The Structures of the H C Fragment of Tetanus Toxin with Carbohydrate Subunit Complexes Provide Insight into Ganglioside Binding*

(Received for publication, August 23, 1999, and in revised form, November 16, 1999)

Paul Emsley‡, Constantina Fotinou‡, Isobel Black§, Neil F. Fairweather§, Ian G. Charles¶, Colin Watts§, Eric Hewitt‡, and Neil W. Isaacs**

From the ²Department of Chemistry, University of Glasgow, Glasgow G12 8QQ, the ⁷Department of Biochemistry, Imperial College of Science, Technology and Medicine, South Kensington, London SW7 2AZ, ³The Cruciform Project, The Rayne Institute, 5 University Street, London WC1E 6JJ, and the ⁷Department of Biochemistry, The University, Dundee DD1 4HN, United Kingdom

The entry of tetanus neurotoxin into neuronal cells proceeds through the initial binding of the toxin to gangliosides on the cell surface. The carboxyl-terminal fragment of the heavy chain of tetanus neurotoxin contains the ganglioside-binding site, which has not yet been fully characterized. The crystal structures of native H C and of H C soaked with carbohydrates reveal a number of binding sites and provide insight into the possible mode of ganglioside binding.

Tetanus toxin (TeNT) ¹ and the botulinum toxins (BoNTs) are members of the family of clostridial neurotoxins, produced by Clostridium tetani and Clostridium botulinum, respectively. These protein toxins are structurally and functionally related, each being synthesized as a 150-kDa single polypeptide, which is processed to give a 50-kDa amino-terminal L chain, disulfide-bonded to a 100-kDa carboxyl-terminal H chain. These toxins cause paralysis by inhibiting release of neurotransmitter from presynaptic nerve terminals. The differences in clinical symptoms that tetanus and botulinum toxins exhibit are due to the distinct sites of action of the toxins. TeNT undergoes retrograde transport from the neuromuscular junction to the central nervous system and targets inhibitory neurons within the spinal cord causing a spastic paralysis. In contrast, BoNTs do not undergo retrograde transport and target peripheral sensory neurons resulting in flacid paralysis.

The mechanism of action of TeNT and BoNTs has been described (1) as a four-stage process: cell binding, vesicular internalization, cytoplasmic translocation, and finally proteolytic cleavage of the neuronal substrate. The 1315 residues of TeNT (2, 3) undergo proteolytic cleavage to yield a 50-kDa amino-terminal L chain, which carries metalloprotease catalytic activity against synaptobrevin (4), an essential component of the exocytosis machinery. The H-chain can be cleaved into two domains: H N and H C. After internalization of the entire toxin into vesicles, the L-chain is translocated into the cytosol, a process apparently requiring the activity of the H N fragment (5, 6).

The H C fragment, also termed fragment C, is the 50-kDa carboxyl-terminal fragment of TeNT (residues 865–1315) and is required for the early stages of intoxication. It has long been recognized that TeNT displays ganglioside binding activity (7). Ganglioside consists of a sialic acid-containing oligosaccharide linked to ceramide. Most have the basic form: Galβ3GalNAcβ4(NeuAcα3)Galβ4GlcβCer to which one or more N-acetylenuraminic (sialic) acids are bound. GM1 and GD1b have mono- and disialic acid residues respectively attached to the internal galactose residue. GT1b and GQ1b have, in addition, mono- and disialic acids attached to the terminal galactose residue.

Binding studies of TeNT to brain membranes and to purified gangliosides show that for the most part both mediated by the H C fragment (8) and found in Ref. 8). Analysis of the binding efficiencies of different gangliosides to both TeNT and the H C fragment has shown an absolute requirement for the disialic acid moiety attached to the internal galactose present in GD1b, GT1b, and GQ1b (10, 11). A single sialic acid residue on this internal gal residue, as found in the monosialic ganglioside GM1, is clearly not sufficient for binding (10, 11). The presence of sialic acid on the terminal Gal residue (as in GT1b and GQ1b) appears to enhance binding only slightly and cannot therefore be considered essential for binding. However, the Galβ3GalNAc disaccharide moiety is necessary for binding (10, 11). Recent studies using surface plasmon resonance to study binding of purified H C fragment have extended these findings and demonstrated a preference for GD1b over GT1b and negligible binding to GQ1b gangliosides (12).

It has long been questioned whether gangliosides represent the sole receptor for TeNT, and much convincing evidence has been presented that argues that a second, protein, receptor exists (1, 9). This putative receptor could determine the specificity of TeNT for certain neuronal cell types and could possibly also be involved in the retrograde transport of TeNT to higher centers of the central nervous system. A two-receptor model, invoking both ganglioside and protein receptors, has been presented (9).

The x-ray crystal structure of H C, at 2.7-Å resolution reported previously (13) shows the protein to have two domains, an amino-terminal jelly roll domain and a carboxyl-terminal β-trefoil domain. A closely homologous structure has been found in the receptor-binding domain of botulinum toxin (14). The amino-terminal jelly roll domain, which is similar in structure to many lectins, is an obvious candidate for the ganglioside bind-
ing of TeNT. However, deletion mutagenesis studies (15, 16) suggest that carboxyl-terminal residues, in particular residues 1306–1310 of the β-trefoil domain, are essential for cell and ganglioside binding activity. The structure of Umland et al. (13) showed that these residues are largely solvent-exposed forming one edge of a shallow pocket. The exception is Val-1306, which is the final residue in a β-strand belonging to the β-trefoil core. A later report (17) of photoaffinity studies using derivatized ganglioside GD1b showed binding to be accompanied by photoactive modification of His-1293, which is located on the other side of the β-trefoil domain from the carboxyl-terminal residues. The apparent discrepancy in these results can be reconciled if the deletion of residues 1306–1309 causes a change in the conformation of the β-trefoil and disrupts the true ganglioside-binding site.

To investigate the structural basis of ganglioside binding we have determined structures from crystals of native Hc and from Hc crystals soaked in solutions of lactose, galactose, N-acetylgalactosamine, and sialic acid. As these sugars are all subunits of ganglioside, it is expected they would bind at locations corresponding to potential ganglioside-binding sites.

**MATERIALS AND METHODS**

Expression and Purification of Hc—DNA from plasmid pETTetac215 (18) was incorporated into the Escherichia coli expression vector pET16b to produce Hc as a fusion protein with a 24-amino acid amino-terminal tag, including 10 His residues. Cells were grown to midlog phase, and expression was induced by the addition of isopropyl-β-D-thiogalactopyranoside to 1 mM for 4 h. Cells were then harvested and
lysed by sonication, and the Hc protein was purified by nickel chelation and gel filtration chromatography.

Crystallization, Data Collection, and Processing—Crystals were produced by vapor diffusion using a sitting drop containing 2 μl of a 3 mg/ml protein solution in 20 mM imidazole (pH 7.0) and 100 mM NaCl, mixed with 3 μl of a well solution, consisting of 200 mM ammonium sulfate with 40% (w/v) PEG 4K and 1% (v/v) 2-methyl-2,4-pentanediol.

Several data sets were collected from native, derivative, and carbohydrate-soaked crystals as detailed in Table I. Crystals were soaked in well solutions containing 100 mM carbohydrate for 24 h (lactose, galactose, sialic acid) and 15 days (N-acetylgalactosamine) before data collection. Diffraction data were collected at the Daresbury synchrotron radiation source under cryocooled conditions (100 K). The lactose- and galactose-soaked crystals used 15% (v/v) glycerol made up with the well solution as the cryoprotectant. The N-acetylgalactosamine- and sialic acid-soaked crystals used 30% (v/v) PEG 400 as the cryoprotectant. Galactose, N-acetylgalactosamine, and sialic acid were purchased from Sigma and lactose from Fluka.
Data from native and galactose- and lactose-soaked crystals were processed using the HKL Suite (19). The native protein crystallizes in space group P2_12_1 with unit cell dimensions $a = 56.7\,\text{Å}, b = 57.0\,\text{Å}, c = 122.2\,\text{Å}$ and one molecule in the asymmetric unit. This cell is about 12% (by volume) smaller than that of the crystals of HC reported previously by us (20) but is still larger than that reported by Umland et al. (13).

The data from the sialic acid- and N-acetylgalactosamine-soaked crystals were processed with Mosflm (21) and Scala (22, 23). Table I shows the data and phasing statistics.

Crystal Structure Determination—The initial mercury heavy atom position was solved using ShelX-90 (24). Cross-phased difference maps were used to find sites in other heavy atom derivative data sets. Initial heavy atom refinement (including anomalous data) proceeded using MLPHARE (23, 25). Subsequent use of SHARP (26) located further minor heavy atom sites. After density modification and phase refinement with DM (27), O (28) was used to build the molecule into the electron density extending in resolution to 2.5 Å. The Cα-trace of the molecule in Ref. 13 facilitated the sequence assignment in the map. This native structure was not fully refined because a structure of similar quality had been published (13), and higher resolution data were obtained with carbohydrate-soaked crystals.

The high resolution (1.8 Å) lactose-soaked data set showed considerable lack of isomorphism with respect to the MIR data sets. Molecular replacement using AMORE (29) was used to reposition the partially refined native structure in the unit cell of the high resolution data set. Further rebuilding and refinement of the lactose-soaked structure proceeded using O, REFMAC (30), and ARP (31). The structure geometry was analyzed using PROCHECK (32). The galactose-, N-acetylgalactosamine-, and sialic acid-soaked structures were refined and validated similarly (Table II). The ligand interactions were displayed using LIGPLOT (33).

RESULTS AND DISCUSSION

The topology of TeNT H$_C$ has already been described (13). It consists of two domains, an amino-terminal jelly roll domain and a carboxyl-terminal $\beta$-trefoil domain linked by a short peptide as shown in Fig. 1. Fig. 2 shows the difference electron density observed for each of the carbohydrate-soaked crystals. Clear electron density shows single sugar-binding sites for the lactose-, galactose-, and sialic acid-soaked crystals (Fig. 2, a, b, and c) and two sites for N-acetylgalactosamine-soaked crystals (Fig. 2, d and e). The locations of these bound carbohydrate ligands on the protein surface are shown in Fig. 5. The structures of the proteins in the four complexes are very similar to...
each other. The most significant difference is a small change in the poorly defined loops 983–985 and 1180–1185, the latter of which moves to accommodate the galactose.

The final models of the lactose-, galactose-, sialic acid-, and N-acetylgalactosamine-soaked crystals consist of 441 residues (875 to 1315 of TeNT) and 692, 352, 226, and 506 water molecules, respectively (Table II). The geometry of the models is adequate, with 87.2% (Lac-soak), 85.5% (Gal-soak), 86.5% (NGA-soak), and 82.5% (sialic acid-soak) of the residues in the most favored areas of the Ramachandran plot as defined by PROCHECK (32). Only one residue in the NGA-soak structure and three in the sialic acid-soak structure are in the disallowed regions.

A comparison of the lactose complex with the structure published by Umland et al. (13) (Protein Data Bank accession code 1AF9) shows that the structure models are largely very similar. There are however a number of regions around residues 943, 984, 1067, 1183, 1220, and 1293 that have tandem differences of greater than 1 Å in the Cα-positions. All these regions are on the surfaces of the domains, and the differences probably reflect large mobility. Excluding these regions, the average deviation in Cα-positions between the two models is only 0.3 Å.

Our studies with carbohydrate soaks do not provide adjacent binding sites for the different carbohydrates. The lactose-soaked crystal shows density for the lactose in the region close to His-1293 and Trp-1289 (Figs. 2a and 3). Hydrogen bonds are formed between the galactose unit and the protein through O-6 and OG Ser-1287, O-6 and OD-2 Asp-1222, and O-4 and carbonyl oxygen of Thr-1270. O-3 forms two water-mediated H-bonds with OH Tyr-1290 and the amide N of Gly-1300. The glucose unit forms two H-bonds between O-3 and O-2 with ND-2 Asn-1220 (Fig. 4). Hydrophobic ring packing interactions are formed between the galactose ring and Trp-1289 (Fig. 3). The glucose unit is oriented with respect to the HC so that a ceramide attached at the C-1 position would be directed away from the protein. In addition, the location of the C-3 of the galactose unit allows space for attached sialic acids.

In the galactose-soaked crystal there is clear density (Fig. 2b) for a galactose molecule near residues 1195, 1179, and 1180. Hydrogen bonds between O-2 and amide NH of Phe-1195, O-3 and both amide NH Tyr-1180 and carbonyl oxygen Phe-1195, O-6, and NE Arg-1179, and O-4 and carbonyl oxygen of Thr-1181 are formed from the galactose to the protein (Fig. 4).

The N-acetylgalactosamine-soaked map contains density corresponding to two separate carbohydrate molecules (Fig. 2, d and e). The observed electron density is of similar quality for

**Fig. 5.** A stereo view, in the same orientation as Fig. 1, of the positions of the carbohydrate units with respect to TeNT HCS. The carbohydrate units bind in four distinct sites, and their positions and orientations make it unlikely that these would correspond to a single ganglioside binding to a single HCS protein.
both molecules. Although the sample of N-acetylglactosamine used in the soak contained a mixture of α- and β-anomers, no significant contribution to the density can be seen from the β-anomer. The NGA1 interactions with HC are between N-2 and OD-1 of Asp-1147, O-3 and OD-2 of Asp-1147, and a water-mediated hydrogen bond between O-4 and carbonyl oxygen of Ile-1275. O-1 forms an H-bond with a sulfate ion that is salt-bridged to Arg-1226 and H-bonded to the amide NH of Asn-1216. The NGA2 interactions are between O-7 and amide NH Ile-1208, O-5 and OH Tyr-1199, and water-mediated interactions between O-1 and the carbonyl oxygen as well as Oε-1 of Glu-1206 (Fig. 4).

The sialic acid-soaked crystals show clear density for a single sialic acid (Fig. 2c). The sialic acid binds in the region of residues 1226 and 1229, adjacent to the position of NGA1. Specifically, the hydrogen bonds formed to the substrate are as follows: O-1 to NH-1 of Arg-1226 and amide NH of Asn-1216 and O-6 to NH-2 of Arg-1226. Water-mediated interactions are formed from O-4 to OD-1 and OD-2 of Asp-1147 and from O-2 to amide NH of Asn-1216 and CO of Asp-1214. The positions of the sialic acid O-1 and O-6 atoms correspond to oxygens of the sulfate ion that forms a salt bridge with Arg-1226 in the NGA-soaked structure.

When these different binding sites are positioned on a common HC molecule (Fig. 5) they are located in four distinct regions (the sialic acid-binding site and the NGA1-binding site are considered a single region). The lactose-binding site appears to be part of the ganglioside-binding site. It is located close to the His-1293 identified from photoaffinity labeling (17) and to other residues identified by mutagenesis studies (16) as a ganglioside-binding site.

Another carbohydrate-binding region has adjacent sites for both NGA and sialic acid. This suggests that this broad region is possibly capable of binding linked carbohydrate units. Crystal-soaking experiments with disaccharide units are being undertaken to investigate this possibility.

The single galactose-binding site occurs in a depression on the protein surface created by a loop from residues 1180 to 1208 (Fig. 4) in the vicinity of Arg and Asp residues, with its apolar β face parallel to an aromatic residue (Tyr). These features are characteristic of galactose-binding sites (34, 35). The galactose is bound in a position where it could mimic a terminal galactose of a ganglioside such as GD1β, because the O-1 is in an open environment.

A comparison of the structures of tetanus toxin HC and botulinum toxin HC (Protein Data Bank accession code 3BTA) shows much similarity in the carbohydrate-binding regions. The strongest similarity is in the lactose-binding site where, with the exception of ND-2 Asn-1220, each of the TeNT HC atoms coordinating the carbohydrate has a structural counterpart in the BoNT HC structure (TeNT Ser-1287 with BoNT Ser-1263, Asp-1222 with Glu-1202, Thr-1270 with Phe-1251, Tyr-1290 with Tyr-1266, Gly-1300 with Gly-1278, Trp-1289 with Trp-1265). These data are in accord with studies on BoNT. Fluorescence quenching experiments by Kamata et al. (36) have been interpreted by Lacy and Stevens (37) to implicate Trp-1265 in the ganglioside-binding site of BoNT. This is supported by neutralizing antibody data from Kubota et al. (38) that implicates residues 1286–1272. Lacy and Stevens (37) find that all these residues are located in a deep, positively charged cleft, which they assign to the ganglioside-binding site in the BoNT structure. In the galactose- and NGA-binding regions the structural coincidences are fewer because of differences in the lengths and positions of loops. It remains to be seen if move-
The Structures of the $H_C$ Fragment of Tetanus Toxin with Carbohydrate Subunit Complexes Provide Insight into Ganglioside Binding

Paul Emsley, Constantina Fotinou, Isobel Black, Neil F. Fairweather, Ian G. Charles, Colin Watts, Eric Hewitt and Neil W. Isaacs

*J. Biol. Chem.* 2000, 275:8889-8894.
doi: 10.1074/jbc.275.12.8889

Access the most updated version of this article at [http://www.jbc.org/content/275/12/8889](http://www.jbc.org/content/275/12/8889)

Alerts:
- When this article is cited
- When a correction for this article is posted

[Click here](http://www.jbc.org/content/275/12/8889.full.html#ref-list-1) to choose from all of JBC's e-mail alerts

This article cites 34 references, 8 of which can be accessed free at [http://www.jbc.org/content/275/12/8889.full.html#ref-list-1](http://www.jbc.org/content/275/12/8889.full.html#ref-list-1)