of the programme are described and its application for the series of linear and branched oligosaccharides is discussed.