PROTEINS: Structure, Function, and Genetics 10:260–269 (1991)

Structure of Ricin B-Chain at 2.5 Å Resolution

Earl Rutenber and Jon D. Robertus
Clayton Foundation Biochemical Institute, Department of Chemistry and Biochemistry, University of Texas, Austin, Texas 78712

ABSTRACT **The heterodimeric plant toxin ricin has been refined to 2.5 Å resolution. The B-chain lectin (RTB) is described in detail. The protein has two major domains, each of which has a galactose binding site. RTB has no regular secondary structure but displays several Ω loops. Each RTB domain is made of three copies of a primitive 40 residue folding unit, which pack around a pseudo threefold axis. In each domain, galactose binds in a shallow cleft formed by a three residue peptide kink on the bottom and an aromatic ring on the top. At the back of the cleft, an aspartate forms hydrogen bonds to the C_3 and C_4 hydroxyls of galactose, whereas a glutamine bonds to the C_4 alcohol, helping to define specific epimer binding. In addition to analyzing the sugar binding mechanism, the assembly of subdomain units around the pseudo threefold axis of each domain is described. The subdomains contribute conserved Trp, Leu, and Ile residues to a compact central hydrophobic core. This tight threefold binding probably drives the peptide folding and stabilizes the protein structure.**

Key words: ricin toxin, B-chain, galactose binding, molecular evolution

INTRODUCTION

Ricin toxin B-chain (RTB) is the sugar binding moiety of the cytotoxic plant protein ricin. It is linked to the A-chain (RTA) by a disulfide bond. RTB binds reversibly to galactose residues presented by cell surface glycopeptides and glycolipids and facilitates uptake of whole ricin by endocytosis.[1] Following reduction of the disulfide bond RTA is released into the cytoplasm, where its action as an N-glycosidase inactivates cellular ribosomes, often resulting in cell death. Recent evidence implicates the Golgi apparatus as the intracellular compartment from which the RTA escapes into the cytosol.[2] It has been suggested that RTB serves to facilitate both toxin transport within the Golgi and translocation of the catalytic A-chain into the cytosol.[3]

Several research groups have investigated the sugar binding properties of ricin using a variety of methods. Equilibrium dialysis revealed two lactose binding sites with association constants 2800 M^{-1} and 35,000 M^{-1} at 4°C and 3,000 M^{-1} and 19,000 M^{-1} at 25°C.[4,5] However, a study using fluorescent galactose analogs showed nearly equal affinities for the two sites.[6] A difference Fourier analysis of ricin with and without bound lactose, also suggested that the sites are occupied unequally.[7] Ultraviolet difference spectroscopy implicated a Tyr residue in the stronger binding site,[8] whereas chemical modification studies suggested a Trp residue was important to sugar binding in the lower affinity site.[9] It was also shown that RTB binds only a single molecule of N-acetylgalactosamine (GalNAc).[7–9]

Finally, it has been shown that the association constants of ricin binding to naturally occurring cell surface galactosides exceeds that for lactose by a factor of about 1,000.[10] The increased binding affinity is observed for saccharides with multiple antennae terminating with a galactoside as well as for saccharides with only a single terminal galactoside.

A 2.8 Å structure of ricin, including bound lactose, has been reported previously.[11] That model was analyzed to reveal that RTB arose by gene triplication and then duplication, from a primitive 40 residue galactose binding peptide.[12] Recently the ricin model has been refined to 2.5 Å resolution,[13] and a thorough description of the revised RTB model and its biochemical implications is presented in this work.

MATERIALS AND METHODS

Difference Fourier analysis of oligosaccharide binding to RTB was performed as follows. Ricin was prepared and crystals grown in the presence of 4 mM lactose, as previously described.[14] A biantennary oligosaccharide (BioCarb), representative of naturally occurring cell surface saccharides, was diffused into native ricin crystals by gradually replacing the artificial mother liquor with artificial mother liquor plus 10 mM oligosaccharide. A single crystal, so prepared, was mounted and data collected by automated diffractometry, as previously described,[11] to a resolution of 5.0 Å.

The least-squares superposition of RTB subdomains was carried out as follows. The 1α unit, de-

Received July 20, 1990; revision accepted February 4, 1991.
Address reprint requests to Jon D. Robertus, Dept. of Chemistry, College of Natural Sciences, University of Texas, Austin, TX 78712-1167.

scribed below, was chosen as the reference subdomain and the five remaining subdomains were superimposed upon it. Each subdomain contributes up to three residues to the hydrophobic core. These and one residue on either side were initially chosen as markers to compute the least-squares rotation and translation operators for each subdomain. This selection provided nine residues total for the α and β subdomains and six residues for the γ subdomains. The routine RIGI, included in the FRODO[15] molecular graphics package, was used to compute the rotation and translocation operators. These were then applied to all atoms in the subdomain. The RMS deviation between the translated coordinates and the target coordinates was 0.4 Å for all five subdomains.

RESULTS

As mentioned above, a refined 2.5 Å model of ricin has been obtained[13] and reflects several important improvements over the previous model.[11] Figure 1 shows the current model of ricin. Each domain of RTB is composed of four peptides designated λ, α, β, and γ. The subdomains of domain 1 (residues 1–135) are labeled 1λ, 1α, 1β, and 1γ; domain 2 is labeled in a corresponding manner. The λ subdomains are homologous to one another,[12] but are unrelated to the others. The α, β, and γ subdomains are each approximately 40 residues in length and are homologs.[12] Figure 2 shows the amino acid sequence alignment for these six major structural subdomains based on maximum sequence identity, weighted with a penalty function for introducing gaps[12] and structural correspondence from the present B-chain model. The λ (or linking) subdomain is the first 16 residues of each domain and is not related to the others. The 1λ subdomain serves to link the B-chain to the A-chain and the 2λ subdomain links the first domain to the second. The main body of each RTB domain is formed from the assembly around a pseudo threefold axis of an α, β, and γ subdomain, whereas the λ-subdomain extends across the top of the domain.

Galactose sugars bind in nonhomologous sites in the two domains. That is, sugar binding site 1 lies in the 1α, whereas site 2 is in the 2γ subdomain. As described previously,[11,12] Asp, Gln, Asn, and aromatic residues are involved in the two sugar binding sites; they are indicated by asterisks in Figure 2. Subdomain residues that are involved in the packing around the pseudo threefold axis are the invariant Trp and strongly conserved aliphatic residues marked with solid triangles. In each domain both the α and β subdomains contain disulfide bridges between conserved Cys residues marked by arrows in Figure 2.

The three subdomain types are topologically similar.[12] We have suggested that the ancestor of the modern B-chain molecule was a galactose-binding peptide approximately 40 residues in length, which resembled the 1α subdomain and may have existed in solution as a trimer stabilized by the hydrophobic interaction of the Trp, Ile, and Leu residues. Gene triplication and fusion produce the main globular body of the domain, whereas addition of the λ subdomain stabilized the threefold structure and complemented the truncated γ subdomain. A recent gene duplication produced the modern B-chain, whereas fusion with the separately evolving A-chain gene created a modern heterodimeric cytotoxin.

Interactions With Carbohydrate

The individual subdomains composing the modern molecule have undergone considerable divergence. However, the sugar binding sites found in subdomains 1α and 2γ bind lactose in a strikingly similar fashion, as seen in Figure 3. Each sugar binding site is composed solely of amino acid residues from one subdomain and may be described as a shallow pocket. The top of the pocket is formed by the side chain of an aromatic residue; the bottom by the peptide backbone of a three-residue kink in the chain. The galactosyl moiety of lactose is oriented and secured by hydrogen bonds to the side chains of several polar residues that lie at the back of the pocket, whereas the glucosyl moiety extends freely into solvent and makes no specific interaction with the protein. In sites 1 and 2, respectively, the aromatic residue is Trp 37 and Tyr 248 and the three-residue kink is formed by residues 24–26 and 236–238.

The residues forming hydrogen bonds to galactose are structurally similar in each site. A conserved aspartic acid residue, Asp 22 in site 1 and Asp 234 in site 2, appears to mediate the primary interaction with the bound sugar. The O_{D1} of these residues accepts a hydrogen bond from the C_3 hydroxyl of the bound sugar and the O_{D2} accepts a hydrogen bond from the C_4 hydroxyl. A hydrogen bond from N_{E2} of a conserved amide, Gln 47 in site 1 and Gln 256 in site 2, to O_{D1} of the aspartic acid orients and stabilizes the alignment of this key residue. Each C_3 hydroxyl of the galactose also forms a strong hydrogen bond with N_{D2} of a conserved amide, Asn 46 in site 1 and Asn 255 in site 2.

The hydrogen bond pattern described above marks a subtle but important change from our view of galactose binding in the unrefined molecule.[11] Originally, it appeared that the principle hydrogen bond made to the sugar was the one from the binding site Asn (46 or 255) to the hydroxyl at C_4, and, indeed, site-directed mutagenesis of Asn255 has confirmed that they are important in galactose binding.[26] The binding site Asp residues (22 and 234) were initially thought to merely orient the crucial asparagines. Adjustment of the model during

refinement, including the bound lactose, has altered the view of the acids to where they appear to form the strongest (shortest) bonds directly to the sugar. The asparagine residues continue to form strong bonds (~2.8 Å) and contribute with the aspartates to create the polar environment conducive to sugar binding.

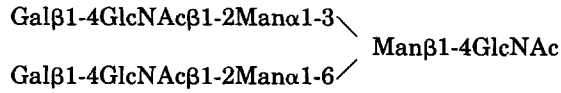
The hydrogen bond network of site 1 is more complex than that of site 2 in that N_{E2} of Gln 35 makes a bifurcated hydrogen bond to the C_4 and C_6 hydroxyls. In site 2, the analog of Gln 35 is Ile 246 and the hydrogen bonding interaction seen in site 1 is replaced by a hydrophobic contact between this residue and C_6 of the sugar. In contrast to the galactose in site 1, the C_6 hydroxyl is oriented toward solvent in site 2 and not involved in hydrogen bonding to the protein.

Epimeric specificity is conferred upon the sugar binding site through hydrogen bond interactions with the C_4 hydroxyl. Glucose is the C_4 epimer of galactose and it does not bind to RTB. A line of nonpolar atoms of the galactose ring, C_4, C_5, and C_6, contact the plane of the aromatic ring found in each site. This juxtaposition of the hydrophobic face of a sugar against an aromatic ring is a common feature emerging in protein-carbohydrate interactions.[16]

Based on spectroscopy,[8] chemical modification studies,[9] and difference Fourier analysis,[7] site 1 is the low affinity and site 2 is the high affinity site for lactose binding. However, it is unclear from the present structure why site 2 binds lactose more strongly. Fewer hydrogen bonds are made to lactose than in site 1 but binding may benefit from the hydrophobic interaction between C_6 of the sugar and Ile 246. The strength of a given hydrogen bond and the free energy of a hydrophobic interaction are difficult quantities to assess, and the measured equilibrium constants reflect a difference in free energy of only about 1.5 kcal/mol.

The x-ray model can provide a rationale for the observation that GalNAc is bound only at site 2 and not at site 1. GalNAc differs from galactose by the presence of an N-acetyl group at the C_2 position. The least-squares superposition of the sugar binding sites (not shown) reveals that the bound galactose residue of site 2 is rotated roughly 15° relative to that of site 1. An N-acetyl group at the galactose C_2 position in site 2 would extend freely into solvent but would encounter steric hindrance upon binding in site 1. In particular, the N-acetyl group would clash with the side chain of Asp 44.

An effort was made to explore the observation that RTB binds roughly 1,000-fold better to oligosaccharides with terminal galactosides.[10] Binding of a high mannose (Man), biantennary oligosaccharide isolated from human urine was examined by difference Fourier methods. The structure of the oligosaccharide is:

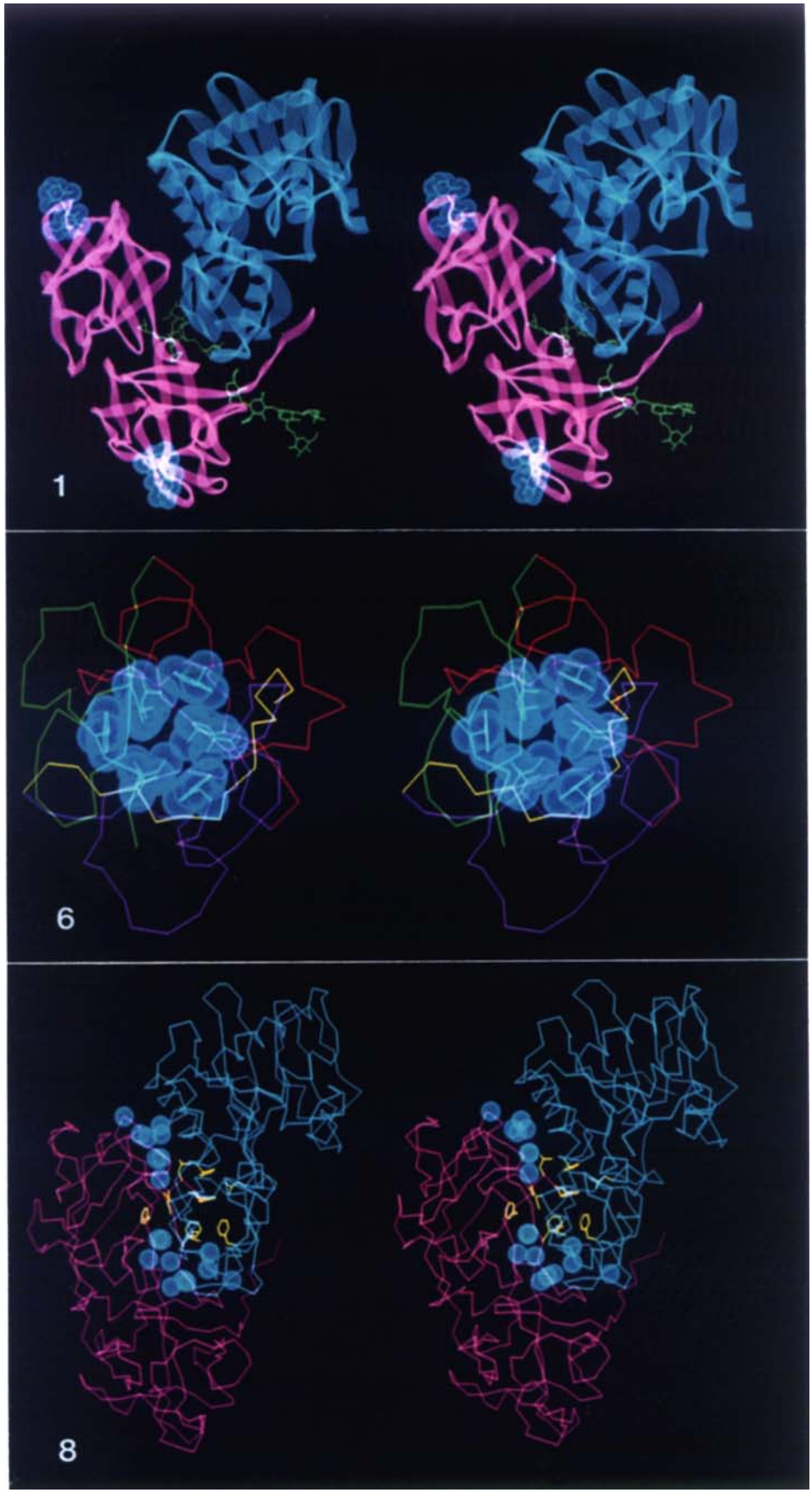


The two long antennae terminate in galactose (Gal), whereas the third terminal residue is N-acetylglucosamine (GlcNAc), which does not bind to ricin. The 5 Å difference electron density map together with a superimposed model of the oligosaccharide is shown in Figure 4. The terminal galactoside of one long antennae is bound by the domain 1 sugar binding site of one ricin molecule, whereas the terminal galactoside of the second long antennae is bound by the domain 2 sugar binding site of an adjacent, symmetry related molecule. The 70 Å span between the two RTB sugar binding sites is too great for a single biantennary oligosaccharide to bind both sites. However, the observed binding suggests that biantennary cell surface saccharides are capable of binding two ricin molecules without steric hindrance. The crystalline environment is artificial and the observed arrangement probably does not reflect any particularly significant mode of interaction. Toxin concentrations in vivo would rarely be high enough to accommodate two ricin molecular per galactoside, and indeed it is likely that RTB generally binds two separate oligosaccharides. The only contacts made between the biantennary sugars and the protein are those seen in the binding of lactose; no additional contacts are made at sugars beyond the galactose. The spacing in the crystal imposes a restricted conformation on the biantennary sugar, which must stretch fairly tightly to bind both ends. It is possible that in solution the sugar chain could lie against the protein and form additional hydrogen bonds, which would stabilize the complex and account for the observed increase in binding affinity. It is interesting to speculate that endocytosis may be triggered when two rather widely separated cell surface oligosaccharides are bound at once. Site-directed alteration of a single site would then be expected to diminish cell uptake more profoundly than it would diminish surface binding.

Fig. 1. The three-dimensional structure of ricin. A ribbon drawing of both the A-chain (blue) and the B-chain (red) of ricin as viewed down the crystallographic *a* axis. The reversibly bound galactose residues are shown in a blue surface and the covalently linked high mannose sugars are shown in green.

Fig. 6. Detail of the hydrophobic core of domain 1. The alpha carbon backbone is color coded; the λ subdomain is yellow, α is purple, β is red, and γ is green. The side chain atoms of the hydrophobic core are shown in white with a blue surface at the van der Waals radius. The stereo view is down the pseudo threefold axis.

Fig. 8. The A–B interface of ricin. The B-chain is shown in blue and the A-chain is in red. The hydrophobic residues involved in the interface are shown in yellow and the trapped solvent molecules are highlighted with a blue surface. For orientation, the view is similar to that in Figure 1, roughly down the crystallographic *a* axis.


1
6
8

a)

```
1α  17  N G L C V D V R D G R F H N G N A I Q L W - - P C K S N T D A N Q L W T L K R D N T I R S  59
2α 148  Y G L C L Q A - - - - - - - - - N S G Q V W I E D C - S S E K A E Q Q W A L Y A D K S I R P 183
1β  60  N G K C L T T - - - Y G Y S P G V Y V - M I Y D C N T A A T D A T R W Q I W D N G T I I N 100
2β 187  R D N C L T S - - - D S N I R E T V V - K I L S C - G P A S S G Q R W M F K N D G T I L N 226
1γ 103  S S L V L A A T S G N S G T T L T V Q T N I Y A V - - - - - - - S Q G W - L P T N 135
2γ 229  S G L V L D V R A - - S D A S L K - Q I I L Y P L - - H G D P N Q I W - L P L F 262
```

b)

```
1α  17  N G L C V D V R D G R F H N G N A - I Q L W P C K S N T D A N Q L W T L K R D N T I R S  59
2α 148  Y G L C L Q A - N - - - S - - G Q - V W I E D C S S E K A - E Q Q W A L Y A D K S I R P 183
1β  60  N G K C L T T Y - G - Y S P G V Y - V M I Y D C N T A A T D A T R W Q I W D N G T I I N 100
2β 187  R D N C L T S - D S N - I R - E T V V K I L S C P A - - S S G Q R W M F K N D G T I L N 226
1γ 103  S S L V L A A T - S G N S - G T T - L T V Q T N I Y A - V - S Q G W L P T N 135
2γ 229  S G L V L D V R A S D A S - L K Q - I I L Y P L H - G - D P N Q I W L P L F 262
```

Fig. 2. Sequence similarity of the α, β, and γ subdomains based on **(a)** maximum identity and **(b)** examination of the present x-ray structure. The solid triangles denote those highly conserved residues that form the hydrophobic core of each RTB domain. The open triangles denote additional, less conservative hydrophobic residues that augment the core. Residues marked with a star are those that participate in galactose binding in the 1α and 2 γ subdomains. The down arrows mark cysteine residues that form disulfides in subdomains α and β.

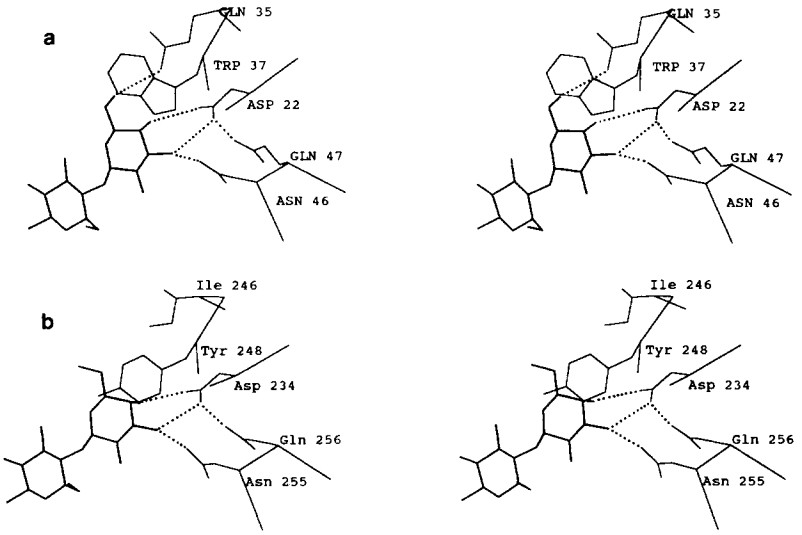


Fig. 3. The lactose binding sites of RTB. The sugar binding sites of **(a)** domain 1 and **(b)** domain 2 are presented in a common orientation as stereo pairs. Key residues on the protein are labeled, and lactose is shown in heavier bonds.

In keeping with our suggestion that each of the six subdomains is descended from a common ancestral sugar binding peptide, it is reasonable to examine the specific structural differences between the subdomains that presently bind sugars (1α and 2γ) and those that do not. From the homologies based on sequence alignment shown in Figure 2a, it is clear that each subdomain not involved in sugar binding lacks one or more of the key polar residues implicated in sugar binding. However, the 2α subdomain possesses a significant number of functional features necessary for sugar binding including Trp 160, which is homologous to Trp 37 in 1α. Indeed, based on statistical considerations alone, 1α more closely resembles 2α than 2γ, the other galactose binding subdomain.[12] In the 2α subdomain, however, the Trp160 side chain is displaced two amino acid positions by the insertion of Glu162 and Asp163 between itself and the fixed Cys164. In a functional binding site, the Trp side chain would project into solvent and form the top of a sugar binding pocket; in its displaced position it lies flat against the protein and partially blocks the potential sugar binding site. The inserted Glu 162 occupies the position in space of the binding site top and projects into solvent.

Subdomain 1β retains the aromatic residue, Tyr78, of the lactose binding site, and the backbone structure is very similar to those of the productive sites, but it lacks all of the specific polar interactions. The remaining two subdomains, 1γ and 2β, lack aromatic residues at the binding site, kink residues, and key polar residues and so their failure to bind galactose is not surprising.

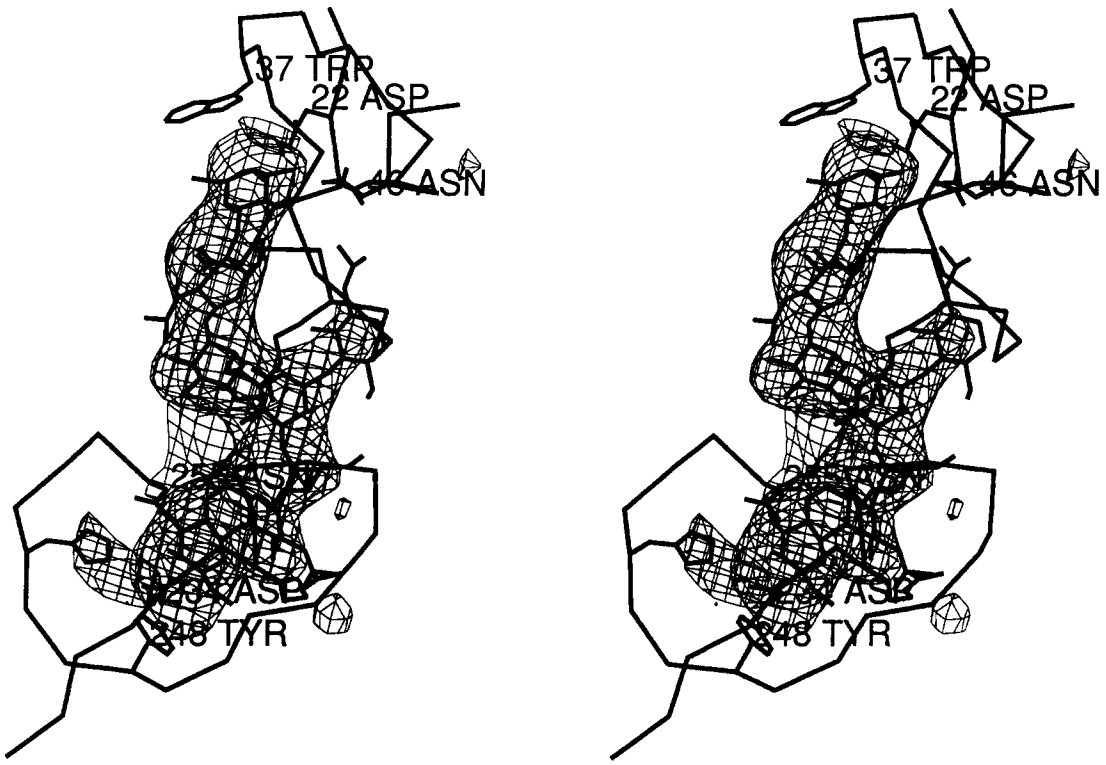


Fig. 4. Difference Fourier density of the biantennary oligosaccharide. The difference density is at 5 Å resolution, with the model of the biantennary oligosaccharide superimposed.

The B-chain of ricin is a glycopeptide and electron density for N-linked oligosaccharides is found confluent with the density for the side chains of Asn 95 and Asn 135, both of which are in the consensus sequence for sugar attachment Asn-X-Thr. Although chemical analysis of the covalently bound sugars of RTB indicates some heterogeneity, the common core of the deduced structures, and the structure fit to the electron density is shown below.[17]

*Man*α1-6\
  *Man*β1-4*GlcNAc*β1-4*GlcNAc*
*Man*α1-3/

The electron density in the current map shows both sugar chains extending into solvent. The density washes out rapidly beyond the mannose branch point, presumably because of high thermal motion. The first GlcNAc residue of the sugar bonded to Asn 95 lies against Trp 93 in a manner reminiscent of the interaction between the sugars and the aromatic residues seen in the reversible sugar binding sites. The acetyl group from the second GlcNAc at Asn 95 seems to be involved in a hydrogen bond with the side chain of Gln 91. The remaining sugar residues extend into solvent from a cleft between the A-chain and the B-chain with no further contacts to the protein. The sugar bound to Asn 135 makes no contact with the protein and extends directly into the solvent.

B-Chain Folding

RTB has no major segments of a helix or β sheet. The only secondary structural element found is the recently defined Ω-loop.[18] An idealized Ω-loop is a compact, contiguous segment of polypeptide that traces a "loop-shaped" path in three-dimensional space; the main chain resembles a Greek omega (Ω). They are invariably situated at the surface of the protein, often have disulfide bonds securing the neck of the loop and are often implicated in molecular function. The loops are often about the size of a typical exon and may serve as modules of evolutionary exchange. Omega loops appear to be formed within each subdomain of RTB and are particularly easy to see in the α and β units in which disulfide bonds secure the neck. In addition, structures that meet the loop criterion can be defined for chain segments joining several subdomains. The omega loops are often stabilized by short, irregular runs of hydrogen bonds between backbone nitrogen and carbonyl oxygen atoms. The backbone hydrogen bonding pattern for RTB is shown in Figure 5; although irregular, the pattern is quite complex and appears to offer a considerable measure of stability to the overall structure.

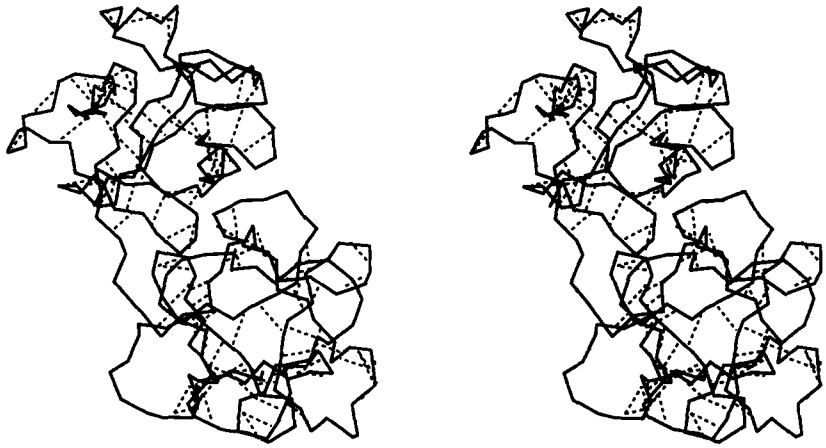


Fig. 5. Ricin B-chain hydrogen bonding pattern. The carbon alpha backbone tracing of the B-chain is a solid line and the hydrogen bonds between carbonyl oxygens and backbone nitrogens are shown as dashed lines. The stereo view is down the pseudo threefold axis of domain 1.

Each domain of the B-chain is composed of three subdomains arranged to form a pseudo threefold rotational axis with the linking λ-subdomain extending across the threefold structure (Fig. 6). At the core of each domain is an interlocking, pseudosymmetric arrangement of hydrophobic residues forming a short cylinder that is capped on both ends by a less regular arrangement of hydrophobic residues. The axis of the cylinder is aligned with the pseudo threefold rotational axis. In general, each subdomain contributes one Leu, one Trp, and one Ile to the core, as indicated in Figure 2. The γ subdomains, however, are truncated and lack the conserved Ile. That residue is provided to the core by the λ subdomain, preserving its pseudo threefold symmetry; Ile13 and Ile144 are contributed to the domain 1 and 2 cores, respectively. As seen in Figure 8, the three Ile side chains extend in toward the pseudo threefold axis, forming the top layer of the core cylinder. Each Leu side chain lies directly below the Ile side chain from an adjacent subdomain; they also project toward the central axis and form the bottom layer of the core. The Trp side chains lie with their long dimension parallel to the core axis. They interdigitate with the lobes formed by the stacked aliphatic side chains. Together, these nine residues form a compact, solid, and interlocking hydrophobic core.

The core described above and shown in Figure 6 is the most readily recognized aspect of the threefold assembly of the domain, but the pattern continues beyond this main core. Figure 2 shows that each subdomain has two other positions that are invariably hydrophobic, although the amino acid sequences differ slightly. They are marked with open triangles. These residues also pack in the hydrophobic molecular center and, indeed, manifest the same pseudo symmetry as the more regular core. For example, in domain 1 Ile34, Val75 and Leu118 are contributed by the α, β, and γ subdomains, respectively. They pack around the pseudo threefold axis on a plane immediately below that shown in Figure 6. This triangle of residues make van der Waals contact with each other and with the core above. Clearly the arrangement is less perfect than the core, since each element of this triangle is chemically different. In the same plane as this triangle of hydrophobic side chains is a second triangle, also centered roughly on the pseudo threefold, but at a larger radius. The respective subdomains contribute Leu36, Ile77, and Val120 to it. They make contact with the first triangle and also some with the edges of the core.

As is typical of globular proteins, the charged residues of the B-chain residue predominantly near the solvent accessible surface. They cluster near the ends of the Ω-loops, which tend toward the surface of each domain. The hydrophobic residues that cap the top and bottom of each core are shielded from solvent by a loose network of charged residues.

As described in Materials and Methods, the six major subdomains of RTB were superimposed in a least-squares sense; this is shown in Figure 7. It is clear that despite considerable divergence in their sequences, the folding of the individual subdomains is very similar. In a given domain three subdomains pack around a pseudo threefold axis. The horizontal loops shown in Figure 7 point in toward the core of the domain; those portions of the subdomains are particularly conservative. The solvent exposed Ω loops, on the right in the figure, are roughly perpendicular to the core loops and are less conservative in structure. The lactose sugar for subdomain 1α is shown on the right side of figure to indicate the position of the binding site with respect to the overall subdomain shape. Galactose sits in a shallow cleft, whereas glucose is in the surrounding solvent. Note that immediately below the bound galactose, subdomains 1α and 2γ exhibit a sharp bend in the chain.

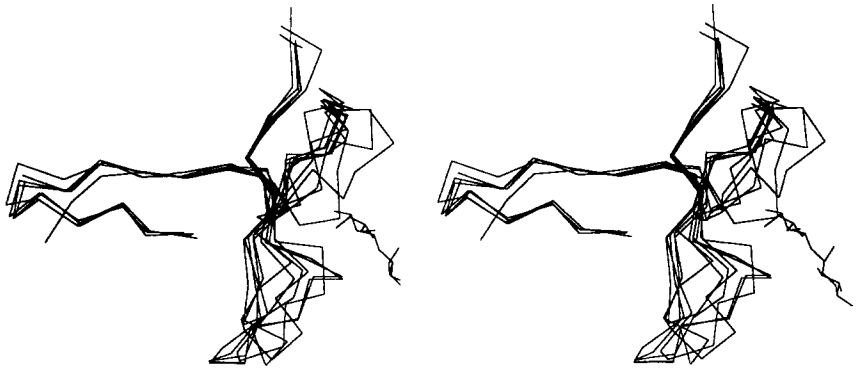


Fig. 7. Superposition of the six subdomains. The alpha carbons of the subdomains were superimposed in a least squares sense on subdomain 1α. Lactose, as bound to subdomain 1α, is shown to the right for orientation.

This is the bottom of the binding pocket, which appears to make nonspecific contacts with the ligand.

RTA-RTB Association

In native ricin, the A- and B-chains are connected by a single disulfide bridge between Cys 259 of RTA and Cys 4 of RTB. This bond can be reduced without loss of toxic activity. The thermodynamic parameters of the association between ricin chains have been evaluated using analytical ultracentrifugation.[19] The free energy governing the micromolar dissociation constant showed positive values for entropy and enthalpy, suggesting that hydrophobic interactions are responsible for chain association. The A-B disulfide bond therefore may function only in maintaining chain association, vital to toxicity, at very low toxin concentrations.

A model for ricin A-chain has been produced based upon the recent crystallographic refinement.[20] Solvent accessibility calculations show the area of contact between RTA and the RTB is 1600 Å^2, roughly 14% of the total surface area of either chain alone. Much of the interface is solvated by two pools of trapped water molecules, as seen in Figure 8. As predicted by the data discussed above, the interface contains various aromatic rings and aliphatic side chains, which interact in a rather disordered manner and presumably secure heterodimer association. The hydrophobic contact area is surprisingly narrow, as shown in Figure 8. The main region of interaction appears to lie near the carboxy terminus of the B-chain. Here the ring of Phe 262 from RTB contacts the side chains of Phe 140 of RTB and Phe 240 of RTA. These aromatic residues are oriented in the energetically favorable manner described by Burley and Petsko[21] as "edge-to-face." In addition to hydrophobic interactions, several polar contacts are formed between the A-chain and the B-chain, such as the ion pair formed between Glu41 of RTA and Lys219 of RTB. Table I summarizes the interactions of specific residues at the interface between the two chains.

DISCUSSION

The architecture of the ricin B-chain is a remarkable and enlightening example of the forces at play in the development of a highly specialized, complex protein from a much simpler functional unit. Rarely is the apparent evolutionary history of a protein so clearly recorded in its structure. The well-known maxim that tertiary structure is more highly conserved than primary sequence is born out in the B-chain; the individual subdomains show a sequence similarity of about 20%, yet have nearly identical three-dimensional folds. The presence of the rather elegant hydrophobic core also supports the notion of a common ancestor. This ancestor was probably quite similar to the 1α subdomain, a unit that retains the ancient sugar binding activity and possesses a full compliment of hydrophobic residues involved in trimer self-assembly. In the natural evolution of the B-chain, gene duplication is the means by which the end is achieved, and we have proposed a series of gene duplication events that we feel most likely gave rise to the modern B-chain.[15] This gene has fused with the independently evolved A-chain gene to create the toxin heterodimer.

At least one more step in the ongoing evolution of the ricin molecule is manifest in the castor plant. The seeds of the castor plant contain, in addition to ricin, the *Ricinus communis* agglutinin (RCA), a dimer composed of two ricinlike molecules in the form $(AB)_2$.[22] RCA, like ricin, binds two galactosides and not four as might be expected from a ricinlike dimer.[23] Also, RCA does not bind GalNAc.[24] This suggests that RCA has lost the ability to bind sugar in the domain 2 binding sites. The gene and corresponding amino acid sequence of the RCA B-chain has been determined.[25] All of the residues implicated in sugar binding are intact in site 1 of the RCA B-chain. There are a number of changes, however, in site 2. Probably the most interesting is the conversion of Tyr248 of RTB to a His residue in RCA B-chain. This residue would be expected to form the

Table I. Interactions Across the A/B Interface

Polar Interactions		Distance
A-chain	B-chain	(Å)
Arg2580	Ala 1N	3.10
Ala260 N	Asp 2 O	3.17
His 40 ND1	Asp 94 OD2	2.93
Arg234 NH1	Val141 O	2.79
Glu 41 OE2	Lys219NZ	2.77
Gln182 NE2	Asn220 O	2.87
Ile 249 O	Asn220 ND2	2.78
Ile 252 O	Asn220 ND2	3.01
Arg235 N	Phe262 O	2.88

Hydrophobic Residues*	
A-chain	B-chain
Tyr183	Phe262
	Pro260
Leu207	Phe262
Phe240	Phe140
	Phe262
Ile 247	Phe140
Pro250	Phe218
	Pro260
Ile 251	Pro260
	Phe262

*Residues are listed when carbon atoms from an RTA side chain make van der Waals contact (distance < 4.2 Å) with side chain carbons from RTB.

top of the galactose binding pocket and was shown above to make hydrophobic contact with the nonpolar face of the sugar. Since the substituted His is charged a considerable fraction of the time, it might be expected to interfere with this type of interaction. The organization of subunits in RCA is unknown, and it is unclear what selective pressures have led to the loss of the site 2 activity. It is possible that the components of RCA associate in such a way as to block the domain 2 sugar binding sites of both molecules leading to the eventual loss of sugar binding function.

A final point worth discussion is our observation concerning the alignment of residues based on sequence overlap and structure as shown in Figure 2. It seems that residues that may have had a common ancestor can have different structural roles. As described above, Trp 160 is a homologue of Trp 37, a functionally relevant residue; their positions in the structures of their respective subdomains are no longer equivalent, however. Structural predictions or mutagenesis experiments based on the assumption that these would retain their structure and function between subdomains would be in error. In general, sequence homologies are a powerful predictive tool for structure-function analysis, but clearly there are limitations.

ACKNOWLEDGMENTS

We are grateful to Michael Ready for helpful discussions and to Raquelle Smalley for her help in preparing the figures. This work was supported by grants GM 30048 and GM 35989 from the National Institutes of Health and by a grant from the Foundation for Research.

REFERENCES

1. Olnes, S., Pihl, A. The molecular action of toxins and viruses. In: "The Molecular Action of Toxins and Viruses," Cohen, P. and Van Heynigen, S. (eds.). New York: Elsevier, 1982:52–105.
2. Youle, R. J., Colombatti, M. Hybridoma cells containing intracellular anti-ricin antibodies show ricin meets secretory antibody before entering the cytosol. J. Biol. Chem. 262:4676–4682, 1987.
3. Johnson, V. G., Youle, R. J. Intracellular routing and membrane translocation of diptheria toxin and ricin. In: "Intracellular Trafficking of Proteins," Clifford J. Steer and John A. Hanover (eds.). Cambridge: Cambridge University Press (in press).
4. Zentz, C., Frenoy, J. P., Bourrillon, R. Binding of galactose and lactose to ricin: Equilibrium studies. Biochim. Biophys. Acta. 536:18–26, 1978.
5. Shimoda, T., Funatsu, G. Binding of lactose and galactose to native and iodinated ricin D. Agr. Biol. Chem. 49:2125–2130, 1985.
6. Houston, L., Dooley, L. Binding of two molecules of 4-methylumbelliferyl galactose or 4-methylumbelliferyl N-acetylgalactosamine to the B chains of ricin and *Ricinus communis* agglutinin and to purified ricin B chain. J. Biol. Chem. 257:4147–4151, 1982.
7. Villafranca, J. E. The crystal structure of ricin at low resolution. Ph.D. dissertation. University of Texas at Austin, 1980.
8. Yamasaki, N., Hatakeyama, T., Funatsu, G. Ricin D–saccharide interaction as studied by ultraviolet difference spectroscopy. J. Biochem. 98:1555–1560, 1985.
9. Hatakeyama, T., Yamasaki, N., Funatsu, G. Identification of the tryptophan residue located at the low-affinity saccharide binding site of ricin D. J. Biochem. 100:781–788, 1986.
10. Baenziger, J. U., Fiete, D. Structural determinants of *Ricinus communis* agglutinin and toxin: Specificity for oligosaccharides. J. Biol. Chem. 254:9795–9799, 1979.
11. Montfort, W., Villafranca, J. E., Monzingo, A. F., Ernst, S., Katzin, B., Rutenber, E., Xuong, N. H., Hamlin, R., Robertus, J. D. The three-dimensional structure of ricin at 2.8 Å. J. Biol. Chem. 262:5398–5403, 1987.
12. Rutenber, E., Ready, M., Robertus, J. D. Structure and evolution of ricin B chain. Nature 326:624–626, 1987.
13. Rutenber, E., Katzin, B. J., Collins, E. J., Mlsna, D., Ernst, S. E., Ready, M. P., Robertus, J. D. Crystallographic refinement of ricin to 2.5 Å. Proteins 10:240–250, 1991.
14. Villafranca, J., Robertus, J. D. Crystallographic study of the anti-tumor protein ricin. J. Biol. Chem. 116:331–335, 1977.
15. Jones, T. A., In: "Computational Crystallography," D. Sayre, (ed.). Oxford: Oxford University Press, 1982:303.
16. Molecular Recognition and Protein-Carbohydrate Interactions. Transactions of the American Crystallographic Association, Vol 25, Einspahr, H. and Ward, K.B. (eds.). Buffalo, NY: ACA, (in press).
17. Kimura, Y., Hase, S., Kobayashi, Y., Kyogoku, Yd., Ikenaka, T., Funatsu, G. Structures of sugar chains of ricin D. J. Biochem. 103:944–949, 1988.
18. Leszczynski, F.J., Rose, G.D. Loops in globular proteins: A novel category of secondary structure. Science 234:849–855, 1986.
19. Lewis, M.S., Youle, R.J. Ricin subunit association: Thermodynamics and the role of the disulfide bond in toxicity. J. Biol. Chem. 261:11571–11577, 1986.
20. Katzin, B.J., Collins, E.J., Robertus, J.D. The structure of ricin A chain at 2.5 Å Proteins 10:251–259, 1991.
21. Burley, S.K., Petsko, G.A. Aromatic-aromatic interaction: A mechanism of protein structure stabilization. Science 229:23–28, 1985.
22. Nicolson, G.L., Blaustein, J., Etzler, M.E. Characterization of two plant lectins from *ricinus communis* and their

quantitative interaction with murine lymphoma. Biochemistry 13:196–204, 1974.
23. Podder, S.K., Surolia, A., Bachhawat, B.K. On the specificity of carbohydrate-lectin recognition. Eur. J. Biochem. 44:151–160, 1974.
24. Nicolson, G.L., Blaustein, J. The interaction of *ricinus communis* agglutinin with normal and tumor cell surfaces. Biocim. Biopys. Acta 266:543–547, 1972.
25. Roberts, L. M., Lamb, F. I., Pappin, D. J. C., Lord, M. J. The primary sequence of *Ricinus communis* agglutinin: Comparison with ricin. J. Biol. Chem. 260:15682–15686, 1985.
26. Vitetta, E.S., Yen, N. Expression and functional properties of genetically engineered ricin B chain lacking galactose binding activity. Biochim. Biophys. Acta 1049:151–157, 1990.