Carbohydrate Research, 28 (1973) 21–29

STRUCTURAL ANALYSIS OF OLIGOSACCHARIDES BY MASS SPECTROMETRY: ACETATES OF *N*-ARYLGLYCOSYLAMINES

O. S. CHIZHOV, N. N. MALYSHEVA, AND N. K. KOCHETKOV

N. D. Zelinsky Institute of Organic Chemistry, Academy of Sciences, Leninsky Prospect 47, Moscow B-334 (U. S. S. R.)

(Received August 3rd, 1972; accepted for publication, October 18th, 1972)

ABSTRACT

The electron-impact mass spectra of the acetylated *N*-arylglycosylamine derivatives of tri-, tetra-, and penta-saccharides allow the determination of molecular weight, the nature and sequence of the monosaccharide residues, and, in certain instances, the position of some of the interglycosidic linkages.

RESULTS AND DISCUSSION

It has been shown[1] that the electron-impact mass spectra of *O*-acetyl-*N*-arylglycosylamines of disaccharides can be used for the determination of molecular weight, identification of the reducing and non-reducing units, and, in some instances, for determination of the position of the glycosidic linkages. We now report an extension of this work to the acetylated *N*-arylglycosylamine derivatives of higher oligosaccharides.

The mass spectra of the acetylated *N*-arylglycosylamine derivatives of the tri-, tetra-, and penta-saccharides (**1–15**) were studied in order to evaluate the main features of the fragmentation pattern under electron impact. The relevant data, shown in Tables I–IV, indicate that for each compound there is a moderately intense peak for M^+. Hence, the molecular weight can be determined directly from the mass spectrum. The abundance of M^+ ions for the higher oligosaccharides is smaller than for disaccharides[1,2], but even so, for pentasaccharides (**13–15**), the M^+ peaks have significant intensity.

TABLE I

MASS-SPECTRAL DATA FOR COMPOUNDS **1–5**

m/e	*Intensity* (%)					m/e	*Intensity* (%)				
	1	**2**	**3**	**4**	**5**		**1**	**2**	**3**	**4**	**5**
81	13	70	17	61	32	99	4	14	4	17	9
91	1.3	24	7	23	21	103	5	14	5	20	16
97	13	22	11	25	28	106	2	45	7	23	12
98	4	28	11	25	23	107	21	67	15	34	16

TABLE I (*Continued*)

m/e	*Intensity* (%)					m/e	*Intensity* (%)				
	1	2	3	4	5		1	2	3	4	5
109	31	80	47	84	65	258	2	3	2	3	0.7
111	4	12	6	5	5	259	10.5	—	—	—	—
112	3	13	6	17	12	271	10	14	12	1.5	7
115	13	21	7	30	28	286	—	—	—	0.5	0.7
118	3	11	7	19	21	289	5	2	5	8	8
120	2	27	—	14	6	304	—	—	—	1	0.7
127	18	30	15	36	28	317	10.5	6	4	14	5
135	9	10	6	17	6	318	5	5	4	5	1.5
136	9	27	5	34	12	331	79	19	33	86	27
139	21	12	8	19	14	335	3	—	—	—	—
140	—	9	5	12	16	346	—	—	—	1	1
141	5	22	6	23	9	364	—	—	—	0.5	0.5
145	8	17	6	19	14	377	—	—	0.7	1.5	0.7
149	5	33	11	17	14	378	2	3	1.8	3	0.6
157	8	10	6	19	16	395	1	—	—	—	—
169	100	100	100	100	100	397	1.6	—	0.8	—	0.1
174	2	6	3	6	3	398	1	2	2.3	—	0.1
178	2	17	2	7	4	406	3	1	0.4	0.7	0.2
186	2	4	—	6	2	423	0.7	1.2	0.8	1.4	0.5
187	5	5	3.5	7	5	437	0.8	—	—	—	—
191	13	11	8	9	5	457	1.6	2.3	2	0.8	0.5
197	2	2	3	11	3	479	1.5	—	—	—	—
198	3	3	8	5	2	531	—	1.5	0.6	0.1	0.1
204	9	—	—	—	—	559	1.6	0.5	2.4	0.2	1.4
211	13	9	12	6	7	606	4	0.9	0.6	0.15	0.2
213	1.5	4	3.5	—	2	619	12	4	3	0.3	1
215	2	4	3	12	3.5	666	0.4	1.5	0.9	0.15	0.23
216	1	4	4	6	3	833	—	—	—	0.3	0.2
229	10	12	7	8	7	893	—	—	—	0.2	0.15
242	3	4	3.5	8	7	907	0.6	0.5	0.6	0.1	0.2
243	3	13	2	8	4	953	0.6	0.5	0.4	0.4	0.2
257	—	4	1.4	8	0.8						

TABLE II

MASS-SPECTRAL DATA FOR COMPOUNDS 6–8

m/e	*Intensity* (%)			m/e	*Intensity* (%)		
	6	7	8		6	7	8
81	4	48	10	107	7	22	37
91	7	9	9	109	—	77	16
97	100	61	14	115	22	26	9
98	14	22	8	118	5	6	6
99	—	13	4	127	16	30	8
103	13	17	4	128	50	—	6
106	4	7	37	135	3	13	8

TABLE II (*Continued*)

m/e	*Intensity* (%)			m/e	*Intensity* (%)		
	6	7	8		6	7	8
136	4	22	8	402	0.3	—	—
139	65	91	24	406	—	2	1
141	6	26	5	415	0.7	—	—
144	3	—	—	423	—	1	1.5
145	6	—	—	433	3.5	—	—
149	5	13	58	462	1	—	—
157	100	43	17	475	6	—	—
169	—	100	25	522	0.5	—	—
170	50	—	11	534	—	1.5	—
178	—	5	3	547	—	2.2	2.5
186	2	—	—	565	0.2	—	—
187	2	4	2	589	0.2	—	—
191	2	11	5	594	—	2	—
199	42	18	10	611	0.06	—	—
211	2	9	2.5	606	—	—	2.5
216	—	7	2.5	617	0.35	—	—
217	42	7	2.5	631	0.4	—	—
229	1.4	7	2.5	641	—	1.3	—
243	1	3	4	649	0.06	—	—
245	11	13	1.3	653	0.1	—	—
246	3	3	0.7	666	—	—	1.2
259	56	—	80	677	0.1	—	—
271	1	9	—	691	0.3	—	—
286	—	4.5	—	695	0.1	—	—
289	1	—	—	737	0.3	—	—
304	—	2	—	754	—	—	0.6
306	1.4	—		755	0.3	—	—
313	3	—	—	761	—	0.4	—
317	1	—	7.5	797	3	—	—
318	—	—	6	821	—	1	2.5
331	6	30	—	839	—	—	1.2
346	—	8	—	881	—	1.3	1.2
373	1.3	—	—	899	—	—	3
377	—	3	—	941	—	20	100
378	—	2	6				

TABLE III

MASS-SPECTRAL DATA FOR COMPOUNDS 9–12

m/e	*Intensity* (%)				m/e	*Intensity* (%)			
	9	10	11	12		9	10	11	12
81	38	58	22	7	99	12	18	7	—
91	17	28	22	15	103	25	33	9	—
97	45	58	22	100	106	48	17	20	21
98	28	60	43	21	107	48	32	29	23

TABLE III (*Continued*)

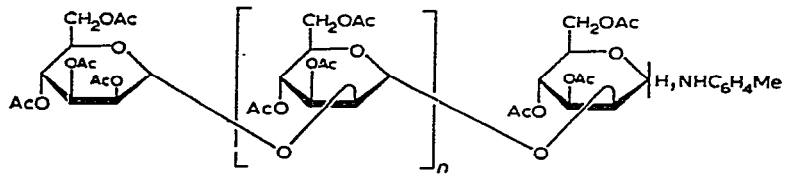
m/e	*Intensity* (%)				m/e	*Intensity* (%)			
	9	10	11	12		9	10	11	12
109	88	75	57	—	355	—	—	—	6
112	8	15	7	—	373	—	—	—	1.4
115	45	37	12	23	377	—	1	0.4	—
118	13	30	31	14	378	2	2.3	1.2	—
120	—	25	11	21	395	0.3	—	—	—
127	45	38	22	9	398	0.4	0.5	1.3	—
128	—	—	—	23	402	—	—	—	0.7
135	—	8	4	12	406	1.2	1	0.2	—
136	13	3	3	3	415	—	—	—	2.3
139	30	17	9	72	423	1	0.2	0.3	—
140	10	20	9.5	—	433	5	8.3	0.3	3
141	8	17	7	5	457	2	1	1.3	—
144	20	28	—	5	462	—	—	—	0.5
145	21	25	9	5	475	—	—	—	3
149	14	8	27	5	479	0.3	—	—	—
157	15	33	11	88	487	—	—	—	0.4
169	100	100	100	—	517	2	—	—	—
170	—	—	—	23	522	—	—	—	0.2
174	5	4	3	—	529	—	—	—	0.4
178	5	6	1.3	—	531	0.25	0.5	0.3	—
182	—	—	3	—	559	0.7	0.4	2.3	—
186	6	10	4	—	571	—	—	—	0.1
187	12	9	5	3	577	1.5	1	—	—
191	5	20	7	—	589	—	—	—	0.2
197	4	4	6	—	606	2	1.5	0.3	—
198	4	4	4	—	607	—	—	—	0.2
199	—	—	—	42	619	6	3	2.2	—
204	13	—	—	4	649	—	—	—	1.2
211	15	12	10	2	666	0.2	0.5	0.4	—
213	3	4	2	—	691	—	—	—	0.2
215	9	13	3	—	707	0.2	—	—	—
216	3	12	2	—	711	0.1	—	—	—
217	3	—	—	28	721	1	0.3	—	—
229	23	12	9	—	738	—	—	—	0.1
242	5	13	4	—	745	0.2	0.2	—	—
243	3	10	3.3	1	805	0.1	0.2	—	—
245	2	—	—	10	847	0.2	0.1	0.32	—
246	—	—	—	2	865	0.3	0.07	—	—
257	2.5	3	2	—	907	0.2	0.3	0.25	—
258	—	1.3	1.7	—	954	0.2	0.2	0.2	—
259	8	—	—	58	971	0.2	—	—	—
271	10	10	13	1	995	0.2	—	—	—
289	17	25	10	—	1013	—	—	—	0.2
306	—	—	—	0.4	1055	0.2	—	—	—
313	—	—	—	2	1195	0.1	—	0.1	—
317	16	12	5	—	1217	0.1	—	—	—
318	4	3	3.3	—	1241	0.1	—	—	—
346	1	—	—	—	1259	0.1	—	—	—
331	72	50	33	3.5	1301	0.2	1	0.6	—

TABLE IV

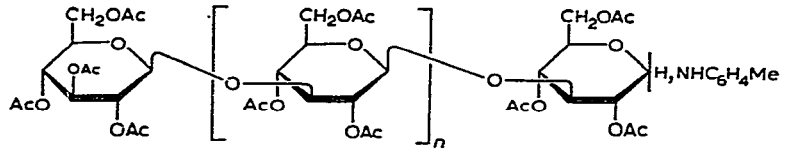
MASS-SPECTRAL DATA FOR COMPOUNDS 13–15

m/e	*Intensity* (%)			m/e	*Intensity* (%)		
	13	**14**	**15**		**13**	**14**	**15**
81	17	42	44	378	1	2	0.6
91	4	40	34	397	1.5	1	1.2
97	28	74	56	398	0.5	8	2
98	15	74	57	406	0.4	1.2	0.3
99	5	4	14	415	1	4	1
103	—	40	20	423	0.7	6	0.5
106	—	16	20	433	3.6	10	2
107	17	28	23	457	2	1	2.3
109	51	64	74	479	0.1	—	—
115	25	44	26	517	1.7	1	1.5
118	4	60	44	531	0.4	0.7	0.6
120	—	20	41	559	1	0.6	3.3
127	28	36	39	577	2	0.4	1.4
135	—	10	7	579	0.2	0.2	0.5
136	1	20	5	606	2	2	0.7
139	17	22	21	619	13.6	2.4	3.4
140	8	20	16	661	0.2	—	—
141	6	16	10	666	0.1	2.4	0.3
144	7	36	7	685	—	1	0.5
145	13	34	21	703	0.2	0.4	0.2
149	5	16	12	711	—	0.7	0.1
157	11	48	25	721	2.5	1	0.1
169	100	100	100	745	0.2	0.3	0.2
174	15	6	4	805	0.4	0.3	0.2
186	4	18	7	847	0.3	0.1	0.3
187	8.5	10	8	865	0.2	0.1	0.2
191	1	20	6	894	0.3	0.1	0.2
197	3	8	7	907	0.1	0.2	0.3
198	3	6	8	954	0.06	0.1	0.1
211	17	9	15	1080	0.4	0.2	0.03
215	3	14	9	1122	0.06	0.1	0.06
216	1	5	4	1135	0.1	—	0.05
229	19	11	17	1182	0.06	0.06	0.1
242	6	13	8	1195	0.2	0.1	0.05
243	4	8	9	1242	0.1	0.1	0.1
257	1.5	3	4	1349	—	0.2	0.01
259	4	—	—	1367	—	0.06	0.06
271	13	7	15	1409	—	0.5	0.1
287	1.5	2	2	1469	—	0.1	0.2
289	13	10	17	1483	0.06	—	0.05
317	13	18	9	1529	0.04	0.2	0.1
318	2.6	4	3	1589	0.1	0.04	0.1
331	51	42	48				

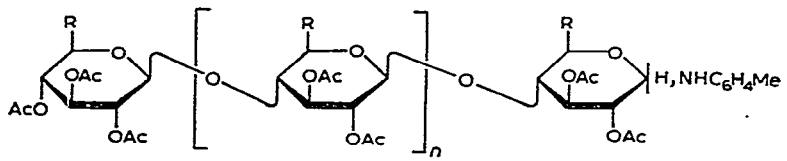
The main fragment ions are formed by cleavage of glycosidic linkages, followed by elimination of acetic acid, ketene, and acetic anhydride. Two series of ions (RS



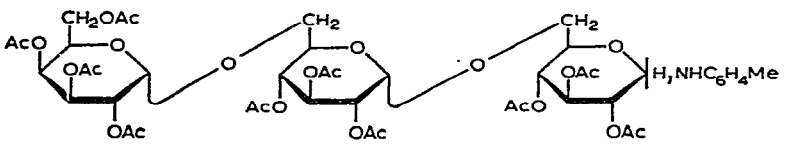
1 $n=1$
9 $n=2$
13 $n=3$



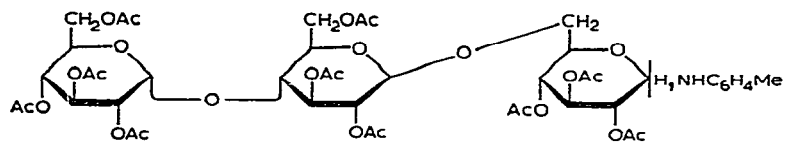
2 $n=1$
10 $n=2$
14 $n=3$



3 $n=1$, R=CH_2OAc
6 $n=1$, R=H
11 $n=2$, R=CH_2OAc
12 $n=2$, R=H
15 $n=3$, R=CH_2OAc

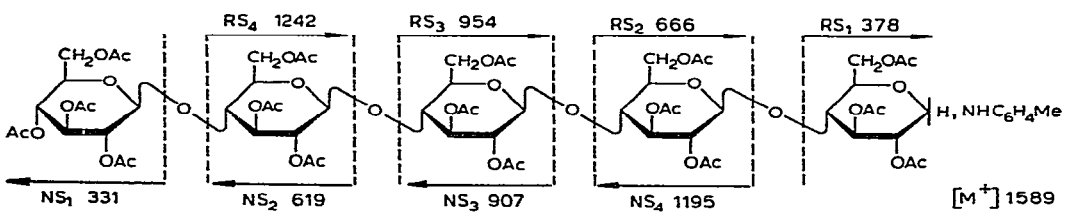


4

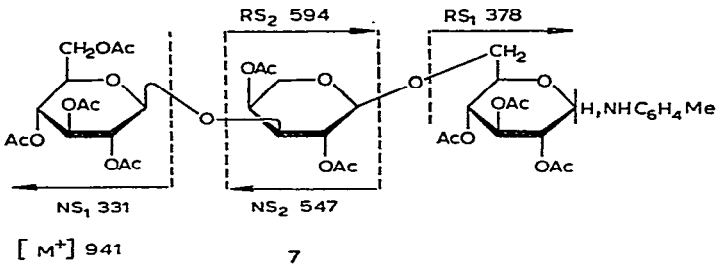


5

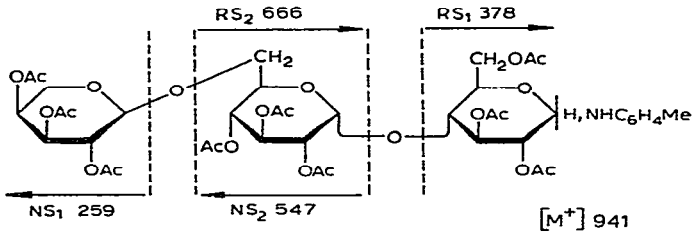
and NS), clearly evident in the high-mass range, reflect the sequence of monosaccharide residues from the reducing and non-reducing ends, respectively. The source of these ions is exemplified below for the cellopentaose derivatives:



The sequence of monosaccharide residues can be seen even more clearly from the mass spectra of trisaccharides (**7,8**) which consist of different monosaccharide moieties:



7



8

In the high-mass range of the mass spectrum of **7**, which has a hexose residue at the non-reducing end, four peaks are present (*m/e* 331, 547, 378, and 594) in addition to that for M^+ (*m/e* 941). The first pair corresponds to the sequence from the non-reducing end, and the second pair corresponds to that from the reducing end. The isomeric trisaccharide **8** exhibits peaks at *m/e* 259 (pentose residue, non-reducing end) and at *m/e* 666 (hexosyl-hexose moiety, reducing end), instead of peaks at *m/e* 331 and 594 as in the mass spectrum of **7**.

In the mass spectra of trisaccharides (**1–5**) consisting of hexose residues, the fragments NS_1 (*m/e* 331) and NS_2 (*m/e* 619) arise by the loss of one or two monosaccharide residues from the non-reducing end, and RS_1 (*m/e* 378) and RS_2 (*m/e* 666) are formed from one or two hexose residues from the reducing end. Homologous fragments NS_3 (*m/e* 907) and RS_3 (*m/e* 954), consisting of three hexose residues, were observed in the mass spectra of both tetra- (**9–11**) and penta-saccharides (**13–15**), whereas the fragments NS_4 (*m/e* 1195) and RS_4 (*m/e* 1242) were found only in the mass spectra of the pentasaccharides.

The structures and mechanisms of formation of the heavy fragments (NS_2–NS_4) are similar to those described in the previous paper[1] for the fragments NS_2 and RS_1, which are formed from corresponding disaccharide derivatives. They differ from the latter only by one, two, or three additional hexose residues.

In the mass spectra of xylose oligosaccharides (**6** and **12**), fragments NS_2 and

RS_2 are shifted to *m/e* 475 and 522, whereas fragments NS_3 and RS_3 in the mass spectrum of **12** give peaks at *m/e* 691 and 733.

In addition to the peaks listed above, in the high-mass of range the mass spectra of oligosaccharide *N*-arylglycosylamine acetates (**1–15**), there are peaks which correspond to the fragments formed from M^+ by the loss of ketene (M−42), acetic acid (M−60), acetic anhydride (M−102), or two molecules of acetic acid (M−120). In the low-mass range, all the peaks are present which correspond to the fragments arising from both reducing and non-reducing terminal units.

A comparison of the mass spectra of oligosaccharides with different positions of glycosidic bonds shows that their fragmentation follows a similar pattern. However, in some instances, the position of the glycosidic bond between the two monosaccharide residues at the reducing end can be deduced from the mass spectrum. Thus, a peak at *m/e* 377, characteristic of a (1→6)-linked reducing moiety, is found in the mass spectra of **4** and **5** in addition to an RS_1 peak (*m/e* 378). However, the relative intensity of these peaks is less than in the mass spectra of disaccharide derivatives. It was shown that the ratio of the peaks at *m/e* 377 and 378 depends on the temperature of the inlet system. The mass spectra of **4** and **5** were recorded at an inlet temperature of 280°, whereas the mass spectra of the disaccharide derivatives were measured at 220°. At 240°, the peak at *m/e* 377 in the mass spectrum of **4** is twice as intense as the peak at *m/e* 378.

Formation of the fragments of the D-series (*m/e* 286, 304, 346, 364) is characteristic of trisaccharides (**4** and **5**) containing a (1→6) linkage between the reducing and penultimate monosaccharide residues[3]. Thus, the presence of fragments of the D-series and a fragment with *m/e* 377 (J_4) in the mass spectra of the compounds **4**, **5**, and **7** allows a (1→6)-linked reducing unit to be distinguished from those having (1→2), (1→3), or (1→4) links.

In the mass spectra of (1→2)-linked oligosaccharides (**1**, **9**, and **13**), peaks at *m/e* 259 and 479 are present, which are characteristic of the (1→2) linkage.

Occasionally, the mass spectra allow the determination of the positions of glycosidic bonds more distant from the reducing unit. Thus, in the mass spectra of compounds **3**, **11**, and **15**, the ratio of peaks NS_2' (*m/e* 559) and NS_2 (*m/e* 619) is approximately unity. The fragment NS_2' is formed by loss of acetic acid from NS_2 (*m/e* 619). In the mass spectra of the isomeric oligosaccharides **1**, **2**, **4**, **9**, **10**, and **13**, this ratio is much smaller than unity.

Thus, the most suitable type of derivative to date for mass-spectrometric determination of the sequence of monosacchaide residues in oligosaccharides is the acetate of an *N*-arylglycosylamine. These derivatives give a moderately intense peak for M^+ and characteristic fragments which reflect the nature and sequence of monosaccharide residues and, to some extent, the positions of the glycosidic linkages. Acetates of *N*-arylglycosylamines can be obtained directly from acetates of oligosaccharides[9] which, in turn, are formed, for example, by the partial degradation of polysaccharides by acetolysis.

EXPERIMENTAL

Mass spectra were measured with a CH-6 Varian MAT instrument at 70 eV, with the ion source at 180–190° and the inlet system at 250–280°. Compounds **1–15** were prepared by the literature procedure[9] and purified by t.l.c. on silica gel.

REFERENCES

1 O. S. CHIZHOV, N. K. KOCHETKOV, AND N. N. MALYSHEVA, *Izv. Akad. Nauk SSSR, Ser. Chim.*, in press.
2 O. S. CHIZHOV, N. K. KOCHETKOV, N. N. MALYSHEVA, AND A. I. SHIYONOK, *Org. Mass Spectrom.*, 5 (1971) 481, 1157.
3 N. K. KOCHETKOV AND O. S. CHIZHOV, *Advan. Carbohyd. Chem.*, 21 (1966) 39.
4 C. S. JOHNSON, W. S. RULIFFSON, AND R. G. COOKS, *Chem. Commun.*, (1970) 587.
5 C. S. JOHNSON, W. S. RULIFFSON, AND R. G. COOKS, *Carbohyd. Res.*, 18 (1971) 233, 243.
6 K. HEYNS AND D. MÜLLER, *Tetrahedron Lett.*, (1966) 6061.
7 J. KARLINER, *Tetrahedron Lett.*, (1968) 3545.
8 J. KÄRKKÄINEN, *Carbohyd. Res.*, 17 (1971) 1.
9 W. T. HASKINS, R. M. HANN, AND C. S. HUDSON, *J. Amer. Chem. Soc.*, 67 (1945) 939.