Ricin B Chain Is a Product of Gene Duplication*

(Received for publication, September 15, 1980, and in revised form, October 29, 1980)

Jesus E. Villafranca and Jon D. Robertus

From the Clayton Foundation Biochemical Institute, Department of Chemistry, The University of Texas, Austin, Texas 78712

The cytotoxin ricin is a heterodimer composed of an A chain which enzymatically inhibits protein synthesis on eu karyotic ribosomes and a lectin B chain which binds to cell surfaces and triggers uptake of the toxin. A low resolution (4 Å) x-ray structure revealed that the B chain is a globular structure and that each domain binds a galactose sugar. These apparent structural similarities suggested that the protein might have internal symmetry and might have arisen by gene duplication. A subsequent search of the amino acid sequence provided strong evidence for homology between the NH2- and COOH-terminal halves of the B chain and suggested that they may have arisen from a common ancestor.

There is no strong relationship between the halves of the A chain and little, if any, significant sequence homology between the A and B chains of ricin.

Ricin is a toxic glycoprotein found in the seeds of Ricinus communis, the castor bean plant. Its extreme toxicity in eukaryotic cells has been attributed to its ability to inhibit protein synthesis. Since Lin et al. (1) reported that ricin showed antitumor activity, the biochemical and biological properties of this protein have been extensively studied by several investigators (2-6). The toxin is composed of two subunits linked together by a single disulfide bond. The A chain (M, = 30,600) has been shown catalytically to inactivate the 60 S ribosomal subunit such that it has a greatly reduced activity. This investigation was supported by United States Public Health Service Grants CA-24059 and AI-13584 from the National Cancer Institute and the National Institute of Allergy and Infectious Diseases. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

EXPERIMENTAL PROCEDURES

X-ray diffraction data were collected by oscillation photography to a resolution of 2.8 Å for native ricin and four isomorphous heavy metal derivatives. Phasing and heavy metal parameter refinement were carried out by conventional crystallographic methods and electron density maps at 4 Å and 2.8 Å resolution were calculated. Details of the crystallographic analysis will be published later, when our high resolution analysis is complete.

The 4 Å three-dimensional electron density map was traced onto plexiglass and displayed in a mounting box. Solvent regions were easily discernible through most of the map, facilitating the assignment of molecular boundaries. A model was constructed from %inch balsa wood using the electron density map as a guide and the corresponding density contour paper tracings as templates. Positions of the lactose binding sites were obtained from difference Fourier calculations, using the ho0, 0h0, 4h0, and h04 projections of native ricin and of a ricin-lactose complex, formed by soaking the crystals in 2 nM sugar. The sugar peaks were 6 and 4 times the root mean square difference density and could be positioned with confidence.

Amino acid sequence alignments were analyzed by the computer program KOF0 provided by Dr. J. L. Fox. The theory and operation of this program have been discussed extensively (17) and originate from ideas proposed by W. M. Fitch (18). In brief, the program scans peptides of a given length for all possible alignments of the compared sequences then lists only those within a given range of statistical probability based on the minimum base change per codon (MBC/C) per scanned peptide.

RESULTS AND DISCUSSION

The 4 Å ricin model is shown in Fig. 1. The molecule appears to be a compact structure with the two elongated subunits oriented in a roughly parallel fashion. The A chain (the lighter subunit in Fig. 1) has an asymmetrical wedge shape and is more tightly associated with the B chain at its larger end (top in Fig. 1). Very likely, the interchain disulfide bond is situated in this region as there is evidence that the bond is partially buried in the interior of the molecule (19). We are presently unable to determine the disulfide position unambiguously from the electron density map, however. The B chain appears to be composed of two similarly shaped domains, each containing a lactose binding site. The lactose sites are shown as white blobs on the darker subunit in Fig. 1. The difference Fourier indicated that one of the lactose sites (upper site) was more highly occupied than the other. Two binding sites for the B chain have previously been suggested (20, 21), which is consistent with the observation that ricin has lectin activity. A prominent groove in the region along the lactose sites indicates that an extended site may exist for oligosaccharide binding. This conclusion is supported by the observation that higher binding constants are obtained for oligosaccharides and glycopeptides than for simple sugars and disaccharides (23). Extended sites have already been proposed for several other lectins (24-26). Although CD measurements indicate that ricin undergoes a small conformational change upon lactose binding (22), no significant conformational difference between the native and the ricin-lactose complex has been detected in our maps.

The similarity in shape of the domains of the B chain and the nearly symmetrical distribution of sugar binding sites suggested that the protein might have arisen by gene duplication, and an analysis of the B chain sequence was therefore...
Ricin B Chain Is a Product of Gene Duplication

FIG. 1. Low resolution model of ricin. The model shown was constructed from balsa wood based upon the 4-Å electron density map of ricin. The lighter region is the A chain, the darker region is the B chain. Lactose binding sites, discovered by difference Fourier, are shown as small white dumbbells on the B chain. Molecular dimensions are $73 \times 58 \times 35$ Å. The scale bar at the upper right = 30 Å.

Fig. 2. Structural homologies within the ricin B chain. The linear sequence of the B chain is represented as a line with disulfide bridges forming loops. The dashed vertical line divides the chain into an NH$_2$-terminal half (1-132) and a COOH-terminal half (133-260). The circled numbers described the number of residues within each loop. Regions of strong amino acid homology are marked off along the chain by $\circ$, $\square$, and $\blacksquare$. A stretch of a given pattern in one half is related to a matching sequence pattern in the second half. The pattern of disulfide bridges, loop size, and sequence homologies suggest the two halves of the B chain are strongly related in a translational sense. Cys 4 of the B chain forms the disulfide bond to the A chain as indicated on the left-hand side of the figure.

Fig. 3. Amino acid sequence comparison of the two halves of ricin B chain. The sequence of the NH$_2$-terminal half (residues 1-132) is compared with the COOH-terminal half (133-260), aligning the 2 pairs of disulfide bridges in each half. Identical residues are enclosed in boxes. Amino acid single-letter symbols are according to Dayhoff (28).
to search for homologies between the A and B chain sequences and within the A chain itself. No strong agreement was found. It is interesting to note, however, that the strongest agreement (probability = $2 \times 10^{-4}$) was a match between A chain sequence 11-30 and B chain sequence 216-235 which differed by 17 base changes over the 20 residues. Since this is apparently a highly conserved part of the B chain sequence, it may indicate that the chains had a common ancestor in the very distant past although we stress that the A to B correlation is very much lower than that between the two halves of the B chain.

The evidence provided here presents a strong case for gene duplication in the B subunit of ricin. That is, the protein shows two distinct folding domains of similar size and shape. The toxins abrin and modeccin also have A and B subunits with almost identical physical and biological properties to those of ricin (2, 30-32). Structural comparisons between these toxins should therefore prove interesting. Areas of high homology in sequence comparisons of the A chains could indicate those regions which participate in A chain-ribosome interactions, while B chain comparisons could pinpoint those areas involved with oligosaccharide binding. The 2.8-A resolution analysis now in progress should reveal to what extent the B chain’s internal homology persists in the secondary and tertiary structure and may provide a more powerful tool than sequence comparison for assessing the relatedness of the A and B chains.

REFERENCES
1. Lin, J., Tseng, K., Chen, C., Lin, L., and Tung, T. (1970) Nature 227, 292-293
2. Olsnes, S., and Pihl, A. (1976) in The Specificity and Action of Animal, Bacterial and Plant Toxins (Cuttrecas, P., ed) Series B, Vol. I, pp. 138-173, Chapman and Hall, London
3. Ishiguro, M., Funatsu, G., and Funatsu, M. (1971) Agric. Biol. Chem. 35, 729-735
4. Gurrieri, L. G., and Horts story, H. J. (1973) Biochim. Biophys. Acta 295, 583-594
5. Nicolson, G. L., Blaustein, J., and Etzler, M. (1974) Biochemistry 13, 196-204
6. Weis, C. H., Hartman, F. C., Prater, R., and Yank, W. K. (1974) J. Biol. Chem. 249, 3061-3067
7. Montanaro, L., Sertoli, S., Mattioli, A., Testoni, G., and Stirpe, F. (1975) Biochem. J. 146, 127-131
8. Nolan, R. D., Gras, H., and Dyewa, J. (1975) Eur. J. Biochem. 64, 69-76
9. Fernandez-Puentes, C., Benson, S., Olsnes, S., and Pihl, A. (1976) Eur. J. Biochem. 64, 437-443
10. Nicolson, G. L., and Blaustein, J. (1972) Biochim. Biophys. Acta 266, 543-547
11. Olsnes, S., Refsnes, K., and Pihl, A. (1974) Nature 248, 627-631
12. Olsnes, S., Sandrig, K., Refsnes, K., and Pihl, A. (1976) J. Biol. Chem. 251, 3985-3992
13. Nicolson, G. L. (1974) Nature 251, 628-630
14. Yoshitake, S., Funatsu, G., and Funatsu, M. (1978) Agric. Biol. Chem. 42, 1267-1274
15. Funatsu, G., Kimura, M., and Funatsu, M. (1979) Agric. Biol. Chem. 43, 2221-2224
16. Villafranca, J. E., and Robertus, J. D. (1977) J. Mol. Biol. 116, 331-335
17. Kobayashi, K., and Fox, J. L. (1978) J. Mol. Biol. 11, 233-243
18. Fitch, W. M. (1966) J. Mol. Biol. 16, 9-16
19. Lappi, D. A., Rasmussen, W., Seigle, J. M., and Kaplan, N. O. (1978) Proc. Natl. Acad. Sci. U. S. A. 75, 106-110
20. Zentz, C., Frey, J. A., and Bourillon, B. (1977) FEBS Lett. 81, 23-27
21. Funatsu, G., Veno, S., and Funatsu, M. (1977) Agric. Biol. Chem. 41, 1737-1743
22. Shimazu, K., Wilkorn, E. F., Jr., Neri, G., and Irgensens, B. (1975) Arch. Biochem. Biophys. 169, 731-736
23. Baenninger, J. V., and Fiete, D. (1979) J. Biol. Chem. 254, 9795-9799
24. Goldstein, L. J., Reichert, C. M., Misaki, A., and Gorin, P. A. J. (1973) Biochim. Biophys. Acta 317, 500-504
25. Lotan, R., Skutelsky, E., Denon, D., and Sharon, N. (1975) J. Biol. Chem. 250, 8518-8523
26. Privat, J. P., De Loef, F., Mialonier, G., Bouchard, P., and Monsigny, M. (1974) Eur. J. Biochem. 47, 5-14
27. Jukes, T. H., and Cantor, C. R. (1972) in Mammalian Protein Metabolism (Munro, H. M., ed) Vol. 22, Academic Press, New York
28. Dayhoff, M. O., ed (1972) Atlas of Protein Sequence and Structure, Vol. 5, National Biomedical Research Foundation, Washington, D. C.
29. Pappenheimer, A. M., Jr., Olsnes, S., and Harper, A. A. (1974) J. Immunol. 113, 835-841
30. Olsnes, S., Sandrid, K., Eklind, K., and Pihl, A. (1978) J. Supramol. Struct. 9, 15-25
31. Gasperi-Compagni, A., Barbiere, L., Lorenzoni, E., Montanaro, L., Sertoli, S., Bonetti, E., and Stirpe, F. (1978) Biochem. J. 174, 491-496
32. Montanaro, L., Sertoli, S., Zomboni, M., Denare, M., Testoni, G., Gasperi-Lampani, A., and Stirpe, F. (1978) Biochem. J. 176, 371-379