# MASS SPECTROMETRY OF GLYCOPROTEINS: *O*-GLYCOSYL DERIVATIVES OF L-SERINE AND L-THREONINE

O. S. CHIZHOV, V. A. DEREVITSKAYA, B. M. ZOLOTAREV, L. M. LIKHOSHERSTOV, O. S. NOVIKOVA, AND N. K. KOCHETKOV

*Institute for Organic Chemistry, Academy of Sciences, Moscow (U. S. S. R.)*

(Received April 6th, 1971; accepted for publication, April 30th, 1971)

ABSTRACT

The mass spectra of acetylated *O*-glycosyl derivatives of serine and threonine methyl esters show characteristic features which allow identification of the amino acid and sugar moieties and the type of linkage between them. Some information concerning the sequence of sugar units and the position of the glycosidic bond in the carbohydrate moiety of disaccharide *O*-glycosyl derivatives of serine methyl ester is provided also by mass spectrometry.

## INTRODUCTION

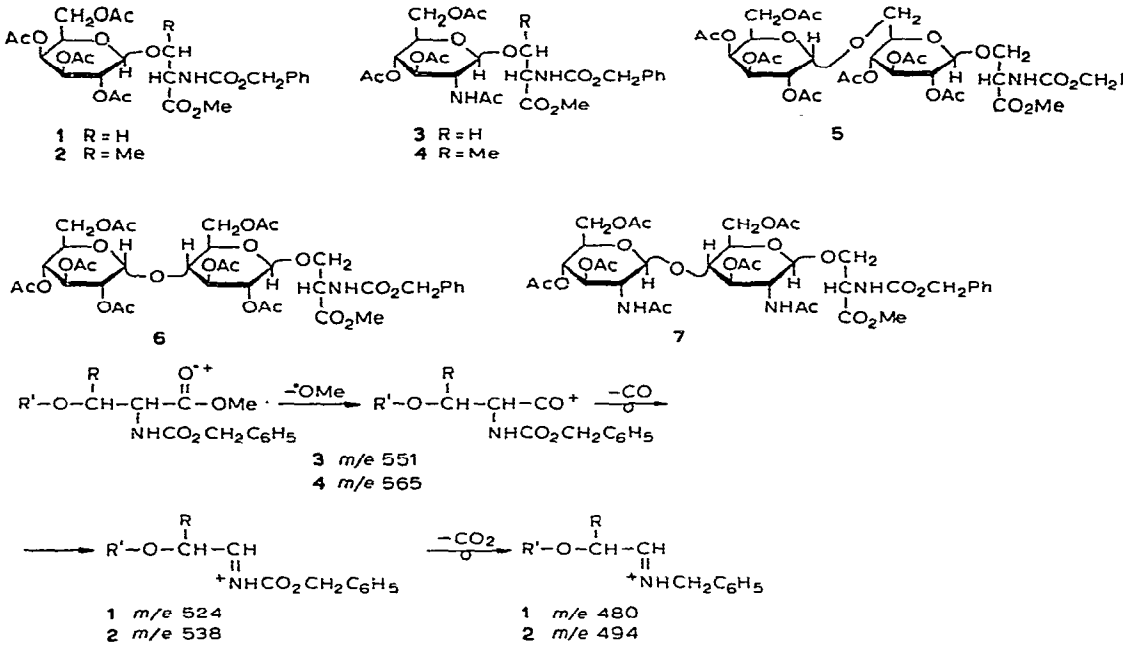
Identification of the type of carbohydrate–peptide linkage, which is an important part of structural studies of glycoproteins, is usually difficult since fragments containing this linkage can be isolated only in very low yield[1]. Determination of the structure of such fragments is facilitated by mass spectrometry because of the high sensitivity of the technique.

The mass spectra of model compounds containing an *N*-acylglycosylamide bond have been described recently[2]. We now report on the fragmentation pattern of L-serine and L-threonine glycosides, which are model compounds containing another type of carbohydrate–peptide linkage known to be present in glycoproteins. Compounds **1–7** have been investigated; the associated synthetic work has been reported elsewhere[3–5].

## RESULTS AND DISCUSSION

### *Mass spectra of monosaccharide derivatives of* N-*benzyloxycarbonyl*-L-*serine and* -L-*threonine methyl esters*

Although each of these compounds displays a molecular ion peak of low intensity, these allow direct determination of the molecular weights of the glycopeptides (Figs. 1–4). In the high mass range, there are only peaks of low intensity. The ions giving rise to these peaks are probably formed as a result of successive loss of MeO, CO, and $CO_2$ (the "amino acid" type of fragmentation[2,6,7]). The $(M-OMe)^+$

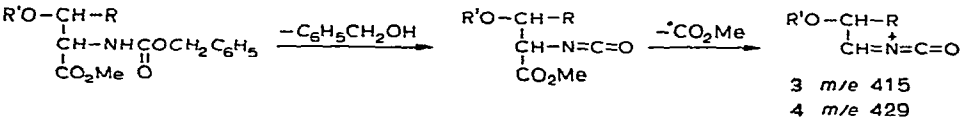


Scheme 1.

fragment was detected only in the mass spectra of the hexosaminides **3** and **4**, and the fragment $(M-COOMe-CO_2)^+$ was clearly observable only in the spectra of hexosides **1** and **2**. The origin of these fragments is shown in Scheme 1.

Also, the mass spectra of the hexosaminides **3** and **4** show peaks at *m/e* 415 and 429, respectively, which are probably due to thermal decomposition* of the benzyloxycarbonyl groups with elimination of benzyl alcohol and subsequent ionization and fragmentation of the resulting isocyanates according to the amino acid pattern[2,6,7] (Scheme 2). The presence of these peaks and the absence of $(M-CH_2O)^+$ and $(M-CH_3CHO)^+$ peaks, which are characteristic of serine and threonine derivatives having a free hydroxyl group[8], clearly indicates that the sugar residue is glycosidically bound to the β-hydroxyl group of the hydroxyamino acid.

The ions formed according to the "sugar" fragmentation pattern produce a number of peaks of high intensity in the range *m/e* 331–90. The ions fall into two

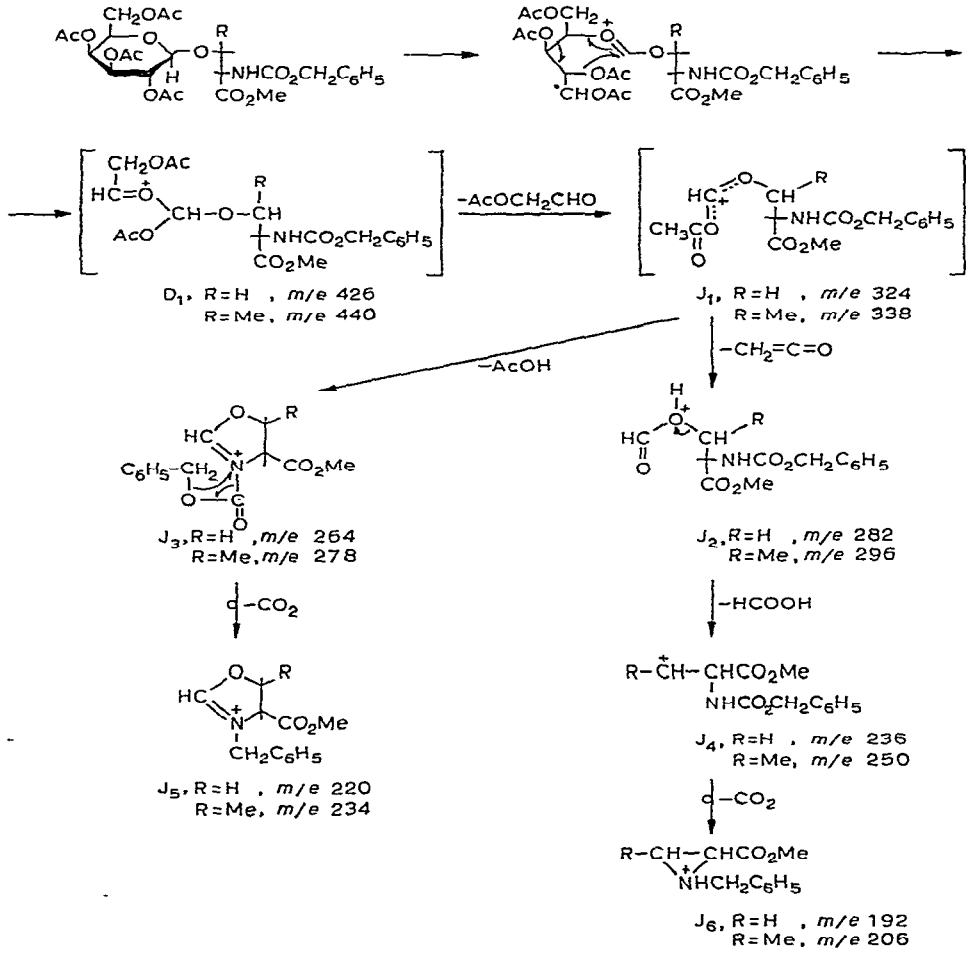


Schéme 2.

*The peaks of $(M-AcOH)^+$ in the mass spectra of **1** and **2** (*m/e* 523 and 537, respectively), and those of $(M-COOMe-AcOH)^+$ in the mass spectra of **3** and **4** (*m/e* 463 and 477) are also probably due to thermal decomposition.

groups according to their origin and analytical significance: (*a*) fragments containing the amino acid moiety and (*b*) other fragments. For purposes of simplicity, group (*a*) will be referred to as "amino acid" fragments and group (*b*) as "sugar" fragments, although it is recognized that these names are not entirely precise.

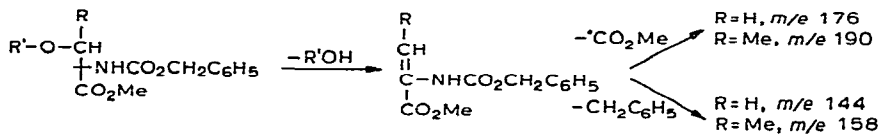
The substituent migration from C-3 to C-1 (Scheme 3) (J-ion series) reported for various sugar derivatives[9,10] may be responsible for the formation of the amino acid fragments, namely, *m/e* 296, 278, 250, 234, and 206 in the glycosides (2 and 4) of *N*-benzyloxycarbonylthreonine and *m/e* 282, 264, 236, 220, and 192 in the glycosides (1 and 3) of *N*-benzyloxycarbonylserine. No peaks corresponding to the $D_1$ and $J_1$ ions have been found, but the subsequent fragments $J_{2-6}$ produce rather intense peaks in the spectra of each of the compounds examined. Their formation can be accounted for by loss, from the hypothetical ion $J_1$, of either a ketene molecule




Scheme 3. Pathways of formation of amino acid fragments.

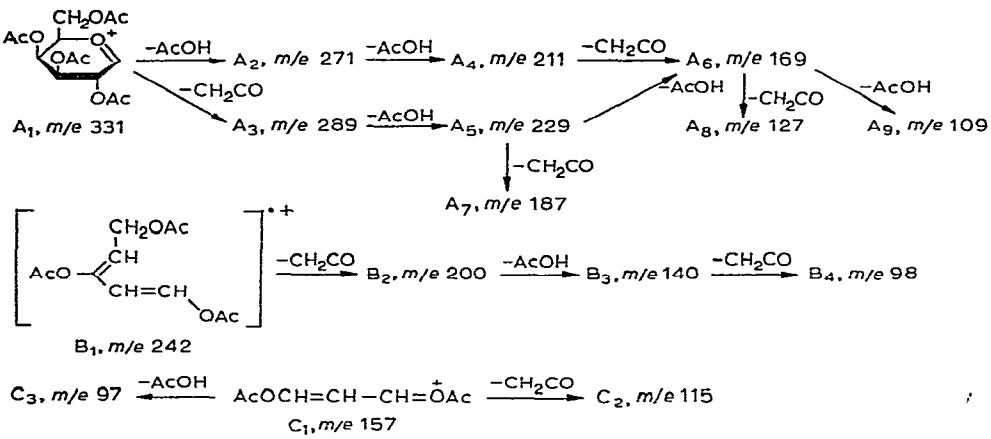
(to form the $J_2$ ion, which then loses formic acid and $CO_2$, successively, to give the $J_4$ and $J_6$ ions, respectively) or an acetic acid molecule (to form the $J_3$ ion which then loses $CO_2$ to give the $J_5$ ion). In the threonine derivatives, the ions are 14 mass numbers higher than those in the serine derivatives, which allows unambiguous identification of the amino acid component in the *O*-glycosyl derivatives of *N*-benzyloxycarbonylhydroxyamino acids.

It is noteworthy that, at the elevated temperatures of the ion-source chamber and inlet system, the mass spectra of **1–4** show peaks that can also be used for the identification of the amino acid component. Ions are formed due to thermal degradation of the molecules by β-elimination of the sugar moiety, and they are more pronounced in the case of the involatile amino acid derivatives **3** and **4** (Scheme 4).




Scheme 4.

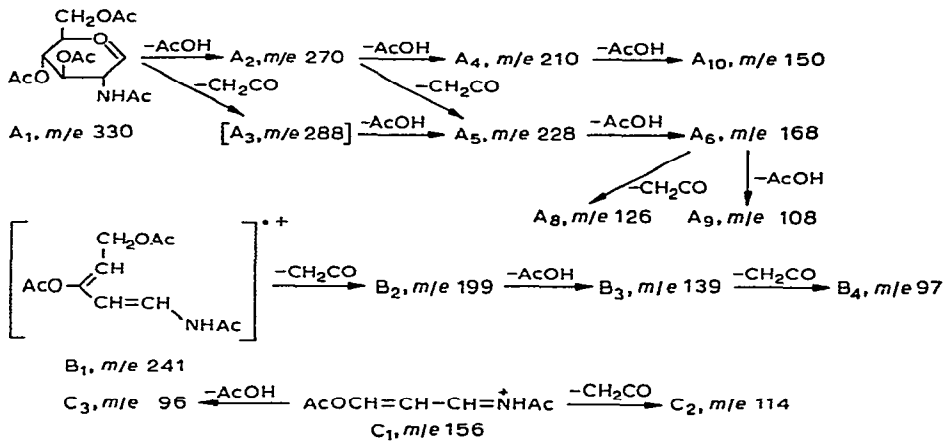
The sugar fragments are those formed from sugar acetates under electron impact[11,12]. These are primarily ions of the A, B, and C series; their structure and the mechanism of formation are presented in Schemes 5 and 6. In the case of the hexosaminides **3** and **4**, the peaks of the A, B, and C series are 1 mass unit lower than those




Scheme 5. Pathways of sugar fragment formation from hexosides **1** and **2**.

of the corresponding hexosides, which makes the hexose derivatives distinguishable from those of hexosamines. Likewise, the derivatives of other monosaccharides (pentoses, 6-deoxyhexoses, etc.) can be identified from the position of the peaks for these series[9,10].

Thus, the mass-spectral analysis shows that monosaccharide derivatives of β-hydroxyamino acids display a number of characteristic peaks which allow the




Scheme 6. Pathways of sugar fragment formation from hexosaminides 3 and 4.

carbohydrate and amino acid units, and the linkage between them, to be identified, and therefore the structure to be established.

*Mass spectra of disaccharide derivatives of* N-*benzyloxycarbonyl-*L-*serine methyl ester*

The fragmentation pattern of the *N*-benzyloxycarbonyl-L-serine methyl ester disaccharide derivatives **5–7** is similar to that of the monosaccharide derivatives, but also shows some characteristic features. Each of the three compounds examined displays a molecular ion peak of low intensity (Figs. 5–7). The amino acid fragments $[(M-CO_2Me)^+$ at *m/e* 812 for **5** and **6** and at *m/e* 810 for **7**; $(M-CO_2Me-CO_2)^+$ at *m/e* 768 for **5** and **6**] give rise to peaks of low intensity, and therefore they are of limited analytical value. On the other hand, peaks (*m/e* 176 and 144) arising because of pyrolytic $\beta$-elimination of the sugar moiety are more intense for the disaccharide derivatives **5–7**.

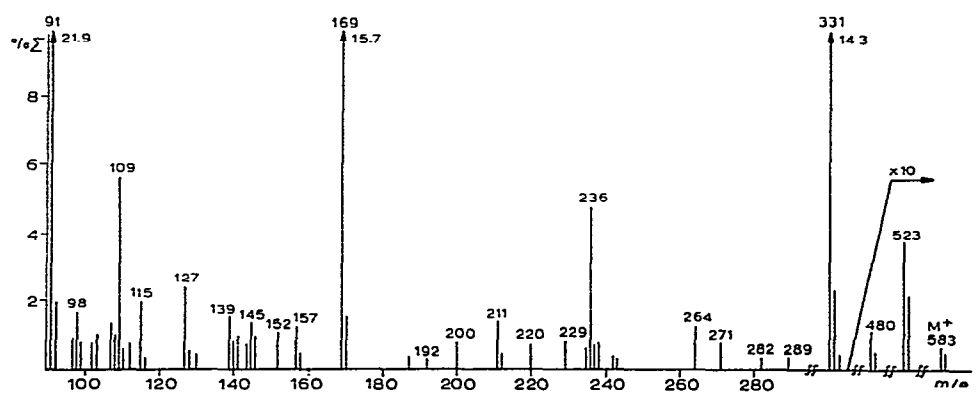



Fig. 1. Mass spectrum of *N*-benzyloxycarbonyl-*O*-(tetra-*O*-acetyl-$\beta$-D-galactopyranosyl)-L-serine methyl ester (**1**, mol. wt. 583).

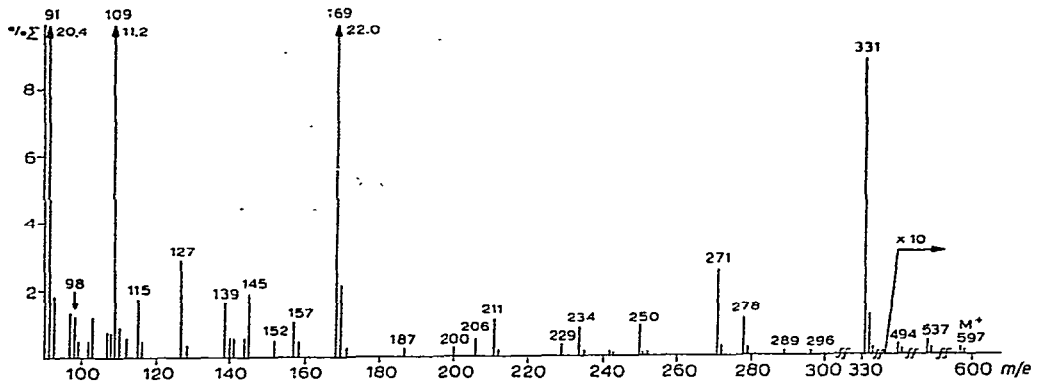



Fig. 2. Mass spectrum of *N*-benzyloxycarbonyl-*O*-(tetra-*O*-acetyl-β-D-galactopyranosyl)-L-threonine methyl ester (2, mol.wt. 597).

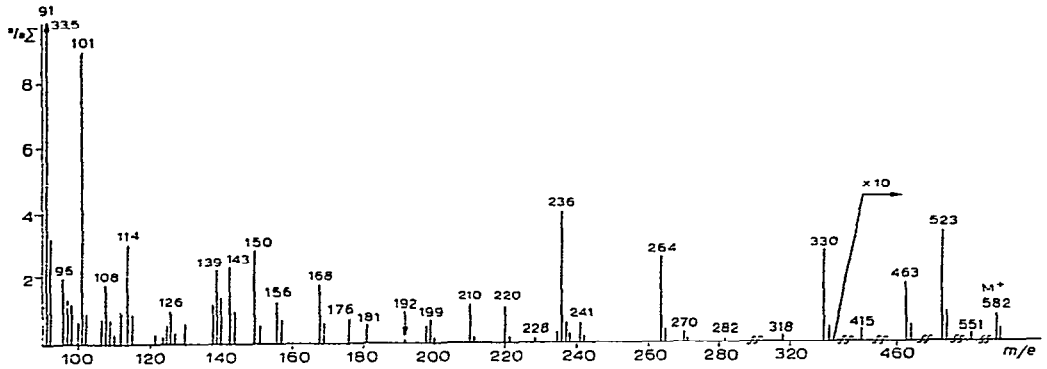



Fig. 3. Mass spectrum of *O*-(2-acetamido-3,4,6-tri-*O*-acetyl-2-deoxy-β-D-glucopyranosyl)-*N*-benzyloxycarbonyl-L-serine methyl ester (3, mol.wt. 582).

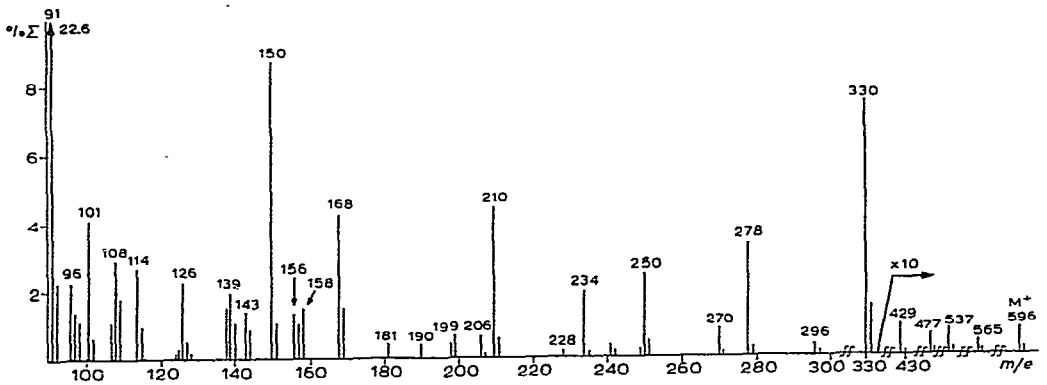



Fig. 4. Mass spectrum of *O*-(2-acetamido-3,4,6-tri-*O*-acetyl-2-deoxy-β-D-glucopyranosyl)-*N*-benzyloxycarbonyl-L-threonine methyl ester (4, mol.wt. 596).

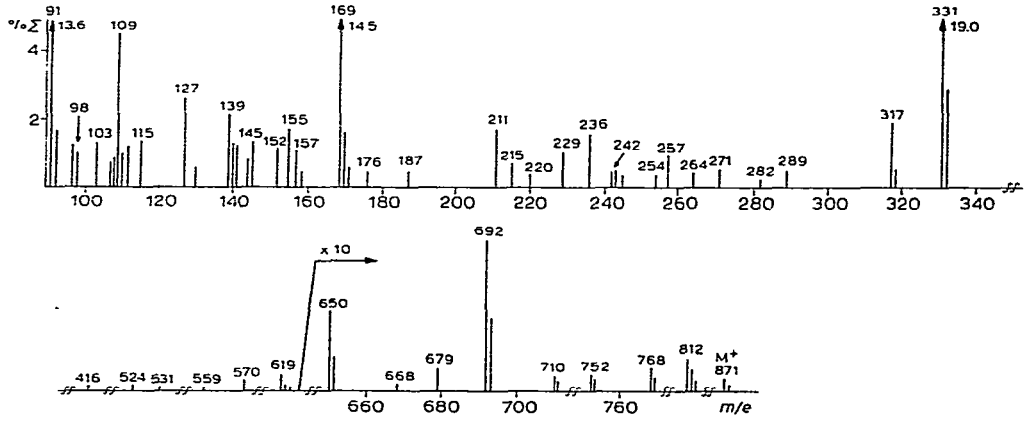


Fig. 5. Mass spectrum of *N*-benzyloxycarbonyl-*O*-[2,3,4-tri-*O*-acetyl-6-*O*-(tetra-*O*-acetyl-α-D-galactopyranosyl)-β-D-glucopyranosyl]-L-serine methyl ester (5, mol.wt. 871).

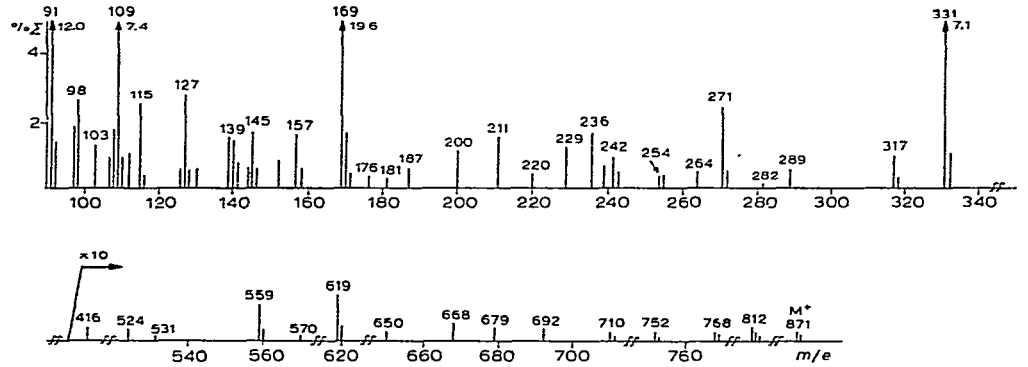


Fig. 6. Mass spectrum of *N*-benzyloxycarbonyl-*O*-[2,3,6-tri-*O*-acetyl-4-*O*-(tetra-*O*-acetyl-α-D-glucopyranosyl)-β-D-glucopyranosyl]-L-serine methyl ester (6, mol.wt. 871).

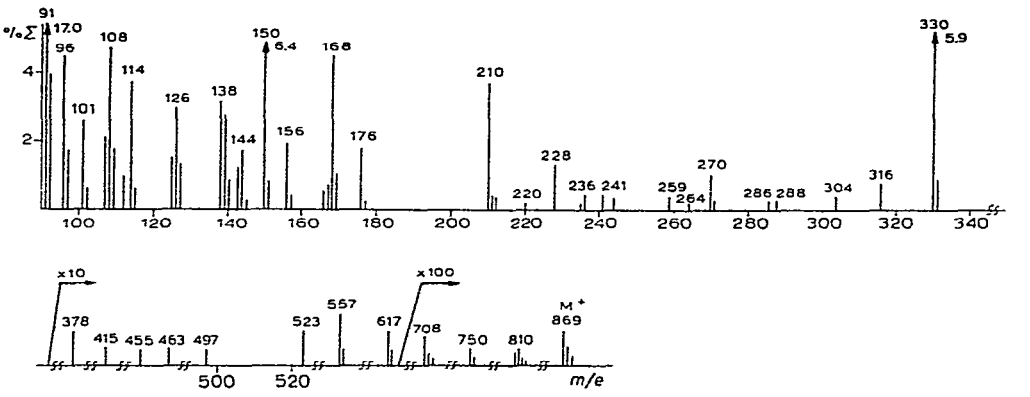
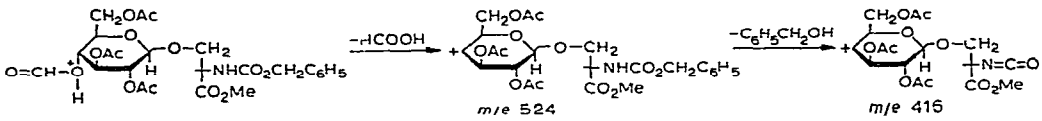


Fig. 7. Mass spectrum of *O*-[2-acetamido-4-*O*-(2-acetamido-3,4,6-tri-*O*-acetyl-2-deoxy-β-D-glucopyranosyl)-3,6-di-*O*-acetyl-2-deoxy-β-D-glucopyranosyl]-*N*-benzyloxycarbonyl-L-serine methyl ester (7, mol.wt. 869).




Scheme 7.

Two series of amino acid fragments of the J-type are produced from the disaccharide derivatives **5–7**. The first series (*m/e* 282, 264, 236, and 220) is similar to that arising from the monosaccharide glycosides **1** and **3**. Fragments of the second series (*m/e* 570, 524, and 416 for **5** and **6**, *m/e* 523 and 415 for **7**) involve the serine moiety and the sugar unit linked to it as is shown in Scheme 7 for **6**.

The presence of all the peaks listed above allows the amino acid component to be identified as serine and the compounds examined as *O*-glycosyl derivatives of serine.

The sugar fragments, primarily those of the A-series, give rise to intense peaks in the lower mass range (*m/e* 331 and 169 for **5** and **6**, *m/e* 330 and 168 for **7**). This allows identification of the terminal, "non-reducing" sugar residue as hexose in **5** and **6**, and as hexosamine in **7**.

The sugar fragments in the high mass range (*m/e* 619 for **5**, *m/e* 619 and 559 for **6**, and *m/e* 617 and 557 for **7**) arise according to the same mechanism as $A_1$ and $A_2$ (see Schemes 5 and 6) and exhibit similar structures. The only difference is that they involve two monosaccharide units. By comparing the mass numbers of these fragments and those of the sugar fragments in the low mass range, it is possible to identify the second sugar unit and to establish the sequence of units. For this purpose, use can also be made of the amino acid fragments of the J-series (*m/e* 570,524, and 416 for **5** and **6**, and *m/e* 523 and 415 for **7**.

It is noteworthy that no fragment of the $A_2$ type (*m/e* 559) is observed for **5** [(1→6)-linkage], whereas for compounds **6** and **7** [(1→4)-linkages] such fragments give rise to peaks of rather high intensity. Additionally, there are peaks of ions (*m/e* 752, 710, 692, 668, and 650 for **5** and **6**, and *m/e* 750 and 708 for **7**) generated by loss of acetic acid and ketene molecules from the ion $(M-CO_2Me)^+$. Compounds **6** and **7**, in which the monosaccharides are (1→4)-linked, display peaks of low intensity, whereas **5** displays peaks of rather high intensity at *m/e* 692 and 650. These data are likely to be of use for distinguishing (1→6)-linked disaccharide derivatives of β-hydroxyamino acids from their (1→4)-linked isomers. It remains to be seen whether or not such differences are of a general character.

Thus, the mass spectrometry of serine disaccharide derivatives enables identification of the amino acid component and the sugar moieties, and determination of the sequence of sugar units and the presence of the *O*-glycosyl linkage between the sugar and amino acid components. It also provides some evidence about the position of the glycosidic bond in the disaccharide part of the molecule.

## EXPERIMENTAL

Mass spectra were recorded on a CH-6 MAT spectrometer (ionizing voltage 70 eV, ionizing current 100 mamp, source temperature 180–190°, inlet-system temperature 140–170° for **1–4** and 200–230° for **5–7**.

The synthesis of compounds **1**, **2**, **5**, **6**, and **7** was described earlier[3–5]. Compounds **3** and **4** were obtained by the oxazoline method[13].

## REFERENCES

1 A. NEUBERGER, A. GOTTSCHALK, AND R. D. MARSHALL, in A. GOTTSCHALK (Ed.), *Glycoproteins*, Elsevier, Amsterdam, 1966, p. 273.
2 L. MESTER, A. SCHIMPL, AND M. SENN, *Tetrahedron Lett.*, (1967) 1697.
3 V. A. DEREVITSKAYA, M. G. VAFINA, AND N. K. KOCHETKOV, *Carbohyd. Res.*, 3 (1967) 377.
4 V. A. DEREVITSKAYA, E. M. KLIMOV, AND N. K. KOCHETKOV, *Carbohyd. Res.*, 7 (1968) 7.
5 N. K. KOCHETKOV, V. A. DEREVITSKAYA, L. M. LIKHOSHERSTOV, V. M. KALINEVICH, AND O. S. NOVIKOVA, *Izv. Akad. Nauk SSSR, Ser. Khim.*, (1969) 2509; (1971) 652.
6 R. T. APLIN, J. H. JONES, AND B. LIBEREK, *J. Chem. Soc., C*, (1968) 1011.
7 E. BRICAS, J. VAN HEYENOORT, M. BARBER, W. A. WOLSTENHOLM, B. C. DAS, AND E. LEDERER, *Biochemistry*, 4 (1965) 2254.
8 K. HEYNS, AND H. F. GRÜTZMACHER, *Fortschr. Chem. Forsch.*, 6 (1966) 536.
9 N. K. KOCHETKOV AND O. S. CHIZHOV, *Advan. Carbohyd. Chem.*, 21 (1966) 39.
10 K. HEYNS, H. F. GRÜTZMACHER, H. SCHARMANN, AND D. MÜLLER, *Fortschr. Chem. Forsch.*, 5 (1966) 448.
11 K. BIEMANN, D. C. DEJONGH, AND H. K. SCHNOES, *J. Amer. Chem. Soc.*, 85 (1963) 1763.
12 K. HEYNS AND D. MÜLLER, *Tetrahedron Lett.*, (1966) 6061.
13 T. S. ANTONENKO, S. E. ZURABYAN, AND A. YA. KHORLIN, *Izv. Akad. Nauk SSSR., Ser. Khim.*, (1970) 1153.