PREPARATION, SPECTRAL AND PHYSICOCHEMICAL CHARACTERISTICS OF METHYLAMIDE N^{α}-PHENYLTHIOCARBAMOYL DERIVATIVES OF NATURALLY OCCURRING AMINO ACIDS

Manfred PAVLÍK[a], Ivan KLUH[a], Františka PAVLÍKOVÁ[b], Soňa VAŠÍČKOVÁ[a] and Vladimír KOSTKA[a]

[a] *Institute of Organic Chemistry and Biochemistry, Czechoslovak Academy of Sciences, 166 10 Prague 6 and*
[b] *Prague Institute of Chemical Technology, 166 28 Prague 6*

Received June 15, 1988
Accepted December 7, 1988

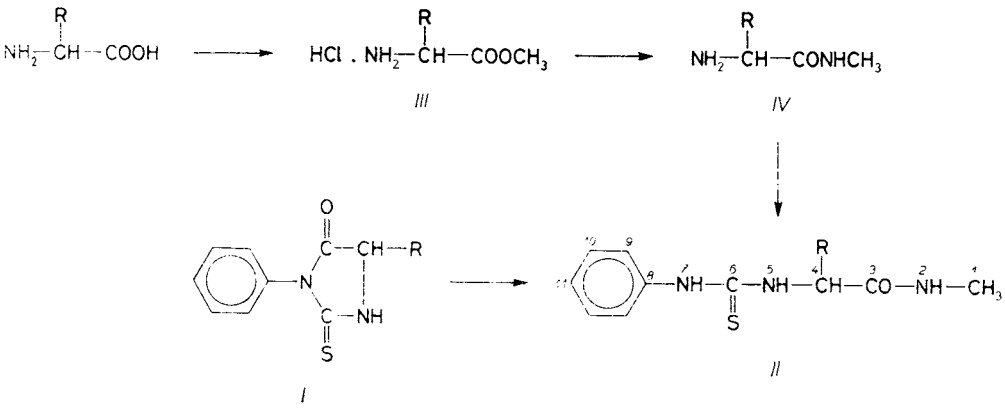
The methylamide N^{α}-phenylthiocarbamoyl derivatives of encoded amino acids *II* were prepared either by the addition of phenylisothiocyanate to amino acid methylamides or by the treatment of amino acid phenylthiohydantoins (5-alkyl-3-phenyl-2-thioxo-4-imidazolinones) *I* with methylamine. The derivatives were prepared of 19 amino acids and their melting points, ^{1}H NMR, ^{13}C NMR, mass, ultraviolet and infared spectra were measured.

The sequence degradation of proteins and peptides[1,2] is the method which at present is employed most frequently for the determination of the primary structures of these substances. The final products of liberation of the N-terminal amino acids from the degraded polypeptide chains are 5-alkyl-3-phenyl-2-thioxo-4-imidazolinones (*I*) usually referred to as phenylthiohydantoins of amino acids (PTH). These derivatives are separated as a rule by HPLC (ref.[3]) and detected by measurement of their UV absorbance at 269 nm, the detection limit being approximately 1 pmol (ref.[4]). During sequence degradation compounds *I* arise by acid conversion of less stable primary products of cleavage of amino acid anilinothiazolinones (ATZ). The conversion is effected in an aqueous solution of trifluoroacetic acid or in a methanolic solution of hydrochloric acid[5]. If the acid conversion is replaced by "alkaline conversion" the methylamides of N^{α}-phenylthiocarbamoyl amino acid derivatives *II*, (refs[6,7]) are obtained as the end products.

This work has been aimed at the preparation of compounds *II*, which are the products of modified sequence degradation of proteins and peptides using methylamine for conversion, and at the determination of the structure and spectral characteristics of *II*.

The standards of *II* were prepared by two different procedures. The first one was based on the reaction of free amino acids with methanol in the presence of $SOCl_2$ yielding amino acid methyl esters (*III*) which were subsequently converted into

methylamides (*IV*) by methylamine in an aqueous medium. Derivatives *IV* were treated with phenyl isothiocyanate (PITC) in an alkaline solution to afford the required amino acid derivatives *II* (Scheme 1).



SCHEME 1

An alternative method of preparation of compounds *II* is based on the treatment of phenylthiohydantoins *I* with methylamine.

EXPERIMENTAL

The system used for high pressure liquid chromatography consisted of two Beckman 114M Solvent Delivery Modules, Beckman 421 Controller, Altex 210 Injection Valve and, if necessary, of Waters-Millipore WISP 470 Autosampler. The chromatography was monitored at 254 or 269 nm by a Shimadzu SPD-2A detector. The effluents were quantitated in a Shimadzu C-R3A Integrator. The separation was carried out in Beckman Ultrasphere ODS columns (5 μm, 4·6 . . 250 mm) or in Dupont Zorbax ODS columns of identical dimensions. An elution gradient was employed of methanol–water or alternatively of acetonitrile–water. The course of column chromatography was monitored by a UVM4 Detector (Vývojové dílny, ČSAV).

The melting points were determined in a Boetius block and are not corrected. The ultraviolet spectra were measured in Specord UV-VIS (Zeiss, Jena) on ethanolic solutions. The infrared spectra were measured in a Perkin–Elmer 580 and 621 Spectrophotometer using KBr pills. The ^{1}H and ^{13}C NMR spectra were measured in Bruker AM 400 Spectrometer using solutions of compounds in $(CD_3)_2SO$ or in $CDCl_3$, respectively, at 297 K with TMS as internal standard. The conditions of the measurement were the following: ^{1}H NMR-working frequency 400·13 MHz, digital resolution 0·2 Hz/point; ^{13}C NMR-working frequency 100·62 MHz, digital resolution 1 Hz/point. The structure was confirmed and the individual signals assigned using homonuclear decoupling[8] in ^{1}H NMR spectroscopy and APT (ref.[9]) in ^{13}C NMR spectroscopy. The mass spectra were measured in AEI MS-902 mass spectrometer (70 eV).

TABLE I

Melting points, yields and elemental analyses of compounds *II*

Compound Amino acid	M.p. °C	Yield, %[a] *A*	*B1*	*B2*	Formula M.w.	Calculated/Found % C	% H	% N	% S
IIa	158–160[b]	30	80	–	$C_{10}H_{13}N_3OS$	53·79	5·87	18·82	14·44
Gly					223·3	53·64	5·49	18·63	14·28
IIb	162[b]	65	90	–	$C_{11}H_{15}N_3OS$	55·67	6·37	17·71	13·51
Ala					237·3	55·58	6·25	17·97	13·66
IIc	154[b]	50	–	90	$C_{13}H_{19}N_3OS$	58·84	7·22	15·83	12·08
Val					265·4	58·64	6·96	15·84	11·90
IId	160–163[b]	50	–	–	$C_{14}H_{21}N_3OS$	60·18	7·57	15·04	11·48
Leu					279·4	60·12	7·40	14·81	11·23
IIe	158–160[b]	60	–	–	$C_{14}H_{21}N_3OS$	60·18	7·57	15·04	11·48
Ile					279·4	60·38	7·39	15·11	11·71
IIf	65–68[c]	55	–	–	$C_{17}H_{19}N_3OS$	65·15	6·12	13·41	10·23
Phe					313·4	64·93	5·94	13·16	9·99
IIg	[d]	35	80	–	$C_{17}H_{19}N_3O_2S$	61·98	5·81	12·76	9·97
Tyr					329·4	62·24	5·62	12·31	9·51
IIh	166–167[c]	70	–	85	$C_{13}H_{19}N_3OS_2$	52·50	6·44	14·13	21·56
Met					297·4	52·65	6·36	14·28	21·03
IIi	[d]	70	75	95	$C_{13}H_{17}N_3OS$	59·29	6·51	15·96	12·17
Pro					263·4	59·42	6·46	16·04	12·12
IIj	179[c]	65	90	80	$C_{19}H_{20}N_4OS$	64·75	5·72	15·90	9·10
Trp					352·5	64·31	5·71	16·08	9·20
IIk	[d]	–	50	90	$C_{13}H_{20}N_4O_3S$	49·98	6·45	17·94	10·26
Asp[e]					312·4	49·87	6·39	17·73	10·07
IIl	[d]	–	60	80	$C_{12}H_{16}N_4O_2S$	51·41	5·75	19·99	11·44
Asn					280·3	52·08	5·29	19·75	10·98
IIm	[d]	–	30	90	$C_{14}H_{22}N_4O_3S$	51·51	6·79	17·16	9·82
Glu[e]					326·4	51·91	6·64	17·29	9·71
IIn	[d]	–	40	80	$C_{13}H_{18}N_4O_2S$	53·04	6·16	19·03	10·89
Gln					294·4	52·65	6·06	19·12	10·85
IIo	144–145[c]	30	–	75	$C_{11}H_{15}N_3O_2S$	52·15	5·97	16·59	12·66
Ser					253·3	52·13	5·73	16·65	11·96
IIp	186–188[c]	35	–	80	$C_{12}H_{17}N_3O_2S$	53·91	6·41	15·72	11·99
Thr					267·3	53·67	6·62	15·63	11·47
IIq	170–172[c]	50	–	–	$C_{21}H_{27}N_5OS_2$	58·71	6·34	16·30	14·93
Lys					429·6	58·93	6·23	16·42	14·80

TABLE I
(*Continued*)

Compound Amino acid	M.p. °C	Yield, %[a] *A*	*B1*	*B2*	Formula M.w.	Calculated/Found % C	% H	% N	% S
IIr	[d]	—	—	80	$C_{14}H_{17}N_5OS$	55·43	5·65	23·08	10·57
His					303·4	55·58	5·60	22·58	10·65
IIs	[d]	—	—	85	$C_{14}H_{22}N_6OS$	52·15	6·89	26·06	9·94
Arg					322·4	—[f]	—[f]	—[f]	—[f]

[a] The yields are given only for those procedures which were used for the preparation of the corresponding derivative *II* and only when the derivative was isolated in pure state. [b] The compound was crystallized from the mixture ethanol–water. [c] The compound was crystallized from the mixture ethanol–chloroform. [d] The compound was prepared in amorphous or oily state. [e] The derivative was isolated as the methylammonium salt. [f] The elemental analysis is not in agreement with the calculated value.

Preparation of Compounds *II*

A: The amino acid (0·1 mol) was dissolved or suspended in 50 ml of dry methanol. The solution was cooled down to −20°C and $SOCl_2$ (0·12 mol) was added. The reaction mixture was set aside for 12—72 h at room temperature and was then concentrated to 25 ml. The product was precipitated by the addition of 500 ml of dry diethyl ether. The precipitate was separated and immediately dissolved in 10—20 ml of 30% methylamine in water (0·5—1·0 ml). The reaction mixture was heated at 40°C for 30 min, was then concentrated three times with 50 ml of methanol in a vacuum evaporator and the crystals of methylamine hydrochloride which had separated were isolated. The methanolic mother liquor was concentrated and dried in vacuo. The crystalline or oily amino acid methylamide was dissolved in 15 ml of methanol, 10 ml of trimethylamine and 13 ml of phenylisothiocyanate (0·11 mol) were added. The mixture was heated 30 min at 40°C, was then concentrated by the evaporation of methanol and triethylamine. The residue was precipitated by water and the precipitate together with the aqueous phase was extracted three times with 50 ml of hexane. The aqueous phase was concentrated together with the precipitate. The crude product was purified by column chromatography on Kieselgel 100 (i.d. 2·5 cm, $l = 50$ cm) using chloroform and up to 20% methanol in chloroform (according to the character of the derivative prepared) for elution. The content of the products in the effluent was monitored by a UV detector at 300 nm. The fractions collected were concentrated and analyzed by TLC on Silufol plates ($CHCl_3$, or 10% methanol in $CHCl_3$, detection by UV light or iodine vapors) or by HPLC. Compounds *II* were crystallized from ethanol–water or from ethanol–chloroform. The yields varied between 30 and 70%.

B1: The PTH derivative of the amino acid (0·01 mol) was dissolved in 2 ml of 30% methylamine in water ($17·4 \cdot 10^{-3}$ mol) and was then heated at 40°C for 30—50 min. The content of the compounds in the reaction mixture was examined by TLC and HPLC. After completion of the reaction the reaction mixture was concentrated, diluted and evaporated 3 times with methanol. The crude product was purified by column chromatography under the conditions described under *A*. The yields of compounds *II* varied between 30 and 95%.

TABLE II

^{1}H NMR shifts (δ, ppm) of compounds *IIa*—*IIs*, measured in $(CD_3)_2SO$. See Scheme 1 for the numbering of the hydrogen atoms

Compound Amino acid	H-1[a]	H-4	H-9	H-10	H-11	H-2	H-5	H-7	R[b]
IIa[c] Gly	2·82 d	4·30 d	7·29 m	7·44 m	7·29 m	6·21 s	7·05 s	8·06 s	
IIb Ala	2·62 d	4·84 k	7·50 d	7·32 t	7·10 t	8·10 d	7·83 d	9·82 s	1·31 d, 3 H (J = 6·9)
IIc[c] Val	2·80 d	4·81 t	7·30 d	7·40 t	7·26 t	6·34 s	7·06 d	8·34 s	0·96 t, 6 H (J = 7·1, 7·9); 2·17 sex, 1 H (J = 6·9)
IId[c] Leu	2·81 d	5·07 q	7·27 m	7·41 t	7·27 m	6·58 d	6·81 d	8·18 s	0·95 d, 6 H (J = 5·8); 1·65 m, 3 H
IIe[c] Ile[d]	2·79 d 2·80 d	4·85 t[e] 4·94 dd[e]	7·31 m	7·40 m	7·26 m	6·10 d	7·05 d	8·40 s	0·91 m, 6 H; 1·15 m, 1·46 m, 1·54 m, 2 H; 1·96 m, 1 H
IIf[c] Phe	2·65 d	5·20 q		7·05—7·40 m		6·09 d	7·05—7·40 m	8·5 s	3·04 dd, 1 H (J = 8·2); 3·26 dd, 1 H (J = 6·3, 13·7); 7·05—7·40 m, 5 H
IIg Tyr	2·58 d	5·08 q	7·42 d	7·29 t	7·09 t	7·90 d	7·56 d	9·74 s	2·81 dd, 1 H (J = 6·5); 3·0 dd, 1 H (J = 6·1, 13·75); 6·65 d, 2 H (J = 8·5); 6·94 d, 2 H (J = 8·5); 9·1 s, 1 H
IIh Met	2·63 d	4·95 q	7·51 d	7·31 t	7·10 t	8·02 d	7·81 d	9·74 s	1·99 m, 2 H; 2·03 s, 3 H; 2·44 t, 2 H (J = 7·9)
IIi[c] Pro	2·76 d	5·11 s	7·32 d	7·37 t	7·20 t	7·05 t	—	7·57 s	2·05 m, 2 H; 2·21 m, 2·30 m, 2 H; 3·58 m, 3·27 m, 2 H

IIj Trp	2·58 d	5·08 q	7·39 d	7·27 t	7·08 m	8·05 q	7·26 d	9·8 s	3·16 dd, 1 H (J = 5·9); 3·31 dd, 1 H (J = 5·7, 14·7); 6·96 t, 1 H (J = 7·5); 7·05 m, 3 H; 7·33 d, 1 H (J = 8·0); 7·56 d, 1 H (J = 7·8); 10·84 s, 1 H
IIk Asp	2·57 d	4·99 t	7·59 d	7·27 t	7·04 t	7·98 s	8·9 bs	8·9 bs	2·26 m, 1 H; 2·31 s, 3 H; 2·62 m, 1 H; 6·0 bs, 3 H
IIi Asn	2·59 d	5·11 s	7·49 d	7·32 t	7·11 t	7·81 m	7·99 t	9·97 s	2·5[f], 2·69 m, 1 H; 6·94 s, 1 H; 7·42 s, 1 H
IIm Glu	2·60 d	4·78 t	7·58 d	7·28 t	7·05 t	8·06 d	8·85 s	10·5 bs	1·9 m, 2 H; 2·05 m, 2 H; 2·33 s, 3 H; 6·5 bs, 3 H
IIn Gln	2·62 d	4·88 t	7·51 d	7·32 t	7·10 t	8·11 q	7·85 s	9·7 s	1·86 m, 1·99 m, 2 H; 2·06 m, 2 H; 6·76 s, 1 H; 7·37 s, 1 H
IIo Ser	2·64 d	4·86 m	7·53 d	7·31 t	7·09 t	7·85 d	7·73 d	9·89 s	3·65 k, 1 H (J = 5·3); 3·78 k, 1 H (J = 4·4, 10·6); 4·93 t, 1 H (J = 5·2)
IIp Thr	2·62 d	4·8 dd[g]	7·56 d	7·33 t	7·10 t	7·77 q	7·61 d	10·01 s	1·06 d, 3 H (J = 6·3); 4·16 m, 1 H; 5·08 d, 1 H (J = 4·4)
IIq Lys	2·62 d	4·87 q	7·52 d[h]	7·31 m	7·09 m	8·10 q	7·8 d[i]	9·78 s[j]	1·29 m, 2 H; 1·55 k, 2 H; 1·72 m, 2 H; 3·43 m, 2 H; 7·31 m[k], 3 H; 7·39 d[h,k], 2 H; 7·73 s[l,k], 1 H; 9·47 s[j,k], 1 H
IIr His	2·57 d	5·05 q	7·45 d	7·31 t	7·10 t	7·94 q	7·88 d	9·85 s	2·98 d, 2 H (J = 5·9); 6·77 s, 1 H; 7·51 s, 1 H; 11·82 s, 1 H
IIs Arg	2·62 d	4·89 t	7·56 d	7·31 t	7·09 t	8·15 q	viz R	10·0 bs	1·49 m, 2 H; 1·67 m, 1·73 m, 2 H; 3·13 t, 2 H (J = 6·85); 7·4 bs, 5 H

[a] J(H-1, H-2) = 4·8—4·9 Hz (in $CDCl_3$) and 4·4—4·6 Hz (in $(CD_3)_2SO$). [b] the assignment of the signals in amino acid chain R is the usual one[10], the values of the interaction constants are given in Hz; [c] measured in $CDCl_3$; [d] two unresolved isomers; [e] 4·85 t (J = 8·2), 4·95 dd (J = 8·8, 8·6); [f] partly overlapped by the solvent; [g] J = 2·2, 8·1; [h,i,j] the signals may be mutually interchanged; [k] the compound was isolated and characterized in the form of an N$^{\omega}$-phenylthiourea derivative.

TABLE III

^{13}C NMR chemical shifts (δ, ppm) of compounds *IIa*—*IIs* measured in $(CD_3)_2SO$. See Scheme 1 for the numbering of carbon atoms

Compound Amino acid	C-1	C-3	C-4	C-6	C-8	C-9	C-10	C-11	R[a]
IIa[b] Gly	26·33	169·06	48·93	180·67	136·06	125·04	130·22	127·47	—
IIb Ala	25·71	172·51	52·62	179·53	139·46	122·96	128·67	124·25	19·38
IIc[b] Val	26·19	171·81	64·22	180·90	136·66	124·96	129·87	127·00	18·76, 19·06, 30·77
IId[b] Leu	26·28	172·56	56·99	180·53	136·32	125·04	129·99	127·20	22·57, 22·82, 24·89, 40·75
IIe[b,c] Ile	26·25 26·22	171·94 171·80	63·12 62·87	180·97 180·68	136·57	125·00 124·91	129·94 129·89	127·11 127·01	11·12, 11·66, 15·12, 15·35, 25·40, 26·05, 36·81, 37·20
IIf[b] Phe	26·18	171·44	59·65	179·96	136·56[d]	124·82	129·82	127·09	38·35, 126·91, 128·70, 129·29, 136·40[d]
IIg Tyr	25·58	171·02	58·40	179·79	139·26	122·89	130·21[d]	124·24	37·37, 115·06, 127·21, 130·82[d], 156·00
IIh Met	25·59	171·04	56·32	180·15	139·36	122·84	128·50	124·14	14·76, 29·28, 32·65
IIi[b] Pro	26·23	171·91	64·78	179·97	139·18	125·85	128·67	126·09	24·99, 28·83, 49·61

IIj Tzp	25·74	171·36	57·76	179·67	139·29	122·96	128·64	124·26	28·21, 109·50, 111·36, 118·37, 118·69, 120·97, 123·72, 127·80, 136·17
IIk Asp	25·69	172·24	55·21	179·73	139·85	122·52	128·11	123·36	24·59, 32·99, 174·50
III Asn	25·71	171·90	54·29	179·78	139·04	122·69	128·48	124·09	37·04, 169·97
IIm Glu	25·27	171·62	56·82	179·93	139·73	122·33	128·03	123·25	24·31, 29·01, 32·96, 176·21
IIn Gln	25·46	171·21	56·30	179·79	139·24	122·66	128·24	124·00	28·58, 30·88, 173·42
IIo Ser	25·65	170·19	59·06	179·97	139·34	122·64	129·43	123·98	61·77
IIp Thr	25·61	170·28	62·31	180·70	139·30	122·48	128·44	123·96	20·13, 66·44
IIq Lys	25·44	171·48	56·63	179·74[d]	139·27	122·59	128·40	123·94	22·32, 28·32, 32·48, 43·69, 122·59[e], 123·94[e], 128·40[e], 139·26[e], 180·20[d,e]
IIr His[f]	25·57	170·97	57·16	179·67	139·02	122·75	128·48	124·05	29·94, 134·50
IIs Arg	25·46	171·38	56·27	180·05	139·49	122·41	128·30	123·79	24·73, 29·64, 40·31, 156·90

[a] The assignment of the signals in amino acid chain R is the usual one; [b] measured in $CDCl_3$; [c] the isomers cannot be resolved; [d] the signals may be mutually interchanged; [e] the compound was isolated and characterized in the form of an N^{ω}-phenylthiourea derivative; [f] signals C-2 and C-5 of the imidazole ring were not found[12].

TABLE IV
Mass spectra of compounds *II*

Compound Amino acid derivative	*m/z* (% of relative intensity)
IIa Gly	M^+ 223 (81·1); 222 (10·8); 192 (27·0); 163 (16·2); 136 (34·2); 135 (54·1); 131 (21·6); 109 (10·8); 105 (10·8); 103 (10·8); 93 (89·2); 88 (23·8); 77 (62·2); 58 (13·5); 51 (24·3); 32 (70·2); 30 (100).
IIb Ala	M^+ 237 (25·3); 206 (87·9); 177 (13·2); 163 (5·5); 152 (15·3); 145 (9·9); 136 (27·5); 135 (48·5); 120 (11·0); 106 (16·5); 93 (62·6); 77 (62·6); 41 (100).
IIc Val	M^+ 265 (16·8); 234 (33·2); 219 (10·5); 207 (11·6); 136 (23·7); 135 (16·8); 105 (18·4); 93 (63·2); 77 (52·6); 72 (100); 58 (31·6); 55 (34·2).
IId Leu	M^+ 279 (52·0); 248 (72·0); 223 (28·0); 205 (72·0); 192 (36·0); 187 (20·0); 136 (56·0); 135 (60·0); 93 (100); 86 (96·0); 77 (88·0); 58 (36·0); 44 (56·0); 43 (48·0).
IIe Ile	M^+ 279 (47·7); 248 (77·3); 246 (12·5); 221 (27·3); 219 (40·9); 192 (15·9); 191 (9·5); 187 (7·9); 166 (8·2); 153 (18·2); 152 (20·5); 151 (18·2), 136 (36·4); 135 (35·2); 128 (27·2); 119 (15·9); 110 (15·9); 109 (13·9); 93 (100); 86 (90·9); 77 (70·5); 58 (34·0); 51 (22·7); 41 (36·4); 32 (54·5); 30 (40·9).
IIf Phe	M^+ 313 (7·5); 282 (65·9); 220 (8·4); 166 (27·3); 162 (22·7); 152 (7·7); 136 (18·2); 135 (90·9); 93 (100); 91 (93·2); 87 (54·5); 77 (17·5); 66 (38·6); 65 (34·1); 58 (15·9); 51 (43·2); 41 (18·2); 38 (27·3); 32 (20·5).
IIg Tyr	M^+ 329 (0·6); 298 (2·6); 236 (2·1); 192 (6·9); 178 (4·4); 177 (7·3); 166 (22·0); 136 (13·2); 135 (48·5); 107 (39·7); 93 (100); 77 (39·7); 66 (47·1); 65 (26·5); 51 (23·5); 40 (23·4).
IIh Met	M^+ 297 (3·3); 266 (26·7); 205 (33·5); 204 (13·3); 192 (13·3); 161 (14·0); 143 (10·7); 136 (14·7); 135 (40·0); 130 (16·7); 93 (100); 77 (53·3); 66 (24·0); 65 (16·0); 56 (24·7); 51 (20·7).
IIi Pro	M^+ 263 (9·5); 232 (9·7); 203 (2·3); 136 (7·5); 135 (43·2); 127 (4·1); 93 (6·7); 77 (3·4); 70 (100); 51 (11·8); 43 (7·1); 32 (27·3).
IIj Trp	M^+ 352 (1·2); 321 (0·6); 201 (4·9); 200 (28·4); 170 (6·0); 166 (6·0); 159 (7·6); 135 (94·0); 130 (100); 93 (77·6); 77 (100).
IIk Asp	264 (1·0); 251 (3·0); 188 (1·7); 166 (11·3); 135 (10·0); 93 (100); 77 (16·7); 74 (8·7); 66 (23·3); 65 (13·3).
IIl Asn	M^+ 280 (18·0); 249 (31·2); 205 (22·0); 187 (28·0); 166 (30·0); 136 (25); 135 (69·0); 93 (100); 87 (46); 77 (62); 66 (48); 64 (33); 52 (35); 44 (35).

TABLE IV
(*Continued*)

Compound Amino acid derivate	*m/z* (% of relative intensity)
IIm Glu	166 (1·7); 135 (2·5); 93 (45·8); 31 (75·0); 30 (100).
IIn Gln	201 (2·1); 184 (2·8); 166 (7·8); 142 (4·3); 135 (8·6); 93 (100); 84 (16·4); 77 (11·4); 74 (7·8); 66 (27·1); 65 (14·3).
IIo Ser	M^+ 253 (0·8); 235 (0·3); 222 (0·3); 166 (3·4); 153 (3·2); 152 (3·0); 136 (4·2); 135 (18·4); 93 (100); 77 (18·4); 66 (19·7); 60 (39·5).
IIp Thr	M^+ 267 (6·0); 223 (5·2); 218 (3·6); 182 (9·6); 175 (18·0); 153 (100); 152 (68·0); 136 (40·0); 135 (13·6); 119 (11·2); 116 (14·0); 93 (56·0); 77 (44·0); 74 (32·0); 58 (16·0); 56 (22·4); 51 (20·4); 32 (100); 30 (24·0).
IIq Lys	398 (0·2); 371 (0·5); 312 (0·2); 302 (1·2); 263 (2·4); 243 (5·0); 201 (11·7); 194 (1·9); 184 (5·7); 166 (21·4); 143 (16·7); 136 (26·2); 135 (100); 93 (66·7); 77 (95·0); 66 (100); 51 (85·7); 39 (83·3).
IIr His	210 (10·8); 166 (10·8); 135 (2·4); 133 (1·8); 110 (4·8); 104 (3·6); 93 (100); 82 (11·0); 81 (6·4); 77 (6·8); 66 (22·0); 65 (14·0); 39 (16·0); 30 (34·0).
IIs Arg	229 (0·1); 166 (5·6); 152 (1·0); 135 (3·7); 110 (1·3); 104 (1·2); 93 (100); 77 (7·9); 74 (4·8); 66 (34·0); 65 (17·6); 39 (13·2).

B2: The PTH derivative of the amino acid (0·01 mol) was condensed at −30°C with methylamine vapors (0·3—0·5 ml, 8·7—14·5 . 10^{-3} mol, dried by KOH), the flask was closed and set aside 2 h at room temperature; methylamine was evaporated off and the product was dried afterwards. This procedure afforded almost pure compounds *II* in a yield of 80—95%.

RESULTS AND DISCUSSION

The melting points and yields of compounds *II* are listed in Table I. When the compounds could not be obtained crystalline they were either oily or amorphous products. The yields of compounds *II* are given only for procedures yielding the corresponding derivatives which were isolated in pure state afterwards. Procedure Bb using treatment of amino acid PTH's with trimethylamine was the most convenient one and afforded up to 95% yields even with polar amino acids.

When compounds *IIl* and *IIn* were synthesized (Asn and Gln derivatives) a simultaneous formation of other products was observed in which the —CO—NH_2

group of the side chain is replaced by the —CO—NH—CH_3 group. Both these compounds were isolated when procedure *A* was used for their preparation in a yield of 20–30%. The characteristics of these products are not listed here. Compounds *IIk* and *IIm* (derivatives of Asp and Glu) were isolated in the form of the ammonium salts of the free carboxyl. Table I lists also the elemental analyses of compounds *II* which are, with the exception of compound *IIs* (Arg derivative), in agreement with calculated values. Considerable differences from the calculated value were observed with compound *IIs* and the analytical data are therefore not listed. The reason for these differences is obviously the hygroscopic character of this derivative. The ^1H and ^{13}C NMR data are summarized in Tables II and III. The data in these tables show that the ^1H and ^{13}C NMR spectra are characteristic of the whole series of amino acid derivatives, i.e. that the values of their ^1H and ^{13}C chemical shifts and of other parameters of the backbone of derivatives *II*, except for C-4 and H-4, do not change any significantly with the change in substituent R. A partial shift can be observed when solvents are switched from $CDCl_3$ to $(CD_3)_2SO$. There are problems with the identification of nitrogen bound protons H-2, H-5 and H-7 whose shape

TABLE V
Ultraviolet spectra of compounds *II*

Compound	Amino acid derivative	Band I		Band II		Band III	
		λ_1, nm	$\log \varepsilon_1$	λ_2, nm	$\log \varepsilon_2$	λ_3, nm	$\log \varepsilon_3$
IIa	Gly	216	2·99	243	3·09	266	3·10
IIb	Ala	216	2·97	246	3·09	267	3·08
IIc	Val	211	2·96	246	3·05	266	3·00
IId	Leu	213	2·99	246	3·11	265	3·06
IIe	Ile	215	3·02	246	3·14	266	3·09
IIf	Phe	216	3·11	248	3·11	265	3·09
IIg	Tyr	225	3·16	250	3·11	267	3·13
IIh	Met	214	3·07	247	3·11	266	3·06
IIi	Pro	223	3·05	246	3·13	—	—
IIj	Trp	226	3·33	—	—	268	3·25
IIk	Asp	211	3·02	247	3·08	267	3·06
IIl	Asn	213	2·99	250	3·06	263	2·99
IIm	Glu	211	3·02	244	3·07	267	2·99
IIn	Gln	214	3·03	246	3·12	267	3·08
IIo	Ser	215	2·99	246	3·09	267	3·09
IIp	Thr	212	2·93	246	2·98	267	2·97
IIq	Lys	216	3·26	247	3·41	264	3·37
IIr	His	216	3·11	247	3·10	267	3·09
IIs	Arg	210	3·01	246	3·03	266	3·00

and multiplicity of signals is affected, in addition to derivative *II* itself, also by the presence of moisture in the solvent. As a result, the shape of the signal changes from a very broad one to a well defined doublet for H-5 (^3J(H-5, H-4) = 7·3–8·3 Hz). a quartet for H-2 (^3J(H-2, H-1) = 4·6 Hz) and a singlet for H-7. This fact together with the prochiral properties of H-1 protons in substituent R again complicate the shape and multiplicity of the H-4 signal which shows a considerable variability from a broadened singlet to a well defined multiplet (cf. Table II). The characterization of substituent R in ^{1}H and ^{13}C NMR spectra is unambiguous (see also Tables II and III). The values of their ^{1}H and ^{13}C NMR chemical shifts do not differ any significantly from the values of the chemical shifts characterizing free amino acids and their interpretation is thus evident[10,11]. The ^{1}H and ^{13}C NMR spectrum of derivative *IIe* (isoleucine derivative) shows the presence of two isomers at a ratio of 1 : 1, as a result of the presence of another asymmetric center the molecule. The signal corresponding to carbons C-2 and C-5 of the imidazole ring was not found in the ^{13}C NMR spectrum of compound *IIc* (histidine derivative) and the signal of the carbon of the CH_2 group (29·94 ppm) was very broadened. These phenomena can be accounted for by the valence isomery of the imidazole substituent and of the quadrupole moment of two nitrogen atoms of the ring[12] .The ^{1}H and ^{13}C NMR spectra of the given class of amino acid derivatives can be used for their characterization since the variability of the values of chemical shifts of the individual signals is 0·2 ppm and 1 ppm for the ^{1}H and ^{13}C NMR spectra, respectively.

Table IV lists the mass spectra of compounds *II*. The molecular ions were found for most compounds *II* with the exception of the derivatives of Asp (*IIk*), Glu (*IIm*), Gln (*IIn*), Lys (*IIq*), His (*IIr*) and Arg (*IIs*). This fact can be explained by the high molecular mass (compound *IIq*), ionic character (compounds *IIk* and *IIm*) or decomposition under the conditions of ionization. The ionization of all compounds *II* results in the removal of methylamine (molecular mass 31); an $M^+ - 31$ ion can be detected in the spectra which in its structure corresponds most likely to the PTH derivative of the given amino acid *I*. Additional cleavage of these PTH derivatives has been reported[13], the presence of characteristic ions has been detected in all cases. Another possible cleavage of compounds *II* can be represented by scission of the ionized molecule to ions $m/z = 58$ (structure CO—NH—CH_3) or $m/z = 166$, which corresponds in its structure to N-methyl-N′-phenylthiourea. This compound is probably a result of the decomposition of the above amino acid derivatives under the conditions of ionization.

Table V lists the main bands in the ultraviolet spectra of compounds *II*. All these compounds show the presence of three bands: the first one in the range of of 210 to 225 nm, the second one in the range of 243–250 nm and the third one in the range of 264–268 nm. An exception represent compounds *IIj* (Trp derivative) and *IIi* (Pro derivative) in which the second and third band, respectively are missing. This fact can be explained by overlapping by intensive neighboring bands.

TABLE VI

Characteristic IR bands (cm^{-1}) of compounds *II*

Compound Amino acid derivative	ν(NH)	Amide I ν(C=O)	Amide II δ(NH)		Aromatic ring ν(C—C)			ν(CS)	δ(NCS)	ν(CN) ν(CS) δ(NCN)	δ(CH_3)
						γ(CH)					
IIa Gly	3 295 3 245 3 215	1 664	1 523 1 563	1 497	1 597	707	762	1 256	1 306	1 414	1 439 1 453
IIb Ala	3 315 3 160	1 648	1 525	1 490	1 595	697 710	750	1 264	1 311	1 408 1 418	1 442 1 450
IIc Val	3 275	1 655	1 541	1 497	1 598	704 714	760	1 253	1 298 1 305	1 409	1 451
IId Leu	3 335 3 290 3 270	1 648	1 539	1 497	1 599	711	767	1 226 1 258	1 317	1 414	1 447
IIe Ile	3 375 3 310 3 190	1 648	1 537	1 495	1 592	699 708	758	1 240	1 314	1 412	1 450
IIf Phe	3 285	1 658	1 528	1 497	1 599	700	751	1 244	1 315	1 411	1 452
IIg Tyr	3 280	1 657	1 515	1 496	1 596	696	747	1 238	1 315	1 411	1 450
IIh Met	3 312 3 272 3 194	1 652	1 540	1 497	1 600	697	722 768	1 236	1 316	1 415	1 442

IIi Pro	3 275	1 65	1 528	1 498	1 599	697	757	1 229 1 254	1 305	1 405	1 448
IIj Trp	3 380 3 265	1 667	1 540	1 502	1 598	695	746 754	1 236	1 314	1 416	1 443 1 457
IIk[a] Asp	3 298 3 175	1 647	1 522 1 500	overlap	1 588	693	744	1 240	1 312	overlap	1 448
IIl Asn	3 390	1 662	1 523	1 498	1 596	698	747	1 240	1 318	1 410	1 448
IIm[b] Glu	3 270	1 650	1 530	1 497	1 590	700	760	1 260	1 315	1 400	1 448
IIn Gln	3 295 3 200	1 655	1 526	1 494	1 594	698	760	1 250	1 300 1 312	1 408	1 448
IIe[c] Ser	3 350 3 290 3 210	1 662	1 521	1 495	1 596	697	751	1 243 1 253	1 310	1 407	1 450
IIp[d] Thr	3 335 3 285	1 646	1 519	1 495	1 594	696	746 760	1 251 1 266	1 315	1 406	1 450
IIq Lys	3 300 3 250	1 635	1 545	1 496	1 598	696	750	1 239	1 317	1 402	1 453
IIr His	3 260	1 658	1 527	1 49	1 597	698	764	1 246	1 315	1 411	1 450
IIs Arg	3 260 3 165	1 660	1 524	1 498	1 595	704	759	1 252	1 316	1 412	1 452

[a] $\nu(COO^-)$ 1 390, 1 404; $\nu(NH_3^+)$ 2 130, 2 528, 2 620, 2 795; [b] $\nu(COO^-)$ 1 400; $\nu(NH_3^+)$ 2 120, 2 520, 2 780; [c] ν(C—O) 1 076; ν(OH) overlap ν(NH); [d] ν(C—O) 1 095; ν(OH) overlap ν(NH).

Table VI lists the characteristic vibrations of atoms in the IR spectra of all compounds *II*. The assignment of the individual vibration bands to individual atoms was carried out with the use of data recorded in literature[14,15]. The spectra of all the compounds confirmed their postulated structure.

The procedure for the derivatization of anilinothiazolinone derivatives of amino acids (ATZ) formed during sequence degradation of peptides and proteins by treatment of ATZ with methylamine was proposed many years ago[6,7]. Even though this method seems to promise a certain advantage compared to the routine conversion into phenylthiohydantoins (*I*) it has not received practical application. Neither has been reported a comparison of these two conversion methods. In our opinion one of the reasons of this situation is the complicated method of preparation of standard N^{α}-phenylthiocarbamoyl derivatives of amino acid methylamides *II*. We have therefore concentrated our efforts in this study on the development of a new procedure for their preparation and on their characterization by physicochemical and spectral methods. Methods of separation of derivatives *II* by HPLC and of the characterization of the products of sequence degradaion converted by treatment with methylamine were also developed.

The authors thank Mrs A. Čermáková and Mrs Z. Šenfederová for their technical assistance in the experimental part of this study. The authors also thank the members of the staff of the analytical laboratory and of the laboratory of mass spectrometry for the analysis of the derivatives described and for the measurement of their spectra.

REFERENCES

1. Edman P., Begg G.: Eur. J. Biochem. *1*, 80 (1967).
2. Allen G. in: *Sequencing of Proteins and Peptides* (T. S. Work and R. H. Burton, Eds). Elsevier, Amsterdam 1981.
3. Lottspeich F.: J. Chromatogr. *326*, 321 (1985).
4. Hewick R. M., Hunkapiller M. W., Hood L. E., Dreyer W. J.: J. Biol. Chem. *256*, 7990 (1981).
5. Ilse D., Edman P.: Aust. J. Chem. *16*, 411 (1963).
6. Inman J. K., Appella E.: Methods Enzmol. *58*, 374 (1977).
7. Inman J. K., Appella E. in: *Solid-Phase Methods in Protein Sequence Analysis* (R. A. Laursen, Ed.), p. 241. Pierce Chemical Company, Rockford, IL, 1975.
8. Martin M. L., Del Puech J.-J., Martin G. J.: *Practical NMR Spectroscopy*. Heyden and Son, London 1980.
9. Att-ur-Rahman: *Nuclear Magnetic Resonance*. Springer, New York 1986.
10. Wuthrich K.: *NMR in Biological Research: Peptides and Proteins*. North Holland, Amsterdam 1976.
11. Pretch E., Seibl J., Simon W.: *Strukturaufklärung organischer Verbindungen*. Springer, Berlin 1981.
12. Horsley W., Sternlicht H., Cohen J. S.: J. Am. Chem. Soc. *92*, 680 (1970).
13. Hagenmaier H., Ebbighausen W., Nicholson G.: Z. Naturforsch. *25b*, 681 (1970).
14. Gosavi R. K., Agarwala V., Rao C. N. R.: J. Am. Chem. Soc. *89*, 235 (1967).
15. Horák M., Papoušek D.: *Infračervená spektra a struktura molekul*. Academia, Praha 1976.

Translated by the author (V.K.).