Modification of dipeptides by alkyl chloroformate—alkanol mixtures for analysis by gas chromatography/mass spectrometry with electron and chemical ionization and collisional activation: differentiation of isomers

R. S. Borisov,[a] V. G. Zaikin,[a]★ B. V. Vas´kovskii,[b] and L. Yu. Sklyarov[c]

[a]*A. V. Topchiev Institute of Petrochemical Synthesis, Russian Academy of Sciences, 29 Leninsky prosp., 119991 Moscow, Russian Federation. Fax: +7 (495) 954 2269. E-mail: zaikin@ips.ac.ru*
[b]*M. M. Shemyakin and Yu. A. Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, 16/10 ul. Miklukho-Maklaya, 117997 Moscow, Russian Federation*
[c]*Institute of Immunology, 24-4 Kashirskoe sh., 115478 Moscow, Russian Federation*

A possibility of dipeptide derivatization by alkyl chloroformate—alkanol (Alk = Me, Et, Pr, Bu) mixed reactants for their determination and differentiation of isomers in mixtures by gas chromatography/mass spectrometry with electron and chemical ionization and collisional activation was studied. Prospects for using derivatization by mixtures of methyl chloroformate with 2,2,2-trifluoroethanol, 2,2,3,3,3-pentafluoropropanol, and 2,2,3,3,4,4,4-heptafluorobutanol were shown. Diagnostic ions that make it possible to determine the amino acid sequence in dipeptides and distinguish residues of isomeric leucine and isoleucine were revealed.

Key words: dipeptides, mass spectrometry, derivatization, electron ionization, chemical ionization, collisional activation.

Although considerable success of new mass spectrometric methods (matrix-activated laser desorption/ionization, electrospray, Fourier transform mass spectrometry) for solving problems of proteomics was achieved, these methods can mainly be applied for studying high-molecular-weight proteins and peptides (determination of the amino acid sequence, conformations, intra- and intermolecular interactions, and others). At the same time, traditional mass spectrometric methods, especially very sensitive gas chromatography/mass spectrometry (GC/MS), remain popular for analysis of amino acids and small peptides (including nonprotein amino acids, peptidomimetics, peptide synthons, and others). New developments in this structural analytical area continuously appear.[1] Interest in analysis of small peptides is increasing, because a series of biologically active dipeptides was found in biological matrices (skeletal muscles, mammalian brain). In addition, specific di- and tripeptides exhibit properties of analgesics or suppress some reactions during blood coagulation and protect mucous membrane of the stomach from stomach ulcer (family of glycoprolines, such as Pro-Gly, Gly-Pro, Pro-Gly-Pro, Hyp-Gly, Gly-Hyp, and *cyclo*-Pro-Gly).[2] A very important problem that can be solved by mass spectrometry is to elucidate the origin of small peptides in the organism of mammals and routes of their transfer to receptors. Small peptides are present in biological matrices in trace amounts and, hence, the use of mass spectrometry (mainly GC/MS and liquid chromatography/mass spectrometry (LC/MS)) and radioimmunological methods for their detection and study of metabolism seems to be especially promising. It can be assumed that some small peptides are biomarkers and their profiles reflect specific diseases. Also note that the amino acid sequences of proteins and oligopeptides are manifested in the dipeptide sequence, which predetermines necessity to develop methods of their analysis in dipeptidase hydrolyzates.[3,4]

Several derivatization approaches used for determination by GC/MS of small peptides in mixtures are known. The earlier used methods most widely included the esterification of terminal carboxyl groups to form methyl, ethyl, and pentafluorobenzyl esters and acylation of terminal amino groups (as well as additional OH and SH groups in residues of several amino acids) with the introduction of *N*-formyl, *N*-acetyl, *N*-trifluoroacetyl, *N*-pentafluoropropionyl, or *N*-heptafluorobutyryl groups (see, *e.g.*, Refs 4—7). The main drawback of this approach is the necessity to isolate nonmodified peptides from aqueous

Published in Russian in *Izvestiya Akademii Nauk. Seriya Khimicheskaya,* No. 12, pp. 2200—2204, December, 2006.
1066-5285/06/5512-2285

solutions, because these reactions can be carried out only in organic solvents. Modification of amino acids and peptides by mixtures of alkyl chloroformates and alkanols (Husek reagents) for further GC/MS analysis seems especially efficient, because these reactions occur in aqueous solutions without[8] additional time expenses to isolation of desired substances. Various alkyl chloroformates (Alk = Me, Et, Pr, Bu^n, Me_2CHCH_2, $C_6F_5CH_2$, *n*-C_6H_{13}, 2,2,3,3,4,4,5,5-octafluoropentyl, 9-fluorenylmethyl) and alkanols (methanol, ethanol, propanol, butanol, 2-methylpropanol, 2,2,2-trifluoroethanol, 2,2,3,3,4,4,4-heptafluorobutanol) are used for the derivatization of amino acids; all of them can be applied for the modification of small peptides. Only several studies on GC/MS analysis of dipeptides using this approach are known. For instance, a combination of ethyl chloroformate and 2,2,2-trifluoroethanol was proposed[9] for fast derivatization and GC/MS analysis with chemical ionization of dipeptides in mixtures. The chemical ionization mass spectra contained only peaks of the $[M + H]^+$ ions and did not characterize the structure of dipeptides. Modification by an ethyl chloroformate—methanol mixture was successfully applied[10] to glutathione tripeptide (L-γ-glutamyl-L-cysteineglycine). The electron ionization mass spectrum of the formed N,S-bis(ethoxycarbonyl) methyl ester contains characteristic peaks that make it possible to determine the amino acid sequence. It is of interest to study in more detail this derivatization technique in combination with GC/MS, in particular, for differentiation of isomeric dipeptides.

In the present work we studied possibilities of the preliminary derivatization of dipeptides by the Husek reagents during their qualitative determination in mixtures by gas chromatography combined with mass spectrometry in the electron ionization (GC/EI-MS), chemical ionization (GC/CI-MS), and chemical ionization with collisional activation (GC/CI-MS/MS) modes. The pairs of isomeric dipeptides were used: Leu-Ile/Ile-Leu, Ala-Ile/Ile-Ala, Gly-Pro/Pro-Gly, Ile-Gly/Gly-Ile, Leu-Gly/Gly-Leu, and Leu-Ala/Ala-Leu, as well as dipeptides Gly-Phe, Ala-Pro, Ala-Val, Gly-Leu, and Hyp-Gly. Symmetric derivatization when the alkyl residues in alkyl chloroformate AlkOCOCl and alcohol AlkOH are the same (Alk = Me, Et, Pr, Bu) was used. In addition, possibilities of derivatization of isomeric dipeptides by combinations of methyl chloroformate with 2,2,2-trifluoroethanol, 2,2,3,3,3-pentafluoropropanol, or 2,2,3,3,4,4,4-heptafluorobutanol were used.

Experimental

Methyl, ethyl, propyl, and *n*-butyl chloroformates, 2,2,2-trifluoroethanol, 2,2,3,3,3-pentafluoropropanol, and 2,2,3,3,4,4,4-heptafluorobutanol were used (Fluka). Dipeptides were synthesized using standard procedures or purchased (Bachem).

Derivatization of dipeptides was carried out according to a described procedure.[9] The corresponding alcohol, pyridine, and alkyl chloroformate in equimolar amounts were sequentially added to a solution of individual dipeptide or a mixture of dipeptides in water. Chloroform was added after stirring for several minutes, and the organic layer was used for GC/MS.

The derivatives were studied by GC/EI-MS and GC/CI-MS on a Finnigan MAT 95XL gas chromatograph/mass spectrometer (energy of ionizing electrons 70 eV, temperature of an ion source 200 °C). A quartz capillary column (30 mm×0.19 mm, stationary liquid phase polydimethylsiloxane containing 5% Ph groups) was used in the chromatographic part. The temperature was programmed from 30 to 120 °C with a rate of 5 °C min^{-1} and then to 290 °C with a rate of 10 °C min^{-1}. Isobutane and methane were used as reactant gases for recording chemical ionization mass spectra.

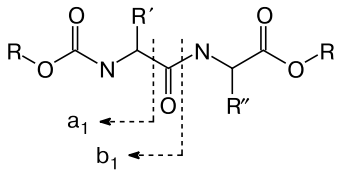
Mass spectra with collisional activation (argon, collisional energy 50 eV, gas pressure in the collisional cell 0.3 Torr) for the $[M + H]^+$ ions generated from modified dipeptides upon chemical ionization were measured on a Finnigan TSQ-70B gas chromatograph/mass spectrometer (chromatographic conditions were the same as in the previous case).

Results and Discussion

The reaction of dipeptides with an alkyl chloroformate—alcohol mixture in an aqueous medium results in the transformation of the terminal carboxyl and amino groups into the ester and *N*-alkoxycarbonyl groups, respectively. The hydroxyl group in hydroxyproline is transformed into the carbonate group.

The electron ionization mass spectra of all the derivatives do not contain at all or contain only insignificant $[M]^{\bullet +}$ peaks. The main fragmentation of the derivatives is caused by dissociations of the "amine" type or amide bond dissociation (Scheme 1) to form the N-terminal ions a_1 and b_1, respectively (the ions are designated according to Biemann´s nomenclature[11]).

Scheme 1



These ions are very characteristic, and the peaks of the a_1 ions are always major in the spectra. Their identification allows one to perform a clear distinction of isomeric dipeptides with different sequences of amino acids (Fig. 1). The intensities of peaks of these ions for a series of the derivatives are given in Table 1.

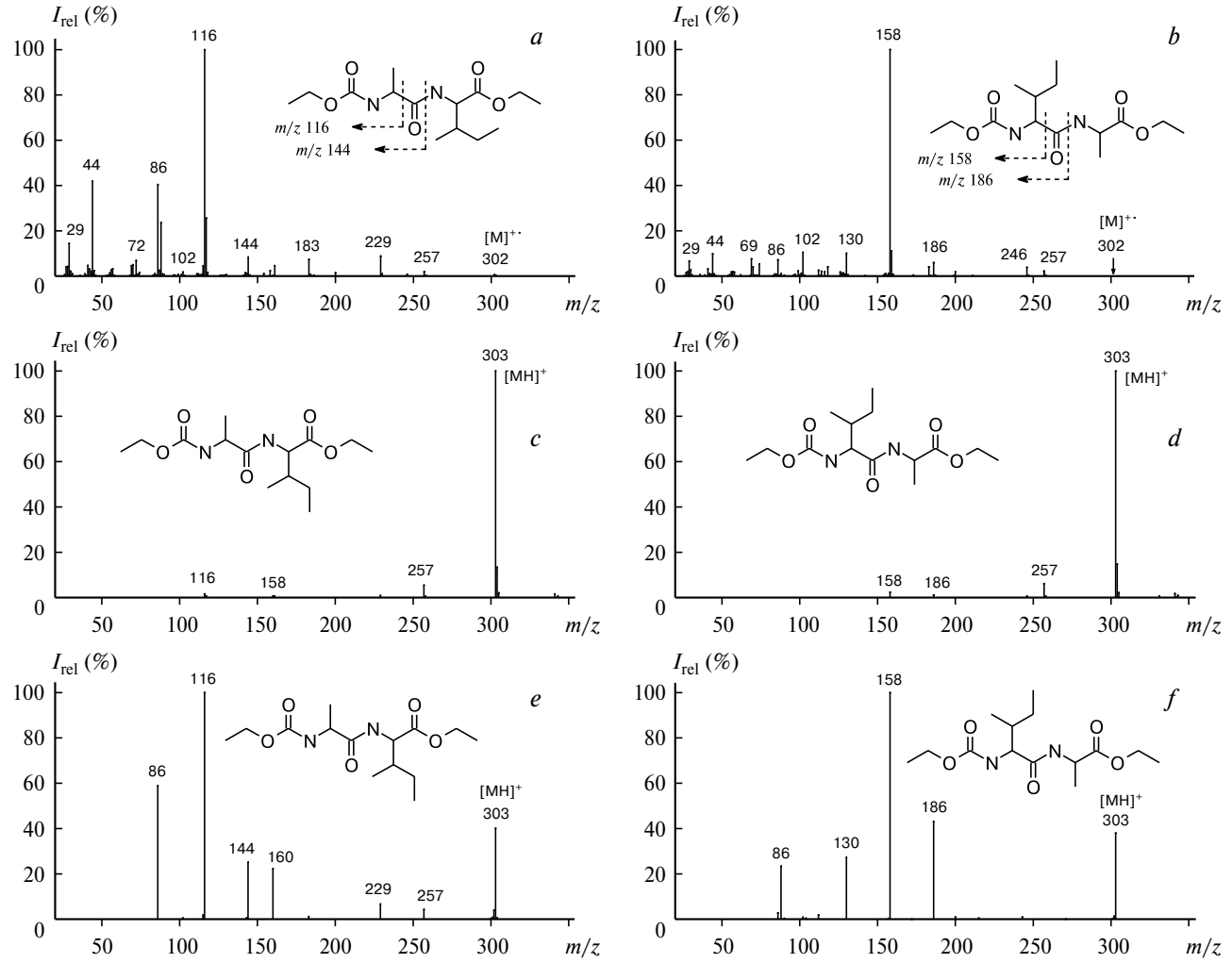
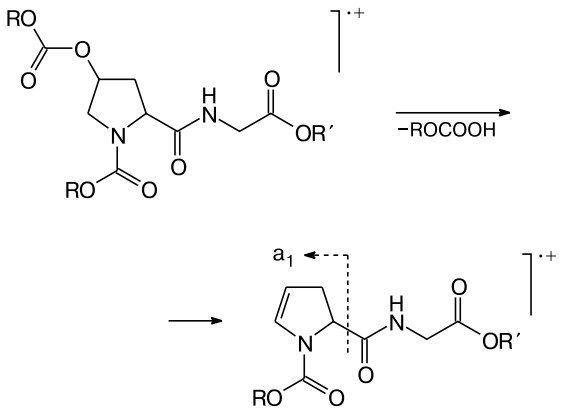


Fig. 1. Mass spectra of ethyl esters of *N*-ethoxycarbonyl-L-alanyl-L-isoleucine and *N*-ethoxycarbonyl-L-isoleucyl-L-alanine measured under conditions of electron ionization (*a*, *b*), chemical ionization (*c*, *d*), and chemical ionization/collisional activation (*e*, *f*).

Note that a similar fragmentation of *N*,*O*-bis(alkoxycarbonyl) derivatives of alkyl esters of dipeptides containing the hydroxyproline residue occurs after the primary elimination of the ROCOOH molecule from the molecular ion (Scheme 2).

Scheme 2



As expected,[9] the chemical ionization mass spectra of the studied derivatives contain almost only peaks of the $[M + H]^+$ ions that make it possible to determine the molecular weight. At the same time, virtually no noticeable peaks of fragment ions reflecting the amino acid derivatives of dipeptides are observed (see Fig. 1, *c*, *d*). In this respect the collisional activation spectra, which were obtained for the $[M + H]^+$ ions generated upon chemical ionization, contain simultaneously the peaks of the initial ions and the peaks characteristic of the electron ionization spectra (see Fig. 1, *e*, *f*). Thus, collisional activation spectra for protonated molecules detected under chemical ionization are more convenient for structural assignments. This fact predetermines the use of chemical ionization in the regime of tandem mass spectrometry for deconvolution of unresolved chromatographic peaks in analysis of complex mixtures of modified dipeptides. The

Table 1. Values of m/z (I_{rel} (%)) for peaks of characteristic ions in the electron ionization mass spectra (70 eV) of alkyl esters of *N*(*O*)-alkoxycarbonyl derivatives of some dipeptides

Initial dipeptide	Alk = Me		Alk = Et		Alk = Pr		Alk = Bu	
	a_1	b_1	a_1	b_1	a_1	b_1	a_1	b_1
Leu-Ile	144 (100)	172 (1.8)	158 (100)	186 (3.1)	172 (100)	200 (4.6)	186 (100)	214 (6.2)
Ile-Leu	144 (100)	172 (1.6)	158 (100)	186 (3.5)	172 (100)	200 (5.7)	186 (100)	214 (6.5)
Ala-Ile	102 (100)	130 (4.8)	116 (100)	144 (8.3)	130 (100)	158 (11.4)	144 (100)	172 (13.9)
Ile-Ala	144 (100)	172 (3.9)	158 (100)	186 (5.9)	172 (100)	200 (6.6)	186 (100)	214 (6.2)
Gly-Pro	88 (15.5)	116 (2.3)	102 (9.4)	130 (2.8)	116 (2.9)	144 (2.2)	130 (1.5)	158 (2.6)
Pro-Gly	128 (100)	156 (1.9)	142 (100)	170 (2.5)	156 (100)	184 (2.0)	170 (100)	198 (1.5)
Ile-Gly	144 (100)	172 (10.8)	158 (100)	186 (9.9)	172 (100)	200 (32.5)	186 (100)	214 (31.3)
Gly-Ile	88 (55.5)	116 (25.3)	102 (34.3)	130 (21.4)	116 (9.0)	144 (11.7)	130 (9.3)	158 (11.3)
Leu-Gly	144 (100)	172 (8.0)	158 (100)	186 (7.4)	172 (100)	200 (11.4)	186 (100)	214 (7.8)
Gly-Leu	88 (45.7)	116 (12.4)	102 (29.0)	130 (13.1)	116 (9)	144 (11.7)	130 (9.7)	158 (11.4)
Ala-Leu	102 (100)	130 (4.5)	116 (100)	144 (7.2)	130 (100)	158 (13.1)	144 (100)	172 (15.3)
Leu-Ala	144 (100)	172 (4.8)	158 (100)	186 (7.4)	172 (100)	200 (6.8)	186 (100)	214 (6.0)
Hyp-Gly*	144 (100)	172 (4.9)	158 (100)	186 (4.7)	172 (100)	200 (13.7)	186 (100)	214 (3.5)
Gly-Phe	88 (43.0)	116 (26.2)	102 (57.6)	130 (36.8)	116 (21.6)	144 (33.0)	130 (12.0)	158 (22.9)
Ala-Pro	102 (100)	130 (3.1)	116 (100)	144 (5.1)	130 (100)	158 (7.4)	144 (100)	172 (9.5)
Leu-Ala	144 (100)	172 (4.8)	158 (100)	186 (7.4)	172 (100)	200 (6.8)	186 (100)	214 (6.0)
Ala-Val	102 (100)	130 (9.1)	116 (100)	144 (11.9)	130 (100)	158 (14.8)	144 (100)	172 (17.1)

* The derivatives contain the unprotected OH group in hydroxyproline.

observed regularities are also applicable for analysis of modified dipeptides by mass spectrometry with electrospray when protonated molecules are the main initial ions.

Further we used the following fluorinated alcohols for derivatization: trifluoroethanol, pentafluoropropanol, and heptafluorobutanol. As expected, the products are more volatile and provide better separation by gas chromatography. Their mass spectra contain no new diagnostic peaks, and their peaks of the a_1 and b_1 ions have the same intensity as those of other derivatives (Table 2).

A problem of differentiation of leucine derivatives in oligopeptide molecules often appears. We have recently shown[12] that the electron ionization mass spectra of *N*-methoxycarbonyl derivatives of alkyl (especially fluoroalkyl) esters of isomeric amino acids differ sharply by relative intensities of the peaks with m/z 102 and 115. The first of them is always more intense for the leucine derivatives, and the second is more intense for the isoleucine derivatives. Different mass spectrometric approaches were used to prove the formation of the ion with m/z 115 from the leucine derivatives by the successive elimination of

Table 2. Values of m/z (I_{rel} (%)) for peaks of characteristic ions in the electron ionization mass spectra (70 eV) of some isomeric dipeptides modified by methyl chloroformate and fluorinated alcohols ROH

Initial dipeptide	R = CH_2CF_3				R = $CH_2CF_2CF_3$				R = $CH_2CF_2CF_2CF_3$			
	a_1	b_1	peak with m/z		a_1	b_1	peak with m/z		a_1	b_1	peak with m/z	
			102	115			102	115			102	115
Leu-Ile	144 (100)	172 (0.9)	(12.5)	(0.7)	144 (100)	172 (0.9)	(10.8)	(0.6)	144 (100)	172 (1.0)	(10.6)	(0.8)
Ile-Leu	144 (100)	172 (0.9)	(1.2)	(1.2)	144 (100)	172 (1.4)	(1.5)	(2.1)	144 (100)	172 (1.2)	(1.2)	(1.7)
Ala-Ile	102 (100)	130 (2.4)	(100)	(2.0)	102 (100)	130 (2.6)	(100)	(2.2)	102 (100)	130 (2.8)	(100)	(2.3)
Ile-Ala	144 (100)	172 (7.8)	(2.2)	(3.7)	144 (100)	172 (6.1)	(2.8)	(2.2)	144 (100)	172 (5.3)	(4.0)	(3.5)
Gly-Pro	88 (26.2)	116 (3.6)	(0)	(0)	88 (19.2)	116 (2.2)	(0)	(0)	88 (26.9)	116 (5.3)	(0)	(0)
Pro-Gly	128 (100)	156 (2.9)	(0)	(0)	128 (100)	156 (3.5)	(0)	(0)	128 (100)	156 (3.5)	(0)	(0)
Ile-Gly	144 (100)	172 (3.9)	(1.7)	(7.6)	144 (100)	172 (5.5)	(0.9)	(8.0)	144 (100)	172 (5.4)	(1.9)	(7.5)
Gly-Ile	88 (100)	116 (23.9)	(0)	(5.7)	88 (100)	116 (32.6)	(0)	(5.1)	88 (100)	116 (36.2)	(0)	(5.0)
Leu-Gly	144 (100)	172 (4.6)	(15.2)	(2.4)	144 (100)	172 (5.7)	(16.7)	(2.6)	144 (100)	172 (3.4)	(13.8)	(2.2)
Gly-Leu	88 (78.2)	116 (19.6)	(0.2)	(1.2)	88 (77.9)	116 (20.3)	(0.3)	(2.0)	88 (78.7)	116 (22.4)	(0.5)	(2.9)
Ala-Leu	102 (100)	130 (1.9)	(100)	(0.9)	102 (100)	130 (2.1)	(100)	(1.1)	102 (100)	130 (2.2)	(100)	(1.2)
Leu-Ala	144 (100)	172 (6.2)	(18.4)	(1.3)	144 (100)	172 (5.9)	(14.8)	(1.0)	144 (100)	172 (5.6)	(16.3)	(1.2)

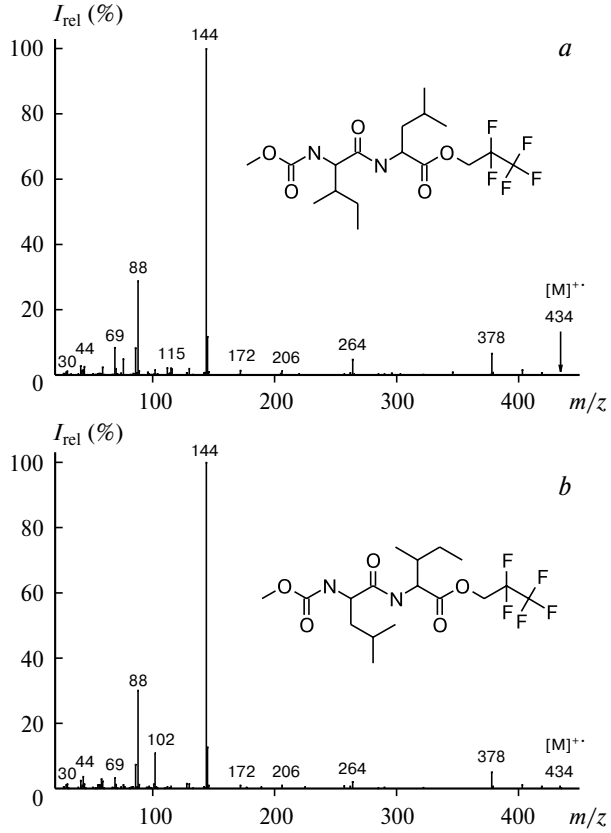


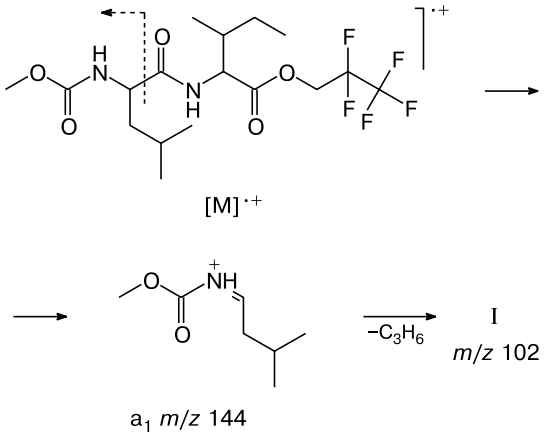
Fig. 2. Electron ionization mass spectra of 2,2,3,3,3-pentafluoropropyl esters of *N*-methoxycarbonyl-L-isoleucyl-L-leucine (*a*) and *N*-methoxycarbonyl-L-leucyl-L-isoleucine (*b*).

butene (side chain) and alkanol (ester group) molecules. The ion with the mass 102 is formed in the case of the leucine derivatives through the "amine" dissociation (formation of the a_1 ion) followed by the elimination of the C_3H_6 fragment due to the side isobutyl group. We analyzed the electron ionization mass spectra of similar dipeptide derivatives including leucine and isoleucine to detect diagnostic ions that make it possible to distinguish these amino acids in dipeptides and determine their sequence.

The mass spectra of *N*-methoxycarbonyl derivatives of pentafluoropropyl esters of isomeric Ile-Leu and Leu-Ile (Fig. 2) demonstrate a clear difference in intensity of the peak with *m/z* 102. This peak is present only in the spectrum of the latter derivative containing the N-terminal Leu residue (see Fig. 2, *b*) and absent in the case of the dipeptide derivative containing the C-terminal Leu residue. In the case of the Ile-Leu derivative, the origin of this ion (I) is caused by the same processes (Scheme 3) that are described for the leucine derivative.

The data presented in Table 2 show that this ion is characteristic of all similar dipeptide derivatives containing the N-terminal Leu residue. As a result, this residue in dipeptides (and, perhaps, in other oligopeptides) can be identified if situated at the N-end.

Scheme 3



No characteristic properties that make it possible to identify and determine the position of the isoleucine residue in the electron ionization mass spectra of the dipeptide derivatives were found. In the case of the dipeptide derivatives, the peak of the diagnostic ion with *m/z* 115 characteristic of the spectra of alkyl esters of *N*-methoxycarbonylisoleucine is not observed regardless of the position of the Ile residue (see Table 2).

The results obtained indicate possibilities of modification of dipeptides by alkyl chloroformate—alkanol mixtures for their identification and quantitative determination by GC/MS in complex mixtures and aqueous media (biological liquids). Analysis by a combination of mass spectrometry with electron and chemical ionization and tandem mass spectrometry is most efficient for structural studies. It should be mentioned that the derivatives of isomeric dipeptides often have close retention times under GC conditions. In these cases, the use of selective ion monitoring makes it possible to resolve overlapping chromatographic peaks and perform qualitative and quantitative analyses of dipeptides in complex mixtures.

The authors are grateful to A. Mandelbaum and C. Denekamp (Technion Israel Institute of Technology, Haifa, Israel) for kindly presented possibility to measure the mass spectra with collisional activation.

References

1. V. G. Zaikin and J. M. Halket, *Eur. J. Mass Spectrom.*, 2005, **11**, 611.
2. G. Samonina, I. Ashmarin, and L. Lyapina, *Pathophysiology*, 2002, **8**, 229.
3. Yu. A. Ovchinnikov and A. A. Kuryshkin, *FEBS Lett.*, 1972, **21**, 300.
4. R. M. Caprioli, W. E. Seifert, Jr., and D. E. Sutherland, *Biochem. Biophys. Res. Commun.*, 1973, **55**, 67.

5. O. S. Reshetova, B. V. Rozynov, M. V. Bezrukov, and I. A. Bogdanova, *Bioorg. Khim.*, 1986, **12**, 1625 [*Sov. J. Bioorg. Chem.*, 1986, **12** (Engl. Transl.)].
6. K. Biemann and S. A. Martin, *Mass Spectrom. Rev.*, 1987, **6**, 1.
7. C. D. Márquez, S. T. Weintraub, and P. C. Smith, *J. Chromatogr. B*, 1994, **658**, 213.
8. P. Husek, *J. Chromatogr. B*, 1998, **717**, 57.
9. P. Cao and M. Moini, *Rapid Commun. Mass Spectrom.*, 1997, **11**, 349.
10. P. Capitan, T. Malmezat, D. Breuille, and C. Obled, *J. Chromatogr. B*, 1999, **732**, 127.
11. K. Biemann, *Methods Enzymol.*, 1990, **193**, 886.
12. R. S. Borisov, A. A. Rychkov, B. V. Vas´kovskii, and V. G. Zaikin, *Mass-spektrometriya* [*Mass Spectrometry*], 2004, **1**, 199 (in Russian).

Received August 4, 2006