Carbohydrate Research, 137 (1985) 31–38
Elsevier Science Publishers B.V., Amsterdam – Printed in The Netherlands

IDENTIFICATION BY MASS SPECTROMETRY OF *N*-(1-DEOXYHEXITOL-1-YL)AMINO ACIDS, REFERENCE COMPOUNDS FOR THE NONENZYMIC GLYCOSYLATION OF PROTEINS

DONALD J. WALTON*, WALTER A. SZAREK,
Carbohydrate Research Institute, Queen's University, Kingston, Ontario K7L 3N6 (Canada)

DAVID B. MACLEAN, AND ROBERT V. GERARD
Department of Chemistry, McMaster University, Hamilton, Ontario L8S 4M1 (Canada)

(Received July 7th, 1984; accepted for publication, September 19th, 1984)

ABSTRACT

Trimethylsilyl derivatives of synthetic *N*-(1-deoxy-D-glucitol-1-yl)- and *N*-(1-deoxy-D-mannitol-1-yl)amino acids have been examined by mass spectrometry and gas–liquid chromatography–mass spectrometry. Electron impact resulted in cleavage between C-1 and C-2, and between the carbonyl carbon and α-carbon atom, to form ions that allow the amino acid portion of the molecule to be identified. Molecular weights were obtained from chemical-ionization spectra, with ammonia as the reagent gas. These methods are to be used for identifying the amino acid portion of *N*-(1-deoxyhexitol-1-yl)amino acids that are formed when *N*-(1-deoxy-D-fructose-1-yl) groups of nonenzymically glycosylated proteins are reduced with sodium borohydride and the product is hydrolyzed in the presence of acid.

INTRODUCTION

The free amino groups of mammalian proteins acquire 1-deoxy-D-fructos-1-yl residues by nonenzymic reactions with glucose. This process, which is called "nonenzymic glycosylation", is of considerable interest as it might account for some of the pathogenesis of the sequelae of diabetes, such as angiopathy[1], neuropathy[2], artherosclerosis[3], and cataractogenesis[4].

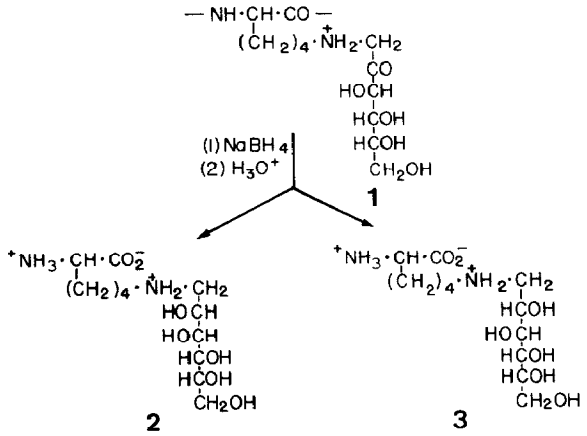
In order to determine which amino acid residues possess *N*-(1-deoxy-D-fructos-1-yl) groups, the latter are reduced with sodium borohydride and the resulting protein is hydrolyzed by acid, to generate *N*-(1-deoxyhexitol-1-yl)amino acids, which must then be identified. This is illustrated by Scheme 1, in which structure **1** is the N^{ϵ}-(1-deoxyfructos-1-yl) derivative of a lysyl residue of a polypeptide. Reduction of **1** and acid-catalyzed hydrolysis of the product results in the formation of N^{ϵ}-(1-deoxy-D-mannitol-1-yl)- and N^{ϵ}-(1-deoxy-D-glucitol-1-yl)-L-lysine (**2** and **3**,

*To whom correspondence should be addressed.

0008-6215/85/$ 03.30

respectively)[5,6]. Similarly, treatment of a polypeptide, such as the β-chain of hemoglobin A_{1c}, in which a 1-deoxyfructos-1-yl group is attached to the *N*-terminus, gives an epimeric pair of *N*-(1-deoxyhexitol-1-yl) derivatives of the *N*-terminal amino acid[7].

Although N^{ϵ}-(1-deoxy-D-glucitol-1-yl)-L-lysine and the *N*-(1-deoxy-D-hexitol-1-yl)-L-valines have been identified by liquid chromatography[5,8] and ^{1}H-n.m.r. spectroscopy[7], respectively, a general method of identification of *N*-(1-deoxyhexitol-1-yl)amino acids is not yet available.



Scheme 1

We have shown already that the trimethylsilyl esters of the penta-*O*-trimethylsilyl derivatives of *N*-(1-deoxyhexitol-1-yl)amino acids may be resolved by gas–liquid chromatography[9]. We now present the results of a study of the mass spectrometry, and gas–liquid chromatography–mass spectrometry (g.l.c.–m.s.) of these compounds, and show that these techniques provide a definitive method of identification that could be used for examining sites of attachment of 1-deoxyfructos-1-yl residues to nonenzymically glycosylated proteins.

RESULTS AND DISCUSSION

Masses and intensities of ions resulting from low-resolution, direct-insertion probe, electron-impact mass spectrometry (e.i.–m.s.) are collated in Tables I and II. Three classes of ions may be recognized in these spectra, namely, (*i*) ions characteristic of the amino acid moiety, which would be of major importance in any analytical investigation of glycosylated proteins, (*ii*) ions derived from the silylated sugar residue by simple cleavage or by cleavage followed by elimination of Me_3SiOH, and (*iii*) ions derived from trimethylsilyloxy groups.

The mass spectra of all compounds listed in Table II, except for the lysine derivatives, may be interpreted by assuming that fragmentation occurs in the manner shown in Scheme 2 for *N*-(1-deoxy-D-glucitol-1-yl)-D-valine (**4**). Three

TABLE I

IONS OF TYPE (*i*) OBSERVED IN ELECTRON-IMPACT MASS SPECTRA OF TRIMETHYLSILYL DERIVATIVES OF *N*-(1-DEOXYHEXITOL-1-YL)AMINO ACIDS[a]

Compound[b]	m/z		
	A	*B*	*C*
ManH-Ala and GlcH-Ala	174	568	56
GlcH-Val and ManH-Val	202	596	84
(^{2}H)GlcH-Val[c]	203	597	85
GlcH-Leu and ManH-Leu	216	610	98
GlcH-Ser[d] and ManH-Ser[d]	262	656	144
GlcH-Thr[d] and ManH-Thr[d]	276	670	158
GlcH-Lys[e] and ManH-Lys[e]	231	625	113

[a]For definitions of ions A, B, and C, see text. [b]GlcH = 1-deoxy-D-glucitol-1-yl; ManH = 1-deoxy-D-mannitol-1-yl. [c]*N*-[1-Deoxy-D-(1-^{2}H)glucitol-1-yl]-L-valine. [d]Molecule includes a trimethylsilyl group attached to the β-hydroxyl group. [e]1-Deoxyhexitol-1-yl group attached to the ε-amino group.

TABLE II

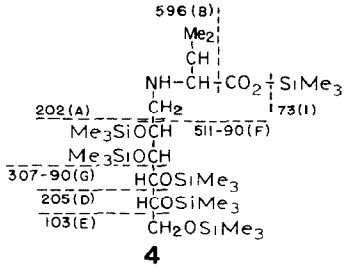
RELATIVE INTENSITIES OF IONS OBTAINED FROM TRIMETHYLSILYL DERIVATIVES OF *N*-(1-DEOXYHEXITOL-1-YL)AMINO ACIDS, USING DIRECT-PROBE E I -M.S

Compound[a,b]	*A* [b]	*B* [b]	*C* [b]	*D* (205)	*E* (103)	*F* (421)	*G* (217)	*H* (147)	*I* (73)
GlcH-Ala	29	11	28	7	14	6	38	18	100
ManH-Ala	41	22	30	10	18	8	63	23	100
GlcH-Val	33	14	17	6	11	5	21	54	100
(^{2}H)GlcH-Val	38	17	20	7	11	6	22	52	100
ManH-Val	36	30	23	11	14	5	35	25	100
GlcH-Leu	24	12	8	8	15	4	30	21	100
ManH-Leu	7	5	6	11	13	1	15	18	100
GlcH-Ser	11	6	3	5	17	2	9	29	100
ManH-Ser	35	7	12	9	15	4	26	27	100
GlcH-Thr	16	8	4	5	6	4	17	12	100
ManH-Thr	18	9	6	4	8	3	9	10	100
GlcH-Lys	39	0	6	8	16	2	26	25	100
ManH-Lys	18	0	4	15	17	2	27	30	100

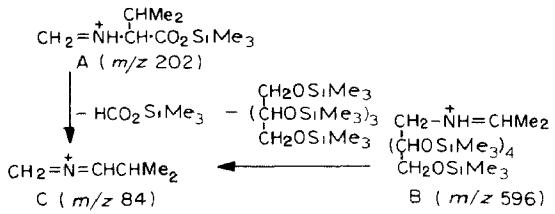
[a]For nomenclature, see footnotes to Table I. [b]For *m/z* values of ions A, B, and C, see Table I. Masses of other ions are shown in parentheses at the top of each column.

ions, which encompass the amino acid moiety in whole or in part, and fall into class (*i*), are formed in each instance. Ion A results from cleavage of the C-1–C-2 bond of the carbohydrate moiety. Ion B arises through cleavage of the bond connecting the carbonyl carbon to the α-carbon atom, in the same manner as that described for trimethylsilyl esters of *N*-trimethylsilylamino acids[10,11]. Ion C is probably derived from A and/or B by one or both of the elimination reactions shown in Scheme 2. For the *N*-(1-deoxyglucitol-1-yl)derivatives of valine, leucine, and serine, the data obtained from high-resolution spectrometry (see Table III) support

the assignments for ions A and C. The inclusion of a deuterium atom at C-1 of *N*-(1-deoxy-D-glucitol-1-yl)-L-valine (**4**) resulted in an increase of one mass unit for the ions of type A, B and C (see Table I), confirming that each ion contains a methylene group derived from position 1 of the carbohydrate chain of the parent molecule.



4



Scheme 2

Thus, the ions of types A, B, and C serve to identify the amino acid moiety of each compound. Although ion B does not appear in the spectra of the N^{ϵ}-(1-deoxyhexitol-1-yl) derivatives of lysine, ions A and C are present, permitting the identification of each compound as a lysine derivative.

The six ions in classes (*ii*) and (*iii*), D–I, originate from the carbohydrate moiety, and arise from processes similar to those described for the per-*O*-trimethylsilylalditols[12,13]. Of the ions of class (*ii*), D and E are formed by direct cleavage of the silylated 1-deoxyhexitolyl group (see Scheme 2). Ions of *m/z* 421 (F) and 217 (G) are formed by elimination of Me_3SiOH from fragments of *m/z* 511 and 307, respectively. The ions in class (*iii*) are $Me_3Si\overset{+}{O}=SiMe_2$ (H, *m/z* 147) and Me_3Si^+ (I, *m/z* 73). Ions D–I are common to all of the spectra (see Table II). Hence, fragmentation of the sugar moiety is not affected appreciably by the structure of the amino acid to which it is attached. E.i.–m.s. does not appear to be suitable for differentiation of *gluco* and *manno* epimers.

Similar electron-impact spectra were obtained by g.l.c.–m.s., as shown in Table IV. As these spectra were not scanned below *m/z* 100, ion C was not observed for derivatives of alanine, valine, and leucine.

All of the compounds were examined by chemical-ionization mass spectrometry (c.i.–m.s.). For the reagent gas, ammonia was used in preference to methane or isobutane, as it caused less fragmention of the substrate than methane and less contamination of the source than isobutane. In all cases, the peak of

TABLE III

HIGH-RESOLUTION ELECTRON-IMPACT SPECTROMETRY OF TRIMETHYLSILYL DERIVATIVES OF *N*-(1-DEOXYHEXITOL-1-YL)AMINO ACIDS

Parent compound[a]	*Ion*[b]	*Formula*	m/z	
			Calc.	*Found*
GlcH-Val	A	$C_9H_{20}NO_2Si$	202.126	202.127
GlcH-Val	C	$C_5H_{10}N$	84.081	84.080
GlcH-Leu	A	$C_{10}H_{22}NO_2Si$	216.142	216.138
GlcH-Leu	C	$C_6H_{12}N$	98.097	98.098
GlcH-Ser	A	$C_{10}H_{24}NO_3Si_2$	262.129	262.130
GlcH-Ser	C	$C_6H_{14}NOSi$	144.085	144.082

[a]For nomenclature, see footnotes to Table I. [b]For definition of ions of types A and C, see text.

TABLE IV

RELATIVE INTENSITIES OF IONS OBTAINED FROM TRIMETHYLSILYL DERIVATIVES OF *N*-(1-DEOXYHEXITOL-1-YL)AMINO ACIDS, USING G L C–M.S

Compound[a,b]	*A* [b]	*B* [b]	*C* [b]	*D* *(205)*	*F* *(421)*	*G* *(217)*	*H* *(147)*
GlcH-Ala	83	30	[c]	14	20	100	35
ManH-Ala	67	41	[c]	14	16	100	34
GlcH-Val	100	55	[c]	47	21	88	65
GlcH-Leu	85	78	[c]	15	21	100	38
ManH-Leu	73	90	[c]	15	15	100	38
GlcH-Ser	100	15	28	10	15	50	30
ManH-Ser	100	20	27	14	13	64	35
GlcH-Thr	100	59	16	17	15	55	45
ManH-Thr	100	100	14	27	17	90	65
GlcH-Lys	100	0	10	10	5	66	40
ManH-Lys	100	0	11	7	3	35	92

[a]For nomenclature, see footnotes *a*, *b*, and *c* of Table I. [b]For *m/z* values of ions A–C, see Table I. Masses of other ions are shown in parentheses at the top of each column. [c]Not recorded, as the spectrum was not scanned below *m/z* 100.

highest mass corresponded to the MH^+ ion, allowing the molecular weight to be verified.

Thus, useful structural information may be obtained by mass spectrometry of trimethylsilyl derivatives of the *N*-(1-deoxyhexitol-1-yl)amino acids. The amino acid portion of each molecule can be defined by e.i.–m.s., and the molecular weight established by c.i.–m.s. The use of g.l.c.–m.s. facilitates the analysis of mixtures of these compounds. These methods will be useful for the identification of the amino acid moieties of *N*-(1-deoxyhexitol-1-yl)amino acids formed by sequential reduction and hydrolysis of nonenzymically glycosylated proteins.

TABLE V

GAS-LIQUID CHROMATOGRAPHIC RETENTION TIMES OF TRIMETHYLSILYL DERIVATIVES OF *N*-(1-DEOXYHEXITOL-1-YL)AMINO ACIDS

Compound[a]	*Retention time (min)*	*Compound*	*Retention time (min)*
GlcH-Ala	7.2	GlcH-Ser	8.8
ManH-Ala	7.0	ManH-Ser	8.5
GlcH-Val	7.4	GlcH-Thr	8.9
ManH-Val	7.3	ManH-Thr	8.7
GlcH-Leu	7.7	GlcH-Lys	12.4
ManH-Leu	7.5	ManH-Lys	12.4

[a]For nomenclature, see footnotes to Table I.

EXPERIMENTAL

Mass spectrometry. — All mass spectrometry experiments were performed with a VG 7070F mass spectrometer equipped with a data system. Data were acquired and processed using a Digital PDP 8a computer; VG software (Data System 2000) was used throughout. Data were stored on hard disks using a Digital RL02 disk drive. Samples were introduced through a direct-insertion probe for high- and low-resolution measurements by electron impact and for chemical-ionization experiments. Probe spectra at low resolution by electron impact were obtained in the following manner. A sample (10 μL), prepared as described next, was introduced into a capillary tube and the solvent (pyridine) was removed under vacuum. The tube was cut to a length of 3 cm, mounted on a probe, and inserted into the mass spectrometer. A mass range of 40–880 m.u. was scanned in 3 s, under the following conditions: source temperature, 200°; filament emission, 100 μA; ionization potential, 70 eV; and accelerating voltage, 3 kV.

Samples for high-resolution measurement were introduced into the mass spectrometer in the same manner as just described. The instrument was set to a resolution of 3000 (10% valley). A mass range was selected that included the ion of interest and perfluorokerosene ions having masses above and below the ion under observation. The appropriate VG software was used for calibration of the instrument and determination of the accurate mass. The scan rate was 10 s, the source temperature, 200°, the filament emission, 100 μA, the ionization potential, 70 eV, and the accelerating voltage, 4 kV.

In the chemical-ionization experiments, a mass range of 200–800 m.u. was scanned in 3 s with ammonia as reagent gas, a source temperature of 150°, a filament emission of 200 μA, an accelerating voltage of 3 kV, and a source pressure 0.5–1.0 torr.

G.l.c.–m.s. — Each *N*-(1-deoxyhexitol-1-yl)amino acid (1 mg) was incubated with a 1.5M solution of *N*-trimethylsilylimidazole in pyridine (0.5 mL) ("Tri-Sil Z" reagent; Pierce Chemical Co., Rockford, IL) for 2 h at 50° in a vial sealed with a

Teflon-lined septum. This procedure resulted in the formation of the trimethylsilyl ester of the 2,3,4,5,6-penta-*O*-trimethylsilyl derivative. An additional trimethylsilylation occurred at the β-hydroxyl group of the derivatives of L-serine and L-threonine.

A sample (1 μL) of the resulting solution was injected into a Varian 3700 gas chromatograph interfaced with the VG 7070F mass spectrometer by a VG jet separator. The carrier gas, helium, flowed through a glass column (180 × 0.2 cm) containing 3% OV-17 on Chromosorb W (100–120 mesh). The temperature was increased from 180 to 300° at 10°/min, and kept constant thereafter. Retention times are shown in Table V. Conditions for e.i.–m.s. were the same as those described for the direct-insertion method, except that a mass range of 100–800 m.u. was scanned.

N-*(1-Deoxy*-D-*hexitol-1-yl)*-L-*amino acids.* — These compounds were synthesized by reductive amination of D-glucose or D-mannose, in the presence of sodium cyanoborohydride, as previously reported[9].

N-[*1-Deoxy*-D-*(1-^{2}H)glucitol-1-yl*]-L-*valine.* — This compound was prepared similarly from L-valine and D-(1-^{2}H)glucose (Merck, Sharp and Dohme, Montreal). Equal quantities of two epimers were formed in which the deuterium atom occupied different diastereotopic positions, 1A and 1B. In the ^{1}H-n.m.r. spectrum, signals for H-1 and H-2 for the two epimers were superposed. The *1A-(^{2}H)* epimer showed δ 3.15 (d, 1 H, $J_{1B,2}$ 10.0 Hz, H-1B) and 4.12 (dd, 1 H, $J_{1B,2}$ 10.0, $J_{2,3}$ 4.9 Hz, H-2); the *1B-(^{2}H)* epimer showed δ 3.25 (d, 1 H, $J_{1A,2}$ 3.3 Hz, H-1A) and 4.12 (dd, 1 H, $J_{1A,2}$ 3.3, $J_{2,3}$ 4.9 Hz, H-2).

ACKNOWLEDGMENTS

Financial support from the Canadian Diabetes Association and from the Natural Sciences and Engineering Research Council of Canada is gratefully acknowledged. Valuable technical assistance was rendered by J. D. McPherson and Iona Mahtab.

REFERENCES

1 B. Chavers, D. Etzwiler, and A. F. Michael, *Diabetes*, 30 (1981) 275–278.
2 B. W. Vogt, E. D. Schleicher, and O. H. Wieland, *Diabetes*, 31 (1982) 1123–1127.
3 J. L. Witztum, E. M. Mahoney, M. J. Branks, M. Fisher, R. Elam, and D. Steinberg, *Diabetes*, 31 (1982) 283–291.
4 V. J. Stevens, C. A. Rouzer, V. M. Monnier, and A. Cerami, *Proc. Natl. Acad. Sci. U.S.A.*, 75 (1978) 2918–2922.
5 R. L. Garlick and J. S. Mazer, *J. Biol. Chem.*, 258 (1983) 6142–6146.
6 G. J. Rucklidge, G. P. Bates, and S. P. Robins, *Biochim. Biophys. Acta*, 747 (1983) 165–170.
7 R. L. Koenig, S. H. Blobstein, and A. Cerami, *J. Biol. Chem.*, 252 (1977) 2992–2997.
8 M. Brownlee, H. Vlassara, and A. Cerami, *Diabetes*, 29 (1980) 1044–1047.
9 D. J. Walton, E. R. Ison, and W. A. Szarek, *Carbohydr. Res.*, 128 (1984) 37–49.
10 W. J. A. Vandenheuvel, J. L. Smith, I. Putter, and J. S. Cohen, *J. Chromatogr.*, 50 (1970) 405–412.

11 K. R. LEIMER, R. H. RICE, AND C. W. GEHRKE, *J. Chromatogr.*, 141 (1977) 355–375.
12 G. PETERSSON, *Tetrahedron*, 25 (1969) 4437–4443.
13 J. LONNGREN AND S. SVENSSON, *Adv. Carbohydr. Chem. Biochem.*, 29 (1974) 41–106.