Amino-acid sequence of Fragment A, an enzymically active fragment from diphtheria toxin

[heterogeneity in length/proteolysis/bromine adduct of 2-(2-nitrophenylsulfonyl)-3-methylindole (BNPS-skatole)/sequenator/ADP-ribosylation]

ROBERT J. DELANGE*, RAE E. DRAZIN†‡, AND R. JOHN COLLIER†‡

* Department of Biological Chemistry, School of Medicine, † Department of Bacteriology, and ‡ Molecular Biology Institute, University of California, Los Angeles, Calif. 90024

Communicated by Emil L. Smith, October 16, 1975

ABSTRACT The amino-acid sequence of Fragment A from diphtheria toxin is reported. Fragment A (molecular weight, Mₐ, 21,145) is the major enzymically active fragment produced upon activation of the intact toxin (M₀ about 60,000) by limited trypsic digestion and reduction. It, or a similar fragment, is believed responsible for the inhibition of protein synthesis in animal cells exposed to the toxin. Fragment A, which corresponds to the amino terminus of the toxin, is shown here to consist of three major forms (190, 192, and 193 residues) resulting from cleavage by trypsin adjacent to any of three closely spaced arginine residues. All three forms are enzymically active.

Diphtheria toxin (molecular weight, M₀, about 60,000) inhibits protein synthesis in sensitive mammalian cells, and this effect is probably responsible for its toxicity for man and certain animals (1, 2). The biochemical mechanism of the inhibition is now well understood. The toxin is a proenzyme which, after activation, catalyzes the covalent attachment of the adenosine diphosphate ribose moiety (ADP-Rib) of nicotinamide adenine dinucleotide (NAD) to elongation factor 2 (EF-2).

EF-2 + NAD → ADP-Rib-EF-2 + nicotinamide + H⁺

This modification inactivates EF-2, thus blocking peptide chain elongation on the ribosomes (1–4).

Activation of the toxin, a single polypeptide chain (M₀ about 60,000), requires proteolytic cleavage as an essential step (5, 6). In fact, the enzymically active entities are large proteolytic fragments from the toxin. These fragments must be dissociated from the remainder of the molecule before activity is expressed. When the active fragment is attached by disulfide linkage, cleavage of this linkage (by reduction, for example) is necessary for dissociation and expression of activity. Presumably similar activation steps occur in vivo, but at present our knowledge rests largely on work in cell-free systems.

The enzymically active fragment which has been studied most extensively is Fragment A (7). This fragment may be isolated in high yield after treatment of intact toxin with low concentrations of trypsin or certain other proteases. Trypsin attacks the toxin preferentially within a single, narrow segment, producing two major fragments, NH₂-terminal A (Mₐ 21,145) and COOH-terminal B (Mₐ about 38,000), which are linked by a single disulfide bridge (5–9). Fragment A becomes enzymically active when released from B after reduction of the disulfide. Fragment B has no known enzymic activity but is essential for toxicity in test animals or cultured cells, and probably acts to bind the whole toxin molecule to cell surface receptors.

Fragment A is the smallest and most active of the enzymically active fragments tested. It is markedly resistant to proteolytic attack, to extremes of pH, and to high temperature. In addition to catalyzing the ADP-ribosylation of EF-2, Fragment A has weak NAD-glycohydrolase activity (EC 3.2.2.5), which results in the release of ADP-Rib, nicotinamide, and H⁺. These properties together with the interesting biological activity of the toxin have led us to determine the primary structure of the fragment.

MATERIALS AND METHODS

Diphtheria Toxin. Partially purified toxin obtained from Connaught Laboratories, Toronto, Canada, was further purified by chromatography on columns of DEAE-cellulose and Sephadex G-100, as described previously (10).

Fragment A. The purified toxin (A₂₅₀ = 12.2) was incubated with trypsin (1 μg/ml) at 25° for 45 min, and digestion was terminated by adding soybean trypsin inhibitor (1.5 μg/ml) (5). The cleaved toxin was then reduced in the presence of 50 mM diithiothreitol and 0.5 M guanidine-HCl, and the free sulfhydryl groups were blocked by dialysis against 30 mM 2,2'-dihydroxyethyl disulfide. Fragment A was subsequently purified by chromatography on Sephadex G-100. The buffer used in all steps was 50 mM Tris-HCl, pH 8.2, containing 1 mM ethylenediaminetetraacetate. About 1.6 g of Fragment A was used for the sequence analysis.

Sequencing Methods. Four types of peptides were obtained to establish the complete sequence of Fragment A: (a) cyanogen bromide peptides, which were initially fractionated on a column of Sephadex G-50 (11); (b) trypic peptides from the maleylated protein, which were initially separated on a column of Sephadex G-50; (c) chymotryptic peptides, which were initially fractionated on a column of Dowex 50; and (d) peptides formed by cleavage of Fragment A at the two tryptophan residues with the bromine adduct of 2-(2-nitrophenylsulfonyl)-3-methylindole (BNPS-skatole) by the procedure of Omenn et al. (12); these peptides were separated on a column of Sephadex G-75. The reduced Fragment A used to obtain peptides (a), (b), and (c) had been treated to form the N-carboxymethylcysteine derivative. The protein used for method (d) was in the form of the mixed disulfide of cysteine and 2-mercaptoethanol. When further purification of peptides was required, gel filtration, paper electrophoresis at pH 1.9 and 3.6, and paper chromatography were utilized. Large peptides were further hydrolyzed...
with trypsin, chymotrypsin, thermolysin, subtilisin, etc., to give smaller peptides which were isolated and studied. Fragment A and several of the largest peptides were subjected to automatic Edman degradation with the aid of a Beckman 860C sequenator. All other sequence methods are as previously described (13, 14).

RESULTS

The amino-acid composition of Fragment A is given in Table 1. For comparison, the numbers of residues found by sequence analysis are also presented. The only disagreements are in the numbers of arginine, threonine, and glutamic acid (Glx). The lower value obtained for arginine by amino-acid analysis is probably due to the shorter molecules (190 and 192 residues) which are found in significant amounts in preparations of Fragment A (see below) and contain one or two fewer arginine residues than the 193-residue molecule. One additional residue of threonine was found in the sequence, presumably because of undercorrection for destruction of this labile residue. Fragment A is distinctive in containing only one residue each of histidine and cysteine, two of tryptophan, and a large number of aromatic residues (19 out of a total of 193 residues).

Five cyanogen bromide peptides which accounted for all of the residues in Fragment A (see Fig. 1) were isolated and extensively studied (11). These were Peptide CNBr-1 (residues 1–14), Peptide CNBr-2 (residues 15–115), Peptide CNBr-3 (residues 116–178), Peptide CNBr-4 (residues 179–182), and Peptide CNBr-5 (residues 183–193). Two other peptides were found which proved to be the same as Peptide CNBr-5, except that one lacked the COOH-terminal arginine, and the other lacked the COOH-terminal Val-Arg-Arg sequence.

The presence of three distinct COOH-terminal peptides indicates that Fragment A is heterogeneous and is formed from diphtheria toxin by tryptic cleavage at any of the three arginine residues in the sequence Arg-Val-Arg-Arg (residues 190–193). The three species formed (190, 192, and 193 residues, respectively) probably correspond to the three major bands observed when Fragment A is electrophoresed in 7.5% polyacrylamide gels in the absence of detergent (7). The presence of an additional, positively charged arginine residue in each successively longer molecule could account for the electrophoretic separation of the three components. All three forms isolated from polyacrylamide gels catalyze both the ADP-ribosylation of EF-2 and the hydrolysis of NAD. They are formed in similar amounts and do not differ markedly from another in specific activity. At least one active, minor electrophoretic form has been detected, but its chemical nature is not yet known.

Six tryptic peptides were isolated from the maleylated protein: Peptide Tm-1 (residues 1–126), Peptide Tm-2 (residues 127–133), Peptide Tm-3 (residues 134–170), Peptide Tm-4 (residues 171–173), Peptide Tm-5 (residues 174–190), and Peptide Tm-6 (residues 191–193).

Table 1. Composition of Fragment A from diphtheria toxin

<table>
<thead>
<tr>
<th>Residue</th>
<th>Number per molecule</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Analysis*</td>
</tr>
<tr>
<td>Lys</td>
<td>15.9 (16)</td>
</tr>
<tr>
<td>His</td>
<td>0.93 (1)</td>
</tr>
<tr>
<td>Arg</td>
<td>6.03 (6)†</td>
</tr>
<tr>
<td>Asx  ‡</td>
<td>22.6 (23)</td>
</tr>
<tr>
<td>Asp</td>
<td>13</td>
</tr>
<tr>
<td>Aan</td>
<td>10</td>
</tr>
<tr>
<td>Thr</td>
<td>8.25 (8)</td>
</tr>
<tr>
<td>Ser</td>
<td>17.0 (17)</td>
</tr>
<tr>
<td>Glx  ‡</td>
<td>22.8 (23)</td>
</tr>
<tr>
<td>Glu</td>
<td>7</td>
</tr>
<tr>
<td>Gln</td>
<td>6</td>
</tr>
<tr>
<td>Pro</td>
<td>6.16 (6)</td>
</tr>
<tr>
<td>Gly</td>
<td>20.0 (20)</td>
</tr>
<tr>
<td>Ala</td>
<td>14.2 (14)</td>
</tr>
<tr>
<td>Val</td>
<td>16.2 (16)</td>
</tr>
<tr>
<td>Met</td>
<td>4.24 (4)</td>
</tr>
<tr>
<td>Ile</td>
<td>5.78 (6)</td>
</tr>
<tr>
<td>Leu</td>
<td>12.4 (12)</td>
</tr>
<tr>
<td>Tyr</td>
<td>9.80 (10)</td>
</tr>
<tr>
<td>Phe</td>
<td>6.96 (7)</td>
</tr>
<tr>
<td>Cys</td>
<td>1.08 (1)§</td>
</tr>
<tr>
<td>Trp</td>
<td>2.1 (2)€</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>192</td>
</tr>
</tbody>
</table>

* Average values (or extrapolated values for serine and threonine) from two 24-hr and two 72-hr hydrolysates in 6 M HCl at 110°.
† Seven arginyl residues are present in the 193-residue Fragment A, but the 190- and 192-residue molecules contain only five and six residues of arginine, respectively.
§ Asx includes aspartic acid and asparagine; Glx includes glutamic acid and glutamine.
€ Determined as S-carboxymethylcysteine.
† Determined spectrophotometrically (16).
More than 40 chymotryptic peptides were purified from all regions of Fragment A except residues 13–17 and 160–167. In addition to the usual hydrolysis at tyrosyl, phenylalanyl, leucyl, and tryptophyl residues, there was limited scission at certain glutaminyl, histidyl, lysyl, and aspartyl residues. The last was probably due to the greater lability of aspartyl bonds in acid (15).

Three peptides were isolated after cleavage of Fragment A with BNPS-skatole: Peptide BNPS-1 (residues 1–50), Peptide BNPS-2 (residues 51–153), and Peptide BNPS-3 (residues 154–193). The reaction proceeded in 30–40% yield, which left some Fragment A and two large, overlapping peptides (residues 1–153 and 51–193) which were not isolated in pure form.

The cleavage with BNPS-skatole facilitated location of the second tryptophan in the sequence. Analysis of tryptophan by the method of Edelhoch (16) had indicated the presence of two such residues instead of 3, as previously estimated (7). One of the tryptophans was clearly placed at residue 153, and the other one was known to be in the region of residues 40–51, but its exact location at residue 50 was established by cleavage with BNPS-skatole.

A second problem was solved with the aid of a Beckman 890C sequencer. We were often unable to detect peptides corresponding to the region around residues 160–167 either in the chymotryptic hydrolysate or in hydrolysates of Peptides CNBr-3 or Tm-3. Automatic Edman degradation of Peptide BNPS-3 with the sequencer provided the previously unknown sequence of residues 162–164 and confirmed the sequence of other residues on either side of this region.

It should be noted that the sequence of the NH2-terminal 17 residues of Fragment A is in agreement with the NH2-terminal sequence determined with a sequenator by Michel et al. (8). We identified residue 16 as asparagine which was not identified by these investigators.

**DISCUSSION**

A major point of interest in the sequence is the heterogeneity at the COOH-terminus of Fragment A. The results suggest that trypsin preferentially attacks the intact toxin at any of three closely spaced arginines (residues 190, 192, and 193). There may be other additional basic residues COOH-terminal to residue 193 which also contribute to the formation of Fragment A, but, if present, these residues are lost as peptides or remain in the NH2-terminal region of Fragment B.

The presence of a cluster of residues which are highly sensitive to proteolysis in the native protein suggests that this region may have been evolutionarily designed for cleavage. There is no direct evidence yet that the toxin is cleaved in an animal host or animal cell cultures, but this seems probable from our knowledge of the activity of the toxin in vitro. The toxin strongly inhibits protein synthesis in cell culture, and this activity is expressed in cell-free systems only by certain proteolytic fragments of the toxin, Fragment A being the smallest and most active (1, 2).

Less well understood is the function of the toxin in the bacterial cell in which it is synthesized. The structural gene for the toxin resides on the genome of bacteriophage β or other similar bacteriophages, but its function in phage replication or control is unknown. Conceivably such a function would require hydrolysis (and reduction) and would aid in explaining the evolution of the toxin and its structure-activity relationships.

The roles of the various amino-acid residues in the catalytic activity of Fragment A have not been extensively studied, but some information is available. Optical studies suggest that at least one of the two tryptophan residues interacts with the nicotinamide moiety of bound NAD (7), and other results indicate that Fragment A is inactivated by destruction of one of the tryptophan residues (J. Kandel and R. J. Collier, unpublished data). There are no data yet on which of the two tryptophans is involved. Also there is a report that nitration of a single tyrosine residue by treatment of the toxin with tetranitromethane results in a 75% loss of the enzymatic activity of Fragment A (17).

The integrity of the sulfhydryl group (residue 186) of reduced Fragment A is not required for catalysis. Although reduction of the disulfide bridge linking Fragment A to B is necessary for enzymic activity, the sulfhydryl of Fragment A may be reacted with iodoacetate or N-ethylmaleimide without significantly affecting its enzymatic activity (5, 6). Therefore, when Fragment B is present in disulfide linkage to Fragment A, the active site of Fragment A is probably masked or not formed.

There are a few interesting features with respect to the distribution of residues in the sequence of Fragment A. All seven arginyl residues are present in the region from residue 126 to 193, with three of them present in the last four residues of the molecule (residues 190–193). Also, the region from residue 66 to 122 contains only one aromatic residue despite an average of 13% aromatic residues in the rest of the molecule.

The molecular weight of Fragment A as calculated from the sequence data (21,145 for the isoelectric longest form) is slightly lower than that estimated by electrophoresis in polyacrylamide gels in the presence of dodecyl sulfate (24,000). Sedimentation equilibrium studies have yielded an intermediate value (22,800 ± 600). The basis of these differences is unknown, but the value from the sequence data may probably be considered the most reliable.

Finally, the sequence presented may be of more widespread interest than originally believed. The notion that the ADP-ribosylation of EF-2 is an activity unique to diphtheria toxin was recently dispelled, when it was found that a protein exotoxin produced by *Pseudomonas aeruginosa* has the same activity (18). We have confirmed this report.

We express our appreciation to Mrs. Barbara Kanegsberg, Mrs. Lynn Williams, Miss Dorothy McNall, and Mr. Richard Igarashi for technical assistance and to Mr. Douglas M. Brown for the sequenator run and molecular weight determination by sedimentation equilibrium. This work was supported in part by Grant GM 11061 of the National Institute of General Medical Science of the U.S. Public Health Service, Grant GB 32406 of the National Science Foundation, and Grant AI-07877 from the National Institute of Allergy and Infectious Disease.