Elderberry (*Sambucus nigra*) contains truncated Neu5Ac(α-2,6)Gal/GalNAc-binding type 2 ribosome-inactivating proteins

Willy J. Peumans[a], Soma Roy[a], Annick Barre[b], Pierre Rougé[b], Fred van Leuven[c], Els J.M. van Damme[a],*

[a]*Laboratory for Phytopathology and Plant Protection, Katholieke Universiteit Leuven, Willem de Croylaan 42, 3001 Leuven, Belgium*
[b]*Institut de Pharmacologie et Biologie Structurale, Université Paul Sabatier, 35 Chemin des Maraîchers, UPR CNRS 9062, 31062 Toulouse Cedex, France*
[c]*Center for Human Genetics, Katholieke Universiteit Leuven, Herestraat 49, 3001 Leuven, Belgium*

Abstract Analysis of affinity-purified preparations of the fetuin-binding proteins from elderberry bark and fruits revealed besides the previously reported Neu5Ac(α-2,6)Gal/GalNAc-specific type 2 ribosome-inactivating proteins (RIP) the occurrence of single chain proteins of 22 kDa, which according to their N-terminal amino acid sequence correspond to the second part of the B chain of the respective type 2 RIP. Both proteins are very similar except that the polypeptides of the fruit lectin are 10 amino acid residues longer than these from the bark lectin. Our findings not only demonstrate the occurrence of carbohydrate-binding fragments of type 2 RIP but also provide further evidence that type 2 RIP genes give rise to complex mixtures of type 2 RIP/lectins in elderberry.

© 1998 Federation of European Biochemical Societies.

Key words: Polynucleotide:adenosine glycosidase; Ribosome-inactivating protein; *Sambucus*

1. Introduction

Ribosome-inactivating proteins (RIP), also called polynucleotide:adenosine glycosidases, are a group of plant enzymes which catalyze the removal of adenosine residues from polynucleotides [1,2]. RIP are classically subdivided in two classes. Type 1 RIP are single chain enzymes composed of a catalytically active subunit of about 30 kDa. Type 2 RIP are built up of protomers consisting of a disulphide-linked A and B chain. These [A-s-s-B] pairs are derived from a single precursor comprising two different domains separated by a linker sequence. After post-translational processing, the N-terminal domain of the precursor eventually yields the A chain with polynucleotide:adenosine glycosidase activity whereas the C-terminal domain is converted into the carbohydrate-binding B chain. This B chain consists of two very homologous domains each of which harbours a carboxyhydrate-binding site and hence exhibits lectin activity [3].

Although type 2 RIP occur in different taxonomic groups only a limited number of these proteins have been isolated and characterized in detail. Interestingly, species of the genus *Sambucus* contain a great variety of type 2 RIP/lectins with different molecular structures and sugar-binding specificities [4–11]. In this report we present evidence for the occurrence in elderberry bark and fruits of small lectins consisting of the second domain of the B chain of the Neu5Ac(α-2,6)Gal/GalNAc-specific type 2 RIP found in the same tissues. Our results not only demonstrate the occurrence of truncated B chains but also provide further evidence that type 2 RIP genes give rise to complex mixtures of type 2 RIP/lectins in elderberry.

2. Materials and methods

2.1. Materials

Ripe fruits were collected from a single *Sambucus nigra* tree, which due to a defective anthocyanin synthesis produces yellow fruits. The berries were processed immediately or stored at −20°C until use. Bark material was collected from the same tree immediately after shedding of the leaves and stored at −20°C until use.

2.2. Isolation of TrSNAIf and TrSNAI

A total preparation of the elderberry fruit Neu5Ac(α-2,6)Gal/GalNAc-binding lectins was isolated using a combination of classical protein purification techniques and affinity chromatography on fetuin Sepharose 4B. Ripe fruits (1 kg) were homogenized gently in 2 l of distilled water using a mixer and the homogenate squeezed through cheese cloth to remove the seeds and fruit walls. Debris were removed by centrifugation at $3000 \times g$ for 10 min and the pH of the diluted juice raised to pH 9.0 with 1 N NaOH. Coagulating materials were removed by recentrifugation ($3000 \times g$ for 10 min) and the cleared supernatant adjusted to pH 3.0 with 1 N HCl. After standing in the cold room overnight, the extract was centrifuged ($3000 \times g$ for 10 min), filtered through filter paper (Whatman 3MM) and loaded on a column (5 cm × 5 cm; 100 ml bed volume) of S Fast Flow (Pharmacia, Uppsala, Sweden) equilibrated with 20 mM acetic acid. After passing the extract the column was washed with 2 l of 20 mM Na-formate (pH 3.8) and the proteins eluted with 500 ml of 0.5 M NaCl in 0.1 M Tris-HCl (pH 7.4). The partially purified protein fraction was applied on a column (2.6 cm × 10 cm; 50 ml bed volume) of fetuin Sepharose 4B equilibrated with phosphate buffer (PB: 20 mM phosphate buffer pH 7.5 containing 0.2 M NaCl). The column was washed with PB until the A_{280} fell below 0.01 and the bound lectins eluted with 20 mM acetic acid. To further purify the lectins, the affinity chromatography step was repeated (after adjusting the pH to 7.5 with 1 N NaOH and adding 0.2 M NaCl). The total preparation of fetuin-binding proteins (SNAIf and TrSNAIf) (about 20 mg per kg fruits) was lyophilized, redissolved in 5 ml of PB and chromatographed on a column (60 cm × 2.6 cm; 300 ml bed volume) of Sephacryl 100 using PB as running buffer. As shown in Fig. 1, the proteins eluted in two peaks. The proteins present in the first and second peak were pooled, lyophilized and reconstituted with an appropriate buffer for further analyses.

TrSNAI was isolated from a total preparation of fetuin-binding bark proteins purified as described previously [5,7] essentially as described above for TrSNAIf.

2.3. Analytical methods

Total neutral sugar was determined by the phenol/H_2SO_4 method [12], with D-glucose as standard.

Analytical gel filtration of the purified proteins was performed on a

Pharmacia Superose 12 using PB containing 0.2 M galactose (to avoid binding to the column) as running buffer.

Lectin preparations were analyzed by SDS-PAGE using 12.5–25% (w/v) acrylamide gradient gels as described by Laemmli [13]. For N-terminal sequencing purified proteins were separated by SDS-PAGE and electroblotted on a PVDF membrane. Individual polypeptides were excised from the blots and sequenced on an Applied Biosystems model 477A protein sequencer.

Agglutination assays were carried out in 96 well V-microtiter plates in a final volume of 100 μl consisting of 20 μl lectin solution and 80 μl of a 1% suspension of trypsin-treated human (type A) or rabbit erythrocytes. Agglutination was measured by visual inspection, 1 h after addition of the erythrocyte suspension.

2.4. Molecular modelling

The hydrophilicity, surface probability and flexibility profiles were calculated with a window size of seven residues [14–16], using Mac Vector (Kodak) run on a Macintosh LC630. Molecular modelling of TrSNAIf and TrSNAI was carried out on a Silicon Graphics Iris 4D25G workstation, using the programmes InsightII, Homology and Discover (Biosym Technologies, San Diego, CA, USA). The atomic coordinates of ricin (code 2AAI) were taken from the Brookhaven Protein Data Bank [17,18] and were used to build the three-dimensional models of TrSNAIf and TrASNAI. Energy minimization and relaxation of the loop regions was carried out by several cycles of steepest descent and conjugate gradient using the cvff forcefield of Discover. The programme TurboFrodo (Bio-Graphics, Marseille, France) was run on a Silicon Graphics Indy R4600SC workstation
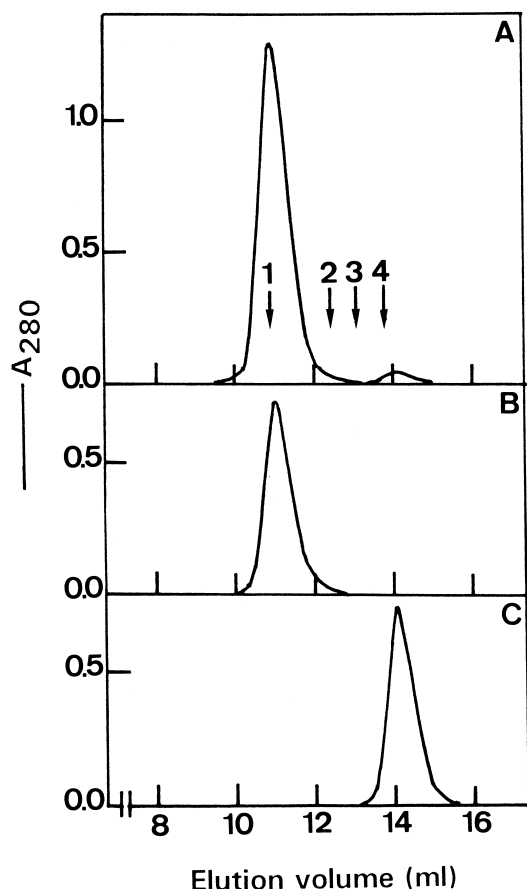

Fig. 2. SDS-PAGE of purified elderberry fruit fetuin-binding proteins. Samples (50 μg each) of the unreduced (lanes 1 and 4) and reduced (with 2-mercaptoethanol) (lanes 2 and 3) proteins were loaded as follows: lanes 1 and 3, SNAIf; lanes 2 and 4, TrSNAIf. Molecular mass reference proteins (lane R) were 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).

to delineate the hydrophilic and solvent accessible regions on the surface of the models.

3. Results

3.1. Purification and characterization of the small fetuin-binding lectins from elderberry fruits and bark

Gel filtration of a total preparation of fetuin-binding fruit proteins[1] yielded two peaks, eluting with an apparent molecular weight of 240 and 25 kDa, respectively (Fig. 1). SDS-PAGE of the 240 kDa protein showed an unusual polypeptide pattern. The reduced protein yielded polypeptides of 35 kDa, whereas the unreduced sample migrated in several distinct high molecular weight bands (Fig. 2). A similar pattern has previously been shown for the Neu5Ac(α-2,6)Gal/GalNAc-specific bark lectin SNAI [5]. To distinguish the 240 kDa Neu5Ac(α-2,6)Gal/GalNAc-specific fruit lectin from the bark lectin SNAI it is further referred to as SNAIf. Reduced and unreduced samples of the small fetuin-binding fruit lectin yielded a single polypeptide band of about 22 kDa, which clearly differed from the polypeptides of SNAIf (Fig. 2). SNAIf and the small lectin account for about 94% and 6%, respectively, of the affinity-purified protein. The total soluble protein content of ripe fruits is around 200 mg/kg (fresh weight) 20 mg of which binds to fetuin. Hence the small fetuin-binding protein represents roughly 0.6% of the total soluble fruit protein.

A similar 22 kDa fetuin-binding lectin was isolated from a preparation of fetuin-binding bark proteins. Two g of fetuin-binding bark proteins yielded about 2 mg of the small protein, indicating that this protein is present only in minute amounts.

N-terminal sequencing of the polypeptides of the small fetuin-binding fruit and bark proteins yielded the sequences SLENNIHAARQGWTVGDVE and QGWTVGDVEPLVT-

Fig. 1. Gel filtration of the Neu5Ac(α-2,6)Gal/GalNAc-binding elderberry fruit protein on a Superose 12 column. Panel A: total preparation of fetuin-binding fruit proteins; panel B: protein eluting in the first peak upon gel filtration on a Sephacryl 100 column; panel C: protein eluting in the second peak upon gel filtration on a Sephacryl 100 column. The elution position of SNAI (1, 240 kDa), SNAV (2, 120 kDa), SNAII (3, 60 kDa) and chymotrypsinogen (4, 25 kDa) is indicated by the arrow heads.
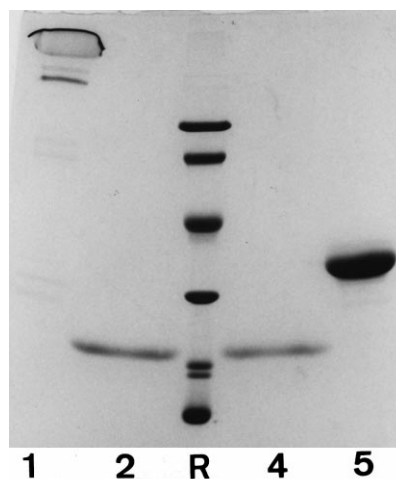
---

[1] SNAI and SNAIf refer to the Neu5Ac(α-2,6)Gal/GalNAc-specific type 2 RIP from bark and fruits, respectively.

```
SNAI-B  :  GGEYEKVCSVVEVTRRISGWDGLCVDVRYGHYIDGNPVQLRPCGNECNQL   50
           ::  :::::::::::::::::::::::::: :::::::: :::::::::::::
SNAIf-B :  GGGYEKVCSVVEVTRRISGWDGLCVDVRDGHYIDGNTVQLGPCGNECNQL


SNAI-B  :  WTFRTDGTIRWLGKCLTASSSVMIYDCNTVPPEATKWVVSIDGTITNPHS  100
           :::::::::::::::::::: ::::::::::::::::::::: ::::::: :
SNAIf-B :  WTFRTDGTIRWLGKCLTTSSSVMIYDCNTVPPEATKWVVSTDGTITNPRS


TrSNAI  :                          QGWTVGDVEPLVTFIVGYKQ
SNAI-B  :  GLVLTAPQAAEGTALSLENNIHAARQGWTVGDVEPLVTFIVGYKQMCLRE  150
           ::::::::::::::::::::::::::::::::::::::::::::::::::: :
SNAIf-B :  GLVLTAPQAAEGTALSLENNIHAARQGWTVGDVEPLVTFIVGYKQMCLTE
TrSNAIf :                    SLENNIHAARQGWTVGDVE


SNAI-B  :  NGENNFVWLEDCVLNRVQQEWALYGDGTIRVNSNRSLCVTSEDHEPSDLI  200
           :::::::::::::::::::: :::::::::::::::::::::::::::::::
SNAIf-B :  NGENNFVWLEDCVLNRVEQEWALYGDGTIRVNSNRSLCVTSEDHEPSDLI


SNAI-B  :  VILKCEGSGNQRWVFNTNGTISNPNAKLLMDVAQRDVSLRKIILYRPTGN  250
           ::::::::::::::::::::::::::::.::::: :::::::: ::::
SNAIf-B :  VILKCEGSGNQRWVFNTNGTISNPNAKLVMDVAQSNVSLRKIILYPPTGN


SNAI-B  :  PNQQWITTTHPA 262
           :::::::::: ::
SNAIf-B :  PNQQWITTTQPA
```

Fig. 3. Comparison of the deduced amino acid sequences of the cDNA clones encoding the B chain of SNAI (SNAI-B) and SNAIf (SNAIf-B), and the N-terminal amino acid sequences of TrSNAI and TrSNAIf. Putative N-glycosylation sites are shown in bold. Colons indicate identical amino acids in the B chains of SNAI and SNAIf, whereas dots represent chemically similar amino acids. The complete sequences of SNAI and SNAIf are available from the Genbank/EMBL Data library under accession numbers U27122 and AF012899, respectively.

FIVGYKQ, respectively. Sequence alignment revealed that the N-terminal sequence of the fruit polypeptide perfectly matches the deduced sequence of the B chain of SNAIf from S^{116} to E^{134} whereas the N-terminus of the bark polypeptide is identical to that of the B chain of SNAI from Q^{126} to Q^{145} (Fig. 3). These results strongly indicate that the small fetuin-binding proteins from fruits and bark correspond roughly to the second domain of the B chain of SNAIf and SNAI, respectively. Therefore, the small proteins have been called truncated SNAIf and SNAI (TrSNAIf and TrSNAI), respectively.

TrSNAIf and TrSNAI contain 14.3 and 14.6% covalently bound carbohydrate, respectively. These values correspond to about 15–16 monosaccharide units per 22 kDa polypeptide indicating that TrSNAIf and TrSNAI possess 2 glycan chains per monomer.

3.2. Carbohydrate-binding specificity and agglutination properties of TrSNAIf and TrSNAI

In contrast to SNAI and SNAIf which yield a clearly visible agglutination at concentrations as low as 5 μg/ml TrSNAIf and TrSNAI did not agglutinate trypsin-treated human (type A) or rabbit erythrocytes. Only at high concentrations (> 1 mg/ml) a slight clumping of the human erythrocytes was observed.

In the absence of a clear agglutination activity, the carbohydrate-binding specificity of TrSNAIf and TrSNAI could not be determined by hapten inhibition assays. To check whether TrSNAIf and TrSNAI exhibit the same specificity as SNAIf and SNAI, their binding to immobilized carbohydrates and glycoproteins was studied. TrSNAIf and TrSNAI

were retained on immobilized fetuin but not on asialofetuin indicating that terminal sialic acid is essential for binding. No binding was observed on immobilized galactose or N-acetylgalactosamine. Binding of TrSNAIf and TrSNAI to immobilized fetuin was inhibited by 20 μM Neu5Ac(α-2,6)Gal(β-1,4)Glc but not by Neu5Ac(α-2,3)Gal(β-1,4)Glc, galactose or lactose at the same concentration. The fact that TrSNAIf and TrSNAI behave like the parent SNAIf and SNAI with respect to their binding to the aforementioned affinity matrices suggests that TrSNAIf and TrSNAI also recognize Neu5-Ac(α-2,6)Gal/GalNAc-containing glycoconjugates.

3.3. Molecular modelling of TrSNAIf and TrSNAI

The hydropathic profiles of the TrSNAIf and TrSNAI polypeptides correspond to the most hydrophilic, accessible and flexible regions of the B chain of the type 2 RIP (results not shown). According to the three-dimensional models built from the X-ray coordinates of the ricin B chain TrSNAIf and TrSNAI consist of a single domain (corresponding to domain 2 of SNAIf and SNAI, respectively). These single domains are exclusively composed of coil structures interconnected by turns and loops (Fig. 4). In TrSNAIf Cys^{32}-Cys^{47} and Cys^{73}-Cys^{90} (corresponding to Cys^{147}-Cys^{162} and Cys^{188}-Cys^{205}, respectively, of the B chain of SNAIf) are linked through disulphide bridges. Similarly, in TrSNAI Cys^{22}-Cys^{37} and Cys^{63}-Cys^{80} (also corresponding to Cys^{147}-Cys^{162} and Cys^{188}-Cys^{205}, respectively, of the B chain of SNAI) are linked through disulphide bridges. SNAIf and SNAI contain three and two putative N-glycosylation sites, respectively in the C-terminal part of their respective B chains (Fig. 3). The putative glycosylation sites at positions 184 and 218 of the
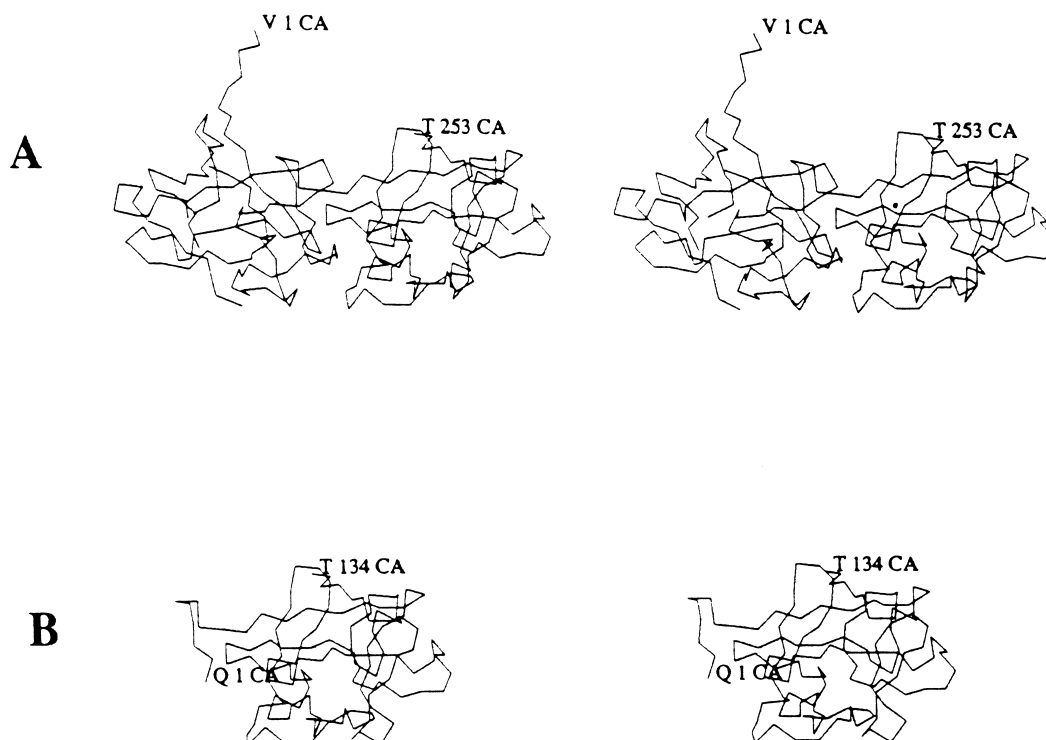
A

B

Fig. 4. Stereoviews of the three-dimensional models of the complete B chain of SNAI (A) and TrSNAI (B).

B chain are located in two exposed loops and are therefore readily accessible for glycosylation, which is in good agreement with the observation that TrSNAIf and TrSNAI contain two glycan chains per monomer. TrSNAIf as well as TrSNAI contain a functional carbohydrate-binding site, similar to that of domain 2 of the B chain of SNAIf and SNAI, respectively. The amino acid residues which built the binding site are Asp[116], Ile[128], Tyr[130], Asn[137] and Gln[138] in TrSNAIf and Asp[10], Ile[118], Tyr[120], Asn[127] and Gln[128] in TrSNAI (corresponding to Asp[231], Ile[243], Tyr[245], Asn[252] and Gln[253] of the B chain of SNAIf and SNAI). Accessible hydrophilic regions are located on the surface of the TrSNAIf and TrSNAI polypeptides. However, the models cannot predict whether interactions between exposed residues in these regions allow a non-covalent association of TrSNAIf and TrSNAI monomers into homodimers.

3.4. There is no evidence for the occurrence of mRNAs encoding C-terminal fragments of SNIf or SNAI

Since evidence has been presented recently that one of the GalNAc/Gal-specific lectins from elderberry fruits is encoded by a type 2 RIP gene with a major deletion in the A domain [11], the occurrence of truncated variants of SNAIf and SNAI mRNAs was carefully examined by Northern blot analysis. Hybridization of fruit and bark poly(A) RNA with a cDNA encoding SNAIf yielded exclusively a signal of 2 kb (results not shown). The absence of smaller hybridizing bands suggests that TrSNAIf and TrSNAI are not encoded by truncated mRNAs but are derived from the same precursors as SNAIf and SNAI.

3.5. C-terminal fragments of the GalNAc/Gal-specific lectins are not detectable in elderberry bark and fruits

The discovery of the C-terminal fragments of SNAIf and SNAI raised the question whether similar fragments occur of the GalNAc/Gal-specific elderberry type 2 RIP/lectins. Using the gel filtration techniques described above, no C-terminal fragments could be isolated from 10 g of a total preparation of the bark GalNAc/Gal-specific elderberry type 2 RIP/lectins (i.e. the mixture of SNAV and SNAII). Similarly, no C-terminal fragments could be recovered from 1 g of a total preparation of fruit GalNAc/Gal-specific elderberry type 2 RIP/lectins (i.e. the mixture of SNAVf and SNAIVf).

4. Discussion

Total preparations of fetuin-binding proteins from elderberry fruits and bark contain besides the previously described Neu5Ac(α-2,6)Gal/GalNAc-specific type 2 RIP also small proteins composed of polypeptide chains which correspond to the second domain of the B chain of the respective type 2 RIP. Since no evidence could be obtained for the occurrence of mRNAs or genes encoding truncated SNAIf or SNAI, TrSNAIf and TrSNAI are most likely proteolytic cleavage products of the parent lectins. TrSNAIf and TrSNAI bind to immobilized fetuin under the same conditions as SNAIf and SNAI. The binding to fetuin not only demonstrates that TrSNAIf and TrSANI have carbohydrate-binding activity but also confirms the previously made prediction (based on molecular modelling) that the Neu5Ac(α-2,6)Gal/GalNAc-binding activity of SNAIf and SNAI resides in site 2 of the respective B chains. In spite of the apparent presence of an active sugar-binding site TrSNAIf and TrSNAI exhibit no clear agglutination activity. This observation indicates that TrSNAIf and TrSNAI are monovalent which confirms that they occur as monomers. It should be noted here that the binding properties and molecular structure of the recombinant C-terminal half of the Neu5Ac(α-2,6)Gal/GalNAc-specific

type 2 RIP from *Sambucus sieboldiana* are virtually identical to these of TrSNAI and TrSNAIf [19].

Sequence comparisons leave no doubt that TrSNAIf and TrSNAI are proteolytic cleavage products of SNAIf and SNAI, respectively. Although the formation of TrSNAIf and TrSNAI most probably relies on a similar mechanism there are some tissue-specific differences because (i) the B chains of the parent type 2 RIP SNAIf and SNAI are cleaved at different positions in an identical sequence and (ii) the relative abundance of TrSNAIf is much higher than that of TrSNAI. These differences suggest that different proteases are involved in the formation of TrSNAIf and TrSNAI and that proteolysis is more extensive in the fruits than in the bark. At present it is not clear whether TrSNAIf and TrSNAI are formed in vivo or are generated during the extraction and isolation procedure.

The discovery of TrSNAIf and TrSNAI not only demonstrates that the C-terminal domain of the B chains of SNAIf and SNAI occurs as a separate Neu5Ac(α-2,6)Gal/GalNAc-binding protein but also provides another example of how elderberry produces different proteins from a single type 2 RIP precursor. Previous work has already shown, indeed, that processing of the precursor of SNAV yields, at least in the bark, a genuine type 2 RIP (SNAV) and a lectin composed of slightly truncated B chains (SNAII) [4].

References

[1] Barbieri, L., Battelli, M.G. and Stirpe, F. (1993) Biochim. Biophys. Acta 1154, 237–282.
[2] Barbieri, L., Valbonesi, P., Bonora, E., Gorini, P., Bolognesi, A. and Stirpe, F. (1997) Nucleic Acids Res. 25, 518–522.
[3] Peumans, W.J. and Van Damme, E.J.M. (1995) Plant Physiol. 109, 347–352.
[4] Van Damme, E.J.M., Barre, A., Rougé, P., Van Leuven, F. and Peumans, W.J. (1996) Eur. J. Biochem. 235, 505–513.
[5] Van Damme, E.J.M., Barre, A., Rougé, P., Van Leuven, F. and Peumans, W.J. (1996) Eur. J. Biochem. 235, 128–137.
[6] Van Damme, E.J.M., Barre, A., Rougé, P., Van Leuven, F. and Peumans, W.J. (1997) J. Biol. Chem. 272, 8353–8360.
[7] Van Damme, E.J.M., Roy, S., Barre, A., Citores, L., Mostafapous, K., Rougé, P., Van Leuven, F., Girbes, T., Goldstein, I.J. and Peumans, W.J. (1997) Eur. J. Biochem. 245, 648–655.
[8] Citores, L., Iglesias, R., Munoz, R., Ferreras, J.M., Jiminez, P. and Girbes, T. (1994) J. Exp. Bot. 45, 513–516.
[9] Citores, K., de Benito, F.M., Iglesias, R., Ferreras, J.M., Jiminez, P., Argüeso, P., Farias, G., Mendez, E. and Girbes, T. (1996) J. Exp. Bot. 47, 1577–1585.
[10] Girbes, T., Citores, L., de Benito, F.M., Iglesias, R. and Ferreras, J.M. (1996) Biochem. J. 315, 343–344.
[11] Van Damme, E.J.M., Roy, S., Barre, A., Rougé, P., Van Leuven, F. and Peumans, W.J. (1997) Plant J. 12, 1251–1260.
[12] Dubois, M., Gilles, K.A., Hamilton, J.K., Rebers, P.A. and Smith, F. (1956) Anal. Chem. 28, 350–356.
[13] Laemmli, U.K. (1970) Nature 227, 680–685.
[14] Kyte, J. and Doolittle, R.F. (1982) J. Mol. Biol. 157, 105–132.
[15] Janin, J. (1979) Nature 277, 491–492.
[16] Karplus, P.A. and Schulz, G.E. (1985) Naturwissenschaften 72, 212–213.
[17] Rutenber, E., Katzin, B.J., Collins, E.J., Mlsna, D., Ready, M.P. and Robertus, J.D. (1991) Proteins 10, 240–250.
[18] Rutenber, E. and Robertus, J.D. (1991) Proteins 10, 260–269.
[19] Minamihara, N., Jordan, E.T., Kakau, H., Minami, E., Hisajima, S. and Shibuya, N. (1997) Glycoconjugate J. 14, (Suppl. 1) S32.