PRIMARY STRUCTURE OF CHOLERA TOXIN SUBUNIT A_1

Isolation, partial sequences and alignment of the BrCN fragments

C. Y. LAI, F. CANCEDDA* and D. CHANG

Department of Physiological Chemistry and Pharmacology, Roche Institute of Molecular Biology, Nutley, NJ 07110, USA

Received 3 January 1979

1. Introduction

Cholera toxin consists of two subunits, A and B, in a molar ratio of 1 A to 5 B [1,2]. Subunit A is responsible for the action of cholera toxin in stimulating the production of cAMP in the cell, and subunit B, for the binding of the toxin to the cell surfaces (reviewed [3]). The A subunit is composed of two polypeptides, A_1 (M_r 22 000) and A_2 (M_r 7500), linked to each other by a single disulfide bond [1]. It has been shown that polypeptide A_1, separated from A_2, is the only component of cholera toxin capable of activating the adenylate cyclase activity in a membrane preparation of pigeon erythrocytes [4].

We report here the isolation of four peptides resulting from BrCN cleavage of the A_1 subunit and the elucidation of their linear alignments. This work has provided not only the necessary information for determination of the complete sequence but also insight into the molecular make-up of subunit A_1.

2. Materials and methods

Cholera toxin was purified from the culture filtrate of *V. cholerae*, strain 569 B, with the modified procedure [5].

Subunits were separated by gel filtration on Sephadex G-75 in 5% HCOOH as in [1]. For the isolation of polypeptide A_1, subunit A was dissolved in 0.4 M Tris–Cl buffer (pH 8.5), containing 0.01 M dithiothreitol (DTT), 0.02 M EDTA and 8 M urea for 2 h and gel filtered on Sephadex G-75 in 5% HCOOH containing 1 mM DTT. *S*-carboxymethyl (CM) A_1 peptide was prepared similarly after reduction and *S*-carboxymethylation of subunit A in 8 M urea [6]. Purity of these preparations was ascertained by acrylamide–gel electrophoresis in 0.1% SDS and 8 M urea, as well as by amino acid analysis [1].

Sequential degradations of peptides were performed according to Edman's procedure [7], manually as in [8], or automatically in a Beckman Model 890C sequencer (Beckman Instrument Co., Palo Alto, CA). The amino acid degraded at each step was identified by amino acid analysis after regeneration in 5.7 N HCl containing $SnCl_2$ [9].

Amino acid analyses were performed with a Joel Model 5AH automatic analyzer (JEOL, USA Inc., Medford, NJ) or with a microcolumn analyzer equipped with the fluorescamine detector system [10] constructed in this laboratory. For the analyses of homoserine, asparagine and glutamine, the column was operated at 32°C.

3. Results

3.1. *Separation and characterization of the BrCN fragments of polypeptide* A_1

Polypeptide A_1 had been found to contain 3 methionine residues/mol. When the BrCN cleavage product [11] of the *S*-carboxymethylated peptide A_1 was chromatographed on Sephadex G-50 in 5%

* Present address: International Institute of Genetics and Biophysics, Naples, Italy

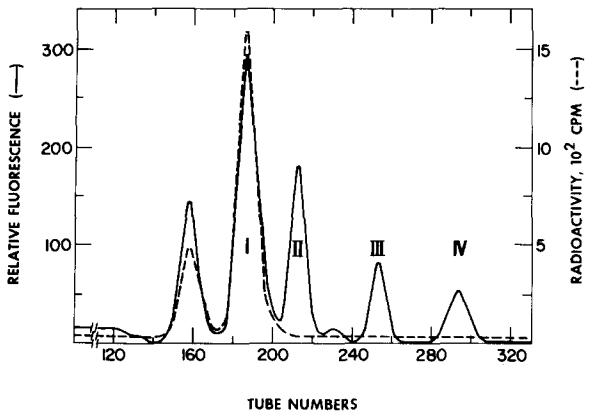

Fig.1. Sephadex G-50F chromatography of BrCN peptides from cholera toxin subunit A_1: *S*-[^{14}C]carboxymethyl polypeptide A_1 (5 mg in 0.5 ml 70% HCOOH) was treated with 2.1 mg BrCN in 50 μl 70% HCOOH for 22 h at room temperature. After dilution with water and lyophilization, the cleavage product was gel filtered on a column (1.5 × 180 cm) of Sephadex G-50F in 5% HCOOH at room temperature. The flow rate was 12 ml/h and fractions (1 ml) were collected. Aliquots (0.1 ml) were analyzed for peptides with fluorescamine after hydrolysis (——) and for radioactivity (30 μl) (- - -).

HCOOH, 4 peptides were separated in excellent yields (fig.1). The small peak appearing before the first major peak was obtained in varying yields from experiment to experiment (0–15% of the total peak area) and was considered an artifact of the BrCN reaction. The four major peaks were designated Cn I, Cn II, Cn III and Cn IV, in the order of their elution from the column (fig.1).

The purity of BrCN peptides Cn I–Cn IV was established by amino acid analyses and the end group analyses; Cn I contained the only CM-cysteine residue in polypeptide A_1 [1,13]. Peptides Cn I and Cn II were devoid of Lys; Cn III of His, Thr and Val; and Cn IV of Lys, His, Ala, Ile and Leu. All but Cn I contained homoserine residue. One-step Edman degradation on each peptide showed a single NH_2-terminal residue for each; Phe for Cn I, Asx for Cn II and Cn III, and Pro for Cn IV. Since the NH_2-terminus of the parent peptide A_1 has been determined as asparagine [14,15], either Cn II or Cn III is indicated as derived from the NH_2-terminus of A_1.

3.2. *Sequence analyses of peptides A_1, Cn I, Cn II, Cn III and Cn IV*

Edman degradation was carried out manually [8] or with an automatic sequencer on about 5 nmol or 100 nmol of the samples, respectively (table 1). The results provided further evidence for chemical purity of these peptides. The NH_2-terminal BrCN peptide has been identified as Cn III.

Peptide Cn I was the only peptide which did not yield homoserine on acid hydrolysis. Digestions of the BrCN peptides with carboxypeptidase A (10 μg for 5–6 nmol peptides in 0.1 ml 0.2 M *N*-ethylmorpholine acetate buffer (pH 8.5) at 37°C) released the following amino acids (expressed in mol fractions) at the specified time intervals:

Polypeptide A_1: 1 h – Ser (0.14), Gly (0.12); 5 h – Ser (0.20), Gly (0.18), Asn (0.1)

Cn I: 1 h – Ser (0.14), Gly (0.12); 5 h – Ser (0.20), Gly (0.18), Ala, Leu (0.13), Asn (0.1)

Cn II: 1 h – Hse (0.133); 7 h – Hse (0.17), Leu (0.077)

Cn III: 10 min – Hse (0.49), Leu (0.51), Gly (0.06), Ser (0.04); 1 h – Hse (0.81), Leu (1.49), Gly (0.09), Ser (0.05)

Cn IV: 30 min – Hse (0.37), Gln (0.30), Thr (0.09), Gly (0.06); 1 h – Hse (0.59), Gln (0.43), Thr (0.14), Gly (0.1).

The results summarized in table 1, confirmed that Cn I was from the COOH-terminus of polypeptide A_1. The arrangement of BrCN peptides is indicated to be (NH_2)Cn III–(Cn II, Cn IV)–Cn I (COOH).

3.3. *Isolation of the peptides containing overlaps for BrCN fragments from the tryptic digest of CM-peptide A_1*

In order to obtain information on the linkage between BrCN peptides, CM-peptide A_1 was digested with trypsin, and the peptides containing methionine were separated (fig.2). These peptides would contain overlaps between the BrCN peptides. Three such peptides, T1 from initial separation on Sephadex G-25 (fig.2A), T3b from DEAE-cellulose chromatography of T3 fraction (fig.2B), and T4h from Dowex 50 chromatography of T4 fraction (fig.2C) were obtained in 82%, 69% and 75% yields, respectively.

3.4. *Alignment of the BrCN peptides*

The methionine-containing peptides (overlap) had the following compositions:

Table 1
Partial sequence of polypeptide A_1 and the BrCN fragments[a]

Polypeptide A_1[b]

NH_2–: Asx–Asx–Asx–Lys– –COOH: –(Asn,Gly)–Ser
78 56 45 6

Peptide Cn I

NH_2–: Phe–Asx–Val–Asx–Asx–Val–Leu–Gly–Ala–Tyr–Ala–Pro–His–Pro–Asx–Glx–Glx–Glx–Val–Ala–Ala–Leu–Gly–Gly–Ile–Pro–Tyr–
73 23 64 33 32 57 21 54 45 41 19 42 6 30 14 23 24 20 12 23 21 14 18 20 9 10 15
–COOH: –Asn–Gly–Ser

Peptide Cn II

NH_2–: Asx–Ile–Asx–Leu–Tyr–Asx–His–Ala–Arg–Gly–Thr–Glx–Thr–Gly–Phe–Val–Arg
59 71 50 60 60 27 21 37 23 39 30 24 31 14 31 23 14
–COOH: Leu–Hse

Peptide Cn III

NH_2–: Asx–Asx–Asx–Lys–Leu– –COOH: –Ser–Gly–Leu–Leu–Hse
76 53 32 7 16

Peptide Cn IV[c]

Pro–Arg–Gly–Ser,Glu,Glu,Tyr–Phe–Asx–Arg–Gly–Thr–Gln–Hse
65 25 17
├–IVT1–┤├–IVT2–┤├–IVT3–┤

[a] The NH_2-terminal sequences were determined by Edman degradation (see section 2). The uncorrected yield after each step (%) is shown under the residue degraded. The COOH-terminal sequences have been deduced from the results of carboxypeptidase A experiments (see text)

[b] Tentative sequence of 17 residues from NH_2-terminus of A_1 has been proposed [12]. The results of manual degradation on 5 nmol polypeptide A_1 are shown here

[c] Partial sequence of Cn IV was determined by isolation and analyses of three tryptic peptides, IVT1, IVT2 and IVT3, in addition to Edman degradation and carboxypeptidase A experiments on the whole peptide

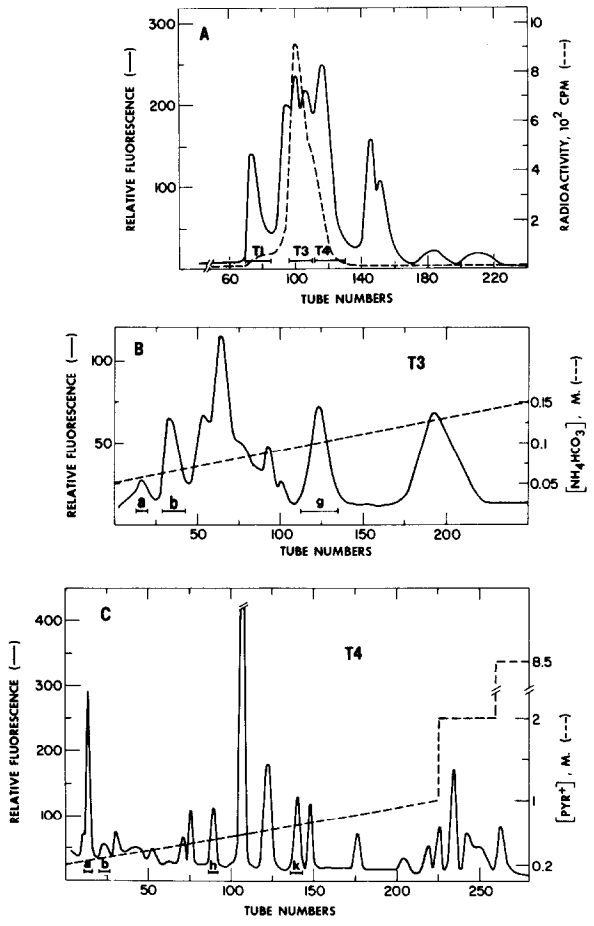

Fig.2. Separation of tryptic peptides containing methionine residues: *S*-[^{14}C]carboxymethyl A_1 (6 mg) was digested with 50 μg TPCK-trypsin in 1 ml 0.05 M Tris–HCl buffer (pH 8.7) for 6 h at 37°C. (A) The digest was gel filtered on a Sephadex G-25F column (1.5 × 180 cm) in 0.1 M NH_4HCO_3 at 14 ml/h. Fractions (2 ml) were collected. For the fluorimetry of peptides (——), 0.1 ml aliquots were used, and for radioactivity (- - -) 20 μl. Amino acid analyses on combined fractions showed that all methionine residues were recovered in fractions T1, T2 and T3. (B) Fraction T3 above was chromatographed on a DEAE-cellulose column (DE-52, 1 × 30 cm) in 0.05 M NH_4HCO_3 (pH 8.5). Elution was carried out with a concentration gradient of the buffer as shown (- - -). The flow rate was 12 ml/h and 1.5 ml fractions were collected. Aliquots (0.1 ml) were used for the fluorescamine analysis (——). (C) The Sephadex G-25 fraction T4 (fig.2A) was chromatographed on a column (0.9 × 27 cm) of Aminex AG50W × 4, with a pyridine–acetate buffer gradient as indicated (- - -). The flow rate was 12 ml/h and fractions (1.6 ml) were collected. Aliquots (0.21 ml) were used for the fluorescamine assay (——).

T1 – $Asp_7Thr_4Ser_5Glu_6Pro_5Gly_7Ala_7Val_7Met_1Ile_5Leu_6Tyr_7Phe_2His_3Arg_1$

T3b – $Asp_3Thr_1Glu_1Gly_1Ala_1Met_1Ile_1Leu_1Tyr_1His_1Arg_1$ (13 residues)

T4h – $Asp_1Ser_1Glu_1Pro_1Gly_2Met_1Leu_1Arg_1$ (9 residues)

In Cn I, no Lys or Arg was found in the NH_2–terminal segment of at least 15 residues (table 1). Only T1 (73 residues) could contain overlap for Cn I.

The sequence of T3b was found to be:

Gly–Thr–Glu–Met–Asx ($Asx_2Ala_1Leu_1Tyr_1His_1$)–Arg

by Edman degradation. On BrCN cleavage, this peptide yielded a peptide, $Thr_1Gly_1Glu_1Hse_1$. The results indicated that T3b overlapped Cn IV and Cn II (table 1).

The above argument indicated that T4h overlapped Cn III and Cn IV. This was confirmed by isolation and analyses of peptide Pro–Arg (NH_2-terminus of Cn IV) on BrCN treatment of T4h and paper electrophoresis. The arrangement of BrCN peptides is therefore:

CnIII– –CnIV– –CnII– –CnI (fig.3).

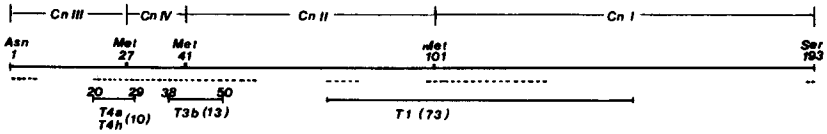

Fig.3. Schematic sketch of the primary structure of cholera toxin subunit A_1. The numbers from the NH_2-terminus are indicated. The lower bars indicate the tryptic peptides containing overlaps between BrCN fragments. The figures in parentheses denote the number of residues in the peptides. The dotted lines indicate the area where the sequences are known.

4. Discussion

Although peptides Cn I and Cn II are still large in size (~92 and 60 residues, respectively), the task of sequence analyses on these peptides is considerably reduced from that on the entire polypeptide A_1. Tentative sequence of altogether 68 residues have been determined (table 1) and shall be of value in the elucidation of the complete primary structure. Knowledge on the alignment of BrCN fragments and their amino acid composition shall also be useful for mapping specific sidechain functions or sites involved in a variety of reactions on polypeptide A_1.

The sum of amino acid residues in the four BrCN peptides is in good agreement with the amino acid composition of polypeptide A_1 [1]. A schematic diagram of the subunit A_1 (fig.3) has revealed interesting features of this molecule. The only two lysine residues are found within a 1/10th segment of the polypeptide chain at the NH_2-terminus. Arginine residues, as much as 14/mol, are distributed among 50% of the chain, which excludes the 73-residue stretch (37% of the chain length) corresponding to peptide T1 situated near the center of the molecule (fig.3). Tryptophan was detected in peptides Cn I and Cn II by spot-test on paper [15]. The single cysteine residue of subunit A_1 is located in the COOH-terminal quarter of the peptide chain, for it is found in peptide Cn I but not in T1 (fig.1,2). This area must contain the active site responsible for the stimulation of adenylate cyclase since the activity of subunit A greatly increases on cleavage of the disulfide bond [4].

Polypeptides A_1 and A_2 have been suggested to derive from a single precursor peptide [16]. In view of the present finding, the precursor peptide would probably be synthesized from the NH_2-terminus of A_2, ending in the COOH-terminus of A_1. Disulfide formation between cysteine residues near both ends would then occur, followed by the proteolysis at the other side of the loop.

Acknowledgements

Dr F. Cancedda was partially supported by NATO Fellowship. We thank Dr A. Kurosky, University of Texas Medical Branch, Galveston and Dr Lawrence K. Duffy for valuable discussions.

References

[1] Lai, C. Y., Mendez, E. and Chang, D. (1976) J. Infec. Dis. 133, S23–S30.
[2] Gill, D. M. (1976) Biochemistry 15, 1242–1248.
[3] Gill, D. M. (1977) Adv. Cyclic. Nucl. Res. 8, 85–118.
[4] Wodnar-Filipowicz, A. and Lai, C. Y. (1976) Arch. Biochem. Biophys. 176, 465–471.
[5] Rappaport, R., Rubin, B. A. and Tint, H. (1973) Infect. Immunol. 9, 294–303.
[6] Crestfield, A. M., Moore, S. and Stein, W. H. (1965) J. Biol. Chem. 238, 622–627.
[7] Edman, P. (1960) Ann. NY Acad. Sci. 88, 602–610.
[8] Tarr, G. E. (1977) Methods Enzymol. 47, 335–357.
[9] Lai, C. Y. (1977) Methods Enzymol. 47, 369–373.
[10] Stein, S., Böhlen, P., Stone, J., Dairman, W. and Udenfriend, S. (1973) Arch. Biochem. Biophys. 155, 203–212.
[11] Gross, E. and Witkop, B. (1961) J. Biol. Chem. 237, 1856–1861.
[12] Lai, C. Y. (1977) Methods Enzymol. 47, 236–243.
[13] Mendez, E., Lai, C. Y. and Wodnar-Filipowicz, A. (1975) Biochem. Biophys. Res. Commun. 67, 1435–1442.
[14] Kurosky, A., Markel, D. E., Touchstone, B. and Peterson, J. W. (1976) J. Infec. Dis. 133, 514–522.
[15] Klapper, D. G., Finkelstein, R. A. and Capra, J. D. (1976) Immunochemistry 13, 605–611.
[16] Reddi, K. K. and Kodicek, E. (1953) Biochem. J. 53, 286–288.
[17] Gill, D. M. and Rappaport, R. (1976) Symp. Cholera, pp. 159–166, Sapporo, Japan.