Elderberry (*Sambucus nigra*) bark contains two structurally different Neu5Ac(α2,6)Gal/GalNAc-binding type 2 ribosome-inactivating proteins

Els J. M. VAN DAMME[1], Soma ROY[1], Annick BARRE[2], Lucia CITORES[3], Kazem MOSTAFAPOUS[4], Pierre ROUGÉ[2], Fred VAN LEUVEN[5], Tomas GIRBÉS[3], Irwin J. GOLDSTEIN[4] and Willy J. PEUMANS[1]

[1] Laboratory for Phytopathology and Plant Protection, Katholieke Universiteit Leuven, Belgium
[2] Institut de Pharmacologie et Biologie Structurale, Université Paul Sabatier, Toulouse, France
[3] Universidad de Valladolid, Facultad de Ciencias, Departemento de Bioquimica, Biologia molecular y Fysiologia, Spain
[4] Department of Biological Chemistry, University of Michigan, Ann Arbor MI, USA
[5] Center for Human Genetics, Katholieke Universiteit Leuven, Belgium

A second NeuAc(α2,6)Gal/GalNAc binding type 2 ribosome-inactivating protein (RIP), called SNAI' has been isolated from elderberry (*Sambucus nigra*) bark. SNAI' is a minor bark protein which closely resembles the previously described major Neu5Ac(α2,6)Gal/GalNAc binding type 2 RIP called SNAI with respect to its carbohydrate-binding specificity and ribosome-inactivating activity but has a different molecular structure. Molecular cloning revealed that the deduced amino acid sequence of SNAI' is highly similar to that of SNAI and that the difference in molecular structure between both proteins relies on a single cysteine residue present in the B chain of SNAI but absent from SNAI'. The isolation of SNAI' not only identifies a minor bark protein as a type 2 RIP but also further emphasizes the complexity of the type 2 RIP/lectin mixture present in the bark of elderberry.

Keywords: Sambucus nigra; elderberry; lectin; ribosome-inactivating protein.

Ribosome-inactivating proteins (RIP) are a group of plant proteins, which possess highly specific rRNA N-glycosidase activity and are capable of catalytically inactivating eukaryotic ribosomes (Barbieri et al., 1993). Based on their molecular structure, ribosome-inactivating proteins are further subdivided into type 1 and type 2 RIP. Type 1 RIP are single-chain proteins built up of a catalytically active subunit of 25−30 kDa. Type 2 RIP are composed of one, two or four pairs of disulphide-bridge-linked A and B chains. Whereas the A chain of the [A-s-s-B] pair has N-glycosidase activity and exhibits sequence similarity to the type 1 RIP, the B chain is catalytically inactive but contains two carbohydrate-binding domains. By virtue of the presence of the latter domain(s), type 2 RIP are also considered as lectins since according to the new definition all proteins possessing at least one non-catalytic domain, which binds reversibly to a specific mono- or oligosaccharide, are regarded as lectins (Peumans and Van Damme, 1995). To emphasize the fact that type 2 RIP possess enzymatic as well as lectin activity they are classified in the subgroup of the 'chimerolectins' (Peumans and Van Damme, 1995).

In contrast to type 1 RIP, which have been isolated from numerous plant species belonging to different taxonomic groupings, only a limited number of type 2 RIP have been described.

Besides the classical examples from *Ricinus communis* (ricin and castor bean agglutinin), *Abrus precatorius* (abrin), *Viscum album* (viscumin and related lectins) and *Adenia digitata* and *A. volkensii* (modeccin and volkensin, respectively) (Barbieri et al., 1993) similar proteins have only been isolated from *Eranthis hyemalis* (Kumar et al., 1993), *Cinnamonum camphora* (Ling et al., 1995) and *Sambucus* (elderberry) species. Interestingly, however, a great variety of type 2 RIP occurs within the genus *Sambucus*. Typical GalNAc-specific type 2 RIP were first isolated from the bark of elderberry (*Sambucus nigra*) (Girbes et al., 1993a; Van Damme et al., 1996a) and leaves of dwarf elder (*S. ebulus*) (Girbes et al., 1993b). More recently, similar type 2 RIP were also found in the seeds and fruits of *S. nigra* (Citores et al., 1994, 1996; Girbes et al., 1996). In addition, molecular cloning of the previously described Neu5Ac(α2,6)Gal/GalNAc-specific agglutinins from *S. nigra* and *S. sieboldiana* (SNAI and SSAI, respectively) revealed that these lectins also are type 2 RIP. Both lectins not only have an unusual specificity but also differ from all other type 2 RIP in that their B chains are pair-wise linked through a disulphide bridge. The (extra) cysteine residue involved in this intramolecular disulphide bridge is typical for the B chains of SNAI and SSAI and has not been found in the B chain of other type 2 RIP (Van Damme et al., 1996b; Kaku et al., 1996).

In the present report we describe the isolation, characterization and molecular cloning of SNAI', a novel Neu5Ac(α2,6)Gal/GalNAc-specific lectin from *S. nigra*. SNAI' is very similar to SNAI but lacks the extra cysteine residue, which is involved in the disulphide bridge formation between the B chains of SNAI. As a result, SNAI' is a dimer of two [A-s-s-B] pairs whereas SNAI is a tetramer of four such pairs. The discovery of SNAI' not only demonstrates that the presence or absence of a single cysteine residue has a profound influence on the molecular

structure of the type 2 RIP but also further illustrates the complexity of the mixture of type 2 RIP/lectins in the elderberry bark.

EXPERIMENTAL PROCEDURES

Materials. Bark material destined for the isolation and characterization of the lectins was collected immediately after shedding of the leaves (around the end of October) and stored at $-20\,°C$ until use.

Radioisotopes were obtained from ICN. A cDNA synthesis kit, the multifunctional phagemid pT7T3 18 U, restriction enzymes and DNA modifying enzymes were obtained from Pharmacia LKB Biotechnology Inc. *Escherichia coli* XL1 blue competent cells were purchased from Stratagene.

Isolation of SNAI and SNAI'. A total preparation of the *Sambucus nigra* Neu5Ac(α2,6)Gal/GalNAc-binding lectins (i.e. SNAI and SNAI') was isolated by affinity chromatography on fetuin-Sepharose 4B as described previously (Van Damme et al. 1996b). To separate SNAI' (which is only a minor protein) from the bulk of SNAI, the difference in molecular mass between the native lectins (120 kDa versus 240 kDa) was exploited. 5 g of total fetuin-binding lectin was dissolved in 200 ml NaCl/P$_i$ (1.5 mM KH$_2$PO$_4$, 10 mM Na$_2$HPO$_4$, 3 mM KCl, 140 mM NaCl, pH 7.4) and 20-ml aliquots chromatographed on a column (40 cm×5 cm; 800 ml bed volume) of Sephacryl 100 equilibrated with NaCl/P$_i$. Due to its very low concentration, SNAI' did not elute as a separate peak or shoulder. Therefore, the last part of the peak (corresponding to about 10% of the total amount of lectin) was collected, the corresponding fractions of all runs were pooled and concentrated by affinity chromatography on fetuin–Sepharose 4B (2.6×10 cm; 50-ml bed volume). Repeated gel filtration on Sephacryl 100 (2.6×40 cm; 200-ml bed volume) and concentration of the lectin by affinity chromatography on fetuin–Sepharose 4B (2.6×10 cm, 50-ml bed volume, for SNAI and 1.6×5 cm, 10-ml bed volume for SNAI') yielded pure lectin preparations for SNAI' and SNAI. The final yield was 18 mg pure SNAI' starting from 5 g of a total preparation of the fetuin-binding bark lectins.

Gel filtration. Analytical gel filtration of the purified proteins was performed on a Pharmacia Superose 12 column (type HR 10/30) using NaCl/P$_i$ containing 0.2 M galactose (to avoid binding of the lectin to the column) as running buffer. Samples (1 mg protein in 0.2 ml NaCl/P$_i$ containing 0.2 M galactose) were eluted at a flow rate of 20 ml/h. Molecular mass reference markers were catalase (240 kDa), *Ricinus communis* agglutinin (120 kDa) and ricin (60 kDa).

Analytical methods. Lectin preparations were analyzed by SDS/PAGE using 12.5−25% acrylamide gradient gels as described by Laemmli (1970). For amino acid sequence analysis, samples of purified SNAI' were separated by SDS/PAGE and electroblotted on a poly(vinylidene fluoride) membrane. Individual polypeptides were excised from the blots and sequenced on an Applied Biosystems model 477A protein sequencer interfaced with an Applied Biosystems model 120A on-line analyzer.

Agglutination assays. Agglutination assays were performed in 96-well V-microtiter plates in a final volume of 100 µl consisting of 33 µl each of NaCl/P$_i$, lectin solution and a 3% suspension of erythrocytes, in that order. Agglutination was measured by visual inspection, 1 h after the addition of the erythrocyte suspension. Carbohydrate-binding specificity was examined by inhibition of agglutination of formaldehyde-stabilized, trypsin-treated human type A erythrocytes. 25 µl lectin solution was placed in each well, then 25 µl of the inhibitor

stock solution was added to well number 2 of a row and serially diluted 1:2, across the row. After 1 h, 25 µl NaCl/P$_i$ followed by 25 µl of an A-cell suspension were added and the results examined after 1 h.

Protein synthesis. *In vitro* protein synthesis was performed using a rabbit reticulocyte lysate and L-[^{3}H]valine as radioactive tracer as described elsewhere (Pelham and Jackson, 1976).

Construction and screening of cDNA library. cDNA clones encoding SNAI' were selected from a cDNA library constructed with poly(A)-rich mRNA from *S. nigra* bark. Details of the construction of this cDNA library have been described previously (Van Damme et al., 1996b). cDNA fragments were inserted into the *Eco*RI site of the multifunctional phagemid pT7T3 18 U. The library was propagated in *E. coli* XL1 blue.

The SNAI' clones were initially screened by colony hybridization using a ^{32}P-labeled synthetic oligonucleotide derived from the N-terminal amino acid sequence of the lectin polypeptide; they were hybridized overnight as described previously (Van Damme et al., 1991). After washing, the filters were blotted dry, wrapped in Saran Wrap and exposed to Fuji film overnight at $-70\,°C$. In a later stage, an additional screening was performed with random-primer-labeled cDNA clones (Van Damme et al., 1991). Positively reacting colonies were selected and rescreened 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 (1987) and sequenced by the dideoxy method (Sanger et al., 1977). DNA sequences were analyzed using programs from PC Gene and Genepro.

Molecular modelling. The hydrophobic cluster analysis (HCA) (Gaboriaud et al., 1987; Lemesle-Varloot et al., 1990) was performed to delineate the structurally conserved regions along the amino acid sequences of the A and B chains of SNAI' and the model lectin ricin. HCA plots were generated on a Macintosh LC using the program HCA-Plot2 (Doriane, Paris).

Molecular modelling of the SNAI' A and B chains was performed on a Silicon Graphics Iris 4D25G workstation, using the programs InsightII, Homology and Discover (Biosym Technologies) as described previously (Van Damme et al., 1996b). The amino acid sequences were aligned on a Micro VAX 3100 (Digital) using the ialign program of the Protein Identification Resource data bank of the National Biomedical Research Foundation (Washington DC).

RESULTS

Nomenclature of the *S. nigra* bark Neu5Ac(α2,6)Gal/GalNAc-specific type 2 RIP and their genes. Several type 2 RIP have been isolated from elderberry. To avoid confusion in the designation of the Neu5Ac(α2,6)Gal/GalNAc-specific type 2 RIP and their genes the following nomenclature is used: SNAI and LECSNAI refer to the tetrameric Neu5Ac(α2,6)Gal/GalNAc-specific type 2 RIP and its corresponding gene, respectively, whereas SNAI' and LECSNAI' refer to the dimeric Neu5Ac(α2,6)Gal/GalNAc-specific type 2 RIP and its corresponding gene, respectively.

cDNA cloning indicates that elderberry bark contains a second Neu5Ac(α2,6)Gal/GalNAc-specific type 2 RIP. Screening of the elderberry bark cDNA library for clones encoding SNAI also yielded a cDNA, which clearly differed from LEC-SNAI. Sequencing of this clone (further referred to as LEC-SNAI') indicated that it contains a 1780-bp open reading frame encoding a polypeptide of 579 amino acids with a possible initiation codon at position 11 of the deduced amino acid sequence

```
                                    Signal Peptide                1  A chain
        LECSNAI  :  asqnkyskMRLVAKLLYLAVLAICGLGIHGALTHPRVTPPVYPSVSFN  12
        LECSNAI' :  ptasqnkyskMKVVATILYLVVLAICGLGIHGAHPTHSAPPTVYPSVSFN  12
                            ********..**..***.************.  ....*.********


        LECSNAI  :  LTGADT..YEPFLRALQEKVILGNHTAFDLPVLNPESQVSDSNRFVLVPL  60
        LECSNAI' :  LTEANSNEYRHFLQELRGKVILGSHRAFDLPVLNPESKVSDSDRFVLVRL  62
                    **.*..  * .**..*..*****.* ***********.****.*****.*


        LECSNAI  :  TNPSGDTVTLAIDVVNLYVVAFSSNGKSYFFSGSTAVQRDNLFVDTTQEE  110
        LECSNAI' :  TNPSRKKVTLAIDVVTFYVVAFAQNDRSYFFSGSSEVQRENLFVDTTQED  112
                    ****  ..********..*****. *..*******..***.*********.


        LECSNAI  :  LNFTGNYTSLERQVGFGRVYIPLGPKSLDQAISSLRTYTLTAGDTKPLAR  160
        LECSNAI' :  LNFKGDYTSLEHQVGFGRVYIPLGPKSLAQSISSLSTYKSSAGDNKRLAR  162
                    ***.*.*****.****************.*.****.**. .***.*.***


        LECSNAI  :  GLLVVIQMVSEAARFRYIELRIRTSITDASEFTPDLLMLSMENNWSSMSS  210
        LECSNAI' :  SLLVVIQMVSEAARFRYIQLRIQASITDAKEFTPDLLMLSMENKWSSMSS  212
                    .****************.***..*****.**************.******


        LECSNAI  :  EIQQAQPGGIFAGVVQLRDERNNSIEVTNFRRLFELTYIAVLLYGCAPVT  260
        LECSNAI' :  EIQQAQPGGAFAQVVKLLDQRNHPIDVTNFRRLFQLTSVAVLLHGCPTVT  262
                    ********* ** **.* *.**..*.********.** .****.**..**

                      Linker sequence        B chain
        LECSNAI  :  SSSYSNNAIDAQIIKMPVFRGGEYEKVCSVVE.VTRRISGWDGLCVDVRY  309
        LECSNAI' :  K.......MPAYIIKMPVFNGGEDEERCSVVEEVTRRIGGRDGFCAEVKN  305
                    . * *******.*** *. ***** *****.*.**.*..*.


        LECSNAI  :  GHYIDGNPVQLRPCGNECNQLWTFRTDGTIRWLGKCLTASSSVMIYDCNT  359
        LECSNAI' :  GDEKDGTPVQLSSCGEQSNQQWTFSTDGTIQSLGKCLTTSSSVMIYNCKV  355
                    *.  **.****..**...** ***.*****. ******.*******.*..


        LECSNAI  :  VPPEATKWVVSIDGTITNPHSGLVLTAPQAAEGTALSLENNIHAARQGWT  409
        LECSNAI' :  VPPESTKWVVSIDGTITNPRSGLVLTAPKAAEGTLVSLEKNVHAARQGWI  405
                    ****.************** .*******.***** .***.*.*******.


        LECSNAI  :  VGDVEPLVTFIVGYKQMCLRENGENNFVWLEDCVLN...RVQQEWALYGD  456
        LECSNAI' :  VGNVEPLVTFIVGYEQMCLETNPGNNDVSLGDCSVKSASKVDQKWALYGD  455
                    **.**********.****  .* .** * *.**  ..   .*.*.******


        LECSNAI  :  GTIRVNSNRSLCVTSEDHEPSDLIVILKCEGSGNQRWVFNTNGTISNPNA  506
        LECSNAI' :  GTIRVNNDRSLCVTSEGKSSNEPIIILKCLGWANQRWVFNTDGTISNPDS  505
                    ******..********.....  *.**** * .********.******..


        LECSNAI  :  KLLMDVAQRDVSLRKIILYRPTGNPNQQWITTTHPA  542
        LECSNAI' :  KLVMHVDQNDVPLRKIILSHPSGTSNQQWIASTHPA  541
                    **.*.*.*.**.****** .*.*. .*****.****
```

Fig. 1. Comparison of the deduced amino acid sequences of the cDNA clones encoding SNAI (LECSNAI) and SNAI' (LECSNAI'). The arrowhead indicates the processing site for the cleavage of the signal peptide. (.) denotes gaps introduced to give maximal similarity. Determined N-terminal amino acid sequences for SNAI and SNAI' are underlined. Putative N-glycosylation sites in LECSNAI and LECSNAI' are shown in bold. Since the first methionine is most probably used as the translation initiation site, the amino acids preceding this methionine are shown in lower case letters. Amino acids that are conserved between SNAI and SNAI' are indicated by asterisks and chemically similar residues are denoted by dots on the bottom line.

(Fig. 1). Translation starting with this codon yields a primary translation product of 569 amino acids (62.6 kDa). Cleavage of the signal peptide between residues 28 and 29 results in a lectin polypeptide of 59.7 kDa (541 amino acids) with an N-terminal sequence identical to the one determined for SNAI'. Further processing of this polypeptide to yield the A and B chains of the mature SNAI' implicates one or more proteolytic cleavages. Since the B chains starts with the sequence GGEDEER (see below), a cleavage must take place between amino acids 275 and 276 of the precursor after cleavage of the signal peptide, resulting in a polypeptide of 275 amino acids encoding both the A chain and the linker sequence (30.9 kDa) and a B chain of 28.8 kDa (266 amino acids). Judging from the sequence similarity with the previously determined type 2 RIP from *Sambucus sieboldiana* (Kaku et al., 1996), the linker sequence will most probably be cleaved around position 263 of the precursor sequence (after cleavage of the signal peptide) leaving an A chain of 29.5 kDa (263 amino acids).

A comparison of the deduced amino acid sequence of LEC-SNAI' and LECSNAI shows that they are closely related. However, in spite of the high percentage of identity (72.3%) and similarity (77.1%), there are some striking differences between SNAI and SNAI'. First, SNAI has eight putative glycosylation sites whereas SNAI' contains only one N-glycosylation site. Second, Cys327, the cysteine residue which is involved in the intramolecular disulphide bridge between two B chains of the tetrameric SNAI (Van Damme et al., 1996b), is absent in SNAI'. Hence, the B chains of SNAI' can not form intermolecular disulphide bridges, which implies that native SNAI' is most likely a dimer instead of a tetramer. As is described below, this presumed difference in molecular mass between SNAI and SNAI' has been exploited to isolate SNAI'.

Isolation and characterization of SNAI'. To isolate SNAI', the occurrence of which was inferred from the cloning of its corresponding cDNA, a total preparation of fetuin-binding lectins was
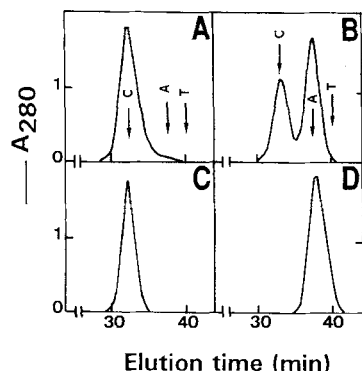
Fig. 2. Gel filtration of the Neu5Ac(α2,6)Gal/GalNAc-binding elderberry bark lectins on a Superose 12 column. (A) Total peparation of Neu5Ac(α2,6)Gal/GalNAc-binding lectins; (B) lectin eluting at the tail of the peak upon gel filtration on a Sephacryl 100 column; (C) purified SNAI; (D) purified SNAI'. The elution position of the molecular mass reference markers catalase (C, 240 kDa) and the *Ricinus communis* hemagglutinin (A, 120 kDa) and toxin (T, 60 kDa) are indicated by the arrows.
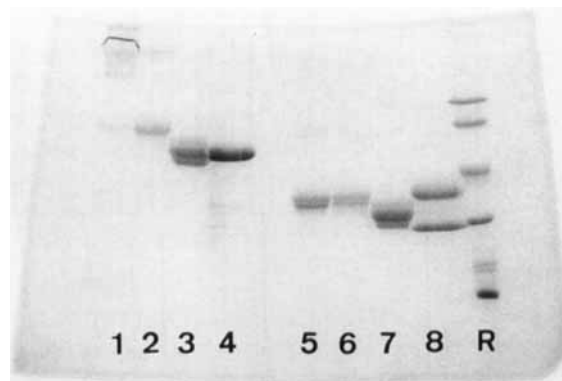

Fig. 3. SDS/PAGE of purified elderberry bark type 2 RIP. Samples (25 μg each) of the unreduced (lanes 1−4) and reduced (with 2-mercaptoethanol) (lanes 5−8) proteins were loaded as follows: lanes 1 and 5, SNAI; lanes 2 and 6, SNAI'; lanes 3 and 7, SNLRP; lanes 4 and 8, SNAV. Molecular mass reference proteins (lane R) were lysozyme (14 kDa), soybean trypsin inhibitor (20 kDa), carbonic anhydrase (30 kDa), ovalbumin (43 kDa), bovine serum albumin (67 kDa) and phosphorylase *b* (94 kDa).

chromatographed on a column of Sephacryl 100. The lectin eluting at the tail of the peak was collected, concentrated by affinity chromatography on a small column of fetuin-Sepharose, and analyzed by gel filtration on a column of Sepharose 12. Since this preparation yielded two peaks eluting with apparent molecular masses of 240 and 120 kDa, the presumed occurrence of a second Neu5Ac(α2,6)Gal/GalNAc-specific type 2 RIP seemed obvious (Fig. 2B). Repeated gel filtration eventually yielded a pure preparation of a 120-kDa protein (Fig. 2D). Starting from 5 g total fetuin-binding lectin, about 20 mg pure SNAI' was obtained. This low yield indicates that SNAI' is about two orders of magnitude less abundant than SNAI.

The purified SNAI' was characterized in some detail with respect to its molecular structure, biochemical properties and RIP activity. SDS/PAGE indicated that the unreduced protein migrated as a single band of about 60 kDa whereas reduced (with 2-mercaptoethanol) SNAI' yielded two polypeptides of 32 and 35 kDa (Fig. 3). It should be mentioned here that the high-molecular-mass bands, which are very typical for unreduced SNAI (Van Damme et al., 1996b), are absent in the polypeptide pattern of unreduced SNAI'. Analytical gel filtration yielded a molecular mass of 120 kDa (Fig. 2D) indicating that native SNAI' is composed of two [A-s-s-B] pairs.

N-terminal sequencing of the A and B chains composing SNAI' yielded the sequences ATPPVYPSVS and GGEDEER, respectively. Since these sequences match the deduced sequence of LECSNAI' at the N-terminus of the A and B domain, respectively (Fig. 1), one can reasonably assume that SNAI' is encoded, indeed, by LECSNAI'.

Comparison of the sugar binding and agglutination properties of SNAI and SNAI'.
Both SNAI and SNAI' were shown to agglutinate human better than rabbit erythrocytes. Among human red blood cells, type A erythrocytes were agglutinated best. Sialidase treatment of the erythrocytes destroyed agglutinability whereas trypsinization and formaldehyde treatment of the cells did not affect their agglutinability. Minimal protein concentrations needed for agglutination by SNAI and SNAI' were 7.5 and 37.5 μg/ml, respectively. At the noted concentrations, both lectins displayed an equal, but relatively weak, clumping activity against cultured Ehrlich ascites tumor cells but not against Chinese hamster ovary cells. Neu5Ac(α2,6)lactose but not Neu5Ac-(α2,3)lactose, at a final concentration of 10 μg/ml, inhibited ag-

glutination by SNAI. Three times as much of the inhibitor was needed for the inhibition of SNAI'. The N-linked type glycoproteins such as fetuin and orosomucoid, at final concentrations of 60 μg/ml and 150 μg/ml, inhibited SNAI and SNAI', respectively, whereas O-linked type glycans such as Neu5Ac(α2,3)Gal-(β1,3)GalNAc (T-antigen) were without effect at the same concentrations.

SNAI' inhibits protein synthesis *in vitro*.
The inhibitory effect of SNAI' on protein synthesis in a reticulocyte lysate was compared to that of SNAI. Since protein was synthesized in the presence of 1 mM dithiothreitol as reductant, the proteins were tested in the reduced form. SNAI and SNAI' inhibited the incorporation of labeled amino acids, the concentration required for 50% inhibition (IC$_{50}$) being about 150 ng/ml.

Molecular modelling of SNAI'.
Because of the high degree of sequence similarity between SNAI' and ricin, it was possible to model the structure of SNAI' using the coordinates of ricin, the three-dimensional structure of which has been resolved by X-ray crystallography (Rutenber and Robertus, 1991; Rutenber et al., 1991). The numbering of the amino acids used in this section refers to the amino acid positions in the mature A and B chains of the type 2 RIP.

Sequence comparison of the A chains of SNAI' and ricin reveals 31% sequence identity and 54% similarity. Furthermore, a comparison of the HCA plots of ricin and SNAI' A chains shows that α helices and strands of β sheets are readily conserved in both proteins. Despite a few insertions and deletions of a few amino acids (Fig. 4), the three-dimensional model of the SNAI' A chain exhibits an overall folding very similar to that of the ricin A chain (Katzin et al., 1991) (Fig. 5). It contains eight α helices and a six-stranded β sheet with a lefthanded twist. A single putative N-glycosylation site Asn12-Leu13-Thr14 occurs at the N-terminal of the A chain of SNAI'.

As demonstrated for ricin, the toxicity of the A chain depends of its N-glycosidase activity towards an adenine residue occurring in an exposed loop of the 28S rRNA of the ribosomal large subunit (Endo et al., 1988). Its active site contains six amino acid residues Tyr80, Tyr123, Glu177, Ala178, Arg180 and Trp211 (Katzin et al., 1991; Kim and Robertus, 1992; Chaddock and Roberts, 1993), which are fully conserved in the A chain of SNAI' (Tyr80, Tyr119, Glu173, Ala174, Arg176 and
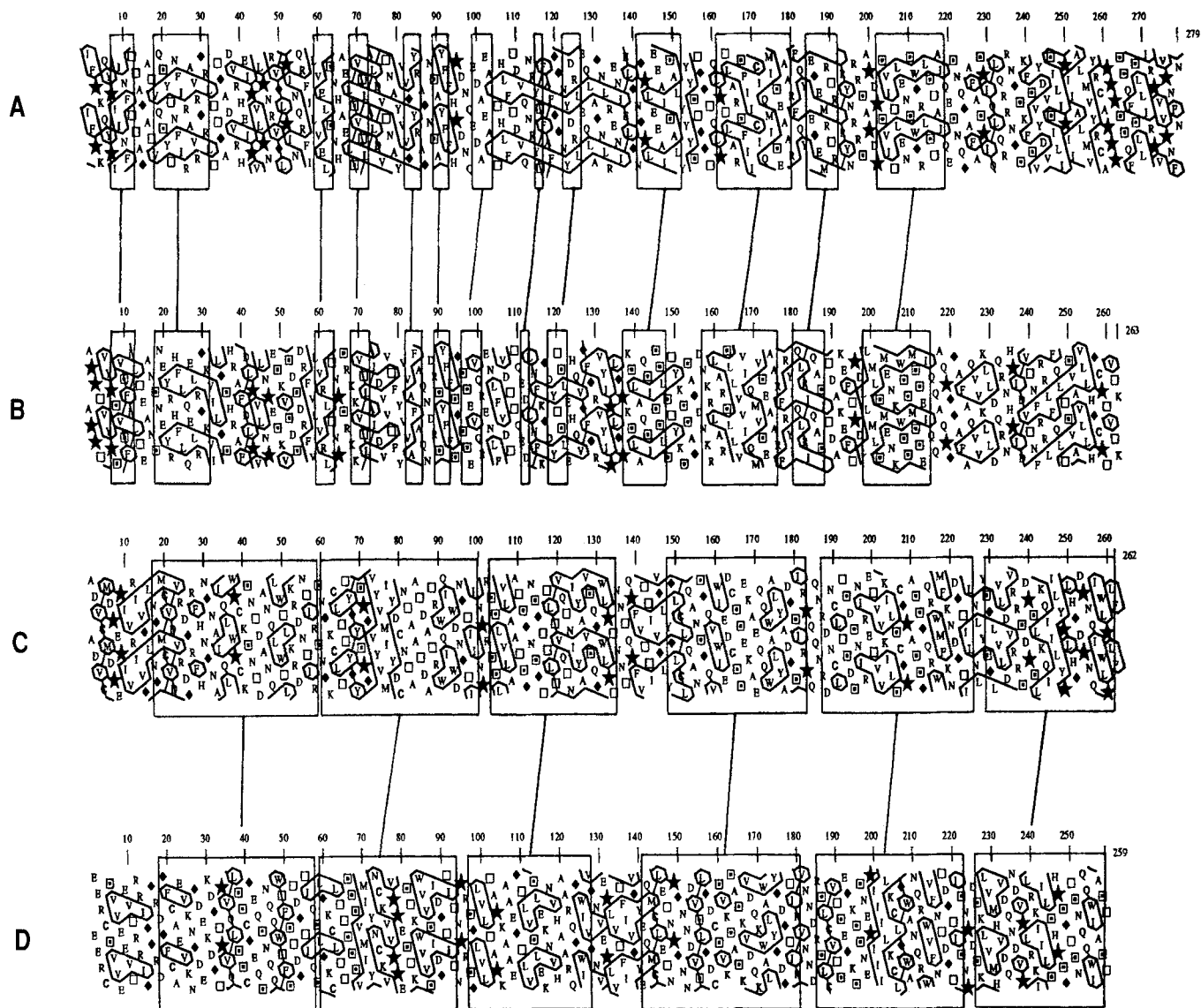
Fig. 4. HCA plots of the ricin A chain (A), SNAI′ A chain (B), ricin B chain (C) and SNAI′ B chain (D). Stretches of amino acids corresponding to secondary structural features (α-helices, β-sheets) and to subdomains delineated along the respective amino acid sequences of the A and B chains of ricin are boxed together with their similar regions in the A and B chains of SNAI′.

Trp207). Similarly, most of the residues located nearby the active site of ricin and probably necessary for the catalytic conformation of the site, i.e. Asn78, Arg134, Gln173, Glu208 and Asn209, are conserved in SNAI′ (Thr78, Arg130, Gln169, Glu204 and Asn205). These observations are in good agreement with the RIP activity of SNAI′.

The B chains of ricin and SNAI′ share 42 % sequence identity and 65 % similarity. Accordingly the four subdomains designated 1γ, 1α, 1β and 1λ and 2γ, 2α, 2β and 2λ which built the two domains of the ricin B chain (Rutenber and Robertus, 1991) are easily recognized on the B chain of SNAI′. Alignment of the sequences shows that the cysteine residues involved in the folding of the ricin B chain are highly conserved in the B chain of SNAI′. A comparison of the HCA plots of ricin and SNAI′ B chains clearly demonstrates the structural similarities between both lectin subunits (Fig. 4) and strongly suggests that the overall folding of SNAI′ B chain must be very similar to that of ricin (Fig. 5). The B chain of SNAI′ does not contain any secondary structures (helices or sheets) and does not possess any putative N-glycosylation sites.

On the analogy to ricin (Rutenber et al., 1991), it is assumed that the two chains of SNAI′ are linked by a disulfide bridge between Cys258 and Cys283, which are conserved in the amino acid sequences of both chains (Fig. 1).

Several amino acid residues in both carbohydrate-binding sites of the SNAI′ B chain are different when compared to ricin. Residues Asp22, Gln35, Trp37, Asn46 and Gln47 of the first binding site of the ricin B chain are replaced by Glu27, Gln40, Ser42, Asn49 and Gln50 in the SNAI′ B chain (two changes). Similarly, the residues forming the second binding site of the SNAI′ B chain are His235, Ile247, Ser251, Asn257 and Gln258 instead of Asp234, Ile246, Tyr248, Asn255 and Gln256 of the ricin B chain (three changes). Taking into consideration the amino acid replacements in the SNAI′ B chain 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 chain of SNAI′, docking experiments performed with Gal or GalNAc showed that steric hindrances occur in both sites between O3 and O4 of Gal or GalNAc and various residues (Glu27 for site 1 and His235 for site 2) of the SNAI′ B chain binding sites. Due
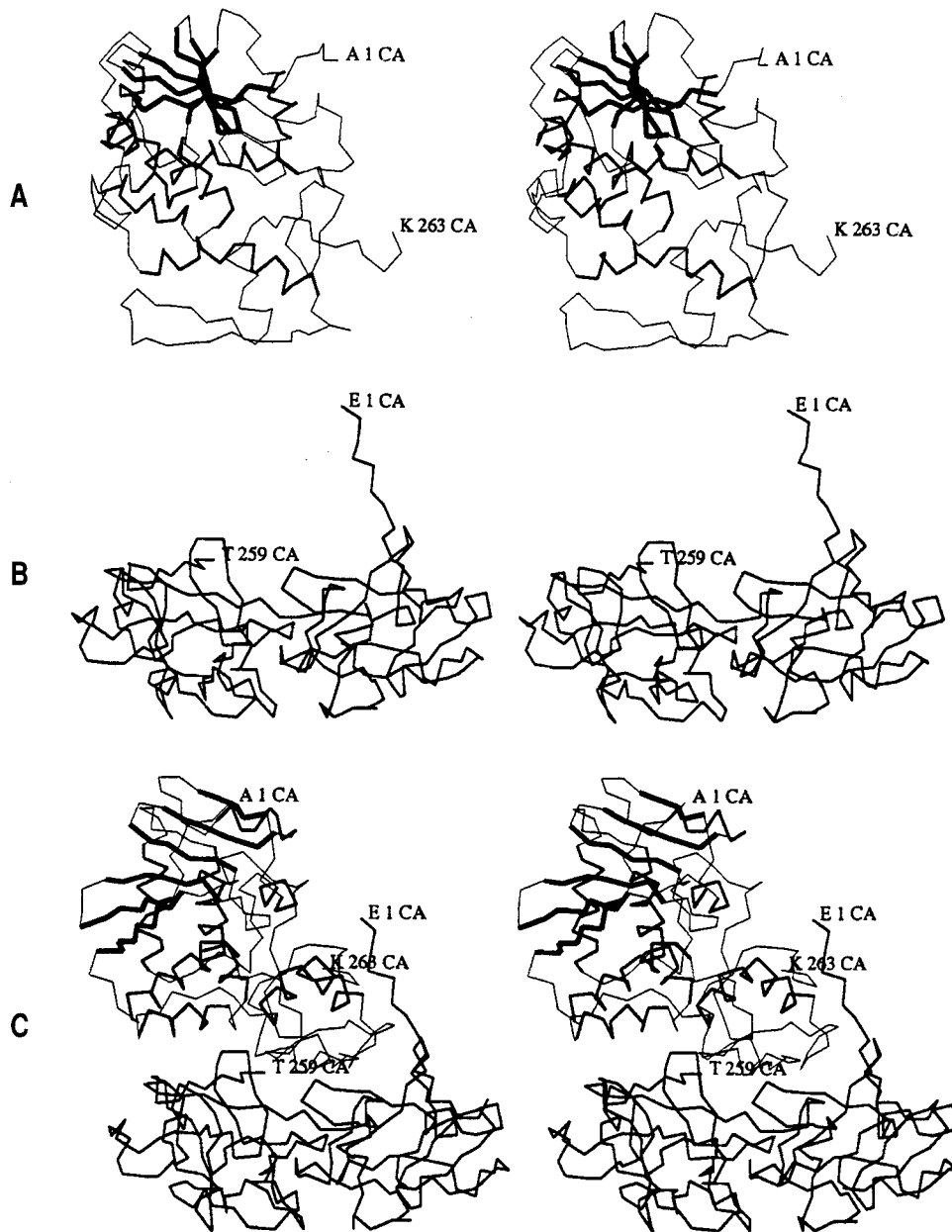
Fig. 5. Stereoviews of the three-dimensional models of the A chain (A), B chain (B) and [A-s-s-B] pair (C) of SNAI′. Helices (thick lines) and left-handed twisted six-stranded β sheet (heavy lines) of the A chain are indicated. It is assumed that, like ricin, SNAI′ results from the covalent association of the A and B chains by a disulfide bridge between Cys258 and Cys283 (Cys8 of the mature B chain).

to these steric hindrances, SNAI′ can not interact with either Gal or GalNAc. Docking performed with a Neu5Ac(α2,6)Gal disaccharide from human lactotransferrin (Bourne et al., 1994), suggests that the carbohydrate-binding site of domain 1 of the SNAI′ B chain can accommodate Neu5Ac since no steric clashes occur with the residues forming this site. However, the binding site of domain 2 cannot accommodate Neu5Ac because of a steric clash between Neu5Ac and His235 (which replaces Asp234 of the ricin B chain).

Superposition of the three-dimensional models of SNAI and SNAI′ shows that both proteins possess a quite similar folding (results not shown). However, some of the amino acid residues forming the carbohydrate-binding sites of the SNAI B chain (Asp26, Gln39, Arg41, Asn48, Gln49 for domain 1; Asp233, Ile245, Tyr247, Asn254, Gln255 for domain 2) are different in

SNAI′ (Glu27, Gln40, Ser42, Asn49, Gln50 for domain 1; His235, Ile247, Ser251, Asn257, Gln258 for domain 2). Docking, performed with the Neu5Ac(α2,6)Gal-disaccharide from human lactotransferrin, suggests that the replacement of Arg41 by Ser42 in the first carbohydrate-binding site of SNAI′ reduces its affinity for Neu5Ac. Similarly, the replacement of Asp233 in the SNAI B chain by His235 in the SNAI′ B chain causes a steric clash which makes the second binding site of SNAI′ unaccessible for NeuAc (whereas the corresponding site of SNAI can accommodate Neu5Ac). It can be expected, therefore, that SNAI′ reacts, like SNAI, with glycoproteins containing terminal Neu5Ac residues (α2,6)-linked to Gal or GalNAc although the binding will be weaker. The latter conclusions are fully supported by the results of the carbohydrate-binding activity and specificity of SNAI and SNAI′.
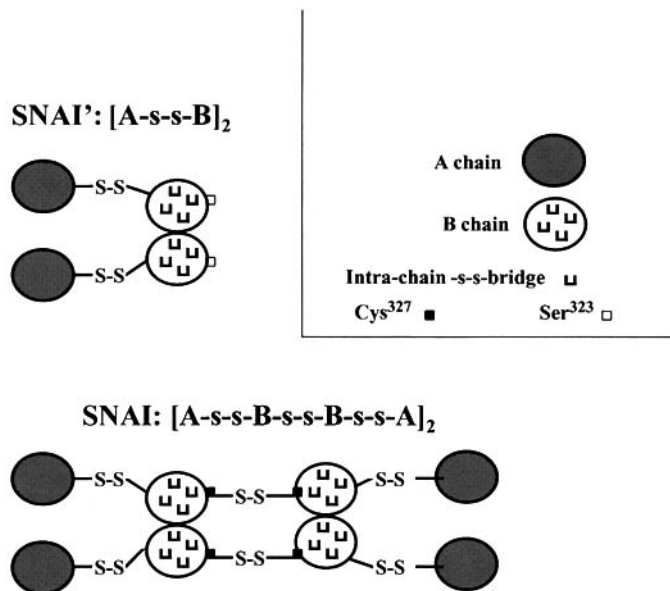
SNAI': [A-s-s-B]₂

SNAI: [A-s-s-B-s-s-B-s-s-A]₂

Fig. 6. Schematic representation of the molecular structure of SNAI′ and SNAI.

DISCUSSION

The occurrence of a second Neu5Ac(α2,6)Gal/GalNAc-specific type 2 RIP in elderberry bark was predicted on the basis of the sequence analysis of cDNA clones which showed a high sequence similarity to LECSNAI but lacked the extra cysteine residue required for the formation of intermolecular disulphide bridges between two B chains. SNAI′ was separated from the bulk of the SNAI present in a total preparation of a Neu5Ac-(α2,6)Gal/GalNAc-specific bark lectins by exploiting the presumed difference in molecular structure between both lectins. Biochemical analyses demonstrated that SNAI′ is composed of two [A-s-s-B] pairs and that its B chains are not pairwise linked through a disulphide bridge. The latter findings confirm that the unusual structure of SNAI relies, indeed, on the presence of an extra cysteine residue in its B chain (Fig. 6).

In spite of their different molecular structures, SNAI′ and SNAI exhibit a similar inhibitory effect on cell-free protein synthesis in a reticulocyte lysate. Agglutination and inhibition studies further indicated that SNAI′ and SNAI have similar, if not identical, properties. The main difference in carbohydrate-binding specificity appears to be quantitative rather than qualitative. SNAI′ binds to cells and inhibitors 3−5 times less strongly than SNAI. SNAI′ has a weaker affinity not only for cells but also for agglutination inhibitors. It takes more of an inhibitor to be effective against SNAI′ than against SNAI agglutinating effect. This may further indicate differences in the number and or the reactivity of the binding sites between these two proteins. The latter assumption is supported by the results of the molecular modelling and docking experiments, which indicated that only carbohydrate-binding site 1 of the B chain of SNAI′ can accommodate Neu5Ac whereas both sites of the B chain of SNAI are accessible for this sugar.

Protein purification and molecular cloning demonstrated that all major elderberry bark proteins are type 2 RIP or proteins derived thereof. Besides the agglutinating type 2 RIP SNAI and SNAV (Van Damme et al., 1996a, b), which bind Neu5Ac(α2,6)-Gal/GalNAc-binding and GalNAc, respectively, the bark also contains large quantities of a non-agglutinating type 2 RIP (called *S. nigra* lectin-related proteins or SNLRP) (Van Damme et al., 1997) and a GalNAc-specific lectin SNAII composed of

two B chains, which are derived from the same precursor as SNAV (Van Damme et al., 1996a). SNAI′ is far less abundant than the other bark type 2 RIP/lectins. Estimates based on the data published by Van Damme et al. (1996a, b, 1997) and the yield of SNAI′ reported here indicate that 1 g lyophilized bark contains about 600 µg SNAI, 2500 µg SNLRP, 2600 µg SNAII, 120 µg SNAV and 3 µg SNAI′. Although the latter estimates have to be interpreted with some care, they clearly indicate that SNAI′ is two to three orders of magnitude less abundant than the other type 2 RIP and lectins present in the elderberry bark.

The isolation of SNAI′ not only increases the complexity of the mixture of type 2 RIP in elderberry bark but also raises the question whether other rare type 2 RIP may occur in this tissue. In addition, the isolation of SNAI′ addresses the question of the possible presence in other plant species of rare type 2 RIP, which because of their low concentration escaped detection. If so, type 2 RIP may be much more widespread in the plant kingdom than is actually believed on the basis of our current knowledge of these proteins.

REFERENCES

Barbieri, L., Battelli, M. G. & Stirpe, F. (1993) Ribosome-inactivating proteins from plants, *Biochim. Biophys. Acta 1154*, 237−282.

Bourne, Y., Mazurier, J., Legrand, D., Rougé, P., Montreuil, J., Spik, G. & Cambillau, C. (1994) Interaction of a legume lectin with the human lactotransferrin N2 fragment or with the isolated biantennary glycopeptide: role of the fucose moiety, *Structure 2*, 209−219.

Chaddock, J. A. & Roberts, L. M. (1993) Mutagenesis and kinetic analysis of the active site Glu177 of ricin-A chain, *Protein Eng. 6*, 425−431.

Citores, L., Iglesias, R., Munoz, R., Ferreras, J. M., Jiminez, P. & Girbes, T. (1994) Elderberry (*Sambucus nigra* L.) seed proteins inhibit protein synthesis and display strong immunoreactivity with rabbit polyclonal antibodies against the type 2 ribosome-inactivating protein nigrin b, *J. Exp. Bot. 45*, 513−516.

Citores, L., de Benito, F. M., Iglesias, R., Ferreras, J. M., Jiminez, P., Argüeso, P., Farias, G., Mendez, E. & Girbes, T. (1996) Isolation and characterization of a new non-toxic two-chain ribosome-inactivating protein from fruits of elder (*Sambucus nigra* L.), *J. Exp. Bot. 47*, 1577−1585.

Endo, Y., Tsurgi, K. & Lambert, J. M. (1988) The site of action of six different ribosome-inactivating proteins from plants on eukaryotic ribosomes: the RNA N-glycosidase activity of the proteins, *Biochim. Biophys. Res. Commun. 150*, 1032−1036.

Gaboriaud, C., Bissery, V., Benchetrit, T. & Mornon, J. P. (1987) Hydrophobic cluster analysis: an efficient new way to compare and analyse amino acid sequences, *FEBS Lett. 224*, 149−155.

Girbes, T., Citores, L., Ferreras, J. M., Rojo, M. A., Iglesias, R., Munoz, R., Arias, F. J., Calonge, M., Garcia, J. R. & Mendez, E. (1993a) Isolation and partial characterization of nigrin b, a non-toxic novel type 2 ribosome-inactivating protein from the bark of *Sambucus nigra* L., *Plant Mol. Biol. 22*, 1181−1186.

Girbes, T., Citores, L., Iglesias, R., Ferreras, J. M., Munoz, R., Rojo, M. A., Arias, F. J., Garcia, J. R., Mendez, E. & Calonge, M. (1993b) Ebulin 1, a nontoxic novel type 2 ribosome-inactivating protein from *Sambucus ebulus* L. leaves, *J. Biol. Chem. 268*, 18195−18199.

Girbes, T., Citores, L., de Benito, F. M., Iglesias, R. & Ferreras, J. M. (1996) A non-toxic two-chain ribosome-inactivating protein co-exists with a structure-related monomeric lectin (SNA III) in elder (*Sambucus nigra*) fruits, *Biochem. J. 315*, 343−344.

Kaku, H., Tanaka, Y., Tazaki, K., Minami, E., Mizuno, H. & Shibuya, N. (1996) Isolation and characterization of a second lectin (SNA II) present in elderberry (*Sambucus nigra* L.) bark, *J. Biol. Chem. 271*, 1480−1485.

Katzin, B. J., Collins, E. J. & Robertus, J. D. (1991) Structure of ricin-A chain at 2.5 Å, *Proteins 10*, 251−259.

Kim, Y. & Robertus, J. D. (1992) Analysis of several key active site residues of ricin A chain by mutagenesis and X-ray crystallography, *Protein Eng. 5*, 775−779.

Kumar, A. M., Timm, D. E., Neet, K. E., Owen, W. G., Peumans, W. J. & Rao, A. J. (1993) Characterization of the lectin from the bulbs of *Eranthis hyemalis* (winter aconite) as an inhibitor of protein synthesis, *J. Biol. Chem. 268*, 25 176−25 183.

Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4, *Nature 227*, 680−685.

Lemesle-Varloot, L., Henrissat, B., Gaboriaud, C., Bissery, V., Morgat, A. & Mornon, J. P. (1990) Hydrophobic cluster analysis: procedure to derive structural and functional information from 2-D-representation of protein sequences, *Biochimie 72*, 555−574.

Ling, J., Liu, W. & Wang, T. P. (1995) Simultaneous existence of two typres of ribosome-inactivating proteins in the seeds of *Cinnamonum camphora* − characterization of the enzymatic activities of these cytotoxic proteins, *Biochim. Biophys. Acta 1252*, 15−22.

Mierendorf, R. C. & Pfeffer, D. (1987) Direct sequencing of denatured plasmid DNA, *Methods Enzymol. 152*, 556−562.

Pelham, H. D. B. & Jackson, R. J. (1976) An efficient mRNA-dependent translation system from reticulocyte lysates, *Eur. J. Biochem. 67*, 247−256.

Peumans, W. J. & Van Damme, E. J. M. (1995) Lectins as plant defense proteins, *Plant Physiol. 109*, 347−352.

Rutenber, E., Katzin, B. J., Collins, E. J., Mlsna, D., Ready, M. P. & Robertus, J. D. (1991) Crystallographic refinement of ricin to 2.5 Å, *Proteins 10*, 240−250.

Rutenber, E. & Robertus, J. D. (1991) Structure of the ricin B-chain at 2.5 Å resolution, *Proteins 10*, 260−269.

Sanger, F., Nicklen, S. & Coulson, A. R. (1977) DNA sequencing with chain terminating inhibitors, *Proc. Natl Acad. Sci. USA 74*, 5463−5467.

Van Damme, E. J. M., Kaku, H., Perini, F., Goldstein, I. J., Peeters, B., Yagi, F., Decock, B. & Peumans, W. J. (1991) Biosynthesis, primary structure and molecular cloning of snowdrop (*Galanthus nivalis* L.) lectin, *Eur. J. Biochem. 202*, 23−30.

Van Damme, E. J. M., Barre, A., Rougé, P., Van Leuven, F. & Peumans, W. J. (1996a) Characterization and molecular cloning of SNAV (nigrin b), a GalNAc-specific type 2 ribosome-inactivating protein from the bark of elderberry (*Sambucus nigra*), *Eur. J. Biochem. 235*, 505−513.

Van Damme, E. J. M., Barre, A., Rougé, P., Van Leuven, F. & Peumans, W. J. (1996b) The NeuAc(α2,6)Gal/GalNAc binding lectin from elderberry (*Sambucus nigra*) bark is a type 2 ribosome inactivating protein with an unusual specificity and structure, *Eur. J. Biochem. 235*, 128−137.

Van Damme, E. J. M., Barre, A., Rougé, P., Van Leuven, F. & Peumans, W. J. (1997) Isolation and molecular cloning of a novel type 2 ribosome-inactivating protein with an inactive B chain from elderberry (*Sambucus nigra*) bark. *J. Biol. Chem. 272*, in the press.