Isolation and Molecular Cloning of a Novel Type 2 Ribosome-inactivating Protein with an Inactive B Chain from Elderberry (Sambucus nigra) Bark*

(Received for publication, July 9, 1996, and in revised form, September 17, 1996)

Els J. M. Van Damme‡§, Annick Barre¶, Pierre Rougeé†, Fred Van Leuven‡, and Willy J. Peumans‡***

From the ‡Laboratory of Phytopathology and Plant Protection, Katholieke Universiteit Leuven, Willem de Croylaan 42 and the ¶Center for Human Genetics, Katholieke Universiteit Leuven, Herestraat 49, 3001 Leuven, Belgium and the †Institut de Pharmacologie et Biologie Structurale, Université Paul Sabatier, 35 Chemin des Maratchers, UPR CNRS 9062, 31062 Toulouse Cedex, France

One of the predominant proteins in the bark of elderberry (Sambucus nigra) has been identified as a novel type 2 ribosome-inactivating protein that exhibits a normal RNA N-glycosidase activity, but is devoid of carbohydrate binding activity. Sequence analysis of the corresponding cDNA clones revealed a striking homology to the previously cloned bark lectins from elderberry, suggesting that the new protein is a lectin-related protein. Molecular modeling of the protein confirmed that its A chain is fully active, whereas its B chain contains two functionally inactive carbohydrate-binding sites. These findings not only demonstrate for the first time the occurrence of a type 2 ribosome-inactivating protein with an inactive B chain, but also offer interesting perspectives for the synthesis of immunotoxins with an improved selectivity.

Seasonally fluctuating bark proteins play an important role in the nitrogen metabolism of deciduous trees of the temperate regions (1). Although these presumed storage proteins have been found in several tree and shrub species, only a few of them have been purified and characterized in some detail. Surprisingly, all bark proteins identified so far are lectins or ribosome-inactivating proteins (RIP)† (1). Lectins have been defined as proteins possessing at least one noncatalytic domain that binds reversibly to specific mono- or oligosaccharides (2). They are widespread in the plant kingdom and form a heterogeneous group of proteins (3). Unlike lectins, RIP are a homologous group of proteins that possess a highly specific RNA N-glycosidase activity and are capable of catalytically inactivating ribosomes (4). Type 1 and 2 RIP are distinguished according to their molecular structure (4). Type 1 RIP consist of a single catalytically active subunit of ~30 kDa, whereas type 2 RIP are composed of A chains with N-glycosidase activity and B chains with a carbohydrate binding activity comparable to that of lectins. Since type 2 RIP are fully capable of agglutinating cells and/or precipitating glycoconjugates, they are also considered as lectins.

Lectins have been identified as major proteins in the bark of the legume trees Robinia pseudoacacia (black locust) (5, 6), Sophora japonica (Japanese pagoda tree) (7), and Cladrastis lutea (yellow wood) (8). Besides in legumes, bark lectins have also been found in different species of the genus Sambucus (elderberry; family Caprifoliaceae) (9–11). Recently, molecular cloning of the cDNAs encoding the Neu5Acα2–6Gal/GalNAcβ2-specific agglutinins from Sambucus nigra and Sambucus sieboldiana revealed that both lectins are type 2 RIP with an unusual specificity and molecular structure (12, 13). Similarly, molecular cloning of Nigrin b (now called SNAV) confirmed that this protein is a typical GalNAc-specific type 2 RIP (14, 15).

A reinvestigation of the bark proteins of elderberry resulted in the isolation of a novel type 2 RIP called SNL RP. Characterization of the protein and molecular cloning of its corresponding cDNA demonstrated that this novel RIP has the same overall structure as the classical type 2 RIP, but contains a B chain that is devoid of carbohydrate binding activity. Our findings not only demonstrate for the first time the occurrence of a type 2 RIP with an inactive B chain, but also provide additional evidence that the bark of elderberry is highly specialized in the accumulation of RIP and related proteins.

EXPERIMENTAL PROCEDURES

Materials—All experiments were carried out with bark samples obtained from a single S. nigra tree. Samples were collected as described previously (12). Radiolabeled riposides were obtained from ICN. A cDNA synthesis kit, the multifunctional phagemid pT7T318U, restriction enzymes, and DNA-modifying enzymes were obtained from Pharmacia Biotech Inc. Escherichia coli XLI Blue competent cells were purchased from Stratagene.

Isolation of SNL RP from Elderberry Bark—Elderberry bark was lyophilized and powdered in a coffee mill. Fifty g of the finely ground tissue were extracted in 500 ml of 0.2 M NaCl containing 0.2 g/liter ascorbic acid (adjusted to pH 6.5) by stirring at room temperature for 1 h. The homogenate was centrifuged at 9000 × g for 15 min, and the resulting supernatant was decanted and filtered through glass wool to remove the floating particles. SNAI and the mixture of SNAIL and SNAILI and

* This work was supported in part by Grant OT/94/17 from the Catholic University of Leuven, by grants from CNRS and the Conseil Régional de Midi-Pyrénées, by Grant G.0223.97 from the National Fund for Scientific Research (Belgium), and by Grant 70.047.90 from the National Funds voor Wetenschappelijk Onderzoek-Levenslijn. 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.

† The nucleotide sequence(s) reported in this paper has been submitted to the GenBankTM/EBI Data Bank with accession number(s) U58357 and U58358.

‡ Senior Research Assistant of the National Fund for Scientific Research (Belgium). To whom correspondence should be addressed. Tel.: 32-16-322379; Fax: 32-16-322976.

§ Senior Research Assistant of the National Fund for Scientific Research (Belgium).

¶ Research Director of the National Fund for Scientific Research (Belgium).

1 The abbreviations used are: RIP, ribosome-inactivating protein(s); SNA, S. nigra agglutinin; SNL RP, S. nigra lectin-related protein; LEC-SNA, cDNA encoding SNA; LRPSN, cDNA encoding SNL RP; PAGE, polyacrylamide gel electrophoresis; HCA, hydrophobic cluster analysis.

2 Neuraminic acid is acetylated at the hydroxyl group at C8.
SNLRP were removed from the partially purified extract by consecutive affinity chromatography on fetuin-Sepharose 4B and GalNAc-Sepharose 4B, respectively, as described previously (12, 15). The lectin-depleted extract (in 1.5 M ammonium sulfate) was loaded onto a column (5 × 5 cm, 100-ml bed volume) of phenyl-Sepharose 4B equilibrated with 5 M ammonium sulfate in 50 mM sodium acetate (pH 5.5). After washing the column with 1.5 M ammonium sulfate until the A_{280} fell below 0.01, the proteins were eluted with 50 mM Tris-HCl (pH 9.0). SDS-polyacrylamide gel electrophoresis (PAGE) showed that the fraction retained on the phenyl-Sepharose column contained almost exclusively a single polypeptide band. Further purification was achieved by ion exchange chromatography and gel filtration. The proteins were dialyzed against 20 mM sodium formate (pH 3.8) and loaded onto a column (15 × 2.6 cm, 75-ml bed volume) of S Fast Flow (Pharmacia, Uppsala) equilibrated with the same buffer. After loading the proteins, the column was washed with 200 ml of buffer, and the proteins were eluted with a linear gradient (500 ml) of increasing NaCl concentration (from 0 to 1 M). The proteins eluting in the main peak were dialyzed against water, lyophilized, and dissolved in 20 ml of phosphate-buffered saline (1.5 mM KH₂PO₄, 10 mM Na₂HPO₄, 3 mM KCl, and 140 mM NaCl (pH 7.4)). Any insoluble material was precipitated by centrifugation at 12,000 × g for 10 min, and the supernatant was chromatographed on a column (40 × 5 cm, 800-ml bed volume) of Sephacryl 100 equilibrated with phosphate-buffered saline. The protein eluting in the main peak was essentially pure SNLRP.

**Gel Filtration—**Analytical gel filtration of the purified proteins was performed on a Pharmacia Superose 12 column using phosphate-buffered saline containing 0.2 mM galactose (to avoid possible binding to the column) as running buffer. Molecular mass reference markers were catalase (240 kDa), Ricinus communis agglutinin (120 kDa), ricin (60 kDa), chymotrypsigenin (25 kDa), and the elderberry lectins SNAI (240 kDa) and SNAV (120 kDa).

**Analytical Methods—**Total neutral sugar was determined by the phenol/H₂SO₄ method (16), with d-glucose as standard. Agglutination assays were conducted using human, rabbit, and pigeon erythrocytes (15).

Lectin preparations were analyzed by SDS-PAGE using 12.5–25% acrylamide gradient gels as described by Laemmli (17). Proteins separated by SDS-PAGE and electrophoretically transferred to a polyvinylidene difluoride membrane were sequenced on an Applied Biosystems Model 477A protein sequencer interfaced with an Applied Biosystems Model 120A on-line analyzer.

Cyanogen bromide cleavage of the proteins (2 mg) was done in 0.1 ml of 70% formic acid containing 10 mg of cyanogen bromide. After incubation for 15 h at 37 °C (in the dark), peptides were recovered by evaporation under vacuum.

**Construction and Screening of a cDNA Library—**A cDNA library was constructed with polya(R)-rich mRNA isolated from *S. nigra* bark using the cDNA synthesis kit from Pharmacia (12). cDNA fragments were inserted into the EcoRI site of the multifunctional phagemid pT7-18U. The library was propagated in *E. coli* XL1 Blue.

The cDNA library was screened using 5′-labelled synthetic oligonucleotide derived from the consensus sequence of the A chain (5′-CC-CTT GCC GCT TCC GAG GAC ATT TGA ATT-3′) of type 2 RIP and oligonucleotides that react exclusively with LECSNAI (5′-GGG GGC TAG TAA AAA AAA-3′) and LECSNAV (5′-GAC GGA ACG TGT ACG-3′) cDNA clones. Hybridization was carried out overnight as described previously (18). After washing, filters were blotted dry, wrapped in Saran Wrap, and exposed to Fuji film overnight at 70 °C.

All colonies that reacted positively with the consensus sequence but were negative with the two other oligonucleotides were selected and re-screened at low density using the same conditions. Plasmids were isolated from purified single colonies on a miniprep scale using the alkaline lysis method as described by Mierendorf and Pfeffer (19) and isolated from purified single colonies on a miniprep scale using the alkaline lysis method as described by Mierendorf and Pfeffer (19) and sequenced by the dideoxy method (20). DNA sequences were analyzed by the dideoxy method (20). DNA sequences were analyzed by the dideoxy method (20). DNA sequences were analyzed by the dideoxy method (20). DNA sequences were analyzed by the dideoxy method (20). DNA sequences were analyzed by the dideoxy method (20). DNA sequences were analyzed by the dideoxy method (20). DNA sequences were analyzed by the dideoxy method (20). DNA sequences were analyzed by the dideoxy method (20). DNA sequences were analyzed by the dideoxy method (20). DNA sequences were analyzed by the dideoxy method (20). DNA sequences were analyzed by the dideoxy method (20). DNA sequences were analyzed by the dideoxy method (20). DNA sequences were analyzed by the dideoxy method (20). DNA sequences were analyzed by the dideoxy method (20). DNA sequences were analyzed by the dideoxy method (20).

**Molecular Modeling of SNLRP—**The hydrophobic cluster analysis (HCA) (21, 22) was performed to delineate the structurally conserved regions along the amino acid sequences of the A and B chains of LRPSN1, LRPSN2, and ricin, which was used as a model. HCA plots were generated on a Macintosh LC using the program HCA-Plot2 (Doriane, Paris).

Molecular modeling of A and B chains of LRPSN1 and LRPSN2 was carried out on a Silicon Graphics Iris 4D25G workstation using the programs InsightII, Homology, and Discover (Biosym Technologies, San Diego, CA). The coordinates of ricin (code 2aai) were taken from the coordinates of ricin (code 2aai).

**RESULTS**

Nomenclature of the *S. nigra* Bark Lectins/RIP and Their Genes—Elderberry bark contains at least four different proteins that have been identified as type 2 RIP or lectins. To avoid confusion in the designation of the proteins and their genes, the following nomenclature has been used. (i) SNAI and LECSNAI refer to the Neu5Acα2–6Gal/GalNAcα2–specific type 2 RIP and the corresponding cDNA clone, respectively (12). (ii) SNAI refers to the GalNAc-specific lectin composed of two subunits that are homologous to the B chains of SNAV (10). (iii) SNAV and LECSNAV refer to the GalNAc-specific type 2 RIP and the corresponding cDNA clone, respectively (15). (iv) SNLRP and LRPSN refer to the novel nonagglutinating type 2 RIP and its corresponding cDNA clone.

**Isolation and Characterization of SNLRP—**SNLRP was purified from a lectin-depleted extract using a combination of hydrophobic interaction chromatography, ion exchange chromatography, and gel filtration. Starting from 50 g of lyophilized bark, 127 mg of SNLRP were obtained. A comparable yield has been reported for SNAI (1 mg/g of wet bark) (9) and SNAII (2.6 mg/g of dry bark) (15), suggesting that SNLRP is as abundant as the latter two bark proteins.

The molecular structure of SNLRP was determined by SDS-PAGE and gel filtration. Unreduced SNLRP yielded two polypeptide bands of −60 and 62 kDa, whereas the reduced protein yielded several bands with molecular masses ranging between 30 and 34 kDa (Fig. 1). It should be mentioned that the reduced samples also yielded the same high molecular mass bands as the unreduced samples probably because of the reassociation of the A and B chains during electrophoresis. Gel filtration of the native protein on a Superose 12 column yielded a single symmetrical peak eluting with an apparent molecular mass of 70 kDa (data not shown). According to these results, SNLRP is composed of two disulfide bridge-linked subunits of −32 kDa.

Determination of the carbohydrate content of SNLRP yielded a value of 3.0% (by mass). Assuming a molecular mass of 170 Da per monosaccharide, the number of sugar residues amounts to −11/native molecule of 60 kDa. In analogy to the ricin glycan formed by several cycles of deepest descent and conjugate gradient using the cvff force field of Discover. The program TurboFrodo (BIOGRAPHICS, Marseille, France) run on a Silicon Graphics Indigo R3000 workstation was used to perform the superimposition of the models.

The amino acid sequence alignments were performed on a MicroVAX 3100 (Digital, Ewy, France) using the ialign program of PIR/NBFRF (Washington, D. C.). MacClade (24) was run on a Macintosh LC 630 to build a parsimony phylogenetic tree relating the different RIP.
SNLRP

CNBR fragment 1: AGAKW ISYRN FLGEL QDLVT

CNBR fragment 2: SKEIQ GATLG GTFAH VVQLQ IVNXD TAI

SNLRP1

34 kDa polypeptide: ATPPN YPSVS LKMAG AKWIS DDEKX TVVVD TRRIS GRDGL

SNLRP2

30 kDa polypeptide: APPNY PSVSL KMAGA KWISY

32 kDa polypeptide: DDEKX TVVVD TRRIS GRDGL

Fig. 2. N-terminal amino acid sequences of SNLRP and cyanogen bromide cleavage fragments.

chains, which consist of 6–8 monosaccharide residues, the native protein SNLRP contains on average 1.5 oligosaccharide side chains/molecule.

N-terminal sequencing of the blotted peptides yielded little conclusive information because the signal was weak, and double peaks were obtained at almost all positions of the sequence (data not shown). Therefore, SNLRP was cleaved with cyanogen bromide, and the fragments were sequenced. Several single and double sequences were obtained (Fig. 2) that exhibited ~50% sequence identity to the deduced amino acid sequences of the A and B chains of SNAI and SNAV. On the basis of these results, it was presumed that SNLRP is a type 2 RIP composed of disulfide bridge-linked A and B chains.

SNLRP Inhibits Protein Synthesis in a Reticulocyte Lysate—The possible ribosome inactivating activity of SNLRP was checked by measuring its inhibitory effect on protein synthesis in a reticulocyte lysate. Native (i.e. unreduced) SNLRP strongly reduced the incorporation of labeled amino acids, with the concentration required for 50% inhibition being ~0.5 μg/ml. These results leave no doubt that SNLRP strongly inhibits mammalian ribosomes. Further details on the ribosome inactivating activity, toxicity, and N-glycosidase activity of SNLRP will be published elsewhere.3

SNLRP Does Not Agglutinate Erythrocytes and Does Not Bind to Immobilized Sugars and Glycoproteins—To assess the possible lectin activity of SNLRP, the agglutination activity of the purified protein was assayed with different types of red blood cells. SNLRP failed to agglutinate untreated as well as trypsin-treated human, rabbit, and pigeon erythrocytes, even at a final concentration of 10 mg/ml. Because this apparent lack of agglutination activity does not necessarily imply that SNLRP has no carbohydrate binding activity, the binding of purified SNLRP to various immobilized sugars (galactose, GalNAc, lactose, mannose, fucose, GlcNAc, and GlcNAcα) and glycoproteins ( fetuin, asialofetuin, mucin, asialomucin, thyroglobulin, and ovomucoid) was checked. Since the protein was not retained on any of these affinity matrices, one can reasonably assume that SNLRP has no carbohydrate binding activity.

Isolation and Characterization of cDNA Clones Encoding SNLRP—The cDNA clones encoding the putative novel RIP were isolated following a strategy whereby only cDNA clones differing from LECSNAI and LECSNAV were recovered. A total cDNA library was screened with an oligonucleotide corresponding to the consensus sequence of the A chain (5'-CTT-GCC GCT TCC GAG ACC ATT TGA AT-3') and oligonucleotides corresponding to the N-terminal sequences of the B chain of SNAI (5'-GGG GCC GGC TAG TAA AAA-3') and SNAV (5'-GAC GGG GAA ACG TGT ACG-3'), which react exclusively with cDNA clones encoding LECSNAV and LECSNAI, respectively. All colonies that reacted positively with the consensus sequence but were negative with the two other oligonucleotides were selected and purified. Subsequent sequence analysis revealed the occurrence of two groups of cDNA clones (LRPSN1 and LRPSN2) encoding proteins containing the N-terminal amino acid sequences of SNLRP as well as the sequences of all the cyanogen bromide fragments (Fig. 3).

LRPSN1 contains a 1734-base pair open reading frame encoding a polypeptide of 578 amino acids with one possible initiation codon at position 13 of the deduced amino acid sequence. Assuming that this methionine is used as the translation initiation site, the primary translation product is a polypeptide of 566 amino acids (62,733 Da) that contains the N-terminal amino acid sequences of both the A and B chains of SNLRP. According to the rules of von Heijne (25), a signal peptide is cleaved between residues 22 and 23 of the RIP precursor. The resulting polypeptide of 60,333 Da is further processed by one or more proteolytic cleavages to yield the A and B chains of mature SNLRP. Since the B chain starts with the sequence DDEKCTVVDV, a cleavage must take place between residues 305 and 306 of the primary translation product, resulting in an A chain of 31,592 Da and a B chain of 28,759 Da. Taking into consideration that the A chain of the SNAI homologue of S. sieboldiana is processed after the sequence VTS (13), it is possible that also the A chain of SNLRP1 is post-translationally processed at this position. The deduced amino acid sequence of LRPSN1 contains five putative N-glycosylation sites at positions 114, 127, 259, 472, and 522 of the precursor (Fig. 3).

LRPSN2 strongly resembles LRPSN1. It encodes a polypeptide of 573 amino acids with one possible initiation codon at position 9 of the deduced amino acid sequence. Translation starting with this methionine results in a polypeptide of 565 amino acids (62,324 Da). A possible cleavage site for the processing of the signal peptide was identified between residues 25 and 26 of the primary translation product, which is in good agreement with the N-terminal amino acid sequence of the A chain. Cleavage of the signal peptide at this site will result in a lectin polypeptide of 59,769 Da. Further processing of this polypeptide to yield the A and B chains of mature SNLRP implies a cleavage between residues 304 and 305, resulting in an A chain of 30,996 Da and a B chain of 28,791 Da. The deduced amino acid sequence of LRPSN2 contains two putative N-glycosylation sites at positions 258 and 521 of the precursor (Fig. 3). To corroborate the possible occurrence of two different isoforms, a total preparation of SNLRP was analyzed by ion exchange chromatography on a Mono-S column. The protein eluted in two peaks (data not shown), which are referred to as SNLRP1 and SNLRP2. SDS-PAGE revealed that SNLRP1 and SNLRP2 are composed of polypeptides with a slightly different size. As shown in Fig. 1, unreduced and reduced isoform 1 yielded a single polypeptide band of 62 and 34 kDa, respectively. Unreduced isoform 2 migrated in a doublet of 60 and 62 kDa, respectively.
FIG. 3. Comparison of the deduced amino acid sequences of the cDNA clones encoding SNLRP1, SNLRP2, SNAI, and SNAV. The arrowhead indicates the processing site for the cleavage of the signal peptide. Dashes denote gaps introduced to obtain maximal homology. The determined N-terminal amino acid sequences of the A and B chains of SNLRP1 and SNLRP2 and their cyanogen bromide cleavage fragments are underlined. Putative glycosylation sites are shown in boldface. Since the first ATG codon is probably used as the translation initiation site, the deduced amino acids preceding this methionine are shown in lower-case letters. The numbers above the sequences refer to the positions of the residues along the A and B chains of SNLRP1.
kDa, whereas the reduced protein yielded several polypeptides with molecular masses ranging between 30 and 34 kDa. N-terminal sequencing of the 34-kDa polypeptide of isoform 1 yielded a double sequence that corresponds to the N terminus of the A and B chains of SNLRP1. The upper and lower bands of isoform 2 yielded unique sequences corresponding to the N-terminal sequences of the B and A chains, respectively, of SNLRP2 (Fig. 2).

**Molecular Modeling of SNLRP**—SNLRP was modeled using the coordinates of ricin, the three-dimensional structure of which has been resolved by x-ray crystallography (23, 26). Although it must be emphasized that the results of these modeling studies have to be interpreted with care, they can give interesting information about structural homologies between related proteins.

A comparison of the HCA plots of the A chains of ricin, LRPSN1, and LRPSN2 indicates that the secondary structural features (α-helices and β-sheets) are readily conserved (Fig. 4). Despite some discrepancies due to the occurrence of a few deletions or insertions, the three-dimensional models of the A chains of LRPSN1 and LRPSN2 built from the coordinates of the A chain of ricin are very similar. The model contains eight α-helices and a six-stranded β-sheet with a left-handed twist similar to that found in the ricin A chain (Fig. 5).

The 6 residues that are believed to constitute the active RNA N-glycosidase site of ricin (Tyr¹⁸⁰, Tyr¹²³, Glu¹⁷⁷, Ala¹⁷⁸, Arg¹⁸⁰, and Trp²¹¹) (27–29) are fully conserved in the A chains of LRPSN1 (Tyr⁷⁸, Tyr¹¹⁷, Glu¹⁷², Ala¹⁷³, Arg¹⁷⁵, and Trp²⁰⁶) and LRPSN2 (Tyr⁷⁷, Tyr¹¹⁶, Glu¹⁶⁸, Ala¹⁶⁹, Arg¹⁷¹, and Trp²⁰²). Similarly, most of the residues that are located in the vicinity of the active site of the A chain of ricin and hence are probably necessary for the catalytic conformation of the site, i.e. Asn⁷⁸, Arg¹³⁴, Gln¹⁷³, Glu²⁰⁸, and Asn²⁰⁹, are conserved in the A chains of LRPSN1 (Asn²⁶, Arg¹²⁹, Gln¹⁶⁸, Glu²⁰³, and Asn²⁰⁶) and LRPSN2 (Asn⁷⁵, Arg¹²⁷, Gln¹⁶⁴, Glu¹⁹⁰, and Asn²⁰⁶). Therefore, the results of these modeling experiments fully confirm that the A chain of SNLRP possesses RNA N-glycosidase activity.

The B chains of ricin and SNLRP share a common three-dimensional structure. Sequence comparisons revealed that the four subdomains designated 1α, 1β, 1γ, and 1δ and 2α, 2β, 2γ, and 2δ, which compose the two respective domains of the ricin B chain (26), are easily recognized along the B chains of LRPSN1 and LRPSN2 on the basis of both sequence alignments and structural features. All cysteine residues involved in the folding of the ricin B chain appear at very conserved positions in LRPSN1 and LRPSN2. HCA plot analysis clearly illustrates the structural similarities between the B chains of ricin and the two isoforms of SNLRP, suggesting that the three-dimensional models of the B chains of LRPSN1 and LRPSN2 are closely related to that of the ricin B chain (Fig. 6).

Since SNLRP has no carbohydrate binding activity, a more detailed comparison was made between the carbohydrate-binding sites of the B chain of ricin and the corresponding parts of the B chain of the elderberry protein. Several amino acid residues that constitute the binding sites of the ricin B chain are apparently changed in the corresponding sites of LRPSN1 and LRPSN2.
LRPSN2. Asp\textsuperscript{22}, Gln\textsuperscript{35}, Trp\textsuperscript{37}, Asn\textsuperscript{46}, and Gln\textsuperscript{47} of the binding site of domain 1 of the ricin B chain are replaced by Asp\textsuperscript{23}, Gln\textsuperscript{36}, Leu\textsuperscript{38}, Ser\textsuperscript{45}, and Gln\textsuperscript{46} in both LRPSN1 and LRPSN2 (i.e., two changes). Similarly, Asp\textsuperscript{234}, Ile\textsuperscript{246}, Tyr\textsuperscript{248}, Asn\textsuperscript{255}, and Gln\textsuperscript{256}, forming the binding site of domain 2 of the ricin B chain, are replaced by Glu\textsuperscript{230}, Ile\textsuperscript{242}, Tyr\textsuperscript{244}, Asn\textsuperscript{251}, and Gln\textsuperscript{252} in the B chains of LRPSN1 and LRPSN2 (i.e., one change).

Taking into consideration the amino acid replacements occurring in the B chains of SNLRP and assuming that the orientation of the galactose moiety of lactose bound to the binding sites of the ricin B chain is conserved in the B chains of LRPSN1 and LRPSN2, docking experiments were performed with Gal and GalNAc. The results indicate that, in SNLRP, a hydrogen bond is lacking between Ser\textsuperscript{45} (which replaces Asn\textsuperscript{46} of the ricin B chain) and O-3 (site of domain 1) and a that steric hindrance occurs between Glu\textsuperscript{230} (which replaces Asp\textsuperscript{234} of the ricin B chain) and O-4 (site of domain 2). In addition, the hydrogen bond interconnecting O-\epsilon\textsubscript{1} of Glu\textsuperscript{230} to N-\epsilon\textsubscript{2} of Glu\textsuperscript{252} is lacking. This should prevent the B chains of LRPSN1 and LRPSN2 from interacting with both sugars and could explain why both proteins do not bind to carbohydrates.

Sequence Homology among SNLRP, SNAI, and SNAV—A comparison of the deduced amino acid sequences of LRPSN1 and LRPSN2 and those of LECSNAI and LECSNAV indicates that nonagglutinating SNLRP is more closely related to SNAI than to SNAV. The dendrogram in Fig. 7 clearly shows that LECSNAI on the one hand and LRPSN1 and LRPSN2 on the other form separate subgroups.
All type 2 RIP from elderberry share a reasonable sequence homology with the toxin and agglutinin of castor bean and abrin, respectively. A closer examination of the sequence homologies between the two domains of elderberry RIP and their homologues from other plants indicates that the B chains of the respective proteins share more homology than the A chains. However, the sequence that is believed to be essential for RIP activity (SEAAR) is well conserved in all type 2 RIP.

**DISCUSSION**

A reinvestigation of the bark proteins from elderberry revealed the occurrence of SNLRP, a predominant RIP that is structurally related to type 2 RIP, but possesses a B chain that is apparently devoid of carbohydrate binding activity. Molecular cloning confirmed that SNLRP is structurally and evolutionarily closely related to the previously cloned elderberry type 2 RIP SNAI and, to a lesser extent, also to SNAV. However, despite its high sequence homology to SNAI, SNLRP is apparently devoid of carbohydrate binding activity. Molecular modeling of SNLRP confirmed the results of the protein synthesis inhibition experiments and agglutination/carbohydrate binding assays. According to the model, the A chain of SNLRP contains a catalytically active N-glycosidase site, whereas both sugar-binding sites of the B chain are functionally inactive because of a few amino acid substitutions that prevent proper binding of carbohydrates to either one of the two sites.

SNLRP is the first documented example of a type 2 RIP with a functionally inactive B chain. Since it is clearly devoid of carbohydrate binding activity, SNLRP cannot be considered as a lectin, but as a lectin-related protein. In the past, several other lectin-related proteins have been isolated and characterized. Well known examples are the *Phaseolus vulgaris* arcelins and α-amylase inhibitor, which are structurally and evolutionarily closely related to the E- and L-type agglutinins from the same species, but exhibit no carbohydrate binding activity (30). Molecular cloning and modeling of an agglutinin from rhizomes of *Polygonatum multiflorum* (Solomon’s seal) revealed that this tissue contains, besides a typical monocot mannose-binding lectin, also a protein that is clearly related to the lectin, but exhibits no carbohydrate binding activity because its sugar-binding sites are not functional (31). Evidently, the isolation of SNLRP demonstrates that the occurrence of lectin-related proteins is not restricted to the legume and monocot mannose-binding lectins, but has to be extended to type 2 RIP. Although no general conclusions can be drawn as yet, the occurrence of at least three different types of lectin-related proteins suggests that several lectin genes evolved into lectin-related genes or vice versa. It should be emphasized, however, that the absence of functionally active carbohydrate-binding chains does not necessarily result in a biologically inactive protein. On the contrary, both the bean α-amylase inhibitor and SNLRP exhibit a well defined inhibitory or catalytic activity.

The isolation and identification of SNLRP demonstrate that elderberry bark accumulates large amounts of at least three different type 2 RIP. Two of these RIP, namely SNAI and SNAV, contain B chains that recognize structurally unrelated sugars (Neu5Ac(a2–6)Gal/GalNAc and GalNAc, respectively), whereas SNLRP possesses a functionally inactive B chain. Evidently, the simultaneous occurrence of large amounts of...
three different type 2 RIP raises the question of the physiological meaning of the apparent specialization of elderberry bark toward the accumulation of type 2 RIP. At present, one can only speculate about the evolutionary advantage of possessing such a mixture of toxic proteins. Most likely, a synergistic effect between the different RIP eventually results in an increased resistance to phytophagous insects and/or herbivorous animals.

REFERENCES
1. Wetzel, S., Demmers, C., and Greenwood, J. S. (1989) Planta (Heidelberg) 178, 275–281
2. Peumans, W. J., and Van Damme, E. J. M. (1995) Plant Physiol. (Bethesda) 106, 347–352
3. Goldstein, I. J., and Poretz, R. D. (1986) in The Lectins: Properties, Functions, and Applications in Biology and Medicine (Liner, I. E., Sharon, N., and Goldstein, I. J., eds) pp. 33–248, Academic Press, New York
4. Barbieri, L., Battelli, M. G., and Stirpe, F. (1996) Biochim. Biophys. Acta 1154, 237–282
5. Neimba-Lubaki, M., and Peumans, W. J. (1986) Plant Physiol. (Bethesda) 80, 747–751
6. Peumans, W. J., Neimba-Lubaki, M., Broekaert, W. F., and Van Damme, E. J. M. (1986) in Molecular Biology of Seed Storage Proteins and Lectins (Shannon, L. M., and Chrispeels, M. J., eds) pp. 53–63, American Society of Plant Physiologists, Rockville, MD
7. Hanks, C. N., Kindinger, J. L., and Shannon, L. M. (1988) Plant Physiol. (Bethesda) 86, 67–70
8. Van Damme, E. J. M., Barre, A., Bemer, V., Rougé, P., Van Leuven, F., and Peumans, W. J. (1995) Plant Mol. Biol. 29, 579–598
9. Broekaert, W. F., Neimba-Lubaki, M., Peeters, B., and Peumans, W. J. (1984) Biochem. J. 221, 163–169
10. Kaku, H., Peumans, W. J., and Goldstein, I. J. (1990) Arch. Biochem. Biophys. 277, 255–262
11. Shibuya, N., Tazaki, K., Song, Z., Tarr, G. E., Goldstein, I. J., and Peumans, W. J. (1989) J. Biochem. (Tokyo) 106, 1098–1103
12. Van Damme, E. J. M., Barre, A., Rougé, P., Van Leuven, F., and Peumans, W. J. (1996) Eur. J. Biochem. 235, 128–137
13. Kaku, H., Tanaka, Y., Tazaki, K., Minami, E., Mizuno, H., and Shibuya, N. (1996) J. Biol. Chem. 271, 1480–1485
14. Girhes, T., Citores, L., Ferreras, J. M., Rejo, M. A., Iglesias, R., Munoz, R., Arias, F. J., Calonge, M., Garcia, J. R., and Mendez, E. (1993) Plant Mol. Biol. 22, 1181–1186
15. Van Damme, E. J. M., Barre, A., Rougé, P., Van Leuven, F., and Peumans, W. J. (1996) Eur. J. Biochem. 237, 505–513
16. Dubois, M., Gilles, K. A., Hamilton, J. K., Rebers, P. A., and Smith, F. (1956) Anal. Chem. 28, 350–356
17. Laemmli, U. K. (1970) Nature 227, 680–685
18. Van Damme, E. J. M., Kaku, H., Perini, F., Goldstein, I. J., Peeters, B., Yagi, F., Decock, B., and Peumans, W. J. (1991) Eur. J. Biochem. 202, 23–30
19. Mierendorf, R. C., and Pfeffer, D. (1987) Methods Enzymol. 152, 556–562
20. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U. S. A. 74, 5463–5467
21. Gaboriaud, C., Bissery, V., Benchetrit, T., and Mornon, J. P. (1987) FEBS Lett. 224, 149–155
22. Lemesle-Varloot, L., Henrisset, B., Gaboriaud, C., Bissery, V., Morgat, A., and Mornon, J. P. (1990) Biochimie (Paris) 72, 555–574
23. Rutenber, E., Katziz, B. J., Collins, E. J., Mlena, D., Ready, M. P., and Robertus, J. D. (1991) Proteins Struct. Funct. Genet. 10, 240–250
24. Maddison, W. P., and Maddison, D. R. (1992) MacClade: Analysis of Phylogeny and Character Evolution, Version 3.0, Sinauer Associates, Sunderland, MA
25. von Heijne, G. (1986) Nucleic Acids Res. 14, 4683–4690
26. Rutenber, E., and Robertus, J. D. (1991) Proteins Struct. Funct. Genet. 10, 260–269
27. Katziz, B. J., Collins, E. J., and Robertus, J. D. (1991) Proteins Struct. Funct. Genet. 10, 251–259
28. Kim, Y., and Robertus, J. D. (1992) Protein Eng. 5, 775–779
29. Chaddick, J. A., and Roberts, L. M. (1993) Protein Eng. 6, 425–431
30. Mirkov, E., Wahlstrom, J. M., Hagiwara, K., Finardi-Filho, F., Kjemtrup, S., and Chrispeels, M. J. (1994) Plant Mol. Biol. 26, 1103–1113
31. Van Damme, E. J. M., Barre, A., Rougé, P., Van Leuven, F., Balzarini, J., and Peumans, W. J. (1996) Plant Mol. Biol. 31, 657–672

FIG. 7. Phylogeny of amino acid sequences encoding SNL RP (LRPSN1 and LRPSN2), SNAI (LECSNAI), SNAV (LECSNAV), the lectin from R. communis (AGGL_RICCO), and the RIP from R. communis (RICI-RICCO) and Abrus precatorius (ABRC_ABRPR). The dendrogram was constructed using the simultaneous alignment and phylogeny program CLUSTAL from the PC Gene Software package (IntelliGenetics).