N-TERMINALLY EXTENDED SOMATOSTATIN: THE PRIMARY STRUCTURE OF SOMATOSTATIN-28

Lucien PRADAYROL, Hans JÖRNVALL, Viktor MUTT and André RIBET
Departments of Chemistry I and Biochemistry II, Karolinska Institutet, S-104 01 Stockholm 60, Sweden and *INSERM U 151 Rangueil, 31054 Toulouse Cedex, France*

Received 18 October 1979

1. Introduction

Somatostatin, first isolated from ovine hypothalamic extracts [1], has been characterized as a tetradecapeptide [2], and the same peptide has been obtained from porcine hypothalamus [3]. Somatostatin-like immunoreactivity and bioactivity are widely distributed in the central nervous system and in the digestive tract tissues [4–7]. Size heterogeneity described in somatostatin-like immunoreactive materials [3,6,8,9] raised the possibility of the existence of prohormonal forms of this molecule [10,11].

During the purification of porcine intestinal somatostatin differences were noticed in the chromatographic behaviour of immunoreactive fractions and synthetic somatostatin-14 [12]. A novel intestinal peptide was then isolated and partly characterized as an N-terminally extended form of somatostatin [13].

We report here the primary structure of this intestinal peptide which has been found to be an octacosapeptide, somatostatin-28. The results establish size heterogeneity in the somatostatin family of peptides. Sequence comparison between somatostatin-28 and prohormones also suggest that it may be a precursor of somatostatin-14.

2. Materials and methods

The extended form of somatostatin was obtained as in [13] except for the final purification which was achieved by reverse-phase high-performance liquid chromatography (Waters Ass. apparatus, 6000 A solvent delivery system, U6K injector, model 450 variable wavelength detector) using a 300 × 3.9 mm μBondapak C_{18} column. A 76% triethylammonium phosphate 0.25 M (pH 3.5)/24% acetonitrile buffer was used in addition to the ammonium acetate/ethanol system in [13].

Batches with about the same degree of purity according to isotachophoresis [13] and thin-layer chromatography [14] have been used for the structural work. Total compositions were determined with a Beckman 121 M amino acid analyzer after hydrolysis at 110°C for 24 h in 6 M HCl containing 0.5% phenol. Tryptic digestion and separation of peptides by high-voltage paper electrophoresis at pH 6.5 and 1.9 was performed as in [15].

The reduced molecule (dithiothreitol 1 μmol/100 nmol peptide) was carboxymethylated with iodo[2-^{14}C]acetate (5 μmol/100 nmol peptide) in Tris–HCl buffer (pH 8.15) [15]. The carboxymethylated peptide was cleaved with cyanogen bromide (900 mg) in 70% formic acid, and the resultant material was fractionated by gel filtration on Sephadex G-25 (1.5 × 100 cm in 30% acetic acid).

N-terminal amino acid analysis was performed by the dansyl method [16], and dansyl amino acids were identified on polyamide layers [17] in four solvent systems [15,16]. Sequential degradations were carried out in a Beckman 890C sequencer using a 0.1 M quadrol peptide program in the presence of polybrene and with two different systems of detection of liberated residues [18]. Phenylthiohydantoin derivatives were analyzed on a Hewlett-Packard 1084B high-performance liquid chromatograph, using a 250 × 4.6 mm RP-8 column; 10 μm particle size (Hewlett-Packard). Gradient elution of PTH amino acids was obtained by increasing acetronitrile from 24–45% in 0.01 M sodium acetate (pH 4.5) [19]. Identification was also

performed by thin-layer chromatography on silica gel with fluorescent indicator. After a 45 min migration in xylene/isopropanol (7/2, v/v) and drying, amino acid derivatives were detected under ultraviolet light and subsequently stained with collidine–ninhydrin [20].

3. Results

3.1. *Total composition*

Results from amino acid analysis of hydrolysates of both oxidized and reduced carboxymethylated intestinal big somatostatin are given in table 1. Tryptophan recovery was low in these analyses but hydrolysis with 4 M methane sulfonic acid clearly demonstrated the presence of this residue. Hydrolytic losses, and possibly minor contaminants, can account for the small deviations from integer values that are observed in the analytical figures.

The data suggest that the larger form of somatostatin is a peptide composed of 28 residues.

3.2. *Structural analysis*

Three different preparations of ~100 nmol of the elongated somatostatin were analyzed in a liquid-phase sequencer. Two batches were reduced and [^{14}C]carboxymethylated prior to analysis. The results are shown in fig.1, together with the whole structure of the molecule, and with the data obtained for different tryptic and CNBr fragments. When the ^{14}C-labelled peptide was analyzed, the radioactivity was recovered in cycles 17 and 28. The corresponding extracts contained ~10-fold increase in radioactivity above the background values in other cycles. The repetitive yield in initial steps was 96% but fell considerably after cycle 15.

The novel part of the structure was further proven by partial sequence determinations of the isolated CNBr fragments (CB1, CB2) obtained after cleavage at the methionine in position 8 (fig.1). Radioactivity from the [^{14}C]carboxymethylation was recovered in cycle 9 of CB2, as expected. Apart from the ordinary CNBr fragments, a peptide derived from a tryptophan cleavage (CB3; fig.1) was also detected.

A tryptic digest of the peptide was separated into 5 components by high voltage paper electrophoresis at pH 6.5. The somatostatin-14 part accounted for two peptides (T4, T5, fig.1). Among the 3 extra fragments, one was highly basic (T3; free lysine), one was neutral (T2; proven as Glu–Arg by both total compo-

Table 1
Total composition of intestinal somatostatin-28

Residue	Acid hydrolysis		Sum of sequence determinations
	After performic acid oxidation	After reduction and carboxymethylation	
Cys	2.0[a]	2.1[c]	2
Asx	2.9	3.2	3
Thr	1.9	1.9	2
Ser	2.7	2.8	3
Glx	0.9	1.1	1
Pro	1.8	2.0	2
Gly	1.2	1.1	1
Ala	3.9	3.9	4
Met	0.8[b]	0.8	1
Phe	3.3	2.8	3
Trp	+	+	1
Lys	3.2	3.3	3
Arg	2.1	2.3	2
Sum			28

[a] Cysteic acid
[b] Methionine sulfone
[c] Carboxymethylcysteine

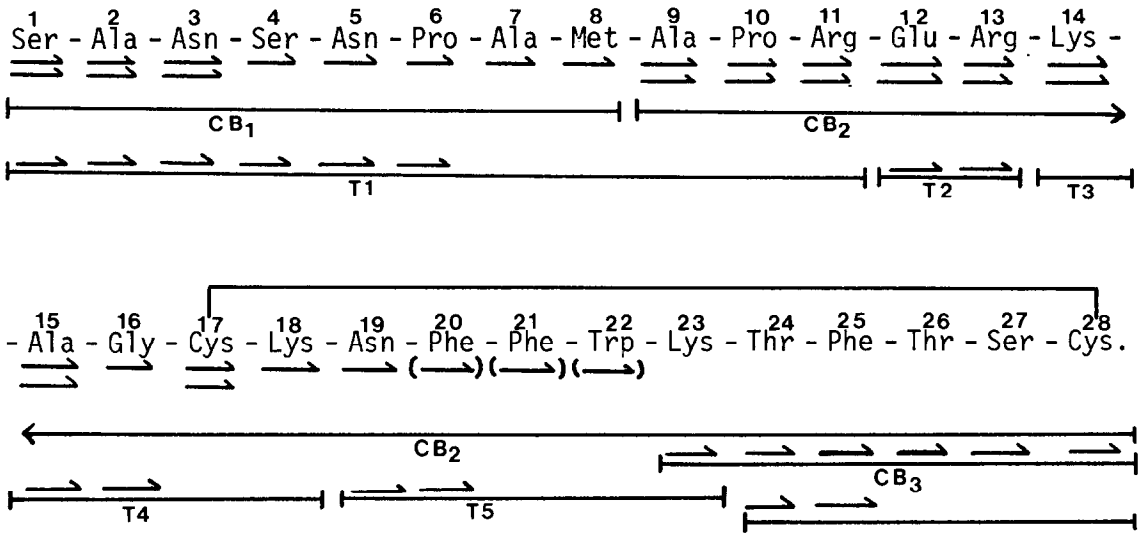

Fig.1. Primary structure of somatostatin-28 as deduced from the present determinations. Top line of arrows show sequencer determinations of the 22 first residues upon degradation of 100 nmol. Initial coupling ~60%; repetitive yield 96% but falling after cycle 15; residue identifications by high-performance liquid chromatography (all cycles), by thin-layer chromatography (all cycles up to 12, except 11), and by radioactivity (cycle 17 on the carboxymethylated derivative). CB1-3 are CNBr-fragments, and T1-5 tryptic fragments, all analyzed as shown. Region 15–28 represents the previously known somatostatin part.

sition and sequence analysis) and one was weakly basic (T1; electrophoretic mobility –0.23 relative to aspartic acid). The latter was incompletely separated from fragment T5 but its total composition and partial sequence analysis are compatible with the structure deduced for the N-terminal region (fig.1).

4. Discussion

The primary structure presently determined establishes the reported partial characterization of the isolated intestinal molecule [13] as an N-terminally extended form of somatostatin. However, the exten-

Table 2
Sequence comparisons of regions preceding cleavage sites at dibasic structures in known proforms of peptide hormones

Proform		Position from cleavage site					Resulting fragment with the N-terminus at the position following –1
		–5	–4	–3	–2	–1	
Gastrin-34	(porcine)	Asp	Pro	Ser	Lys	Lys	Gastrin-17
Proglucagon	(porcine)	Met	Asn	Thr	Lys	Arg	Large glucagon fragment
Proinsulin	(porcine)						
	C–A site	Pro	Pro	Gln	Lys	Arg	Insulin A-chain
	B–C site	Pro	Lys	Ala	Arg	Arg	C-peptide
Proparathyrin	(bovine)	Ser	Val	Lys	Lys	Arg	Parathyrin
Lipotropin	(porcine)						
	site 36→40	Ala	Ala	Glu	Lys	Lys	MSH
	site 56→60	Pro	Lys	Asp	Lys	Arg	β-Endorphin
Somatostatin-28	(porcine)	Pro	Arg	Glu	Arg	Lys	Somatostatin-14

Structures from [22] and this work

sion is longer than previously estimated.

The two consecutive basic residues located just before the cleavage site that produces somatostatin-14 may be noticed. Such dibasic structures are of frequent occurrence in biosynthetic peptide hormone precursors. In the porcine intestinal somatostatin-28, the N-terminal extension is linked to the tetradecapeptide hormone through an Arg–Lys sequence.

Comparisons of known activation sites in prohormones are of interest. As shown in table 2, these sites in prohormones are largely preceded by non-hydrophobic residues, as in [21], often with proline in positions –4 or –5. It seems possible that these characteristics could account for the specificity of the activating enzymes. In any event, the structures suggest that somatostatin-28 may be a prohormone form. The present determination furthermore confirms the presence of somatostatin-like peptides in the gut and provides conclusive evidence of larger forms of somatostatin.

Acknowledgements

Technical assistance by Pierrette Fagot and Henry Lindberg is gratefully acknowledged. This work was supported by grants from INSERM (78.5.238.7), DGRST (77/7/1978), the Swedish Medical Research council (13X-1010, 13F-5380 and 13P-4444), and the Knut and Alice Wallenberg's Foundation.

References

[1] Brazeau, P., Vale, W., Burgus, R. and Guillemin, R. (1974) Can. J. Biochem. 52, 1067–1072.
[2] Burgus, R., Ling, N., Butcher, M. and Guillemin, R. (1973) Proc. Natl. Acad. Sci. USA 70, 684–688.
[3] Schally, A. V., Dupont, A., Arimura, A., Redding, T. W., Nishi, N., Linthicum, G. L. and Schelesinger, D. H. (1976) Biochem. J. 15, 509–514.
[4] Vale, W., Brazeau, P., Rivier, C., Brown, M., Boss, B., Rivier, J., Burgus, R., Ling, N. and Guillemin, R. (1975) Rec. Prog. Horm. Res. 31, 367–397.
[5] Broenstein, M., Arimura, A., Sato, H., Schally, A. V. and Kizer, J. S. (1975) Endocrinol. 96, 1456–1461.
[6] Arimura, A., Sato, H., Dupont, A., Nishi, N. and Schally, A. V. (1975) Science 189, 1007–1009.
[7] Luft, R., Effendic, S., Hökfelt, T., Johansson, O. and Arimura, A. (1974) Med. Biol. 52, 428–430.
[8] Vale, W., Ling, N., Rivier, J., Villareal, J., Rivier, C., Douglas, C. and Brown, M. (1976) Metabolism 25, Suppl. 1, 1491–1494.
[9] Kronheim, S., Berelowitz, M. B. and Pimstone, B. L. (1968) Diabetes 27, 523–529.
[10] Millar, R. P. (1978) J. Endocrinol. 77, 429–430.
[11] Millar, R. P., Denniss, P., Tobler, C., King, J. C., Schally, A. V. and Arimura, A. (1979) in: Biologie cellulaire des processus neurosécretoires hypothalamiques (Vincent, J. D. and Kordon, C. eds) ch. 30, pp. 487–510, CNRS Paris.
[12] Pradayrol, L., Chayvialle, J. A. and Mutt, V. (1978) Metabolism 27, suppl. 1, 1197–1200.
[13] Pradayrol, L., Chayvialle, J. A., Carlquist, M. and Mutt, V. (1978) Biochem. Biophys. Res. Commun. 85, 701–708.
[14] Waley, S. G. and Watson, J. (1954) Biochem. J. 57, 529–538.
[15] Jörnvall, H. (1970) Eur. J. Biochem. 14, 521–534.
[16] Hartley, B. S. (1970) Biochem. J. 119, 805–822.
[17] Woods, K. R. and Wang, K. T. (1967) Biochim. Biophys. Acta 133, 369–370.
[18] Jörnvall, H. and Philipson, L. (1980) Eur. J. Biochem. in press.
[19] Zimmerman, C. L., Appella, E. and Pisano, J. J. (1977) Anal. Biochem. 77, 569–573.
[20] Inagami, T. and Murakami, K. (1972) Anal. Biochem. 47, 501–504.
[21] Geisow, M. J. (1978) FEBS Lett. 87, 111–114.
[22] Dayhoff, M. O. (1972) Atlas of Protein Sequence and Structure, vol. 5; (1978) suppl. 3, National Biomedical Research Foundation, Silver Spring.