THE AMINO TERMINAL SEQUENCE OF CHOLERA TOXIN SUBUNITS

J.W. Jacobs*, H.D. Niall*, and G.W.G. Sharp†

Endocrine Unit* and the Biochemical Pharmacology Unit†
Massachusetts General Hospital and Harvard Medical
School, Boston, Massachusetts 02114

Received September 20, 1974

Summary

The N-terminal amino acid sequence of Cholera toxin, molecular weight 84,000 daltons, has been established. A high sensitivity sequencing procedure, employing ^{35}S-labelled phenylisothiocyanate as the coupling reagent in the automated Edman degradation was used. The toxin was found to consist of two polypeptide chains in the approximate molar ratio of 4:1. The amino-terminal twenty residues of each subunit will be reported here.

INTRODUCTION

Cholera toxin was recently purified and crystallized, and estimated to have a molecular weight of 84,000 daltons (1,2). In recent investigations of subunit structure, Lonroth and Holmgren provided evidence for light (L) and heavy (H) components having molecular weights of 8,000 and 28,000 respectively, suggesting that the toxin might be composed of 7 light units and 1 heavy unit (3). In a contrasting report, van Heyningen estimated the subunit molecular weights to be 25,000 daltons (defined as subunit A) and 15,000 (subunit B), respectively (4). Hence, based on the latter results, the toxin might be composed of 4 light (B) polypeptide chains and 1 heavy (A) chain.

In the present studies, direct determination of subunit molar ratios was possible by subjecting cholera toxin to automated Edman degradation (5). A radioactive micro-sequencing procedure was employed (6) that allowed accurate quantitation of subunit amino acid residues. The amino-terminal 20 residues of each subunit were identified and found to exist in a molar ratio of 4:1.

METHODS

Cholera toxin (lot#1071) was prepared under contract for the National Institutes of Allergy and Infectious Diseases by R.A. Finkelstein and analyzed without further purification. 200 nanomoles of the toxin was degraded for twenty-five continuous cycles in a Beckman Model 890 C Sequencer (Palo Alto, Cal.). Sequencer reagents and solvents were obtained from Beckman Instruments. ^{35}S-labelled phenylisothiocyanate

TABLE I

Amino Terminal Sequence and Molar Yields of the Major (Light) and Minor (Heavy) Subunits of Cholera Toxin

Degradation Cycle	Major Residue	Yield (nM)*	Minor Residue	Yield (nM)*
1	Thr	173.0	Asn	60.0
2	Pro	138.0	Asp	35.3
3	Gln	93.6	Asp	24.1
4	Asn	107.2	Lys	32.3
5	Ile	224.2	Ile	**
6	Thr	75.3	Tyr	36.1
7	Asp	135.2	Arg	23.3
8	Leu	122.0	Ala	32.1
9	Ser	52.1	Asp	21.3
10	Ala	160.7	Ala	**
11	Glu	109.0	Arg	22.3
12	Tyr	65.3	(Cys)	†
13	His	61.2	Pro	19.8
14	Asn	121.3	Asn	**
15	Thr	61.2	Arg	13.3
16	Gln	58.4	Ile	16.2
17	Leu	60.1	Gly	16.8
18	His	31.2	(Cys)	†
19	Thr	24.2	Arg	9.8
20	Leu	42.1	Gly	10.1

*These molar yields are not corrected for recovery losses (see text).

**Identical PTH-residues found in both major and minor chains at this cycle; thus separate molar yield determinations for both chains are not possible.

† PTH-Cys; a very unstable residue; cannot be quantitated (see text).

(Amersham-Searle, Arlington Heights, Ill.) was used as the coupling reagent in the Edman degradation at a specific activity of 1.0 mCi/mmol. Amino acid residues, phenylthiohydantoins (PTH'S) released at each cycle of the degradation were identified by thin-layer chromatography (tlc)

on silica-gel plates (Analtech, Inc., Newark, Del.) using solvent systems previously described (6,7). The ^{35}S-labelled residues were located on the tlc plates by overnight autoradiography, scraped from the plate, and quantitated in a Packard Model 3375 Liquid Scintillation Counter (Packard Instr. Co., Downers Grove, Ill.). Residue identifications were confirmed by gas-liquid chromatography using a Beckman Model 45 Gas Chromatograph with a 10% DC-560 support (8). All identifications were based on a threefold-or-greater rise in yield of the particular residue above the background level of amino acid phenylthiohydantoin at that stage of the degradation. The results of this degradation were confirmed by two additional sequence analyses.

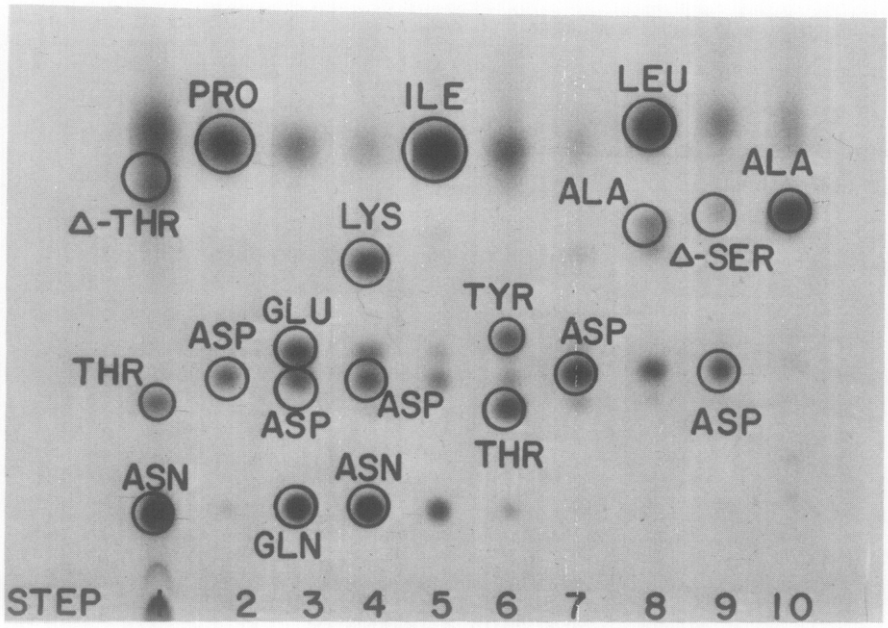

Fig. 1 Autoradiography of $^{(35}S)$PTH residues released during automated Edman degradation of cholera toxin. Residues were separated by tlc on a silica-gel plate employing ethylene dichloride : glacial acetic acid, 30:7 v/v, as the solvent system. Breakdown of various PTH-residues (see text) are apparent from this autoradiograph, e.g., step 1, where PTH-threonine breaks down to a second product PTH-dehydro-threonine (Δ-Thr), and step 4, where PTH-asparagine partially deaminates to form aspartic acid. PTH-arginine (cycle 7) is identified by a separate tlc system and does not appear on the above autoradiograph. For more-definitive identification and quantitation of the PTH-amino acids, two-dimensional tlc systems (6) and gas-liquid chromatography were employed (8).

RESULTS AND DISCUSSION

Table I lists the amino acids identified and their molar yields at each cycle of the degradation. Unique residue assignments were possible for the first 20 positions of both the major (light) and minor (heavy) subunits. To illustrate the use of the high-sensitivity sequencing procedure in identifying subunit residues, an autoradiograph of the first 10 cycles of the degradation is shown in Fig. 1.

No residue could be identified at position 12 and 18 of the minor subunit, but it is extremely probable that these positions are half-cysteine residues since any other PTH-residue would have been readily identifiable. The phenylthiohydantoin derivatives of cysteine and cystine are very unstable and cannot be normally identified (9).

The sequence data show that, in general, there exists a 4:1 molar ratio between the subunits. This 4:1 ratio is best seen at degradation cycles where stable phenylthiohydantoins (PTH-leucine, isoleucine, valine, glycine, alanine, proline, aspartic acid, or glutamic acid) (9) appear in both subunits at one particular cycle, e.g., positions 2,8,17, and 20. One such position, at cycle 8, is depicted in Fig. 2, which compares the yield of PTH-leucine (major chain) to PTH-alanine (minor

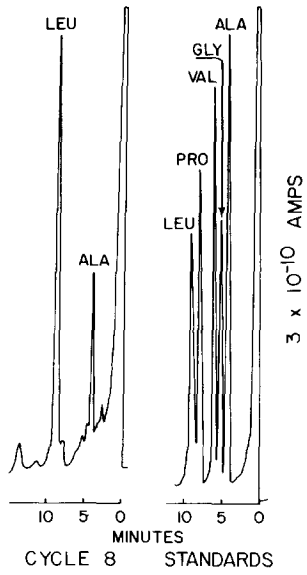

Fig. 2 Gas-liquid chromatography of cycle 8 of cholera toxin degradation. PTH-amino acid standards are shown on right; cycle 8 on left shows PTH-leucine (major chain) and PTH-alanine (minor chain). The molar yield of each residue is given in Table I. The molar ratio of PTH-Leu : PTH-Ala in this chromatogram is approximately 4:1.

chain) as determined by gas-liquid chromatography. At other cycles, the ratio fluctuates in a predictable manner to either (i) the normal breakdown of certain residues during Edman degradation, i.e., PTH-serine, threonine, glutamine, and asparagine (30-50% breakdown rate) (9), or (ii) poor recovery of phenylthiohydantoins during automated sequencing due to incomplete extraction of the residue i.e., PTH- arginine and histidine (50-70% recovery rate) from the sequenator's reaction vessel (9). When the molar yields of poor recovery or unstable residues are corrected, the 4:1 ratio is more closely approximated for each cycle of the degradation. The accurate quantitation of residues and placement into their proper subunit was due, in large part, to the use of the micro-sequencing procedure (6). By employing a radioactive coupling reagent in the Edman degradation, amino acid derivatives were labelled at a high specific activity, allowing identifications of trace quantities of residue.

This study establishes the ratio of subunit chains as 4:1 and supports the estimation of van Heyningen (4) that cholera toxin is composed of four polypeptide chains of approximately 15,000 daltons and one polypeptide chain of approximately 25,000 daltons. In addition, the amino terminal sequence reported here is a first step toward the determination of the primary structure of cholera toxin and subsequently will allow the detailed mode of its action to be understood in molecular terms.

ACKNOWLEDGEMENTS

We wish to thank Phillip Dee for the technical art work. This work was supported in part by grants from the John A. Hartford Foundation, Inc., and by the Public Health Service research grants, HE-06664 from the National Heart Institute and AM-4501 from the National Institute of Arthritis and Metabolic Diseases.

REFERENCES

1. LoSpalluto, J.J., and Finkelstein, R.A. (1972) Biochem. Biophys. Acta 257, 158-161.
2. Finkelstein, R.A., and LoSpalluto, J.J. (1972) Science 175, 529-533.
3. Lonnroth, I., and Holmgren, J. (1973) J. Gen. Microbiol. 76, 417-424.
4. van Heyningen, S. (1974) Science 183, 656-660.
5. Edman, P., and Begg, G. (1967) European J. Biochem. 1, 80-91.
6. Jacobs, J.W., Sauer, R.T., Niall, H.D., Keutmann, H.T., Aurbach, G.P., O'Riordan, J.L.H., and Potts, J.T., Jr. (1973) Federation Proc. 32, 2445.
7. Edman, P. (1970) in Protein Sequence Determination (Needleman, S., ed.), pp. 211-248, Springer-Verlag, New York.
8. Pisano, J.J., and Bronzert, T.J. (1969) J. Biol. Chem. 244, 5597-5605.
9. Niall, H.D. (1973) Methods Enzymol. 22, 942-975.