



БИООРГАНИЧЕСКАЯ ХИМИЯ

Том 19 * № 10 * 1993

УДК 543.422.25 : 547.458.057 : 577.114.012

© 1993 Н. Э. Нифантьев, А. С. Шашков,
Г. М. Липкинд, Н. К. Кочетков, Б. Янн *, К. Янн *

КОМПЬЮТЕРНЫЙ СТРУКТУРНЫЙ АНАЛИЗ О-СПЕЦИФИЧЕСКИХ ПОЛИСАХАРИДОВ *Escherichia coli* O1A, O1B И O1C

Институт органической химии им. Н. Д. Зелинского РАН, Москва;
* Институт иммунобиологии общества Макса Планка, Фрайбург, ФРГ

Проведен компьютерный анализ спектров ^{13}C -ЯМР О-специфических полисахаридов *Escherichia coli* O1A, O1B и O1C, основанный на данных моносахаридного состава и результатах анализа методом метилирования. Анализ спектра полисахарида O1A позволил однозначно установить структуру его повторяющегося звена. В случае полисахарида O1C компьютерный расчет выделил группу структурных кандидатов, из которых правильная структура была определена на основании результата единого эксперимента по ядерным эффектам Оверхаузера. Компьютерный анализ спектра полисахарида O1B был неуспешным из-за присутствия в его структуре фрагмента α -L-Rha-(1 \rightarrow 2)- α -D-Gal-(1 \rightarrow 3)-D-GlcNAc с терминальной (1 \rightarrow 2)-связью, спектральные данные для которого не могут быть рассчитаны по аддитивной схеме только с учетом эффектов гликозилирования. Показано, что для расчета спектров полисахаридов, содержащих подобные фрагменты, необходимы дополнительные данные по отклонениям от аддитивности величин химических сдвигов сигналов углеродов в спектрах ^{13}C -ЯМР линейных трисахаридов с терминальной (1 \rightarrow 2)-связью.



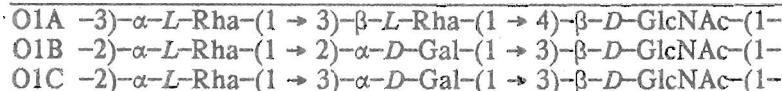
2

1

1



R

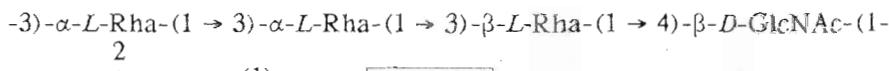


Недавно был предложен метод компьютерного структурного анализа регулярных вицинально разветвленных полисахаридов [1, 2]. Этот метод базируется на компьютерной обработке спектра ^{13}C -ЯМР исследуемого полисахарида, проводимой с учетом данных по моносахаридному составу и результатов анализа методом метилирования. Метод включает три основных этапа: 1) генерирование всех возможных структур полисахарида с данным моносахаридным составом и типами замещения (определяются с помощью метилирования), 2) расчет спектров ^{13}C -ЯМР для каждой из этих структур, 3) поиск структуры, чей рассчитанный спектр лучше всего совпадает с экспериментальным. Такой поиск основан на сравнении S-факторов для каждой генерированной структуры. S-Фактором является сумма квадратов различий величин химических сдвигов соответствующих сигналов в

рассчитанном и экспериментальном спектрах, нормированная на один моносахаридный остаток [1, 2].

Данный подход был недавно использован для определения структуры повторяющегося звена О-специфического полисахарида *Pseudomonas solanacearum*, штамм 7944 [3]. В настоящей работе мы рассматриваем результаты компьютерного структурного анализа О-специфических полисахаридов *Escherichia coli* O1A, O1B и O1C, проведенного параллельно со структурным анализом этих полисахаридов традиционными химическими методами [4, 5]. Недавно [6] строение полисахарида O1A было установлено с использованием химических методов и компьютерного расчета по программе CASPER [7, 8], что позволило нам также провести и сравнение обоих методов компьютерного анализа.

Анализ спектра ^{13}C -ЯМР полисахарида O1A был проведен с учетом того, что его повторяющееся звено содержит один 2,3-дизамещенный остаток *L*-Rha, два 3-замещенных остатка *L*-Rha, один 4-замещенный остаток *D*-GlcNAc и один незамещенный терминальный остаток *D*-ManNAc [4]. Компьютерный расчет выявил группу структур (1–10), которые удовлетворяют данным анализа методом метилирования. Среди кандидатов структура (1) характеризовалась существенно меньшим *S*-фактором (0,4), чем другие структуры (его спектр ^{13}C -ЯМР см. в таблице), и это свидетельствовало о том, что структура (1) — повторяющееся звено полисахарида O1A. Этот вывод совпадает с результатами химического анализа данного полисахарида [3, 6].

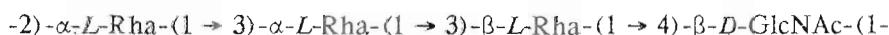


2

(1)

$\boxed{\text{S } 0,4}$

$\beta-D-\text{ManNAc}$

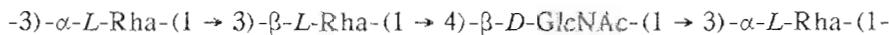


3

(2)

$\boxed{\text{S } 0,8}$

$\beta-D-\text{ManNAc}$

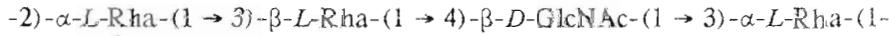


2

(3)

$\boxed{\text{S } 0,9}$

$\beta-D-\text{ManNAc}$

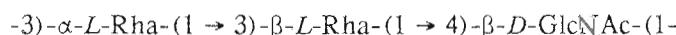


2

(4)

$\boxed{\text{S } 1,0}$

$\beta-D-\text{ManNAc}$

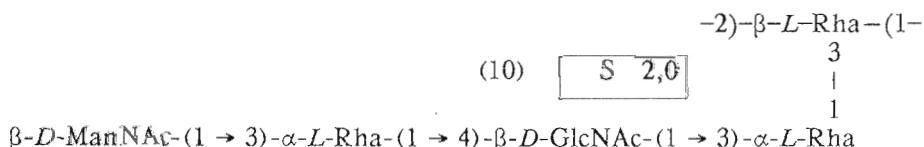
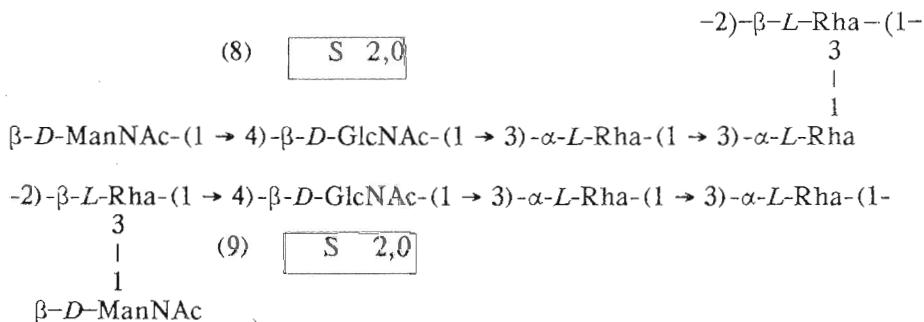
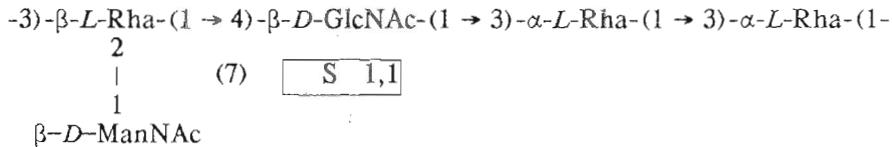
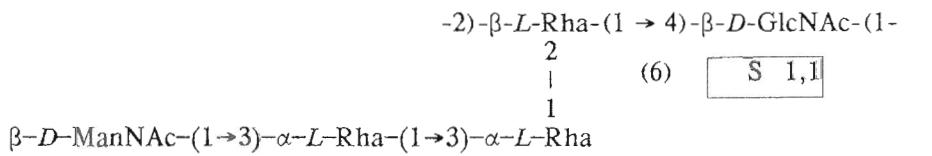


2

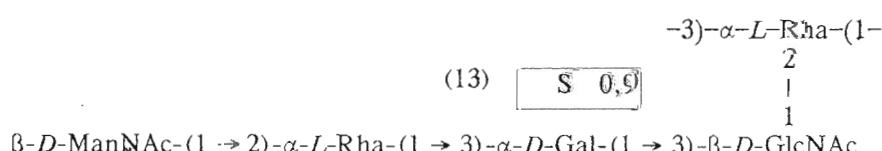
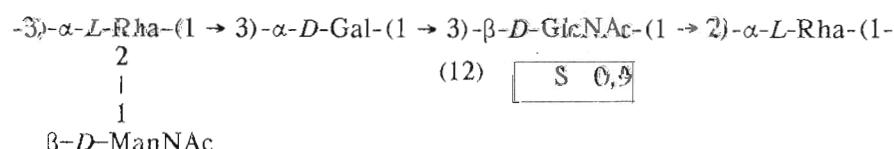
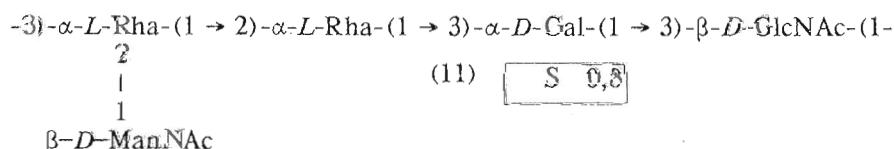
(5)

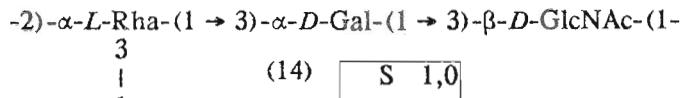
$\boxed{\text{S } 1,0}$

$\beta-D-\text{ManNAc}-(1 \rightarrow 3)-\alpha-L-\text{Rha}$

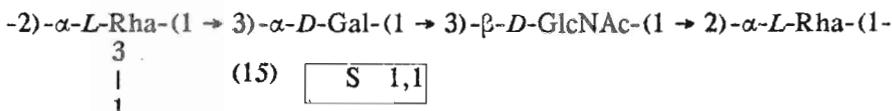


Структуры (1)–(10), как возможные повторяющиеся звенья полисахарида О1А, были также предсказаны и при расчете по программе CASPER [6]. Интересно, однако, отметить, что вероятность этих структур уменьшалась в несколько ином порядке, $(1) > (6) > (2) = (7) > (3) = (8) > (4) = (10) > (9) > (5)$, а также то, что различие между вероятностями наиболее предпочтительных структур (1) и (6) не было столь значительным, как при расчете по методу, использующемуся в данной работе.

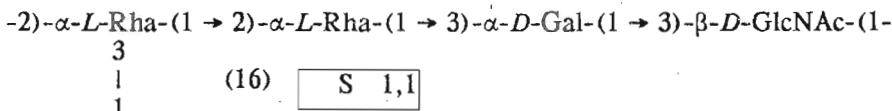




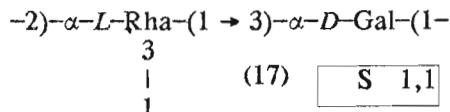
$\beta-D\text{-ManNAc}-(1 \rightarrow 2)-\alpha-L\text{-Rha}$



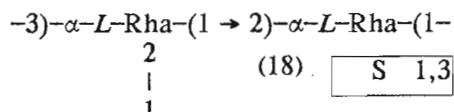
$\beta-D\text{-ManNAc}$



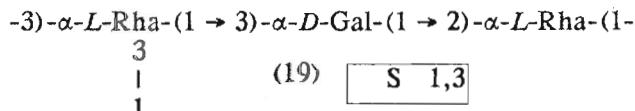
$\beta-D\text{-ManNAc}$



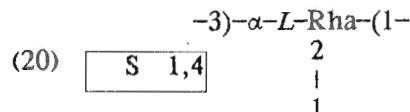
$\beta-D\text{-ManNAc}-(1 \rightarrow 3)-\beta-D\text{-GlcNAc}-(1 \rightarrow 2)-\alpha-L\text{-Rha}$



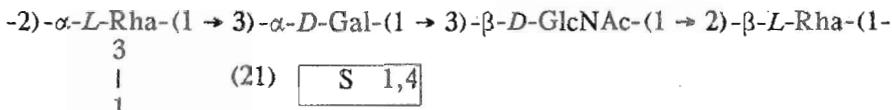
$\beta-D\text{-ManNAc}-(1 \rightarrow 3)-\alpha-D\text{-Gal}-(1 \rightarrow 3)-\beta-D\text{-GlcNAc}$



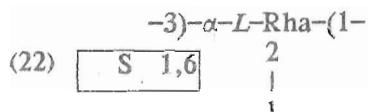
$\beta-D\text{-ManNAc}-(1 \rightarrow 3)-\beta-D\text{-GlcNAc}$



$\beta-D\text{-ManNAc}-(1 \rightarrow 3)-\alpha-D\text{-Gal}-(1 \rightarrow 3)-\beta-D\text{-GlcNAc}-(1 \rightarrow 2)-\alpha-L\text{-Rha}$



$\beta-D\text{-ManNAc}$



$\beta-D\text{-ManNAc}-(1 \rightarrow 3)-\beta-D\text{-GlcNAc}-(1 \rightarrow 2)-\alpha-L\text{-Rha}-(1 \rightarrow 3)-\alpha-D\text{-Gal}$

При компьютерном анализе данных спектра ^{13}C -ЯМР полисахаридов О1С принималось во внимание, что его повторяющееся звено содержит один 2,3-дизамещенный остаток $L\text{-Rha}$, один 2-замещенный остаток $L\text{-Rha}$, по одному 3-замещенному остатку $D\text{-Gal}$ и $D\text{-GlcNAc}$ и один незамещенный терминальный остаток ManNAc . В результате в качестве наиболее вероятных была выявлена

Наблюдаемые [4, 5] и рассчитанные (в скобках) величины химических сдвигов (δ , м. д.) сигналов в спектрах ^{13}C -ЯМР О-специфических полисахаридов *Escherichia coli* O1A, O1B и O1C

Моносахаридное звено	C1	C2	C3	C4	C5	C6
<i>E. coli</i> O1A						
-2,3)- α -L-Rha-(1-	102,5 (102,9)	78,1 (78,2)	80,3 (80,3)	72,5 (72,1)	70,5 (70,2)	17,8 (18,1)
-3)- α -L-Rha-(1-	103,2 (103,5)	71,1 (71,4)	79,5 (79,1)	72,6 (72,9)	70,4 (70,0)	17,9 (18,0)
-3)- β -L-Rha-(1-	101,6 (101,9)	71,7 (71,4)	81,5 (81,5)	72,3 (72,5)	73,3 (73,2)	17,9 (18,0)
-4)- β -D-GlcNAc-(1-	103,2 (103,6)	57,3 (57,0)	74,8 (75,2)	78,2 (77,9)	75,8 (75,8)	61,7 (61,8)
β -D-ManNAc-(1-	101,1 (101,3)	54,3 (54,3)	73,3 (73,3)	68,0 (67,9)	77,3 (77,6)	62,2 (62,1)
<i>E. coli</i> O1B *						
-2,3)- α -L-Rha-(1-	102,2 (102,7)	78,1 (78,2)	80,0 (80,3)	72,4 (72,1)	70,6 (69,9)	17,9 (18,1)
	[102,7]	[79,2]	[80,3]	[72,1]	[69,9]	[18,1]
-2)- α -L-Rha-(1-	101,2 (101,3)	78,8 (78,8)	71,2 (71,3)	73,4 (73,5)	70,4 (70,0)	17,9 (18,0)
	[101,7]	[79,8]	[71,3]	[73,5]	[70,0]	[18,0]
-2)- α -D-Gal-(1-	98,6 (100,3)	74,8 (78,2)	70,2 (69,5)	70,8 (70,6)	72,1 (72,2)	61,9 (62,4)
	[99,1]	[75,1]	[70,2]	[70,6]	[72,2]	[62,4]
-3)- β -D-GlcNAc-(1-	102,8 (103,9)	55,9 (55,8)	78,6 (81,9)	72,7 (71,9)	76,8 (77,2)	61,9 (62,1)
	[103,9]	[55,8]	[79,2]	[71,9]	[77,2]	[62,1]
β -D-ManNAc-(1-	101,0 (100,8)	54,2 (54,3)	73,3 (73,3)	68,1 (67,9)	77,4 (77,6)	61,8 (62,1)
	[101,3]	[54,3]	[73,3]	[67,9]	[77,6]	[61,8]
<i>E. coli</i> O1C						
-2,3)- α -L-Rha-(1-	102,2 (102,7)	78,2 (78,2)	80,0 (80,3)	72,5 (72,1)	70,3 (69,9)	18,0 (18,1)
-2)- α -L-Rha-(1-	101,7 (102,1)	78,7 (78,8)	71,4 (71,3)	73,7 (73,5)	70,7 (70,0)	18,0 (18,0)
-3)- α -D-Gal-(1-	100,6 (101,0)	69,4 (68,9)	77,7 (78,5)	70,1 (70,2)	72,2 (72,2)	61,9 (62,4)
-3)- β -D-GlcNAc-(1-	103,0 (103,9)	55,7 (55,8)	81,5 (81,9)	72,2 (71,9)	76,7 (77,2)	61,8 (61,8)
β -D-ManNAc-(1-	101,0 (101,3)	54,3 (54,3)	73,4 (73,3)	68,3 (67,9)	77,5 (77,6)	62,1 (62,1)

* В квадратных скобках приведены химические сдвиги, полученные при расчете с учетом отклонений от аддитивности в спектре трисахарида (29) (см. выше).

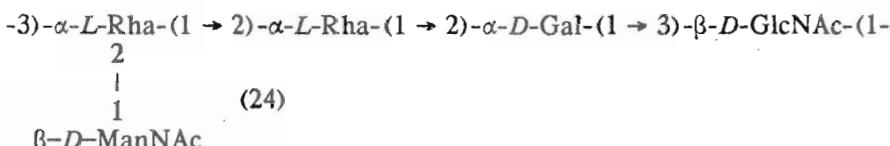
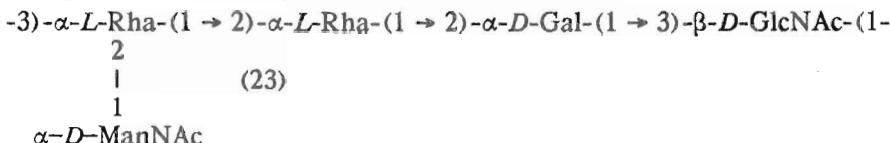
группа повторяющихся звеньев (11)—(22), которые удовлетворяли данным анализа методом метилирования.

Выявленные структуры характеризовались очень близкими значениями S-факторов, что не позволяло достаточно корректно выделить среди них наиболее вероятную без привлечения дополнительных данных. Легко видеть, что среди структур (11)—(22) только повторяющееся звено (11), характеризующееся наименьшим S-фактором, содержит (1 \rightarrow 3)-связанный глюказаминил-рамнозидный

фрагмент. Поэтому представлялось рациональным подтвердить в первую очередь присутствие такого фрагмента в структуре повторяющегося звена полисахарида O1C. Для этого в спектре ^1H -ЯМР полисахарида O1C был идентифицирован сигнал H-1 глюкозаминового остатка (полностью данные спектров ЯМР полисахарида O1C приведены в работе [5]). Предоблучение H-1GlcNAc выявило [5] его межмоносахаридный контакт только с H-3 рамнозного остатка, что указывало на присутствие фрагмента GlcNAc-(1 → 3)-Rha в полисахариде O1C и однозначно подтверждало структуру (11) в качестве его повторяющегося звена.

Таким образом, компьютерный анализ данных спектров ^{13}C -ЯМР полисахаридов O1A и O1C был успешным. В первом случае компьютерный расчет позволил установить повторяющееся звено, а во втором значительно сократить перечень возможных структур, из которых правильная была выделена на основании результата одного эксперимента по ядерным эффектам Оверхаузера.

Иная ситуация наблюдалась при компьютерном анализе данных спектра ^{13}C -ЯМР полисахарида O1B, повторяющееся звено которого, согласно данным моносахаридного анализа и анализа методом метилирования, содержит [5] один 2,3-дизамещенный остаток *L*-Rha, по одному 2-замещенному остатку *L*-Rha и *D*-Gal, один 3-замещенный остаток *D*-GlcNAc и незамещенный терминальный остаток *D*-ManNAc. Компьютерный расчет данных спектра полисахарида O1B не выявил ни одной структуры, чей рассчитанный спектр совпадал бы удовлетворительно [2] с экспериментальным. Расхождение экспериментального и рассчитанных спектров было столь значительным, что даже повторяющееся звено (23), характеризующееся наименьшим *S*-фактором (2,2), не могло рассматриваться как вероятное. В целом в результате компьютерного анализа было симулировано более 50 структурных кандидатов с $S < 7,5$, среди которых была и правильная структура (24) ($S 6,6$), установленная химическими методами [5].



Наибольшая разница в рассчитанном и экспериментальном спектрах полисахаридов O1B отмечена (см. таблицу) в случае сигналов галактозного ($\Delta\delta C_2 -3,4$ м.д.) и глюкозаминового ($\Delta\delta C_3 -3,3$ м.д.) остатков, которые включены в трисахаридный фрагмент (25) с терминальной ($1 \rightarrow 2$)-связью. Аналогичные отклонения от аддитивности отмечались ранее для трисахаридных фрагментов полисахаридов *Streptococcus pneumoniae* 18A [8] (26), *Proteus penneri* 42 [9] (27) и *Salmonella arizonaiae* 045 [10] (28), которые также содержат терминальную ($1 \rightarrow 2$)-связь.



Проведенное нами ранее [10] исследование ядерных эффектов Оверхаузера для метилгликозида трисахарида (28) и составляющих его дисахаридов показало, что наблюдаемые спектральные аномалии в трисахаридах рассматриваемого типа, как и в случае винильально разветвленных трисахаридов (см., например, [11]),

определенены различиями конформаций трисахаридов и образующих их дисахаридных фрагментов.

Таким образом, компьютерный анализ спектра ^{13}C -ЯМР полисахарида O1B еще раз показал [8—10, 12], что присутствие в полисахаридной цепи ($1 \rightarrow 2$)-связи может ограничивать в некоторых случаях применение рассмотренного метода компьютерного структурного анализа. Можно, однако, ожидать, что модификаций компьютерной программы и дополнение баз спектральных данных, используемых в расчетах, величинами отклонений от аддитивности в линейных трисахаридах с терминальной ($1 \rightarrow 2$)-связью позволят в дальнейшем проводить структурный анализ полисахаридов, как линейных, так и разветвленных, содержащих фрагменты типа (25)–(28).

Сказанное можно проиллюстрировать результатами расчета спектра полисахарида O1B, проведенного с учетом отклонений от аддитивности, наблюдавшихся в спектре трисахарида (29) [13]



содержащего, как и полисахарид O1B, фрагмент (25). Использованные в расчете величины отклонений от аддитивности (для C1 Rha +3,4 м. д., для C1, C2 Gal -2,7 и -3,3 м.д., а также для C3, C4 GlcNAc -2,7 и 1,8 м. д. соответственно) получены сравнением данных спектра трисахарида (29) и данных теоретически рассчитанного для него спектра по методу [14]. Полученные в этом случае величины химических сдвигов атомов углерода для полисахарида O1B приведены в таблице в квадратных скобках. Необходимо отметить, что проведенный расчет был успешным — правильное повторяющееся звено (24) характеризовалось в данном случае наименьшим S-фактором, равным 1,1, а повторяющееся звено (23), вероятность которого была на втором месте, имело S-фактор 2,0.

В заключение отметим, что практическое использование рассмотренной модификации компьютерного структурного анализа полисахаридов будет возможно после накопления достаточной базы спектральных данных для линейных трисахаридов с терминальной ($1 \rightarrow 2$)-связью. Недавно были уже опубликованы синтез и спектральные исследования некоторых из необходимых модельных трисахаридов [13, 15–17].

Экспериментальная часть

Компьютерный анализ данных спектров ^{13}C -ЯМР полисахаридов проводился в недиалоговом режиме с использованием персонального компьютера IBM PC/AT 286 (Tandon), как описано ранее [2]. Программа написана на языке Турбо Паскаль 5.5.

СПИСОК ЛИТЕРАТУРЫ

1. Lipkind G. M., Shashkov A. S., Kochetkov N. K. // Carbohydr. Res. 1990. V. 198. № 2. P. 399—402.
2. Lipkind G. M., Nifant'ev N. E., Shashkov A. S., Kochetkov N. K. // Carbohydr. Res. 1992. V. 237. P. 11—22.
3. Kocharova N. A., Knirel Y. A., Shashkov A. S., Nifant'ev N. E., Kochetkov N. K., Varbanets L. D., Moskalenko N. V., Brovarskaya O. S., Muras V. A., Young J. M. // Carbohydr. Res. 1993. In Press.
4. Jann B., Shashkov A. S., Gupta D. S., Panasenko S. M., Jann K. // Carbohydr. Polymers. 1992. V. 18. P. 51—57.
5. Gupta D. S., Shashkov A. S., Jann B., Jann K. // J. Bacteriol. 1992. V. 174. № 24. P. 7963—7970.
6. Baumann H., Jansson P.-E., Kenne L., Widmalm G. // Carbohydr. Res. 1991. V. 211. № 2. P. 183—190.
7. Jansson P.-E., Kenne L., Widmalm G. // Carbohydr. Res. 1987. V. 168. № 1. P. 67—77.
8. Jansson P.-E., Kenne L., Widmalm G. // Carbohydr. Res. 1989. V. 188. N 1. P. 169—191.
9. Sidorszyk Z., Vinogradov E. V., Shashkov A. S., Zych K., Knirel Y. A., Kochetkov N. K. // Eur. J. Biochem. In press.

10. Shashkov A. S., Vinogradov E. V., Knirel Y. A., Nifant'ev N. E., Kochetkov N. K., Dabrowski J., Kholodkova E. V., Stanislavsky E. S. // Carbohydr. Res. 1993. V. 241. P. 177–188.
11. Kochetkov N. K., Lipkind G. M., Shashkov A. S., Nifant'ev N. E. // Carbohydr. Res. 1991. V. 221. P. 145–168.
12. Кочетков Н. К., Виноградов Е. В., Книрель Ю. А., Шашков А. С., Липкинд Г. М. // Биоорганическая химия. 1992. Т. 18. № 1. С. 116–125.
13. Kovac P., Edgar K. J. // J. Org. Chem. 1992. V. 57. N 8. P. 2455–2467.
14. Lipkind G. M., Shashkov A. S., Knirel Y. A., Vinogradov E. V., Kochetkov N. K. // Carbohydr. Res. 1988. V. 175. P. 59–75.
15. Adeyeye A., Jansson P.-J., Kenne L., Magnusson G. // J. Chem. Soc. Perkin Trans. 2. 1991. P. 963–973.
16. Nifant'ev N. E., Shashkov A. S., Kochetkov N. K. // Carbohydr. Res. 1992. V. 226. P. 331–336.
17. Торгов В. И., Ненаев О. А., Усев А. И., Шибаев В. Н. // Биоорганическая химия. 1981. Т. 17. № 3. С. 424–426

Поступила в редакцию
25.III.1993

N. E. Nifant'ev, A. S. Shashkov, G. M. Lipkind,
N. K. Kochetkov, B. Jann *, K. Jann *

COMPUTER-ASSISTED STRUCTURAL ANALYSIS
OF THE *Escherichia coli* O1A, O1B, AND O1C SPECIFIC
POLYSACCHARIDES

N. D. Zelinsky Institute of Organic Chemistry,
Russian Academy of Sciences, Moscow;

* Max-Planck Institute of Immunobiology, Freiburg, Germany

A computer evaluation of ^{13}C -NMR data for the title polysaccharides based on the monosaccharide and methylation analysis data led to the structure of the repeating unit of the O1A polysaccharide as well as to several probable structures of the O1C polysaccharide, of which the correct one was inferred by means of a single NOE experiment. The analysis of the spectrum of the O1B polysaccharide was unsuccessful, due to the presence in its structure of the fragment $\alpha\text{-L-Rha-(1} \rightarrow 2\text{)-}\alpha\text{-D-Gal-(1} \rightarrow 3\text{)-D-GlcNAc}$ with the terminal (1 \rightarrow 2)-linkage, whose spectral data could not be calculated by additive schemes using only glycosylation effects. However in reevaluation of the O1B spectral data by taking into account the deviations from additivities of the chemical shifts values in spectra of the related trisaccharides, to reveal the most probable structure of the O1B's repeating unit.



2

1

1



R

