Complete Amino Acid Sequence of the *Aspergillus* Cytotoxin Mitogillin†

Jose Luis Fernandez-Luna, Carlos López-Otin, Fernando Soriano, and Enrique Méndez*

Servicio de Endocrinologia, Centro Especial, Ramon y Cajal, Carretera de Colmenar Km 9100, Madrid 34, Spain
Received April 10, 1984

ABSTRACT: The complete amino acid sequence of the cytotoxin mitogillin has been determined by sequencing the intact chain and peptide fragments produced by cleavage at methionyl, arginyl, lysyl, and tryptophanyl residues and at one aspartic acid–proline bond. The protein consists of 149 amino acid residues with alanine at the NH_2 terminus and histidine at the COOH terminus. The calculated M_r of the native mitogillin was 16 867. The native molecule presents two disulfide bridges, one between cysteine residues at positions 5 and 147 and another one between cysteine residues at positions 75 and 131. The amino acid sequence of mitogillin shows 86% homology with another cytotoxic protein called α-sarcin.

The cytotoxic protein mitogillin produced by *Aspergillus restrictus* is one of a family of functionally related proteins that inhibit protein synthesis in eukaryotic and procaryotic cells. Such proteins include abrin, ricin, PAP, phenomycin, crotin II, curcin II, modeccin, enomycin, and another *Aspergillus* cytotoxin called α-sarcin. They have been reported to block protein synthesis in eukaryotic cells and microorganisms by catalytically inactivating the 60S ribosomal subunits, apparently without any additional cofactor requirement (Vazquez, 1979). The inactivation produced by mitogillin and α-sarcin on protein synthesis is due to the cleavage of a fragment from 28S rRNA in a large subunit of eukaryotic ribosomes (J. L. Fando et al., unpublished results; Endo & Wool, 1982). The cleavage site of the ribosomal RNA by mitogillin (J. L. Fando et al., unpublished results) is identical with the corresponding one for α-sarcin (Endo & Wool, 1982), and it is located on a purine-rich highly conserved sequence present in yeast 25S rRNA (Veldman et al., 1981), *Escherichia coli* 25S rRNA (Brosins et al., 1980), and *Zea mays* chloroplast 23S rRNA (Edwards & Kossel, 1981). The cytotoxin mitogillin produces inhibition of the protein synthesis in picornavirus infected cells, suggesting that the molecule passes into the cell once the membrane permeability has been modified by viral infection (Fernandez-Puentes & Carrasco, 1980). Mitogillin also has an antitumor activity against three different tumor systems, namely, sarcoma 180, adenocarcinoma 755, and leukemia 1210, induced in mice (Olson & Goerner, 1965; Olson et al., 1965; Roga et al., 1971).

Mitogillin has also been characterized as a basic polypeptide chain with a M_r of about 16 500 calculated by electrophoresis on polyacrylamide gel containing sodium dodecyl sulfate (Roga et al., 1971; Fernandez-Puentes & Vazquez, 1977; Schindler & Davies, 1977; Conde et al., 1978). Chemical similarlity (Rodriquez et al., 1982; Gavilanes et al., 1983) between mitogillin (Olson & Goerner, 1966), α-sarcin (Olson, 1963), and the cytotoxin named restrictocin produced by *A. restrictus* (Olson et al., 1965; Roga et al., 1971) has been recently described.

Furthermore, according to circular dichroism and difference spectroscopy studies, the secondary and tertiary structures of mitogillin are very similar to those of restrictocin and α-sarcin (Gavilanes et al., 1983). In addition, it has been shown that mitogillin antiserum cross-reacts totally with restrictocin and partially with α-sarcin (Fando et al., 1983).

In order to complete the structural characterization of mitogillin and to find the similarities and differences with restrictocin and α-sarcin, we report in the present work the primary structure of the cytotoxic protein mitogillin.

Materials and Methods

Protein mitogillin was kindly provided by Dr. D. M. Schuurmans (Department of Public Health, Lansing, MI) and Dr. D. Vazquez (Centro de Biologia Molecular, Madrid, Spain). Trypsin (treated with Tos-Phe-CH_2Cl),[1] carboxypeptidase A, fluorescent silica gel plates (DC-Alufolien), cellulose plates (20 × 20 cm), pyridine, acetic acid, and other reagents not specified were purchased from Merck. Pepsin, phenylthiohydantoin amino acids, iodoacetic acid, polybrene, and *o*-phthalaldehyde were from Sigma Chemical Co. BNPS-skatole was obtained from Pierce Eurochemie. Cyanogen bromide was from Serva. Fluorescamine was from Hoffman-La Roche. All reagents and solvents for the sequencer and amino acid analyzer were from Beckman.

Reduction and S-Carboxymethylation. Native protein (5 mg) mitogillin in 0.15 mL of 1 M Tris-HCl buffer, pH 8.5, containing 2 mM EDTA, and 6 M guanidine hydrochloride was incubated with 100 mM DTT for 90 min at 37 °C. Radioactivity labeling was achieved by adding 150 μCi of iodo[^{14}C]acetic acid (54 Ci/mol) and incubating the mixture for 15 min at room temperature in the absence of light. Unlabeled iodoacetic acid was then added to a final concentration of 0.2 M, and excess of reagents was removed by gel filtration on a Sephadex G-50 column (2.5 × 50 cm).

Performic Acid Oxidation. Native protein or native tryptic–peptic peptides (10–20 nm) were dissolved in 40 μL of 97% (v/v) formic acid to which was added 60 μL of chilled performic acid. Oxidation was carried out for 2 h in an ice–water bath, after which the sample was diluted with cold distilled

† This work was supported by a grant from the Fondo de Investigaciones Sanitarias de la Seguridad Social (11/83). J.L.F.-L. and C.L.-O. were fellows of the Fondo de Investigaciones Sanitarias.

[1] Abbreviations: Tos-Phe-CH_2Cl, L-1-(tosylamido)-2-phenylethyl chloromethyl ketone; Quadrol, *N,N,N',N'*-tetrakis(2-hydroxypropyl)-ethylenediamine; OPA, *o*-phthalaldehyde; Pth, phenylthiohydantoin; BNPS-skatole, 2-[(2-nitrophenyl)sulfenyl]-3-methyl-3'-bromoindolenine; CNBr, cyanogen bromide; HPLC, high-performance liquid chromatography; Tris-HCl, tris(hydroxymethyl)aminomethane hydrochloride; EDTA, ethylenediaminetetraacetic acid; DTT, dithiothreitol; TFA, trifluoroacetic acid.

0006-2960/85/0424-0861$01.50/0

water and lyophilized (Moore, 1963).

Cleavage at Methionines. Lyophilized, salt-free *S*-carboxymethylated protein was dissolved in 70% (v/v) formic acid (20 mg/mL). After addition of a 25-fold wt excess of cyanogen bromide over the methionine content, the mixture was incubated for 24 h at room temperature. The reaction mixture was then diluted 5-fold with cold distilled water and lyophilized.

Cleavage at Tryptophans. S-Carboxymethylated mitogillin (5 mg/mL) was dissolved in 80% (v/v) acetic acid. After addition of an 8-fold wt excess of solid BNPS-skatole, the solution was incubated for 30 h at room temperature in the dark. After addition of 10 mL of water, excess reagents were extracted 3 times with 8 mL of ethyl ether. The aqueous phase was then diluted 10 times with water and lyophilized.

Cleavage at Lysines and Arginines. S-Carboxymethylated mitogillin was dissolved in 0.2 M *N*-methylmorpholineacetate buffer, pH 8.2 (10 mg/mL), and incubated with trypsin treated with Tos-Phe-CH_2Cl (enzyme to substrate wt ratio 1/100) for 3 h at 37 °C. The digested material was then lyophilized.

Carboxypeptidase A Digestion. S-Carboxymethylated mitogillin was dissolved in 0.2 M *N*-methylmorpholineacetate buffer, pH 8.2 (3 mg/mL). The material was digested with carboxypeptidase A (enzyme to substrate wt ratio 1/30) at 37 °C. Aliquots of 25 μL were taken at different incubation times, and the reaction was stopped by addition of the same volume of 30% (v/v) trichloroacetic acid. The undigested protein was sedimented by centrifugation, and the supernatant was lyophilized. This lyophilized material was then resuspended in 0.2 M sodium citrate buffer, pH 2.2, and applied to the amino acid analyzer.

High-Performance Liquid Chromatography. Tryptic peptides (T) from reduced and S-carboxymethylated mitogillin were fractionated by a high-performance liquid chromatography system consisting of a Waters 6000 pump and one 510 pump, a Waters 680 automated gradient controller, and a Water 480 λ_{max} variable-wavelength absorbance detector. Samples (20 nmol in 40 μL of 0.1% TFA) were applied on a μBondapak C_{18} column (3.9 mm × 30 cm) fitted with a guard column (270-μL volume) packed with μBondapak C_{18}/Corasil. The effluent was monitored by absorbance at 214 nm. Peptides were eluted by increasing concentrations of acetonitrile containing 0.1% trifluoroacetic acid. Runs were carried out at room temperature at a constant flow rate of 0.5 mL/min.

Purification of Fragments from Chemical Cleavage. Cyanogen bromide (CN) and BNPS-skatole (BN) fragments from reduced and S-carboxymethylated mitogillin were fractionated by gel filtration chromatography on a Sephadex G-50 column (1.5 × 200 cm) equilibrated and eluted with 10% (v/v) formic acid.

Amino Acid Analysis. Hydrolyses were carried out at 110 °C for 20 h with 200 μL of 5.7 M HCl containing 0.05% (v/v) 2-mercaptoethanol in evacuated and sealed tubes. The analyses were performed on a Beckman 121-M analyzer equipped with a Beckman integrator 126 data system.

Sequencing Procedure. The whole S-carboxymethylated protein and selected fragments were sequenced in a Beckman sequencer (Model 890 D) according to the method of Edman & Begg (1967). The amino terminal sequence of the S-carboxymethylated mitogillin and the large fragments were determined by using the standard Beckman 1.0 M Quadrol program. Small peptides were sequenced with the standard Beckman 0.1 M Quadrol program. To reduce peptide washout, 3 mg of polybrene was added when the sample was applied to the cup (Tarr et al., 1978). The amino acid thiazolinones were converted to phenylthiohydantoin (Pth) amino acids (Guyer & Todd, 1975). Pth-amino acids were identified by thin-layer chromatography and by amino acid analysis after back-hydrolysis. Regeneration of the corresponding amino acids was performed by hydrolysis in constant-boiling HCl, containing 0.1% stannous chloride at 150 °C for 4 h under vacuum (Méndez & Lai, 1975a). All amide assignments were made by thin-layer chromatography of the Pth-amino acids. The thin-layer chromatography on silica gel plates was carried out in a first system of *n*-heptane/propionic acid/1,2-dichloroethane (140:40:100 v/v) and a second system of chloroform/methanol (261:39 v/v). Aliquots of Pth-amino acids were also taken for ^{14}C measurements for identification of (carboxymethyl)cysteine. Manual procedure (Edman & Henschen, 1975) for stepwise degradation was also used for some peptides. N-Terminal residues of tryptic peptides were determined by the dansyl chloride technique (Gray, 1967).

Location of Disulfide Bridges. Native mitogillin was dissolved in 0.2 M *N*-methylmorpholine, pH 8.2 (10 mg/mL), and incubated with trypsin treated with Tos-Phe-CH_2Cl (enzyme to substrate wt ratio 1/100) for 3 h at 37 °C. After incubation, the material was lyophilized and then resuspended in 1 mL of 5% (v/v) formic acid. This solution was incubated with pepsin (enzyme to substrate wt ratio 1/100) for 18 h at 37 °C. The reaction was stopped by lyophilization. The dried material was dissolved in 0.4 mL of 0.01 M pyridine/acetate buffer, pH 2.1, and applied to a Dowex M-71 column (0.3 × 20 cm) equilibrated with 0.01 M pyridine/acetate buffer, pH 2.1. The peptides were eluted with a gradient of increasing concentrations of pyridine (from 0.01 to 2 M) at 50 °C and at a constant flow rate of 6 mL/h. Fractions of 0.6 mL were collected and aliquots (5%) from every second fraction were analyzed with OPA after alkaline hydrolysis (Méndez & Gavilanes, 1976). Equivalent amounts were oxidized with performic acid and applied to the amino acid analyzer to detect the cystine-containing fractions, which were pooled and dried. The cystine-containing peptides were purified on cellulose thin-layer plates in acetic acid/pyridine/1-butanol/water (6/20/30/24 v/v). The peptides were detected with fluorescamine (0.001–0.005% in acetone) (Méndez & Lai, 1975b). The purified cystine-containing peptides were then oxidized with performic acid and separated on thin-layer chromatography as above. The released cysteic acid containing peptides were detected with fluorescamine and after elution with 10% (v/v) formic acid were identified by amino acid analysis.

RESULTS

Amino-Terminal Sequence. Automated degradation was performed with 130 nmol of the S-carboxymethylated mitogillin. The repetitive yield calculated from alanine residues at position 1 and 27 was 93%. The assignment of the first 48 residues is shown in Table I.

Carboxyl-Terminal Sequence. Upon carboxypeptidase treatment of S-carboxymethylated mitogillin, the following amino acids were liberated (values are moles of residue per mole of S-carboxymethylated mitogillin): after 5 min, His (1.0); after 10 min, His (1.1) and Ser (0.4). From these results, the carboxyl-terminal sequence -Ser-His was deduced. This order was confirmed by direct sequence analysis of the tryptic peptide T-20 (see below).

Products of Tryptic Cleavage. The tryptic digest (5 mg) was fractionated by HPLC on a μBondapak C_{18} column (Figure 1). Numbers on the peaks designate the position of

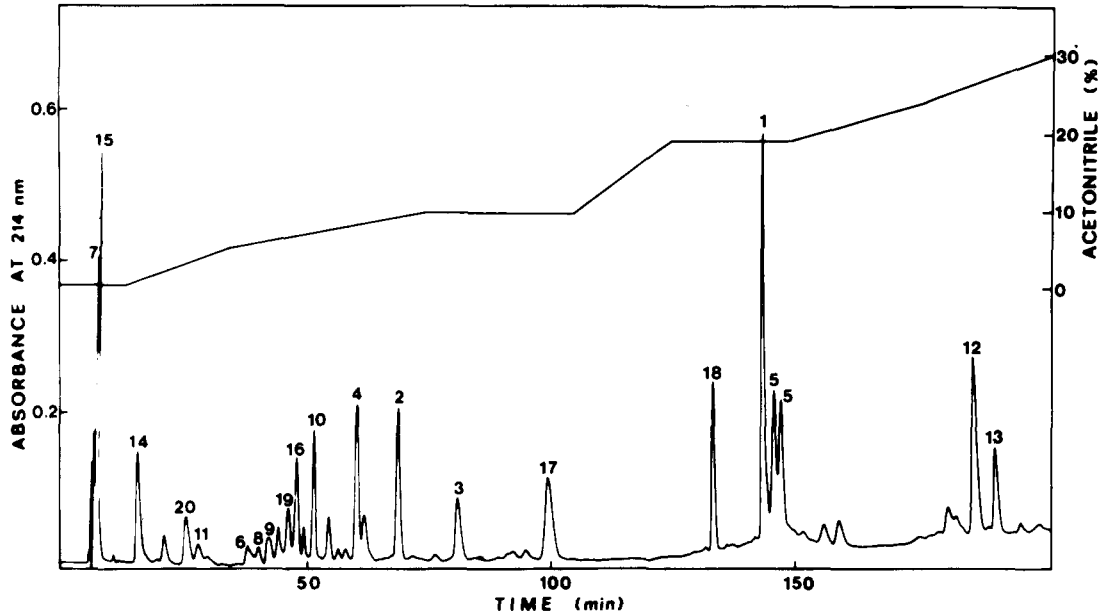

FIGURE 1: Reverse-phase HPLC of tryptic peptides derived from reduced and S-carboxymethylated mitogillin. The sample (40 nmol) was dissolved in 0.1% (v/v) trifluoroacetic acid and applied to a μBondapak C_{18} column (0.39 × 30 cm), equilibrated with 0.1% (v/v) trifluoroacetic acid. The elution of peptides was carried out by a gradient of acetonitrile in 0.1% (v/v) trifluoroacetic acid as indicated in figure.

each peptide in the protein sequence. The amino acid composition of the tryptic peptides corresponding to the numbered peaks is given in the supplementary material (see paragraph at end of paper regarding supplementary material). The other peaks shown in Figure 1 correspond mainly to peptides with very low yield. A total of 20 pure peptides was obtained, and only 12 of them were completely sequenced. Some of these peptides, designated as T-1, T-5, T-10, T-11, T-18, and T-20, were sequenced by automatic degradation (Table I), while T-6, T-7, T-8, T-9, T-14, and T-15 were sequenced by manual degradation. The remaining tryptic peptides were only subjected to N-terminal residue identification. Peaks designated as T-5 (Figure 1) showed the same amino acid composition, probably due to a different amidation degree of acidic residues. Peptide T-1 was identified as the N-terminal tryptic peptide of the molecule by comparison of its sequence with that obtained for the intact protein (Table I). Peptide T-20 corresponded to the C-terminal tryptic peptide of mitogillin, established by comparison of their amino acid sequence with that determined after carboxypeptidase A digestion of the protein. T-1, T-10, T-18, and T-20 (Table I) represent the four different (carboxymethyl)cysteine-containing tryptic peptides; T-11 was the only methionine-containing tryptic peptide in the molecule.

Chemical Cleavage. The amino acid analysis of mitogillin (Table II) showed the presence of a unique methionine residue; therefore, the CNBr cleavage should provide two fragments. Fragments generated by cleavage of S-carboxymethylated protein with CNBr as described under Materials and Methods were fractionated by gel filtration on Sephadex G-50 (Figure 2A). The first peak contained uncleavaged material as deduced by polyacrylamide gel electrophoresis, and it was discarded. The second peak was identified as the N-terminal CNBr fragment, but it was not a necessary fragment for the sequencing strategy. The two following peaks designated as 2 and 3 were pooled and lyophilized. The material of these peaks was rechromatographed on the same column resulting in the isolation of two pure CNBr fragments (CN-2 and CN-3). The amino acid composition of these fragments is given in the supplementary material, both being devoid of homoserine lactone. When these two purified CN-2 and CN-3

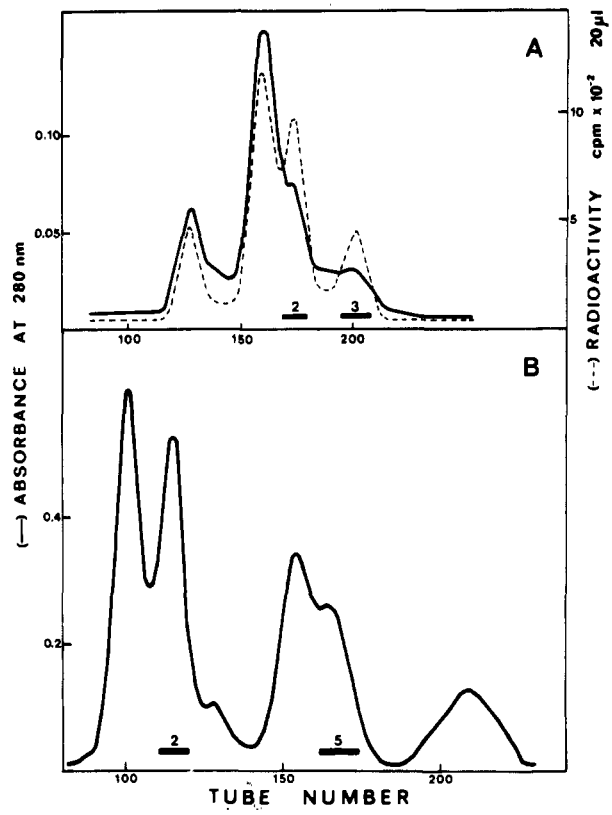

FIGURE 2: Isolation of CNBr (A) and BNPS-skatole (B) fragments obtained from reduced and S-carboxymethylated mitogillin. Separation was on a Sephadex G-50 column (1.5 × 200 cm) in 10% (v/v) formic acid. Flow rate was 10 mL/h, and 1.5 mL/fraction was collected. The effluent fraction was monitored by absorbance at 280 nm (—) and by ^{14}C radioactivity (– –).

fragments were subjected to automatic Edman degradation (Table I), a total of 33 residues could be identified in CN-2 and 32 residues in CN-3. This sequential data showed that CN-3 was included in CN-2 fragment and was obtained by cleavage at an Asp–Pro bond probably due to the action of the formic acid during incubation. The S-carboxymethylated protein was also cleavaged with BNPS-skatole as described

Table I: N-Terminal Sequence of Mitogillin (M) and Its Fragments[a]

	1								10						
M	Ala -	Thr -	Trp -	Thr -	Cys -	Ile -	Asn -	Gln -	Gln -	Leu -	Asn -	Pro -	Lys -	Thr -	Asn -
	50.0	42.2	38.0	32.6	b	21.1	23.5	32.7	23.6	32.5	32.4	27.0	28.0	21.3	12.4
					20										30
	Lys -	Trp -	Glu -	Asp -	Lys -	Arg -	Leu -	Leu -	Tyr -	Asn -	Gln -	Ala -	Lys -	Ala -	Glu -
	7.1	9.5	10.0	8.9	4.3	3.9	3.2	3.6	3.6	4.3	4.0	4.1	4.6	2.6	2.2
										40					
	Ser -	Asn -	Ser -	His -	His -	Ala -	Pro -	Leu -	Ser -	Asp -	Gly -	Lys -	Thr -	Gly -	Ser -
	2.5	1.9	1.2	0.9	0.7	1.0	1.7	1.2	1.4	0.8	0.9	0.7	0.5	0.7	
			48												
	Ser -	Tyr -	Pro ...												
	1.2	0.9	1.0												
	87			90										100	
CN-2	Gly -	Lys -	Asp -	Asp -	His -	Tyr -	Leu -	Leu -	Glu -	Phe -	Pro -	Thr -	Phe -	Pro -	Asp -
	17.4	18.8	14.5	15.5	13.5	12.0	15.5	11.3	11.7	12.5	12.4	6.1	9.9	10.5	6.2
									110						
	Gly -	His -	Asp -	Tyr -	Lys -	Phe -	Asp -	Ser -	Lys -	Lys -	Pro -	Lys -	Glu -	Asn -	Pro -
	7.0	4.1	4.8	3.3	2.8	3.5	3.9	1.4	1.0	1.2	1.5	1.7	2.0	2.7	1.2
			119												
	Gly -	Pro -	Ala ...												
	1.0	0.9	1.1												
	116				120										130
CN-3	Pro -	Gly -	Pro -	Ala -	Arg -	Val -	Ile -	Tyr -	Thr -	Tyr -	Pro -	Asn -	Lys -	Val -	Phe -
	7.8	27.3	9.8	29.9	10.8	9.9	5.6	6.8	7.5	9.3	7.7	6.0	5.8	5.4	1.9
										140					
	Cys -	Gly -	Ile -	Val -	Ala -	His -	Gln -	Arg -	Gly -	Asn -	Gln -	Gly -	Asp -	Leu -	Arg -
	b	2.6	2.4	4.6	1.2	0.8	3.5	0.9	1.6	3.7	4.4	4.4	4.1	1.8	0.5
				149											
	Leu -	Cys -	Ser -	His -	COOH										
	1.7	b													
	51									60					
BN-2	Phe -	Thr -	Asn -	Gly -	Tyr -	Asp -	Gly -	Asn -	Gly -	Lys -	Leu -	Ile -	Lys -	Gly -	Arg -
	12.0	3.0	12.4	13.6	10.7	11.9	13.1	6.5	10.0	4.6	5.9	4.2	5.6	7.1	2.5
					70										80
	Thr -	Pro -	Ile -	Lys -	Phe -	Gly -	Lys -	Ala -	Asp -	Cys -	Asp -	Arg -	Pro -	Pro -	Lys -
	1.1	3.0	1.5	3.3	4.3	3.5	2.6	3.2	2.1	b	2.3	0.6	0.8	1.6	1.1
										90					
	His -	Ser -	Gln -	Asn -	Gly -	Met -	Gly -	Lys -	Asp -	Asp ...					
	0.8	0.5	0.6	0.5	1.9	0.5	2.1	0.8							
			20										30		
BN-5	Glu -	Asp -	Lys -	Arg -	Leu -	Leu -	Tyr -	Asn -	Gln -	Ala -	Lys -	Ala -	Glu -	Ser -	Asn -
	8.5	5.6	5.9	6.7	5.6	6.1	4.2	4.3	1.8	1.6	1.4	1.4	0.8	0.5	1.5
		34													
	Ser -	His ...													
	0.6	1.2													
	43							50							
T-5	Thr -	Gly -	Ser -	Ser -	Tyr -	Pro -	His -	Trp -	Phe -	Thr -	Asn -	Gly -	Tyr -	Asp -	Gly -
	0.8	1.9	0.6	0.5	1.0	0.8	1.0	0.4	0.9	0.6	0.4	0.7	0.4	0.5	0.7
			60												
	Asn -	Gly -	Lys -	COOH											
	0.6	0.8	0.5												
	81							88							
T-11	His -	Ser -	Gln -	Asn -	Gly -	Met -	Gly -	Lys -	COOH						
	1.3	0.8	2.4	2.3	1.6	1.6	1.8	1.0							
	73							80							
T-10	Ala -	Asp -	Cys -	Asp -	Arg -	Pro -	Pro -	Lys -	COOH						
	11.5	8.7	b	7.6	1.3	1.9	1.1	1.0							
		130								138					
T-18	Val -	Phe -	Cys -	Gly -	Ile -	Val -	Ala -	His -	Gln -	Arg -	COOH				
	7.4	6.3	b	8.8	2.5	3.7	3.1	4.0	2.3	1.8					
	1									10					
T-1	Ala -	Thr -	Trp -	Thr -	Cys -	Ile -	Asn -	Gln -	Gln -	Leu -	Asn -	Pro -	Lys -	COOH	
	30.0	9.2	7.1	16.0	b	22.3	19.8	15.7	16.1	12.0	6.2	7.3	4.4		
	146			149											
T-20	Leu -	Cys -	Ser -	His -	COOH										
	22.4	b	7.8	5.3											

[a] Numbers under residues are yields in nanomoles of the amino acids recovered from the Pth derivatives by back-hydrolysis. Fragments that were not sequenced completely end with (...). Numbers above residues indicate their position in the mitogillin sequence. [b] Cysteine residues were identified as *S*-(carboxymethyl)cysteine by radioactivity measurements.

under Materials and Methods, and the fragments obtained were fractionated by gel filtration on Sephadex G-50 (Figure 2B). The first peak contained uncleavaged material as deduced by polyacrylamide gel electrophoresis; the last one did not contain any detectable peptide material as determined by amino acid analysis. Fractions of the peaks 2 and 5 were pooled and lyophilized. The material of each peak was rechromatographed on the same column, resulting in the isolation of two pure fragments, BN-2 and BN-5. The amino acid composition of these fragments is given in the supplementary

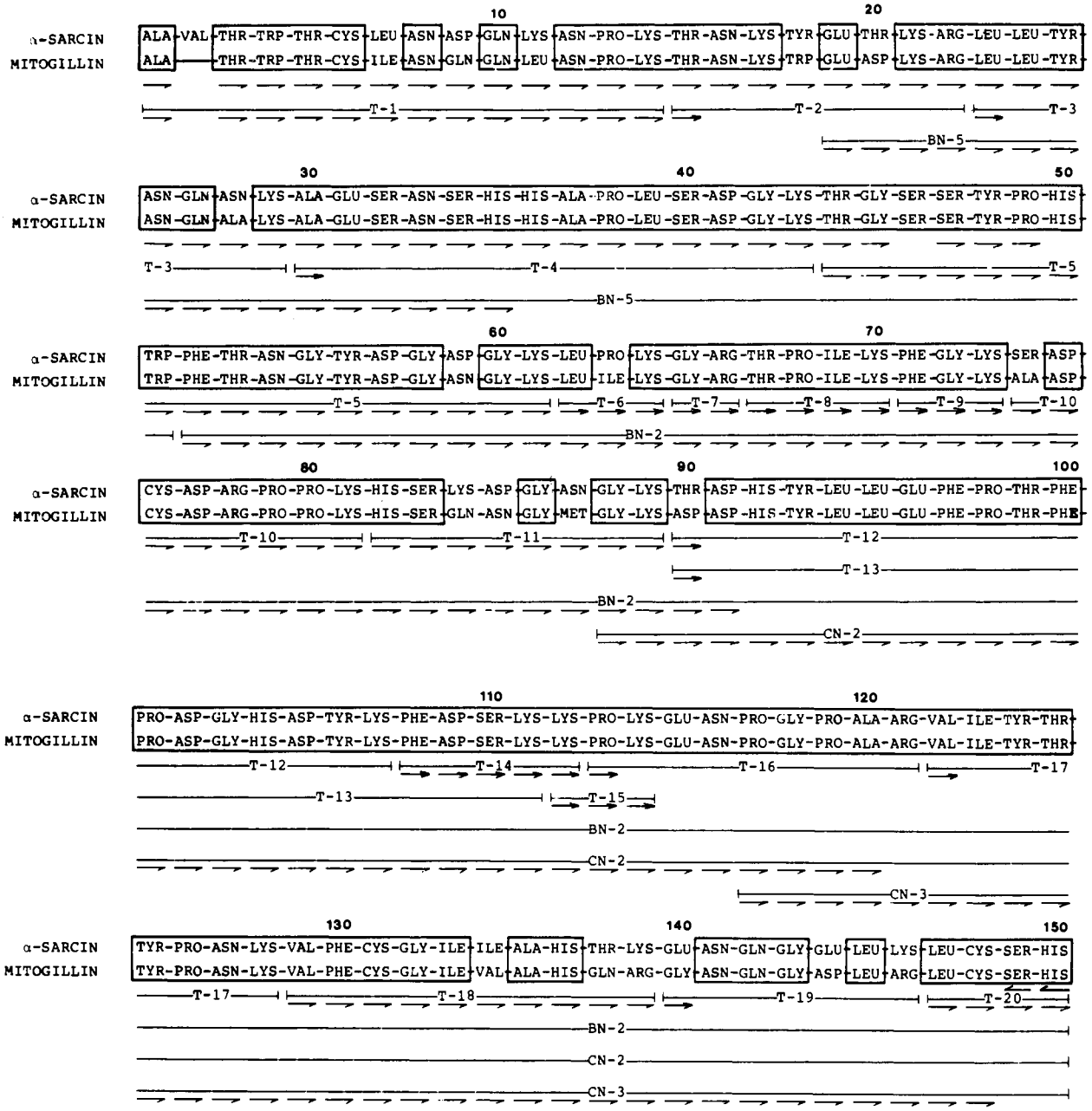

FIGURE 3: Amino acid sequence of mitogillin aligned with that of α-sarcin. Symbols used are trypsin (T), cyanogen bromide (CN), BNPS-skatole (BN), (forward half-arrow) sequencer, (→) sequence determined by manual degradation, and (reverse half-arrow) sequence determined by carboxypeptidase A. Homologies between both molecules are boxed.

Table II: Amino Acid Composition of Native Mitogillin[a]

amino acid		value from sequencing	amino acid		value from sequencing
Asp[b]	21.2	22	Ile	4.7	5
Thr[c]	7.5	8	Leu	9.5	9
Ser[c]	7.7	8	Tyr	7.2	7
Glu[b]	10.6	10	Phe	6.0	6
Pro	10.9	12	His	8.5	8
Gly	14.3	14	Lys	16.4	16
Ala	7.2	7	Arg	5.3	6
Cys[d]	3.4	4	Trp	nd[e]	3
Val	3.5	3			
Met	0.8	1	total		149

[a] Values are given as residues per mole. [b] Aspartic and glutamic acid values include those of Asn and Gln, respectively. [c] Values obtained by extrapolation to zero-time hydrolysis. [d] Determined as *S*-(carboxymethyl)cysteine. [e] nd, not quantitatively determined.

material. Automatic Edman degradation of the fragment BN-2 allowed the identification of its first 37 NH_2-terminal residues (Table I). Fragment BN-2 had a methionine residue in its amino acid composition, and it could be a useful fragment to establish the sequence around the methionine residue. The first 17 NH_2-terminal residues from the fragment BN-5 were also established by automatic degradation (Table I), and they were located near the N-terminal region of the molecule.

Sequence Analysis. Automatic or manual Edman degradations were performed on the S-carboxymethylated protein and peptides derived from mitogillin after chemical (CNBr, BNPS-skatole) or enzymatic (trypsin) cleavage. The complete primary structure of mitogillin is presented in Figure 3. The sequence of the first 48 amino acids of the molecule, with the exception of a serine residue at position 45, was established by automatic degradation of the S-carboxymethylated mitogillin. Portions of this sequence were confirmed by the tryptic peptides T-1 and T-5 and the fragment BN-5. Peptide T-5 overlapped with fragment BN-2 and allowed the identification of the serine residue at position 45. Edman degradation of the BN-2 fragment extended the amino acid sequence up to residue 90. Confirmation of this sequence was obtained with

the degradation of peptides from T-5 to T-11. Fragment BN-2 and peptide T-11 allowed an overlap to be established with fragment CN-2. Automatic degradation of CN-2 extended the amino acid sequence of the protein up to residue 119. The complete sequence of peptides T-14 and T-15 overlapped each other and were included in the CN-2 fragment, which also overlapped with the other CNBr fragment CN-3. The sequence of fragment CN-3 extended the sequence up to residue 147. The last two residues of the protein, 148 and 149, were identified by the sequence of the tetrapeptide T-20, which overlapped with the carboxyl-terminal end of the CN-3 fragment. In addition, the last two residues were also identified by carboxypeptidase A digestion. In accordance with the sequence data (Figure 3), the location of the four cysteine residues in the polypeptide chain of mitogillin has been determined at positions 5, 75, 131, and 147, respectively.

Characterization of Disulfide Bridges. Native mitogillin (10 mg) was digested with trypsin and pepsin as described under Materials and Methods and then fractionated on a Dowex M-71 column (see supplementary material). Only peaks 36 and 45 present a high content of cystine as deduced by amino acid analysis. The cystine-containing peptides were purified by thin-layer chromatography. Spots designated as 36B and 45A correspond to the two purified cystine-containing peptides. After performic acid oxidation of 36B and 45A, four different cysteic acid containing peptides, 36B-1, 36B-2, 45A-1, and 45A-2, were released. Peptides 36B-1 ($Ser_{0.8}$Cys-$A_{1.0}Leu_{0.4}His_{0.8}$) and 36B-2 ($Asp_{1.7}Thr_{1.8}Glu_{1.7}Pro_{0.5}Ala_{0.6}$-Cys-$A_{1.0}Ile_{0.5}Leu_{0.6}Lys_{1.1}$) contained cysteine residues 147 and 5, respectively, and peptides 45A-1 ($Asp_{1.9}Pro_{1.4}Ala_{0.6}$Cys-$A_{1.1}Lys_{1.0}Arg_{0.9}$) and 45A-2 ($Gly_{0.9}Ala_{0.6}$Cys-$A_{1.0}Val_{1.4}Ile_{0.6}Phe_{0.8}His_{1.0}$) contained cysteines 75 and 131, respectively.

Discussion

The complete amino acid sequence of the cytotoxic protein mitogillin is shown in Figure 3. The protein consists of 149 residues, and the M_r calculated from the amino acid sequence is 16 867, close to the value (16 500) obtained by polyacrylamide gel electrophoresis in the presence of sodium dodecyl sulfate. The amino acid composition of the protein calculated by analysis after acid hydrolysis (Table II) is in agreement with the number of residues of each amino acid obtained from the sequence (Figure 3). The primary structure of the whole protein has been established by alignment of fragments obtained by chemical and enzymatic cleavages from the polypeptide chain. A total of 12 tryptic peptides, 2 CNBr fragments, and 2 BNPS-skatole fragments was required to deduce the complete sequence (Figure 3). Mitogillin has a great number of aspartic acid and asparagine (22), lysine (16), glycine (14), and proline (12) residues and has few valine (3), tryptophan (3), methionine (1), and cystene (4) residues. Since native protein does not contain any free sulfhydryl groups as could be deduced from titration with 5,5′-dithiobis(2-nitrobenzoic acid) (Gavilanes et al., 1983), two intrachain disulfide bridges should be expected. In fact, two disulfide bridges have been identified, one between two cysteines located near the NH_2-terminal and COOH-terminal ends at positions 5 and 147, respectively, and another between cysteine-75 located in the middle part of the molecule and cysteine-131, 14 residues from the COOH end.

The amino acid sequence of mitogillin is nearly identical with that of the cytotoxic protein called restrictocin (C. López-Otin et al., unpublished results) also isolated from the same *Aspergillus* strain. The primary structure presented for mitogillin in this work shows a high degree of homology (86%) with another recently reported cytotoxic protein, α-sarcin (Sacco et al., 1983). The sequence alignment between mitogillin (149 residues) and α-sarcin (150 residues), presented in Figure 3, shows only a difference of 20 amino acids, distributed within three regions of the molecule. The first region is included between residue 1 and residue 27, the second is within the last 16 residues, and the third is located in the middle part of the polypeptide chain around the single methionine residue in mitogillin. These small differences in the amino acid sequence between mitogillin and α-sarcin may explain the existance of common antigenic determinants and the partial crossed-immunoreaction of the α-sarcin antiserum with mitogillin (Conde et al., 1978; Fando et al., 1983), as well as the slightly different spectroscopic properties recently described between mitogillin and α-sarcin (Gavilanes et al., 1983).

In addition, the common mechanism of action observed for mitogillin and α-sarcin based in the inhibition of protein synthesis (Fernandez-Puentes & Vazquez, 1977) by cleaving a fragment from rRNA in the large subunit of ribosomes (Shindler & Davies, 1977) may also be related to the presence of an identical active site containing the homologous sequence. In fact the sequence of the 5′-end of the fragment derived from rRNA by cleavage of α-sarcin (Endo & Wool, 1982) is identical with that obtained by cleavage with mitogillin (J. L. Fando et al., unpublished results).

Recently, it has been described that the enzyme RNase U_2 (113 residues), isolated from the mold *Ustilago sphaerogena*, presents 34% homology with α-sarcin (Wool, 1984). In addition, the two disulfide bridges described in this paper for mitogillin are identical with the one reported for α-sarcin (Sacco et al., 1983) and can be aligned with two of the three disulfide bridges in RNase U_2 (Sacco et al., 1983). On the basis of these homology comparisons, it could be possible that the enzyme RNase U_2 is also related to mitogillin. However, the relationship between the structural and the biological functions of RNase U_2, mitogillin, and α-sarcin must be subjected to further investigations. The knowledge of the primary structure of mitogillin could serve in understanding its mechanism of action and in determining the domains that play a critical role in producing a functionally active conformation in the protein.

Acknowledgments

We thank Dr. D. Vazquez, Dr. J. Gavilanes, Dr. C. Y. Lai, and Dr. J. Sulner for reviewing the manuscript and Shirley McGrath for her secretarial work.

Supplementary Material Available

One figure illustrating the ion-exchange chromatography on Dowex M-71 and thin-layer chromatography to characterize the disulfide bridges and two tables showing the amino acid composition of the tryptic peptides and CNBr and BNPS-skatole fragments (5 pages). Ordering information is given on any current masthead page.

Registry No. Mitogillin (*Aspergillus restrictus*), 94483-17-7; mitogillin, 1403-99-2.

References

Brosins, J., Dull, T. J., & Noller, H. F. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 201–204.

Conde, F. P., Fernandez-Puentes, C., Montero, M. T. V., & Vazquez, D. (1978) *FEMS Microbiol. Lett. 4*, 349–355.

Edman, P., & Begg, G. (1967) *Eur. J. Biochem. 1*, 80–91.

Edman, P., & Henschen, A. (1975) *Protein Sequence Determination*, pp 232–279, Springer-Verlag, Berlin.

Edwards, K., & Kossel, H. (1981) *Nucleic Acids Res. 9*, 2853–2869.
Endo, Y., & Wool, I. G. (1982) *J. Biol. Chem. 257*, 9054–9060.
Fando, J. L., Fernandez-Luna, J. L., Alaba, I., Barber, D., Salinas, M., & Méndez, E. (1983) Antibiotics Symposium, Jarandilla de la Vera, Spain, May 1983, Abstract, p 5.
Fernandez-Puentes, C., & Vazquez, D. (1977) *FEBS Lett. 78*, 143–146.
Fernandez-Puentes, C., & Carrasco, L. (1980) *Cell (Cambridge, Mass.) 20*, 769–775.
Gavilanes, J. G., Vazquez, D., Soriano, F., & Méndez, E. (1983) *J. Protein Chem. 2*, 251–261.
Gray, W. R. (1967) *Methods Enzymol. 2*, 139–151.
Guyer, R. L., & Todd, C. W. (1975) *Anal. Biochem. 66*, 400–404.
Méndez, E., & Lai, C. Y. (1975a) *Anal. Biochem. 65*, 281–292.
Méndez, E., & Lai, C. Y. (1975b) *Anal. Biochem. 68*, 47–53.
Méndez, E., & Gavilanes, J. G. (1976) *Anal. Biochem. 72*, 472–479.
Moore, S. (1963) *J. Biol. Chem. 238*, 235–237.
Olson, B. H. (1963) U.S. Patent 3 104 204.
Olson, B. H., & Goerner, G. L. (1965) *Appl. Microbiol. 13*, 314–321.
Olson, B. H., & Goerner, G. L. (1966) U.S. Patent 3 230 153.
Olson, B. H., Jennings, J. C., Roga, V., Junek, A. J., & Schuurmans, D. M. (1965) *Appl. Microbiol. 13*, 322–326.
Rodriguez, R., Lopéz-Otin, C., Barber, D., Fernandez-Luna, J. L., Gonzalez, G., & Méndez, E. (1982) *Biochem. Biophys. Res. Commun. 108*, 315–321.
Roga, V., Hedeman, L. P., & Olson, B. H. (1971) *Cancer Chemother. Rep., Part 1 55*, 101–113.
Sacco, G., DrickKamer, K., & Wool, I. G. (1983) *J. Biol. Chem. 258*, 5811–5818.
Schindler, D. G., & Davies, J. E. (1977) *Nucleic Acids Res. 4*, 1097–1110.
Tarr, G. E., Beecher, J. F., Bell, M., & McKean, D. J. (1978) *Anal. Biochem. 84*, 622–627.
Vazquez, D. (1979) *Inhibitors of Protein Biosynthesis*, p 80.
Veldman, G. M., Klootwijk, J., de Regt, V. C. H. F., Planta, R. J., Branlant, C., Krol, A., & Ebel, J. P. (1981) *Nucleic Acids Res. 9*, 6935–6952.
Wool, I. G. (1984) *Trends Biochem. Sci. (Pers. Ed.) 9*, 14–17.

Unusual Chemical Properties of N-Terminal Histidine Residues of Glucagon and Vasoactive Intestinal Peptide†

Mary Alice Hefford, Robyn M. Evans, George Oda, and Harvey Kaplan*

Department of Biochemistry, University of Ottawa, Ottawa, Ontario, Canada K1N 9B4
Received May 29, 1984

ABSTRACT: An N-terminal histidine residue of a protein or peptide has two functional groups, viz., an α-amino group and an imidazole group. A new procedure, based on the competitive labeling approach described by Duggleby and Kaplan [Duggleby, R. G., & Kaplan, H. (1975) *Biochemistry 14*, 5168–5175], has been developed by which the chemical reactivity of each functional group in such a residue can be determined as a function of pH. Only very small amounts of material are required, which makes it possible to determine the chemical properties in dilute solution or in proteins and polypeptides that can be obtained in only minute quantities. With this approach, the reactivity of the α-amino group of histidylglycine toward 1-fluoro-2,4-dinitrobenzene gave an apparent pK_a value of 7.64 ± 0.07 at 37 °C, in good agreement with a value of 7.69 ± 0.02 obtained by acid–base titration. However, the reactivity of the imidazole function gave an apparent pK_a value of 7.16 ± 0.07 as compared to the pK_a value of 5.85 ± 0.01 obtained by acid–base titration. Similarly, in glucagon and vasoactive intestinal peptide (VIP), apparent pK_a values of 7.60 ± 0.04 and 7.88 ± 0.18, respectively, were obtained for the α-amino of their N-terminal histidine, and pK_a values of 7.43 ± 0.09 and 7.59 ± 0.18 were obtained for the imidazole function. Rothgeb et al. [Rothgeb, T. M., England, R. D., Jones, B. N., & Gurd, R. S. (1978) *Biochemistry 17*, 4564–4571] used proton titration and ^{13}C NMR to assign pK_a values of 7.32 to the α-amino group and 5.32 to the imidazole function of *S*-methylglucagon at 25 °C. It is concluded that the large difference in the apparent pK_a values of the imidazole moiety obtained by proton titration and chemical reactivity is an inherent property of an N-terminal histidine which results from a large decrease in the inductive effect on the imidazole moiety by the α-amino group when the latter loses its proton. While the apparent ionization behavior of the N-terminal histidine in glucagon and VIP does not differ from that in model compounds, the reactivity of the imidazole function in both is substantially greater than that in histidylglycine, and the reactivity of the α-amino group in glucagon is a factor of 2 lower. In addition, the pK_a and reactivity of the solitary lysine in glucagon are very different from those of a normal ε-amino group in solution. These results indicate that both glucagon and VIP have sufficient organized structure in dilute solution (10^{-6} M) to alter the chemical properties of these groups.

Most proteins have only one ionizable group at the amino-terminal residue. There are several biologically active peptides, viz., glucagon, vasoactive intestinal peptide (VIP),[1] PHI, and secretin, which have an N-terminal histidine residue and therefore have two ionizable groups at their amino-ter-

† This work was supported by the Medical Research Council of Canada.

[1] Abbreviations: VIP, vasoactive intestinal peptide; PHI, peptide with N-terminal histidine and C-terminal isoleucine amide; N_2phF, 1-fluoro-2,4-dinitrobenzene; N_2ph, 2,4-dinitrophenyl; Ala-ala, alanylalanine; imlac, L-β-imidazolyllactic acid; TLC, thin-layer chromatography.

0006-2960/85/0424-0867$01.50/0