JOURNAL OF BACTERIOLOGY, Jan. 1986, p. 21–27
0021-9193/86//010021-07$02.00/0

Vol. 165, No. 1

Cloning, Nucleotide Sequencing, and Expression of Tetanus Toxin Fragment C in *Escherichia coli*

NEIL F. FAIRWEATHER,[1]* VALERIE A. LYNESS,[1] DEREK J. PICKARD,[1] GEOFFREY ALLEN,[1] AND ROBERT O. THOMSON[2]

Departments of Molecular Biology[1] and Bacteriology,[2] Wellcome Biotechnology Limited, Beckenham, Kent BR3 3BS, Great Britain

Received 9 August 1985/Accepted 21 October 1985

The amino acid sequence of the first 30 residues of fragment C of tetanus toxin was determined, and a mixture of 32 complementary oligonucleotides, each 17 bases long, was synthesized. A 2-kilobase (kb) *Eco*I fragment of *Clostridium tetani* DNA was identified by Southern blotting and was cloned into the *Escherichia coli* plasmid vector pAT153 with the ^{32}P-labeled oligonucleotide mixture as a probe. A second 3.2-kb *Bgl*II fragment was identified and cloned with the 2-kb *Eco*RI fragment as a probe. The nucleotide sequence of 1.8 kb of this DNA was determined and was shown to encode the entire fragment C and a portion of fragment B of tetanus toxin. The tetanus DNA was expressed in *E. coli* with pWRL507, a plasmid vector containing the *trp* promoter and a portion of the *trpE* gene. The trpE-tetanus fusion proteins were visualized by sodium dodecyl sulfate-polyacrylamide gel electrophoresis and were shown to react with anti-fragment C antibody.

Tetanus toxin is a potent inhibitor of the central nervous system. It causes spastic paralysis by blocking the release of inhibitory transmitters from inhibitory synapses (for a review, see reference 1). The exact mode of action of tetanus toxin at the molecular level remains unknown. The toxin, produced by *Clostridium tetani*, is synthesized as a 150,000-dalton polypeptide (Fig. 1). Upon lysis of the bacterium, the toxin is released from the cells, concomitantly with proteolytic cleavage of the molecule by an endogenous protease. The resulting molecule, termed "extracellular toxin," is composed of two fragments designated the light and heavy chains (13). These chains are held together by one or more disulfide bonds. Purified heavy and light chains are by themselves virtually nontoxic, but when they are reassociated, toxicity is restored.

Papain digesion of tetanus toxin results in cleavage of the heavy chain to give two fragments, B and C (8) (Fig. 1). Purified fragment C is completely nontoxic in animals, whereas fragment B retains some residual toxicity at high doses, although this activity is manifest as a flaccid paralysis in mice rather than the spastic paralysis characteristic of tetanus toxin (7, 8).

Immunity to tetanus toxin is provided by the administration of formaldehyde-treated toxin (tetanus toxoid). Fragments B and C have also been used to successfully immunize animals against tetanus, indicating that the entire molecule is not essential for protection (6). However, the preparation of these fragments is time-consuming and not commercially feasible. Studies with monoclonal antibodies have also shown that all three domains of the tetanus toxin molecule (the light chain, the amino-terminal half of the heavy chain, and the C fragment) may induce neutralizing antibodies (10, 25).

To characterize tetanus toxin further at the molecular level, we undertook the cloning, nucleotide sequencing, and expression in *Escherichia coli* of DNA encoding fragment C of tetanus toxin. We used synthetic oligonucleotides complementary to the amino acid sequence of fragment C to generate hybridization probes for identifying specific *C. tetani* DNA fragments and for screening recombinant clones.

MATERIALS AND METHODS

Bacterial strains. *C. tetani* CN3911, a derivative of the Harvard strain (14), was used as the source of DNA for cloning. *E. coli* K-12 strains DH1 and JM101 have been described previously (12).

***C. tetani* DNA production.** CN3911 was grown for 30 h in 600 ml of Mueller medium (14). The cells were harvested by centrifugation at 5,000 rpm for 10 min with a Sorvall GS3 rotor and suspended in 20 ml of 50 mM Tris hydrochloride (pH 8.0), 5 mM EDTA, and 50 mM NaCl (TES). The cells were harvested by centrifugation at 10,000 rpm for 10 min in a Sorvall SS34 rotor and suspended in 8 ml of TES containing 25% (wt/vol) sucrose. Lysozyme was added to a final concentration of 2 mg/ml, and the cells were incubated at 37°C for 20 min. EDTA (3.2 ml, 0.25 M) was added, and incubation continued at 37°C for 25 min. The cells were lysed by the addition of 7.2 ml of 2% (wt/vol) Sarkosyl in TES followed by incubation for 10 min at 37°C and for 10 min at 4°C. Protease K was added to 10 mg/ml, and the lysate was left overnight at 50°C.

Lysate (8 ml) was added to 35 ml of solution containing 69.6 g of CsCl and 55.2 ml of 10 mM Tris hydrochloride (pH 8.0)–1 mM EDTA. Polymethylsulfonyl fluoride was added to 50 μg/ml, and the solution was centrifuged at 36,000 rpm for 48 h at 20°C in a Beckman 70.1 Ti rotor. The *C. tetani* DNA, which was visible as opaque lumps in the clear solution, was withdrawn from the gradient with a wide-bore syringe and was dialyzed extensively against 10 mM Tris hydrochloride (pH 8.0)–1 mM EDTA (TE). The DNA was extracted once with phenol and three times with ether, followed by precipitation with ethanol at −20°C. The DNA was resuspended in TE to 1 mg/ml and stored at −20°C.

Preparation of crystalline fragment C. Tetanus toxin was produced by using 1 M NaCl to lyse organisms obtained

* Corresponding author.

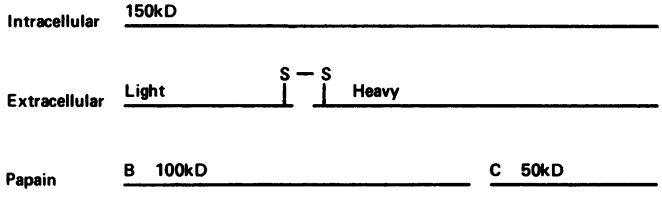

FIG. 1. Structure of tetanus toxin and fragments generated by proteolytic cleavage (adapted from reference 3). kd, Kilodaltons.

from a 3-day culture of Harvard strain of *C. tetani* grown in modified Mueller medium. The toxin was purified by fractional precipitation with potassium phosphate (1.75 M, pH 8.0) followed by adsorption to DEAE-cellulose equilibrated with 0.01 M sodium phosphate (pH 7.2)–0.2 M EDTA. The purified fraction was equilibrated by dialysis against 0.1 M sodium phosphate (pH 6.5)–0.1 M EDTA. Fragment C was prepared from this material by the procedure of Helting and Zwisler (8) by using crystalline papain derived from crude enzyme by the method of Kimmel and Smith (11).

Crystalline fragment C was obtained by concentration by vacuum dialysis of either the appropriate fraction obtained from the Sephadex G-100 column or the digestion mixture, followed by dialysis against water. Several batches were obtained on further dialysis. Recrystallization was achieved by dissolution of fragment C in 0.5 M sodium chloride, followed by dialysis against water. It was found that crystallization reduced the residual toxicity of fragment C but did not eliminate it altogether even after 12 recrystallizations. It was always necessary to further purify fragment C with a column of adsorbed tetanus antitoxin coupled to Sepharose 4B as described by Helting and Zwisler (8). This column adsorbed all the contaminating tetanus toxin together with a small percentage of fragment C. The flowthrough was pure fragment C and was finally dialyzed as a crystalline suspension and freeze-dried.

DNA techniques. Preparation of plasmid DNA and transformation were carried out as described by Maniatis et al. (12). Restriction enzymes, T4 ligase, Bal31 nuclease, and polynucleotide kinase were obtained from Boehringer Corp., London, and were used according to the instructions of the manufacturer. S1 nuclease was from New England Nuclear Corp., Boston, Mass.

Southern blotting and DNA hybridization. Restriction endonuclease-cleaved DNA was transferred onto nitrocellulose filters as described by Southern (20). Labeling of the mixed oligonucleotide with [^{32}P]ATP and hybridization to immobilized DNA on filters was carried out as described by Wallace et al. (26). Filters were prehybridized in 6× NET (1× NET is 0.15 M NaCl–0.015 M Tris hydrochloride [pH 7.5]–1 mM EDTA), 5× Denhardt solution, 0.5% sodium dodecyl sulfate (SDS), and 200 μg of salmon sperm DNA per ml at 55°C for 2 h. Hybridization was in the same buffer except that transfer RNA (100 μg/ml) was used instead of salmon sperm DNA. ^{32}P-labeled oligonucleotide mixture (20 ng; 3 × 10^7 cpm) was added, and hybridization was carried out at 37°C overnight. The filters were washed in 6× SSC (1× SSC is 0.15 M sodium chloride plus 0.015 M trisodium citrate) at room temperature, blotted dry, and exposed to Kodak X-ray film.

Expression of trpE-tetanus fusion proteins. Recombinant plasmids expressing trpE-tetanus toxin fusion proteins were constructed with the plasmid vector pWRL507 (a gift from M. Winther). pWRL 507 (see Fig. 5) is a derivative of pATtrp (18) and contains the tryptophan promoter and a portion of *trpE*, the structural gene for anthranilate synthetase (27). To express fusion proteins, *E. coli* DH1 containing recombinant plasmids was grown overnight at 37°C in 10 ml of M9 medium (12) containing 0.25% (wt/vol) Casamino Acids (Difco Laboratories, Detroit, Mich.) and 50 μg of ampicillin per ml. The cultures were diluted 1 in 5 into fresh medium containing 10 μg of indolacrylic acid per ml and were grown for 4 h with shaking. The cells were harvested by centrifugation at 10,000 rpm for 5 min, and the polypeptides were visualized by SDS-polyacrylamide gel electrophoresis as described elsewhere (4). Western blotting was carried out as described previously (22), with 3% (wt/vol) hemoglobin to block nonspecific binding to nitrocellulose. Anti-fragment C antibody was used at a dilution of 1 in 50, and proteins were visualized with 50,000 cpm of [^{125}I]protein A per ml.

Oligonucleotide synthesis. A set of oligonucleotides, comprising all possible nucleotides that could code for amino acids 6 through 11 of fragment C, was synthesized by a manual solid-phase method similar to that described by Sproat and Banwarth (21).

Determination of amino acid sequence of tetanus toxin fragment C. Fragment C (0.5 mg) was taken up in 300 μl of trifluoroacetic acid, and the amino acid sequence was determined in the presence of 3 mg of Polybrene and 100 μg of glycylglycine (Pierce Warriner Chemicals) with a Beckman 590C sequencer with a Sequimat SC510 controller and a P6 autoconvertor. The first 30 residues were identified by high-pressure liquid chromatography.

Containment. Part of this work was carried out under category II containment facilities with appropriate host-vector combinations as advised by the Genetic Manipulation Advisory Group.

RESULTS

Design of synthetic oligonucleotides. The sequence of the first 30 amino acids of purified fragment C was obtained by automated Edman degradation. The sequence obtained was Lys-Asn-Leu-Asp-Cys-Trp-Val-Asp-Asn-Glu-Glu-Asp-Ile-AsP-Val-Ile-Leu-Lys-Lys-Ser-Thr-Ile-Leu-Asn-Leu-Asp-Ile-Asn-Asn-Asp. This sequence was analyzed for the longest possible stretch of amino acids giving the least-degenerate oligonucleotide mixture. Figure 2 shows the sequence of residues 6 through 11 and a family of 32 oligonucleotides, each 17 bases long. One of these oligonucleotides would be expected to be complementary to the DNA and mRNA encoding fragment C. This family of oligonucleotides was synthesized with mixtures of bases included where indicated.

Identification and cloning of *C. tetani* DNA fragments. Total

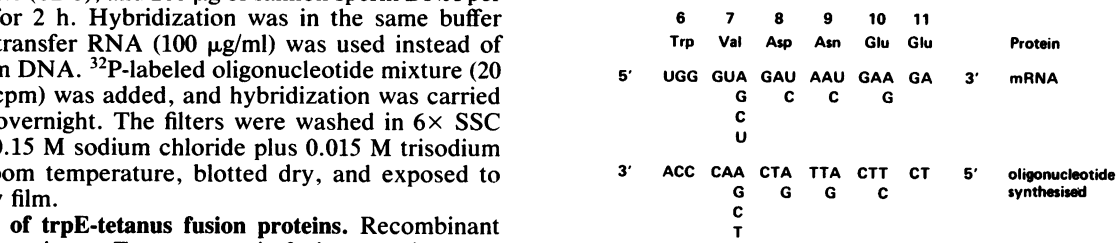

FIG. 2. Amino acid sequences of residues 6 through 11 of fragment C of tetanus toxin (top), the encoding mRNA including degenerate bases where indicated (middle), and the mixed oligonucleotide used as a probe (bottom).

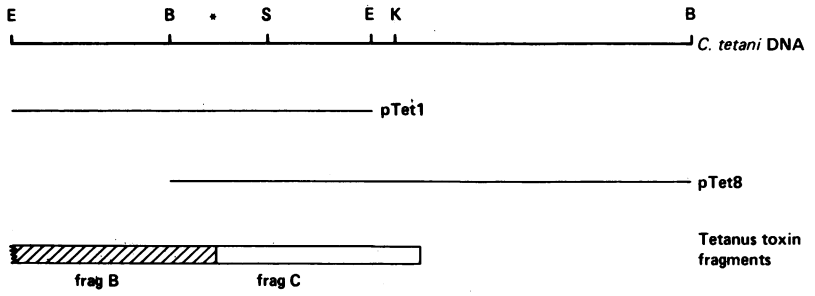

FIG. 3. Restriction map of the cloned *C. tetani* DNA in plasmids pTet1 and pTet8. The top line is a map of *C. tetani* DNA and is derived from digests of both pTet1 and pTet8 (see text for details of their construction). Only the inserts of the plasmids are shown. The location of the fragments of tetanus toxin encoded by the DNA is shown in the bottom line. Restriction endonuclease sites are E, *Eco*RI; B, *Bgl*II; S, *Sac*II; and K, *Kpn*I. The open box indicates the location of fragment C, and the hatched box indicates a portion of fragment B. *, Site of binding of the synthetic oligonucleotide.

cellular DNA from *C. tetani* was digested with several restriction enzymes, transferred to nitrocellulose, and hybridized with the oligonucleotide mixture labeled with [^{32}P]ATP. A prominent band of 2 kilobases (kb) was identified in the *Eco*RI digest, and fainter, higher-molecular-weight bands were identified with *Pst*I, *Kpn*I, and *Hin*dIII (data not shown). *C. tetani* DNA (100 μg) was cleaved with *Eco*RI and electrophoresed on a 0.7% agarose gel, and the DNA was purified from the region containing fragments of approximately 2 kb. This pool of 2-kb *Eco*RI fragments was cloned into *Eco*RI-cleaved and dephosphorylated plasmid pAT153, a nonmobilizable derivative of pBR322 (23).

One hundred recombinant clones were picked and were screened by colony hybridization to the mixed oligonucleotide probe. Seven clones were identified showing various degrees of reaction with the probe (data not shown). Plasmid DNA was prepared from these seven clones, and *Eco*RI digests were probed by Southern blotting. Four plasmids had 2.0-kb *Eco*RI fragments which hybridized to the mixed oligonucleotide, and one of these plasmids, named pTet1, was studied further. A restriction map of the insert in pTet1 is presented in Fig. 3. Further Southern blot experiments showed that only fragments containing the central 300 base pairs of the 2-kb insert hybridized to the probe. Thus the sequence encoding the amino terminus of fragment C was located toward the center of the 2-kb *Eco*RI fragment.

DNA sequence analysis (see below) revealed that the 2-kb *Eco*RI fragment in pTet1 did not contain the entire coding region of fragment C. Therefore, adjacent restriction enzyme-generated fragments were identified by Southern blotting with pTet1 as a probe (data not shown). A 3.2-kb *Bgl*II *C. tetani* fragment which hybridized to pTet1 was identified and cloned into the vector pWRL507 at the *Bgl*II site to generate pTet8. This fragment contains the 1.4-kb *Bgl*II-*Eco*RI fragment present in pTet1 and the adjacent 1.8-kb *Eco*RI-*Bgl*II fragment (Fig. 3).

DNA sequence of the cloned *C. tetani* DNA. The nucleotide sequence of the 1.4-kb *Bgl*II-*Eco*RI fragment of pTet1 and of the adjacent 300 base pairs of pTet8 was determined by the method of Sanger et al. (17, 19). The DNA sequence is shown in Fig. 4. Translation of the entire sequence in both strands revealed only one open reading frame encoding a protein of 63,000 daltons. Translation of nucleotides 367 through 457 gives an amino acid sequence identical to that determined for the first 30 amino acids of fragment C, confirming that the clones do indeed encode a portion of tetanus toxin. The amino-terminal residue of fragment C is lysine, which is in agreement with the results of Neubauer and Helting (15). The calculated molecular weight of fragment C is 51,562, a value in close agreement with our own measurements (data not shown) and with those of Helting and Zwisler (8).

Expression of *C. tetani* DNA in *E. coli*. DNA fragments encoding all or part of fragment C were expressed in *E. coli* as fusion proteins using the expression vector pWRL507 (Fig. 5A), which contains the *E. coli trp* promoter and part of the *trpE* gene. The fusion proteins expressed would be expected to consist of amino-terminal residues of the *trpE* product (anthranilate synthetase) and carboxy-terminal residues of tetanus toxin. The 1.4-kb *Bgl*II-*Eco*RI fragment of pTet1 was cloned into the *Bgl*II-*Eco*RI sites of PWRL507 to generate pTet4. Because the *Bgl*II sites in pTet1 and pWRL507 are not in phase, no fusion protein containing tetanus toxin sequences would be obtained (Fig. 5C).

To generate fusion proteins in the same reading frame, two approaches were used. First, pTet4 was cut with *Bgl*II and digested with exonuclease Bal31 for various times. After ligation and transformation into *E. coli* DH1, 100 colonies were picked and analyzed by a solid-phase immune screen for the presence of induced proteins reacting with anti-fragment C antibody. One clone containing a plasmid designated pTet6 was identified as reacting strongly with antibody (data not shown). The fusion protein encoded by pTet6 was visualized by SDS-polyacrylamide gel electrophoresis as a stained band (Fig. 6A, track 1) and by Western blotting (Fig. 6B, track 1). The size of the fusion protein, with a molecular weight of 70,000, is consistent with a deletion in pTet4 of approximately 300 base pairs, which was confirmed by restriction mapping (data not shown). DNA sequence analysis showed the exact size of the deletion to be 400 base pairs, the fusion protein being encoded by nucleotides 163 through 1,128 of *trpE* (18) and nucleotides 399 through 1,446 of tetanus DNA (Fig. 5C). The structures of pTet4 and pTet6 ar shown in Fig. 5B. The 70,000 molecular weight of the protein produced by pTet6 is close to the expected value of 77,000 molecular weight (Fig. 5B). The phase around the *Bgl*II side of pTet4 was altered in another way by using nuclease S1. S1 preferentially degrades single-stranded DNA and so should digest away the sticky ends of the *Bgl*II fragment to generate blunt ends. Examination of the sequence around the *Bgl*II site of pTet4 shows that S1 treatment, i.e., removal of 4 bases, followed by ligation should place the *trpE* and tetanus sequences in the same reading frame and should result in a trpE-tetanus fusion protein (Fig. 5C). pTet4 was treated with *Bgl*II, S1 nuclease, and T4 ligase and was transformed into DH1. One transformant, contain-

```
                    15                  30                  45                  60
AGA TCT TTA GAA TAT CAA GTA GAT GCA ATA AAA AAA ATA ATA GAC TAT GAA TAT AAA ATA
ARG SER LEU GLU TYR GLN VAL ASP ALA ILE LYS LYS ILE ILE ASP TYR GLU TYR LYS ILE

                    75                  90                  105                 120
TAT TCA GGA CCT GAT AAG GAA CAA ATT GCC GAC GAA ATT AAT AAT CTG AAA AAC AAA CTT
TYR SER GLY PRO ASP LYS GLU GLN ILE ALA ASP GLU ILE ASN ASN LEU LYS ASN LYS LEU

                    135                 150                 165                 180
GAA GAA AAG GCT AAT AAA GCA ATG ATA AAC ATA AAT ATA TTT ATG AGG GAA AGT TCT AGA
GLU GLU LYS ALA ASN LYS ALA MET ILE ASN ILE ASN ILE PHE MET ARG GLU SER SER ARG

                    195                 210                 225                 240
TCA TTT TTA GTT AAT CAA ATG ATT AAC GAA GCT AAA AAG CAG TTA TTA GAG TTT GAT ACT
SER PHE LEU VAL ASN GLN MET ILE ASN GLU ALA LYS LYS GLN LEU LEU GLU PHE ASP THR

                    255                 270                 285                 300
CAA AGC AAA AAT ATT TTA ATG CAG TAT ATA AAA GCA AAT TCT AAA TTT ATA GGT ATA ACT
GLN SER LYS ASN ILE LEU MET GLN TYR ILE LYS ALA ASN SER LYS PHE ILE GLY ILE THR

                    315                 330                 345                 360
GAA CTA AAA AAA TTA GAA TCA AAA ATA AAC AAA GTT TTT TCA ACA CCA ATT CCA TTT TCT
GLU LEU LYS LYS LEU GLU SER LYS ILE ASN LYS VAL PHE SER THR PRO ILE PRO PHE SER

                  B|C
                    375                 390                 405                 420
TAT TCT|AAA AAT CTG GAT TGT TGG GTT GAT AAT GAA GAA GAT ATA GAT GTT ATA TTA AAA
TYR SER|LYS ASN LEU ASP CYS TRP VAL ASP ASN GLU GLU ASP ILE ASP VAL ILE LEU LYS

                    435                 450                 465                 480
AAG AGT ACA ATT TTA AAT TTA GAT ATT AAT AAT GAT ATT ATA TCA GAT ATA TCT GGG TTT
LYS SER THR ILE LEU ASN LEU ASP ILE ASN ASN ASP ILE ILE SER ASP ILE SER GLY PHE

                    495                 510                 525                 540
AAT TCA TCT GTA ATA ACA TAT CCA GAT GCT CAA TTG GTG CCC GGA ATA AAT GGC AAA GCA
ASN SER SER VAL ILE THR TYR PRO ASP ALA GLN LEU VAL PRO GLY ILE ASN GLY LYS ALA

                    555                 570                 585                 600
ATA CAT TTA GTA AAC AAT GAA TCT TCT GAA GTT ATA GTG CAT AAA GCT ATG GAT ATT GAA
ILE HIS LEU VAL ASN ASN GLU SER SER GLU VAL ILE VAL HIS LYS ALA MET ASP ILE GLU

                    615                 630                 645                 660
TAT AAT GAT ATG TTT AAT AAT TTT ACC GTT AGC TTT TGG TTG AGG GTT CCT AAA GTA TCT
TYR ASN ASP MET PHE ASN ASN PHE THR VAL SER PHE TRP LEU ARG VAL PRO LYS VAL SER

                    675                 690                 705                 720
GCT AGT CAT TTA GAA CAA TAT GGC ACA AAT GAG TAT TCA ATA ATT AGC TCT ATG AAA AAA
ALA SER HIS LEU GLU GLN TYR GLY THR ASN GLU TYR SER ILE ILE SER SER MET LYS LYS

                    735                 750                 765                 780
CAT AGT CTA TCA ATA GGA TCT GGT TGG AGT GTA TCA CTT AAA GGT AAT AAC TTA ATA TGG
HIS SER LEU SER ILE GLY SER GLY TRP SER VAL SER LEU LYS GLY ASN ASN LEU ILE TRP

                    795                 810                 825                 840
ACT TTA AAA GAT TCC GCG GGA GAA GTT AGA CAA ATA ACT TTT AGG GAT TTA CCT GAT AAA
THR LEU LYS ASP SER ALA GLY GLU VAL ARG GLN ILE THR PHE ARG ASP LEU PRO ASP LYS

                    855                 870                 885                 900
TTT AAT GCT TAT TTA GCA AAT AAA TGG GTT TTT ATA ACT ATT ACT AAT GAT AGA TTA TCT
PHE ASN ALA TYR LEU ALA ASN LYS TRP VAL PHE ILE THR ILE THR ASN ASP ARG LEU SER

                    915                 930                 945                 960
TCT GCT AAT TTG TAT ATA AAT GGA GTA CTT ATG GGA AGT GCA GAA ATT ACT GGT TTA GGA
SER ALA ASN LEU TYR ILE ASN GLY VAL LEU MET GLY SER ALA GLU ILE THR GLY LEU GLY

                    975                 990                 1005                1020
GCT ATT AGA GAG GAT AAT AAT ATA ACA TTA AAA CTA GAT AGA TGT AAT AAT AAT AAT CAA
ALA ILE ARG GLU ASP ASN ASN ILE THR LEU LYS LEU ASP ARG CYS ASN ASN ASN ASN GLN

                    1035                1050                1065                1080
TAC GTT TCT ATT GAT AAA TTT AGG ATA TTT TGC AAA GCA TTA AAT CCA AAA GAG ATT GAA
TYR VAL SER ILE ASP LYS PHE ARG ILE PHE CYS LYS ALA LEU ASN PRO LYS GLU ILE GLU

                    1095                1110                1125                1140
AAA TTA TAC ACA AGT TAT TTA TCT ATA ACC TTT TTA AGA GAC TTC TGG GGA AAC CCT TTA
LYS LEU TYR THR SER TYR LEU SER ILE THR PHE LEU ARG ASP PHE TRP GLY ASN PRO LEU

                    1155                1170                1185                1200
CGA TAT GAT ACA GAA TAT TAT TTA ATA CCA GTA GCT TCT AGT TCT AAA GAT GTT CAA TTG
ARG TYR ASP THR GLU TYR TYR LEU ILE PRO VAL ALA SER SER SER LYS ASP VAL GLN LEU

                    1215                1230                1245                1260
AAA AAT ATA ACA GAT TAT ATG TAT TTG ACA AAT GCG CCA TCG TAT ACT AAC GGA AAA TTG
LYS ASN ILE THR ASP TYR MET TYR LEU THR ASN ALA PRO SER TYR THR ASN GLY LYS LEU

                    1275                1290                1305                1320
AAT ATA TAT TAT AGA AGG TTA TAT AAT GGA CTA AAA TTT ATT ATA AAA AGA TAT ACA CCT
ASN ILE TYR TYR ARG ARG LEU TYR ASN GLY LEU LYS PHE ILE ILE LYS ARG TYR THR PRO

                    1335                1350                1365                1380
AAT AAT GAA ATA GAT TCT TTT GTT AAA TCA GGT GAT TTT ATT AAA TTA TAT GTA TCA TAT
ASN ASN GLU ILE ASP SER PHE VAL LYS SER GLY ASP PHE ILE LYS LEU TYR VAL SER TYR

                    1395                1410                1425                1440
AAC AAT AAT GAG CAC ATT GTA GGT TAT CCG AAA GAT GGA AAT GCC TTT AAT AAT CTT GAT
ASN ASN ASN GLU HIS ILE VAL GLY TYR PRO LYS ASP GLY ASN ALA PHE ASN ASN LEU ASP

                    1455                1470                1485                1500
AGA ATT CTA AGA GTA GGT TAT AAT GCC CCA GGT ATC CCT CTT TAT AAA AAA ATG GAA GCA
ARG ILE LEU ARG VAL GLY TYR ASN ALA PRO GLY ILE PRO LEU TYR LYS LYS MET GLU ALA

                    1515                1530                1545                1560
GTA AAA TTG CGT GAT TTA AAA ACC TAT TCT GTA CAA CTT AAA TTA TAT GAT GAT AAA AAT
VAL LYS LEU ARG ASP LEU LYS THR TYR SER VAL GLN LEU LYS LEU TYR ASP ASP LYS ASN

                    1575                1590                1605                1620
GCA TCT TTA GGA CTA GTA GGT ACC CAT AAT GGT CAA ATA GGC AAC GAT CCA AAT AGG GAT
ALA SER LEU GLY LEU VAL GLY THR HIS ASN GLY GLN ILE GLY ASN ASP PRO ASN ARG ASP

                    1635                1650                1665                1680
ATA TTA ATT GCA AGC AAC TGG TAC TTT AAT CAT TTA AAA GAT AAA ATT TTA GGA TGT GAT
ILE LEU ILE ALA SER ASN TRP TYR PHE ASN HIS LEU LYS ASP LYS ILE LEU GLY CYS ASP

                    1695                1710                1725                1740
TGG TAC TTT GTA CCT ACA GAT GAA GGA TGG ACA AAT GAT TAA ACA GAT TGA TAT GTT CAT
TRP TYR PHE VAL PRO THR ASP GLU GLY TRP THR ASN ASP £££

                    1755                1770                1785
GAT TAC TCT ATA TAA AAA ATT AAA TAA TAT AAC AAT CTA GCT ATA TTA TTT
```

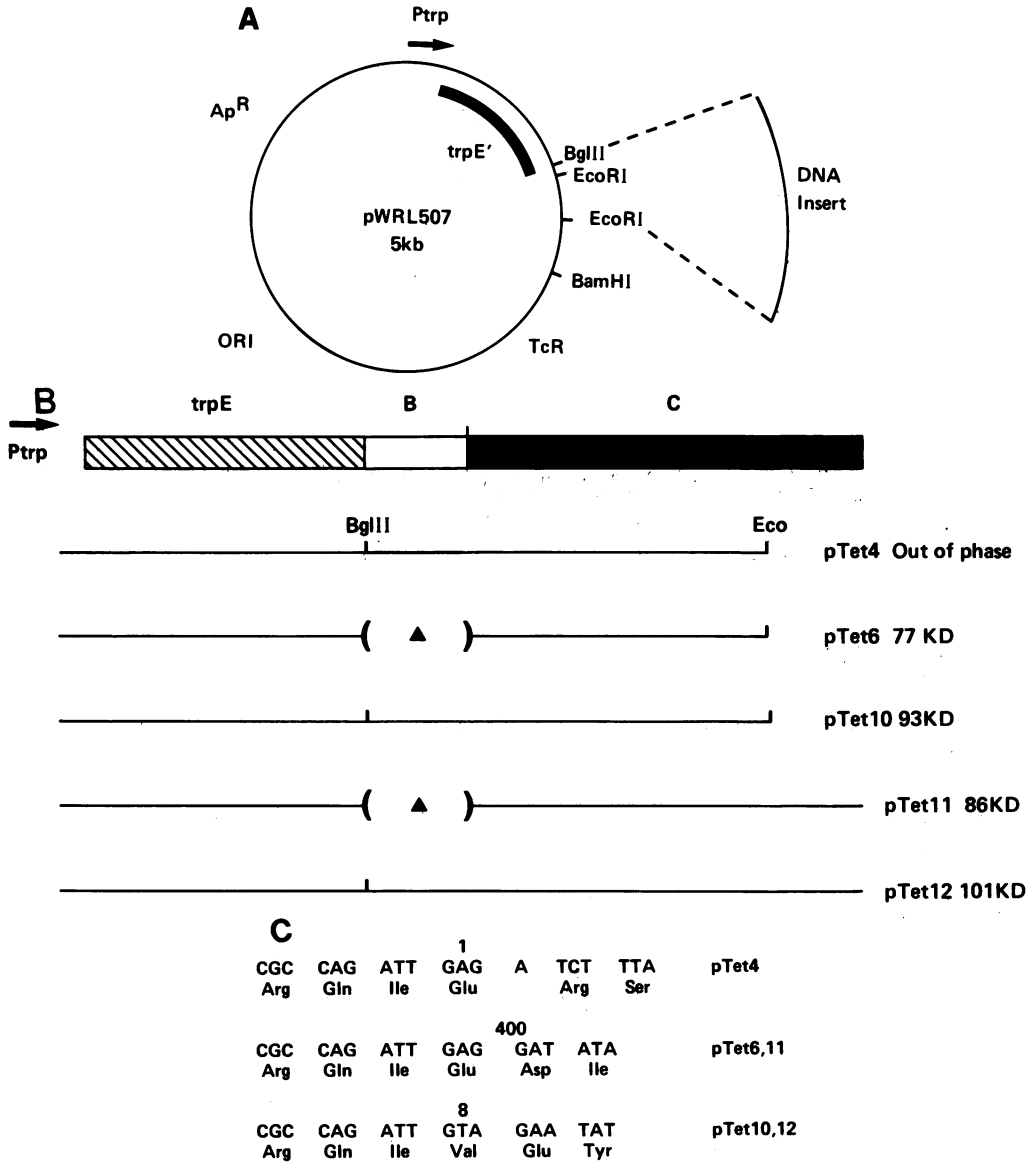

FIG. 5. Plasmids expressing the trpE-tetanus fusion proteins. (A) Restriction map of pWRL507 showing the region at the end of the *trpE* gene into which tetanus DNA was cloned to construct expression plasmids. (B) Partial restriction maps of plasmids pTet4, pTet6, pTet10, pTet11, and pTet12, showing the regions of trpE and tetanus proteins which are encoded. The top line indicates the order of the proteins expressed in the recombinant plasmids below. The hatched box represents the amino-terminal portion of the trpE protein, the open box represents the carboxy residues of tetanus toxin fragment B, and the black portion represents tetanus fragment C. The restriction maps show the extent of the DNA present in the plasmids. The calculated sizes of the fusion proteins (in kilodaltons [KD]) are indicated alongside each plasmid except pTet4, which does not produce a protein in phase. (▲) represents the deletion generated by Bal31 (see text). The deletion of the *Bgl*II sites in pTet10 and pTet12 is indicated by a short line at the former position of the site. (C) Nucleotide sequences around the junctions of the plasmids pTet4, pTet6, pTet10, pTet11, and pTet12. The numbers above the bases of the tetanus DNA correspond to those in Fig. 4. DNA sequencing was carried out on appropriate fragments from pTet6 and pTet10.

ing a plasmid, pTet10, which lacked a *Bgl*II site, was examined for production of a hybrid protein upon induction. Figure 6A shows a band of molecular weight 92,000 produced by pTet10 (track 2) which reacts with anti-fragment C antibody (Fig. 6B, track 2). DNA sequencing of the junction of *trpE* and tetanus DNA in pTet10 revealed that the *trpE* and tetanus DNA were now in phase as expected (Fig. 5C), although the S1 treatment had actually degraded 3 more base pairs than expected.

The DNA in pTet4, pTet6, and pTet10 does not encode the

FIG. 4. Nucleotide sequence of 1.75 kb of *C. tetani* DNA encoding fragment C and a portion of fragment B of tetanus toxin. The coding strand of the DNA is presented in the 5′ to 3′ direction along with the deduced amino sequence of the only open reading frame. The junction between fragments B and C is indicated at nucleotide 366.

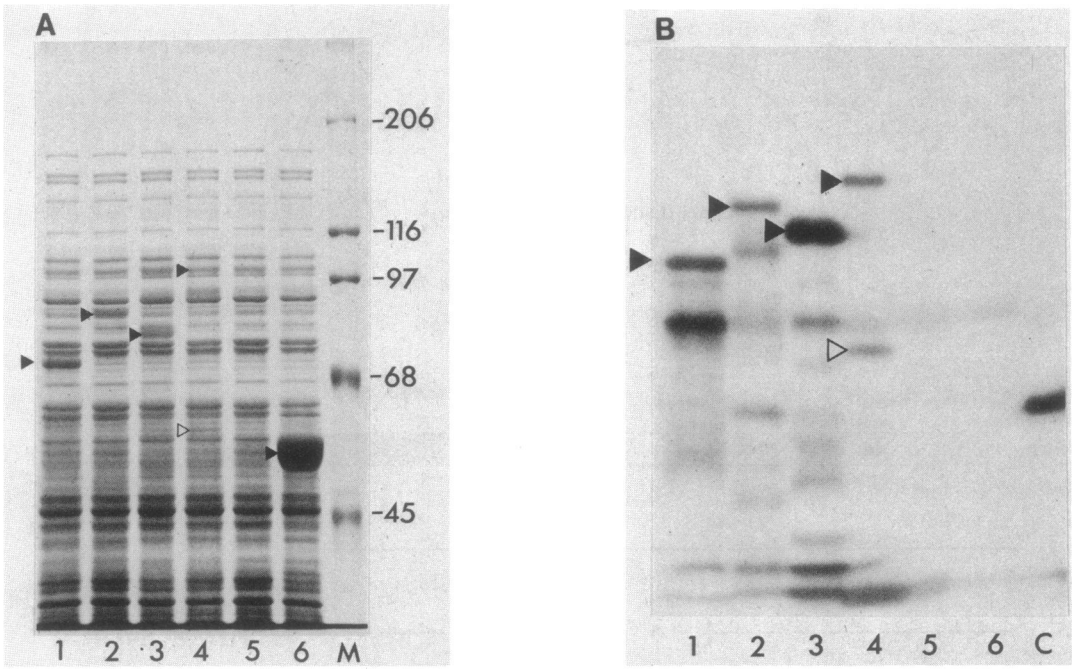

FIG. 6. SDS-polyacrylamide gel electrophoresis of *E. coli* DH1 carrying plasmids expressing the trypE-tetanus fusion proteins. Strains were grown in minimal medium overnight, diluted 1 in 5 in fresh medium containing 10 μg of indoylacrylic acid per ml, and grown for 4 h. The cells were lysed and run on a 7.5% SDS-polyacrylamide gel as described in Materials and Methods. Track 1, pTet6; track 2, pTet10; track 3, pTet11; track 4, pTet12; track 5, pAT153; track 6, pATtrp. C, 1 μg of tetanus toxin fragment C. The positions of the mature forms of the induced proteins are indicated (▶). The position of one of the degradation proteins of pTet12 is also marked (▷). (A) Coomassie blue-stained gel. M, Molecular weight markers (10^6). (B) Autoradiograph of Western blot.

entire C fragment because these plasmids are derived from pTet1, which contains DNA only up to the *Eco*RI site which lies within the coding sequence for fragment C (Fig.3). Derivatives of pTet6 and pTet10 which do contain the DNA encoding the entire C fragment were constructed as follows. pTet8 was cleaved with restriction enzymes *Pst*I and *Sac*II, and the 5.35-kb fragment was purified. This fragment encodes the end of the C fragment and contains noncoding tetanus DNA and a portion of the vector (Fig. 3). Similarly, pTet6 and pTet10 were cleaved with *Pst*I and *Sac*II, and fragments of 2.35 and 2.75 kb were purified. These fragments encode the *trpE* and tetanus DNA 5′ of the *Sac*II site and the portion of the vector missing from the pTet8 fragment. The *Pst*I-*Sac*II fragments of pTet8 and pTet6 were ligated and transformed into *E. coli* DH1 to form pTet11 (7.7 kb). Similarly, the *Pst*I-*Sac*II fragments of pTet8 and pTet10 were ligated to form pTet12 (8.1 kb). Partial restriction maps of pTet11 and pTet12 are presented in Fig. 5B. *E. coli* DH1 carrying these plasmids produced fusion proteins of 78,000 and 98,000 molecular weights, respectively (Fig. 6A, tracks 3 and 4), which cross-reacted with anti-fragment C antibody (Fig. 6B, tracks 3 and 4). These sizes are close to the expected values of 86,000 and 101,000 for molecular weight (Fig. 5B) and are consistent with utilization of the TAA termination codon at nucleotide 1,720 (Fig. 4).

Western blot analysis showed that in all cases, expression of the fusion proteins was accompanied by partial degradation of the mature forms. Thus, additional bands of low molecular weight are seen. Indeed, in the case of pTet12, degradation is sometimes so great that the breakdown products can be seen on a stained gel (Fig. 6A, track 4). This degradation may be another reason for the low amounts of the proteins synthesized (see Discussion).

DISCUSSION

We describe here the cloning and characterization of *C. tetani* DNA encoding the entire fragment C and a portion of fragment B of tetanus toxin. The following two criteria are used to confirm the identity of the cloned DNA: (i) nucleotide sequence of the cloned DNA which encodes exactly the first 30 amino acids of fragment C and (ii) expression of the cloned DNA in *E. coli* as fusion proteins which react with the anti-fragment C antibody. The predicted molecular weight of fragment C from the nucleotide sequence is 51,562, which is in close agreement with molecular weight estimates of 45,000 to 50,000 obtained by SDS-polyacrylamide gel electrophoresis (8; unpublished results).

The cloned DNA was expressed in *E. coli* to give four different fusion proteins, each containing different amounts of tetanus toxin sequences. The levels of each fusion protein produced were not identical; plasmid pTet12 gave lower levels than pTet10 and pTet11, while pTet6 seemed to give the highest levels. The levels of fusion proteins obtained are low compared with that obtained for the intact trpE protein (Fig. 6A). This difference could be due to the presence of more tryptophan residues in the fusion proteins than in the *trpE* product. As expression in this system is under the control of the tryptophan operon regulatory sequences, induction of expression is achieved by addition of the inducer indoylacrylic acid during a period of tryptophan starvation. It has been observed previously that under this control system, the expression of fusion proteins containing many tryptophan residues is consistently reduced compared with those having few or no tryptophan residues (16). The *trpE* product, anthranilate synthetase, has only two tryptophan residues and may obviously be expressed at very

high levels. pTet6 has five tryptophan residues, while pTet10, pTet11, and pTet12 have 6, 8, and 9 tryptophan residues, respectively. This difference may explain the higher levels of the pTet6 protein and the lower levels of the pTet12 protein. Other reasons for poorer expression of fusion proteins may be the instability of these proteins or nonoptimal codon usage specified by the foreign DNA. Instability of all of the proteins was seen by Western blotting, when the degradation products were visualized.

Evidence has been presented showing that the structural gene for tetanus toxin is encoded by a high-molecular-weight plasmid (5). We have been unable to show the presence of such a plasmid in several *C. tetani* strains, including strain CN3911, the one used in this study (unpublished observations). However, we do not rule out the possibility that our strains do harbor a plasmid and that our methods so far have been unable to detect its presence. Our purification method for *C. tetani* chromosomal DNA would not exclude high-molecular-weight plasmid DNA, and it is possible that our cloned DNA is plasmid derived. We intend to reexamine our strains for the presence of plasmid DNA by using our cloned DNA as a probe.

It has been suggested that because of the gross similarity in structure to other toxins which have an intracellular site of action, i.e., a dichain structure linked by disulfide bonds, the tetanus toxin may have a structure-function relationship similar to that of other toxins (2, 24). Thus, by analogy, one might expect the amino-terminal region to contain the toxic determinant and the heavy chain to be involved in the binding of the toxin to the cell membrane and to facilitate the transfer of a toxic fragment located in the amino-terminal region across the membrane to its site of action. Indeed, fragment C has been shown to specifically bind G Dlb and G Tlb gangliosides (9, 24), and experimental evidence suggests that the tetanus receptor may be either a ganglioside or a glycoprotein with gangliosidelike polysaccharide structures (for a review, see reference 3).

The availability of cloned DNA encoding fragment C will facilitate the characterization of the binding of gangliosides to this fragment. In addition, the binding of antibodies, including monoclonal antibodies (10, 25), to different fragments of tetanus toxin may be studied with fragments of defined length constructed by genetic manipulation. We are currently cloning and sequencing fragment B of tetanus toxin. The availability of the entire sequence of tetanus toxin will allow a detailed study of the toxin and may contribute to a fuller understanding of the mode of action of tetanus toxin at the molecular level.

ACKNOWLEDGMENTS

We thank D. Stone and S. Patterson for amino acid sequence analysis and G. Dougan for advice and encouragement.

LITERATURE CITED

1. **Bizzini, B.** 1979. Tetanus toxin. Microbiol. Rev. **43**:224–240.
2. **Boquet, P., and E. Duflot.** 1982. Tetanus toxin fragment forms channels in lipid vesicles at low pH. Proc. Natl. Acad. Sci. USA **79**:7614–7618.
3. **Eidels, L., R. L. Proia, and D. A. Hart.** 1983. Membrane receptors for bacterial toxins. Microbiol. Rev. **47**:596–620.
4. **Fairweather, N., S. Kennedy, T. J. Foster, M. Kehoe, and G. Dougan.** 1983. Expression of a cloned *Staphylococcus aureus* α-hemolysin determinant in *Bacillus subtilis* and *Staphylococcus aureus*. Infect. Immun. **41**:1112–1117.
5. **Finn, C. W., R. P. Silver, W. H. Habig, and M. C. Hardegree.** 1984. The structural gene for tetanus toxin is on a plasmid. Science **224**:881–884.
6. **Helting, T. B., and H. H. Nau.** 1984. Analysis of the immune response to papain digestion products of tetanus toxin. Acta Pathol. Microbiol. Scand. Sect. C **92**:59–63.
7. **Helting, T. B., H. J. Ronnenberger, R. Vollerthun, and V. Neuberger.** 1978. Toxicity of papain-digested tetanus toxin. Pathological effect of fragment B in the absence of spastic paralysis. J. Biol. Chem. **253**:125–129.
8. **Helting, T. B., and O. Zwisler.** 1977. Structure of tetanus toxin. I. Breakdown of the toxin molecule and discrimination between polypeptide fragments. J. Biol. Chem. **252**:187–193.
9. **Helting, T. B., O. Zwisler, and H. Weigandt.** 1977. Structure of tetanus toxin. II. Toxin binding to ganglioside. J. Biol. Chem. **252**:194–198.
10. **Kenimer, J. G., W. H. Habig, and M. C. Hardegree.** 1983. Monoclonal antibodies as probes of tetanus toxin structure and function. Infect. Immun. **42**:942–948.
11. **Kimmel, J. R., and E. L. Smith.** 1954. Crystalline papain. I. Preparation, specificity and activation. J. Biol. Chem. **207**: 515–531.
12. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
13. **Matsuda, M., and M. Yoneda.** 1975. Isolation and purification of two antigenically active, "complementary" polypeptide fragments of tetanus neurotoxin. Infect. Immun. **12**:1147–1153.
14. **Mueller, J. H., and P. A. Miller.** 1945. Production of tetanal toxin. J. Immunol. **50**:377–384.
15. **Neubauer, V., and T. B. Helting.** 1979. Structure of tetanus toxin. N-terminal amino acid analysis of the two molecular forms of tetanus toxin and its composite chains. Biochem. Biophys. Res. Commun. **86**:635–642.
16. **Nichols, B. P., and C. Yanoksky.** 1983. Plasmids containing the *trp* promoters of *Escherichia coli* and *Serretia marcescens* and their use in expressing cloned genes. Methods Enzymol. **104**:155–164.
17. **Norrander, J., T. Kempe, and J. Messing.** 1983. A new pair of M13 vectors for selecting either strand of a double-digest restriction fragment. Gene **26**:101–106.
18. **Odink, K. G., M. J. Lockyer, S. C. Nicholls, Y. Hillman, R. R. Freeman, and A. A. Holder.** 1984. Expression of cloned cDNA for a major surface antigen of *Plasmodium falciparum* merozoites. FEBS Lett. **173**:108–112.
19. **Sanger, F., S. Nickle, and A. R. Coulson.** 1977. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
20. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **89**:503–517.
21. **Sproat, B. S., and W. Banwarth.** 1983. Improved synthesis of oligonucleotides on controlled pore glass using phosphotriester chemistry and a flow system. Tetrahedron Lett. **24**:5771–5774.
22. **Towbin, H., T. Staehelin, and J. Gordon.** 1979. Electrophoretic transfer of proteins from acrylamide gels to nitrocellulose sheets: procedure and applications. Proc. Natl. Acad. Sci. USA **76**:4350–4354.
23. **Twigg, A. J., and D. Sherratt.** 1980. Trans-complementable copy-number mutants of plasmid ColE1. Nature (London) **283**:216–218.
24. **van Heyningen, S.** 1976. Binding of ganglioside by the chains of tetanus toxin. FEBS Lett. **68**:5–7.
25. **Volk, W. A., B. Bizzini, R. M. Snyder, E. Bernhard, and R. R. Wagner.** 1984. Neutralization of tetanus toxin by distinct monoclonal antibodies binding to multiple epitopes on the toxin molecule. Infect. Immun. **45**:604–609.
26. **Wallace, R. B., M. J. Johnson, T. Hirose, T. Miyake, E. H. Kawashima, and K. Hakura.** 1981. The use of synthetic oligonucleotides as hybridisation probes. II. Hybridisation of oligonucleotides of mixed sequence to rabbit β-globin DNA. Nucleic Acids Res. **9**:879–894.
27. **Yanofsky, C. T., T. Platt, I. P. Crawford, B. P. Nichols, G. E. Christie, H. Horowitz, M. van Cleeput, and A. M. Wu.** 1981. The complete nucleotide sequence of the tryptophan operon of *Escherichia coli*. Nucleic Acids Res. **9**:6647–6666.