Biochem. J. (1983) **209**, 897–899
Printed in Great Britain

Similarities in the heavy and light chains of tetanus toxin suggested by their amino acid compositions

Clare F. TAYLOR, Paul BRITTON* and Simon VAN HEYNINGEN
Department of Biochemistry, University of Edinburgh, Hugh Robson Building, George Square, Edinburgh EH8 9XD, U.K.

(*Received 29 November 1982/Accepted 12 January 1983*)

Quantitative comparison of the amino acid compositions of the heavy and light chains of tetanus toxin by the method of Cornish-Bowden [(1983) *Methods Enzymol.* **91**, 60–75)] suggests strongly that there is sequence homology between the two chains and that the heavy chain has two similar halves. Examination (by electrophoresis in polyacrylamide gels in the presence of sodium dodecyl sulphate) of peptides produced from the chains by proteolytic cleavage supports this idea.

Tetanus toxin is a potent inhibitor of neurotransmitter release in the central nervous system; it is secreted by *Clostridium tetani* after infection of a wound [for general reviews see van Heyningen (1980) and Mellanby & Green (1981)]. The toxin is encoded on a plasmid (Laird *et al.*, 1980). The intracellular form of the toxin is a single chain of relative molecular mass (M_r) 150000 that is converted by proteolysis into an extracellular form having two polypeptides linked by one or more disulphide bonds: the heavy (H-) chain (M_r 100000) and the light (L-) chain (M_r 50000). The H-chain is the site of interaction with gangliosides in the cell membrane (van Heyningen, 1976; Helting *et al.*, 1977), but no function has yet been determined for the L-chain, perhaps because of ignorance about the intracellular target of the toxin. There has been no evidence for any similarities between the two chains; they do not interconvert or cross-react immunologically (although there are some technical difficulties and solubility problems with these experiments).

In spite of this, superficial inspection of the amino acid composition of the H- and L-chains shows them to be surprisingly similar (Table 1). Cornish-Bowden (1979, 1983) has developed a quantitative technique for comparing the amino acid compositions of different proteins to see what the chances are that similarities in the composition reflect similarities in sequence. He calculates a compositional index, $S\Delta n$, which is an estimator of the number of sequence differences when two chains each about the same M_r and having N residues, are aligned. When $S\Delta n < 0.42N$, there is a strong indication that the proteins are related; when $0.42N < S\Delta n < 0.93N$, there is a

Abbreviation used: SDS, sodium dodecyl sulphate.

* Present address: Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB2 1QW, U.K.

Table 1. *Amino acid composition of the chains of tetanus toxin*

The amino acid compositions are taken from Britton (1981). The reduced and carboxymethylated heavy and light chains were eluted from SDS/polyacrylamide gels. Results for the heavy chain are averages of three analyses and for the light chain, of two analyses.

	Composition (mol of residue/mol of protein)			
Amino acid	H-chain	Half H-chain	L-chain	$(n_{IH}-n_{IL})$
Cys	6	3	4	1
Asx	129	64.5	68	12.25
Thr	53	26.5	24	6.25
Ser	56	28	26	4
Glx	84	42	41	1
Pro	40	20	25	25
Gly	59	29.5	32	6.25
Ala	61	30.5	25	30.25
Val	54	27	27	0
Met	23	11.5	13	2.25
Ile	59	29.5	27	6.25
Leu	86	43	40	9
Tyr	45	21.5	22	0.25
Phe	27	13.5	16	6.25
His	13	6.5	7	0.25
Lys	63	31.5	29	6.25
Arg	22	11	11	0
Total	880	440	437	116.5

0306-3275/83/030897-03$2.00

weak indication; and when $S\Delta n > 0.93N$, there is little chance that they are related. The two chains of tetanus toxin are not the same length. However, it is noticeable that the H-chain (880 residues) is almost exactly twice the L-chain (437 residues). This suggests that the H-chain might have a sequence with two similar halves, and the useful comparison is between the amino acid composition of the L-chain and half that of the H-chain. Such a comparison suggests strong similarity.

Experimental

Calculation of compositional indices

The amino acid compositions of the reduced and carboxymethylated H- and L-chains were taken from Britton (1981). That of the L-chain was compared with that of the H-chain by using the following equation taken from Cornish-Bowden (1979, 1983):

$$S\Delta n = 0.5\sum(n_{iH}-n_{iL})^2 - 0.035(N_H-N_L)^2 + 0.535(N_H-N_L)$$

N_H is half the total number of residues in the H-chain and N_L the whole number in the L-chain; n_{iH} is half the number of residues of a particular amino acid i in the H-chain, and n_{iL} the number of such residues in the L-chain.

Structural analysis

Tetanus toxin was purified as described previously (van Heyningen, 1976) from material kindly given to us by Dr. R. O. Thomson of the Wellcome Research Laboratories, Beckenham, Kent, U.K. All the normal precautions were taken in handling this toxic material. SDS/polyacrylamide-gel electrophoresis (Laemmli, 1970) showed three bands under reducing conditions: light chain, heavy chain and a small amount of un-nicked toxin (i.e. the single-chain form). A portion (5 mg) of toxin was dissolved in 0.8 ml of 0.1 M-Tris/HCl buffer, pH 8.2, containing 8 M-urea, 2 mM-EDTA, 1 mM-benzamidine, 0.1 mM-phenylmethylsulphonyl fluoride and 2 mM-dithiothreitol, and incubated for 1 h under N_2 at room temperature. The solution was made 5 mM in iodoacetic acid and incubated for a further 1 h. The reaction was stopped with dithiothreitol (7 mM). After dialysis against 50 mM-NH_3, the modified protein was subjected to electrophoresis in 8% polyacrylamide gels (Laemmli, 1970), and the chains detected by staining with Coomassie Brilliant Blue R. Gel slices containing un-nicked toxin or the isolated chains were subjected to partial proteolysis with the V8 proteinase of *Staphylococcus aureus* (Cleveland *et al.*, 1977). Before cleavage with CNBr the three proteins were eluted electrophoretically from the gel, and the stain and the SDS were removed (Walker *et al.*, 1982). The chains were dissolved in 70% (v/v) formic acid, an equal weight of CNBr was added, and the samples left for 24 h. The products were analysed by SDS/polyacrylamide-gel electrophoresis and detected by silver staining (Wray *et al.*, 1981). This very sensitive technique was necessary as the amounts of toxin and of purified chains were limited.

Results and discussion

Table 1 shows the amino acid composition of the H- and L-chains (Britton, 1981). The compositional index, $S\Delta n$, calculated from the composition of the L-chain and half that of the H-chain is 60; $N = 437$ and $0.42N = 184$. Thus the result satisfies the most stringent test for sequence similarity which "...is a strong indication, amounting almost to a certainty, that the two proteins are related." (Cornish-Bowden, 1983). In other words, the data suggest that there will be found to be considerable sequence homology between the L-chain and the H-chain, and, further, that the H-chain may have two similar halves each of which has homologies with the L-chain.

It is possible that this surprising result could be due to some error in the amino acid compositions, although this is unlikely, since both analyses were done at the same time and under the same conditions. We therefore did the same calculations by using a quite different set of data, namely those published by DiMari *et al.* (1982). $S\Delta n$ in this case is 162; still below $0.42N$, though not as strikingly so. The similarity is not therefore likely to be due to any particularities in our own experiments.

There is one intriguing difference in the composition of the two chains: an inverse relationship between proline and alanine contents. The H-chain has five fewer proline residues and five more alanine residues than the L-chain per molecule of M_r 50000. This could be due to a single base change (CCN for proline to GCN for alanine), but because of the special properties of proline, it would produce a marked change in structure and hence in function.

A simple experimental test for any similarity is to look at the products of proteolytic cleavage. Fig. 1(*a*) shows the result of polyacrylamide-gel electrophoresis of the intact toxin and the isolated chains after partial digestion with V8 proteinase from *Staph. aureus*. The L-chain shows several peptides with the same mobilities as peptides from the H-chain. This degree of similarity is significant, as the technique tends to emphasize differences between proteins. Furthermore, the similarity between the digests of the H-chain and whole toxin is marked. One-third of the whole toxin is L-chain, so

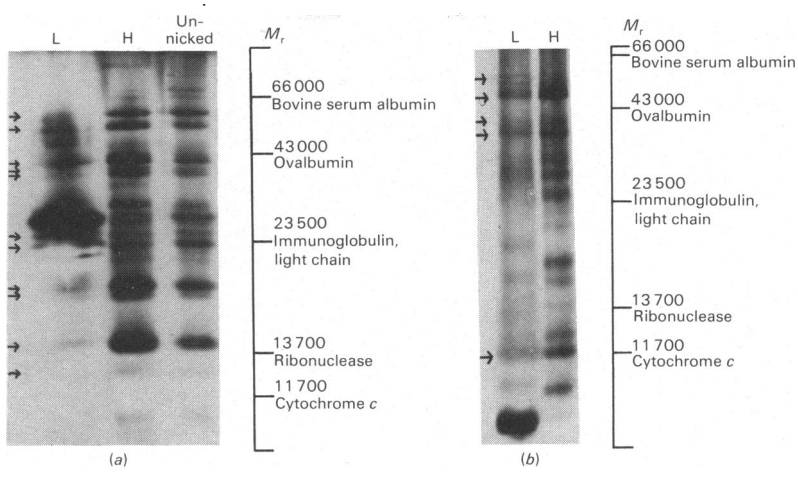

Fig. 1. *SDS/polyacrylamide-gel electrophoresis of peptides from the chains of tetanus toxin*
(*a*) An SDS/12%-polyacrylamide gel showing the products of limited cleavage with V8 proteinase (see the text). There was no visible staining in a parallel track containing the same amount of proteinase only. (*b*) A 13% gel showing the products of total cleavage with CNBr. Peptides common to the H- and L-chains are marked.

one would expect to see a greater difference if there were no similar regions in the two chains.

Fig. 1(*b*) shows a similar gel after total cleavage at methionine residues with CNBr. There are five peptides common to both chains.

Our suggestion of sequence homologies would be accounted for if the 150000-M_r tetanus toxin molecule had arisen by gene-duplication events involving an ancestral gene coding for a 50000M_r protein that now produces three different but related domains, one in the L-chain and two in the H-chain. The conclusive test will come when either the amino acid sequence of the protein or the nucleotide sequence of its gene have been determined.

We are grateful to Dr. Athel Cornish-Bowden, Department of Biochemistry, University of Birmingham, for his helpful advice. C. F. T. holds, and P. B. held, Research Studentships from the Medical Research Council and Science Research Council respectively.

References

Britton, P. (1981) Ph.D. Thesis, University of Edinburgh
Cleveland, D. W., Fischer, S. G., Kirschner, M. W. & Laemmli, U.K. (1977) *J. Biol. Chem.* **252**, 1102–1106
Cornish-Bowden, A. (1979) *J. Theor. Biol.* **76**, 369–386
Cornish-Bowden, A. (1983) *Methods Enzymol.* **91**, 60–75
DiMari, S. J., Cumming, M. A., Hash, J. H. & Robinson, J. R. (1982) *Arch. Biochem. Biophys.* **214**, 342–353
Helting, T. B., Zwisler, O. & Wiegandt, H. (1977) *J. Biol. Chem.* **252**, 194–198
van Heyningen, S. (1976) *FEBS Lett.* **68**, 5–7
van Heyningen, S. (1980) *Pharmacol. Ther.* **11**, 141–157
Laemmli, U. K. (1970) *Nature (London)* **227**, 680–685
Laird, W. J., Aaronson, W., Silver, R. P., Habig, W. H. & Hardegree, M. C. (1980) *J. Infect. Dis.* **142**, 623
Mellanby, J. & Green, J. (1981) *Neuroscience* **6**, 281–300
Walker, J. E., Auffret, A. D., Carne, A., Gurnett, A., Hanisch, P., Hill, D. & Saraste, M. (1982) *Eur. J. Biochem.* **123**, 253–260
Wray, W., Boulikas, T., Wray, V. P. & Hancock, R. (1981) *Anal. Biochem.* **118**, 197–203