

Biosynthesis, primary structure and molecular cloning of snowdrop (*Galanthus nivalis* L.) lectin

Els J. M. Van DAMME¹, Hanae KAKU², Fulvio PERINI², Irwin J. GOLDSTEIN², Ben PEETERS³, Fumio YAGI², Benny DECOCK⁴ and Willy J. PEUMANS¹

¹ Laboratory for Phytopathology and Plant Protection, Catholic University of Leuven, Belgium

² Department of Biological Chemistry, University of Michigan, Ann Arbor, USA

³ Department of Biochemistry, Catholic University of Leuven, Belgium

⁴ Rega Institute for Medical Research, Catholic University of Leuven, Belgium

(Received May 17, 1991) – EJB 91 0652

Poly(A)-rich RNA isolated from ripening ovaries of snowdrop (*Galanthus nivalis* L.) yielded a single 17-kDa lectin polypeptide upon translation in a wheat-germ cell-free system. This lectin was purified by affinity chromatography. Translation of the same RNA in *Xenopus laevis* oocytes revealed a lectin polypeptide which was about 2 kDa smaller than the *in vitro* synthesized precursor, suggesting that the oocyte system had removed a 2-kDa signal peptide.

A second post-translational processing step was likely to be involved since both the *in vivo* precursor and the *Xenopus* translation products were about 2 kDa larger than the mature lectin polypeptide. This hypothesis was confirmed by the structural analysis of the amino acid sequence of the mature protein and the cloned mRNA.

Edman degradation and carboxypeptidase Y digestion of the mature protein, and structural analysis of the peptides obtained after chemical cleavage and modification, allowed determination of the complete 105 amino acid sequence of the snowdrop lectin polypeptide. Comparison of this sequence with the deduced amino acid sequence of a lectin cDNA clone revealed that besides the mature lectin polypeptide, the lectin mRNA also encoded a 23 amino acid signal-sequence and a C-terminal extension of 29 amino acids, which confirms the results from *in vitro* translation experiments.

Lectins are a heterogeneous class of (glyco)proteins grouped together based on their ability to recognize and bind carbohydrate moieties of glycoconjugates. Although numerous plant lectins have been isolated and characterized in detail, the physiological function of these proteins remains unclear (Etzler, 1986). A study of the biosynthesis of lectins and their subsequent subcellular deposition can be helpful in obtaining a better understanding of the molecular biology of lectins. In combination with physiological studies of the lectins concerned (e.g. occurrence and abundance of lectin in different plant tissues) it is possible to obtain important information in a search for the normal biological function of these proteins.

A few years ago a mannose-specific lectin was isolated from bulbs of snowdrop (*Galanthus nivalis*) (Van Damme et al., 1987). The *G. nivalis* agglutinin is a tetrameric protein composed of identical 12.5-kDa subunits, which is devoid of carbohydrate. In recent years similar lectins have been found in bulbs of all representatives of the plant family Amaryllidaceae, e.g. *Narcissus* (Van Damme et al., 1988). Furthermore we have shown that these lectins are found in

almost all plant tissues where they represent the most prominent proteins at certain developmental stages of the plant (Van Damme and Peumans, 1990). Amaryllidaceae lectins have been found to be very useful tools for the analysis and purification of glycoproteins (Shibuya et al., 1988). Moreover they are very potent and selective inhibitors of retroviruses and cytomegalovirus *in vitro* (Balzarini et al., 1990).

A recent study of the *in vivo* biosynthesis of the snowdrop lectin in ripening ovaries has shown that it is synthesized on the rough endoplasmic reticulum as a higher molecular-mass precursor (15 kDa) which is post-translationally converted into the mature 12.5-kDa lectin polypeptide (Van Damme and Peumans, 1988). We report here the *in vitro* translation and processing of the snowdrop lectin, and the isolation of the *in vivo* synthesized lectin precursor. Furthermore we show that the complete amino acid sequence of the mature protein as obtained from structural protein analysis and the sequence of the precursor as deduced from the nucleotide sequence of its cloned mRNA.

EXPERIMENTAL PROCEDURES

Materials

Flowering plants and snowdrop bulbs (*G. nivalis* L.) were collected from a local garden and purchased from a local store, respectively. Oligo-deoxythymidine cellulose and an SDS molecular-mass kit were purchased from Sigma Chemical

Correspondence to E. J. M. Van Damme, Catholic University Leuven, Laboratory for Phytopathology and Plant Protection, Willem de Croylaan 42, B-3001 Leuven, Belgium

Abbreviations: GNA, *Galanthus nivalis* agglutinin (lectin); Pth, phenylthiohydantoin.

Note. The novel nucleotide sequence reported in this paper has been deposited with the Genbank™/EMBL Data library and is available under accession number M55556.

Co. (St Louis, MO). Radioisotopes were obtained from the Amersham Corp. A cDNA synthesis kit, the multifunctional phagemid pT₇T₃ 18U, restriction enzymes and DNA-modifying enzymes were obtained from Pharmacia LKB Biotechnology Inc. *Escherichia coli* XL1 Blue competent cells were purchased from Stratagene (La Jolla, CA). Cyanogen bromide was purchased from Kodak Laboratory and Research products (Rochester, NY). Trypsin treated with tosylphenylalanylchloromethane was obtained from Worthington Biochemical Co. (Freehold, NJ). Carboxypeptidase Y was obtained Boehringer Mannheim GmbH (Federal Republic of Germany).

Lectin isolation

The snowdrop lectin (GNA) was isolated from extracts of bulbs by affinity chromatography on immobilized mannose as reported previously (Van Damme et al., 1987) followed by gel filtration on Sephacryl S-200 (1.5 × 120 cm), eluting with 6 M guanidinium/HCl. The lectin fraction was dialyzed against distilled water and lyophilized. Sequence analysis was performed on a sample of snowdrop lectin purified by affinity chromatography on Synsorb—mannose followed by gel filtration on Sephacryl S-200 in 6 M guanidine/HCl.

RNA isolation

Total cellular RNA was prepared from ovary tissue (stored at -80 °C) essentially as described by Finkelstein and Crouch (1986). Poly(A)-rich RNA was enriched by chromatography on oligo-deoxythymidine cellulose as described by Siflow et al. (1979) except that poly(A)-rich RNA was eluted at room temperature. A second chromatography of the RNA on the same column was performed after Me₂SO treatment of the RNA.

Protein synthesis in the wheat-germ cell-free extract

Poly(A)-rich RNA was translated in a wheat-germ cell-free extract which was prepared as described previously (Peumans et al., 1982) except that the homogenization buffer contained 6 mM 2-mercaptoethanol and an additional 15 mM magnesium acetate. In the presence of a higher concentration of these divalent cations, most of the endogenous messenger ribonucleoproteins from large aggregates which will sediment during the subsequent centrifugation of the homogenate (Peumans et al., 1980). The incubation mixture for cell-free translation contained/ml: 0.4 ml extract, 20 mM Hepes/KOH pH 8.0, 90 mM potassium acetate, 2.5 mM magnesium acetate, 1 mM ATP, 20 μM GTP, 25 μg/ml creatine kinase, 8 mM creatinephosphate, 2 mM dithiothreitol, 0.4 mM spermidine, 19 unlabeled amino acids (100 μM each) and 50 μCi [³⁵S]methionine or 100 μCi [³H]leucine. Prior to adding the mRNA (30 μg/ml final concentration) and the labeled amino acid, the cell-free translation system was incubated for 15 min at room temperature to reduce the endogenous template activity of the system. Subsequently the reaction mixtures were incubated for 1 h at 25 °C and the incorporation of labeled amino acid into trichloroacetic-acid-insoluble material determined.

Isolation and characterization of the in vitro synthesized lectin polypeptides

Lectin isolation from the *in vitro* translation mixture was accomplished either by affinity chromatography on man-

nose—Sephacryl, single or double immunoprecipitation or immunoaffinity chromatography. Whenever microsomal membranes or organelles were present in the incubation mixture they were disrupted using 1% Triton X-100 prior to the isolation of the lectin. Affinity chromatography on mannose—Sephacryl was carried out in the presence of 0.5 M ammonium sulphate (as described by Van Damme and Peumans, 1988) or NaCl/Pi (1.5 mM KH₂PO₄/10 mM Na₂HPO₄/3 mM KCl/140 mM NaCl, pH 7.4). The lectin was desorbed using unbuffered 20 mM 1,3-diaminopropane.

Small aliquots were withdrawn for determination of [³H]leucine or [³⁵S]methionine incorporation. The lectin was subsequently analyzed by SDS/PAGE on 12.5–25% acrylamide gradient gels using a discontinuous system as described by Laemmli (1970). After fixing and destaining, gels were immersed in 1 M sodium salicylate for 30 min, dried under partial vacuum and exposed to X-ray films (Fuji RX, Japan) at -80 °C.

Protein synthesis after injection of X. laevis oocytes

Frogs were anaesthetized by immersion in a 0.1% solution of MS-222 or Metacaini Mesylas in water for 30 min. An ovary lobe was surgically removed and transferred to Barth's solution. Oocytes were manually defolliculated and stored in Barth's solution at 18 °C. Groups of ten oocytes were microinjected with 1 μg poly(A)-rich RNA each and incubated 100 μl Barth's medium containing 25 μCi [³H]leucine. After incubation for 24–28 h at 18 °C, oocytes were homogenized in 10 mM Tris/HCl pH 7.6 containing 150 mM NaCl, 1 mM EDTA, 0.5% sodium desoxycholate, 0.5% Triton X-100, 0.1% SDS and 0.2 μg/ml Trasylol. Afterwards the suspension was centrifuged at 10000 g for 10 min and the lectin isolated by affinity chromatography on mannose—Sephacryl in the presence of 0.5 M ammonium sulphate.

Isolation of the in vivo lectin precursor

Snowdrop ovaries were homogenized as reported previously (Van Damme and Peumans, 1988). The organelles were separated from the soluble proteins and small molecules on a Sepharose 4B column (5 × 30 cm) as described by Van der Wilden et al. (1980). Fractions containing the organelles were pooled and brought to 0.5% Triton X-100 and 0.5 M ammonium sulphate. The lectin was isolated by affinity chromatography on mannose—Sephacryl.

Isolation of microsomal membranes

Microsomal membranes from dog pancreas were prepared essentially as described by Shields and Blobel (1978).

Amino acid sequence analysis

Protein sequencing was conducted on an Applied Biosystems (Foster City, California, USA) Model 470A protein sequencer interfaced with an Applied Biosystems model 120A on-line analyzer.

Cyanogen bromide cleavage and isolation of peptides

Snowdrop lectin (10 mg) dissolved in 0.5 ml 70% (by vol.) formic acid was added to a cyanogen bromide solution (10 mg/0.7 ml formic acid) and the reaction mixture was incubated for 24 h at 25 °C. After degassing by water vacuum, the solution

was diluted tenfold with distilled water and lyophilized. The resulting peptides were dissolved in 0.2 M acetic acid, the insoluble material was removed by centrifugation and the solution applied to a Bio-Gel P-6 gel-filtration column (2.5 × 105 cm), eluting with 0.2 M acetic acid. The peptide fractions were monitored by ultraviolet absorption at 230 nm and 280 nm. The combined fractions M1 (tubes 65–70) and M2 (82–95, see Fig. 5) were dialyzed against distilled water using spectrapor MWCO:1000 and lyophilized.

Acetylation of M1 peptide

Acetylated M1 (0.5 mg) was prepared by the addition of pyridine (300 µl) and acetic anhydride (150 µl) followed by incubation for 15 h at 25°C. After removal of pyridine by repeated evaporation with the concomitant addition of water (4 times), the acetylated M1 peptide was lyophilized.

SDS/urea/polyacrylamide gel electrophoresis

12.5% SDS/urea/PAGE was performed using the procedure of Swank and Munkres (1971). Each fraction, (M1 and M2) was applied to the 12.5% SDS/urea/acrylamide gel using the SDS molecular-mass-marker kit as standard (myoglobin, 16950; 14400; 8160; 6210; 2510 Da). The polyacrylamide gel was stained with Coomassie brilliant blue.

Carboxypeptidase-Y digestion

Carboxypeptidase Y was added to GNA and the M1 peptide solution in 10 mM sodium phosphate buffer, pH 7.0. After incubation at 25°C for 5 min and 30 min, the reaction was stopped by adding 10% trichloroacetic acid (5% final concentration) and the precipitate was removed by centrifugation. The free amino acid composition in the supernatant was analyzed.

Reduction and alkylation

GNA and the M1 peptide were dissolved in 0.1 M ammonium bicarbonate containing 6 M guanidine/HCl, pH 8.1. A tenfold excess of 2-mercaptoethanol over the estimated cysteine concentration was added to the protein solution and the reaction mixture was incubated at 50°C for 3 h. A 2.5-fold excess of 4-vinylpyridine or iodoethane [20% (by vol.) in methanol] over 2-mercaptoethanol was added to the reaction mixture and alkylation was allowed to proceed for 1 h. The samples were dialyzed against distilled water (3 ×) and lyophilized. The pyridylethylated protein was used for sequencing while the methylated form was used for amino acid analysis.

Amino acid analysis

Aliquots of reduced and methylated sample were hydrolyzed in teflon-lined screw-capped culture tubes, in 6 M HCl for 24 h under an atmosphere of nitrogen. The acid was evaporated under a stream of nitrogen. The sample was reconstituted in 0.2 M sodium citrate, pH 2.6. Separation was achieved on a Kratos amino acid analysis system with Beckman 110 A pumps, on a 46 × 120 mm Interaction (Los Altos, CA) cation-exchange column (sodium form). The buffers for the three-step gradient (pH 3.28, 4.25 and 7.40) were purchased from Pickering Labs (Mountain View, CA).

Post-column derivatization with sodium hypochlorite and *o*-phthaldehyde was according to Fujiwara et al. (1987).

Construction and screening of the cDNA library

A cDNA library was constructed from total poly(A)-rich mRNA isolated from snowdrop ovaries using the cDNA synthesis kit from Pharmacia. cDNA fragments were inserted into the *Eco*RI site of the multifunctional phagemid pT₇T₃ 18U. The library was propagated in *E. coli* XL1 Blue.

Recombinant lectin clones were screened by colony hybridization using a ³²P-end-labeled degenerate oligonucleotide probe (15-mer, 5' TGT GTT TGT T/A/GGC CCA 3') derived from residues 41–45 of a partially known amino acid sequence for the lectin (Fig. 4). Hybridization was carried out overnight at 38°C in 0.9 M sodium chloride containing 90 mM Tris/HCl pH 7.5, 6 mM EDTA, 10 × Denhardt's solution, 0.1% SDS, 180 µg/ml hydrolyzed yeast RNA and 2 × 10⁶ cpm/ml ³²P-labeled probe. Hybridization filters were then washed four times in 6 × NaCl/Cit (1 × NaCl/Cit, 0.9 M sodium chloride and 0.09 M sodium citrate, pH 7.0) at room temperature for 15 min, followed by a 5-min wash at hybridization temperature in 6 × NaCl/Cit. Filters were blotted dry, wrapped in Saran Wrap and exposed to Kodak X-Omat S film at –80°C. Colonies that produced positive signals 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).

Computer analyses

Comparison of the amino acid sequence of the snowdrop lectin to the lectins from several mannose-binding proteins from plants, animals and bacteria was conducted with protein sequence data NBRF (Protein Sequence Database of the protein identification resource, National Biomedical Research Foundation). DNA sequences were analyzed using programs from PC Gene and Genepro.

RESULTS AND DISCUSSION

In vitro translation of poly(A)-rich mRNA in a wheat-germ cell-free system

Poly(A)-rich RNA isolated from ovaries of snowdrop was translated in a wheat-germ cell-free system. Afterwards the lectin was isolated by affinity chromatography on mannose–Sephacrose. As shown in Fig. 1a single polypeptide with an apparent molecular mass of approximately 17 kDa is retained on the column. Compared to the *in vivo* precursor which has a molecular mass of 15 kDa (Van Damme and Peumans, 1988) it seems likely that the lectin is synthesized *in vitro* with a signal peptide.

Fractionation of the total wheat-germ translation mixture on a gel-filtration Superose 12 column (Pharmacia LKB Biotechnology Inc.) revealed the highest concentration of *in vitro* lectin precursor in fractions 33–34 whereas the mature lectin (50 kDa) elutes in fraction 39, as determined in a separate run, indicating that the lectin precursor has a higher molecular mass (Fig. 2). From this experiment it can be concluded that the *in vitro* synthesized precursor of the snowdrop lectin does not occur as a monomer but is an oligomer containing four precursor chains.

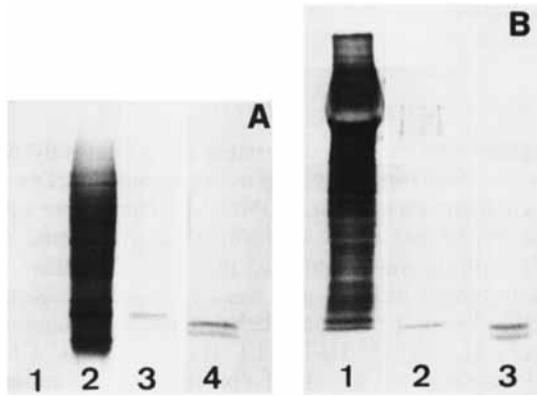


Fig. 1. Fluorograph of translation products of poly(A)-rich RNA from snowdrop in a wheat-germ cell-free system (A) and *Xenopus oocytes* (B). (A) Lane 1, endogenous products of the wheat-germ cell-free system; lane 2, total homogenate of translation products synthesized in response to poly(A)-rich RNA from snowdrop; lane 3, affinity-purified lectin polypeptide from translation products of a wheat-germ system; lane 4, *in vivo* labeled precursor and lectin polypeptides. (B) Lane 1, total homogenate of proteins from *Xenopus oocytes* injected with poly(A)-rich mRNA isolated from snowdrop; lane 2, affinity purified lectin from *Xenopus oocytes*; lane 3, *in vivo* labeled precursor and lectin polypeptides

Isolation of the lectin synthesized *in vitro*

In a previous paper we have shown that the lectin synthesized in *in vivo* labeling experiments could easily be isolated by affinity chromatography on mannose–Sephacryl in the presence of 0.5 M ammonium sulphate (Van Damme and Peumans, 1988). However, for some unknown reason the *in vitro* synthesized lectin was poorly retained on the column in the presence of ammonium sulphate. When affinity chromatography was performed in NaCl/Pi the *in vitro* lectin polypeptides readily bound to the affinity matrix. It appears that both the *in vivo* and the *in vitro* synthesized precursor exhibit the same carbohydrate-binding activity as the lectin itself. Therefore the sequence which is cleaved off when the precursor is processed into the lectin is not essential for the carbohydrate-binding properties of the lectin precursor.

Besides affinity chromatography on mannose–Sephacryl, the lectin can also be isolated by immunoprecipitation or immunoaffinity chromatography. However, when precipitating lectin polypeptides from a wheat-germ extract, non-specific reactions often occur. By adding detergents to the cell-free system and the washing solutions it was possible to reduce the non-specific binding but not to abolish it completely.

Cell-free protein synthesis in the presence of microsomes

Many secretory proteins or proteins which are found in specific locations in the cell are synthesized with an N-terminal signal sequence which is proteolytically cleaved off *in vivo* before the synthesis of the polypeptide is completed. To investigate whether the snowdrop lectin is also synthesized with such an extra signal sequence, *in vitro* translation experiments in the presence of microsomes were carried out.

When a wheat-germ cell-free system was supplemented with dog pancreas microsomes there was no visible change in molecular mass of the lectin precursor on the fluorogram. However, when a wheat-germ extract was prepared from developing embryos (which have an extended network of rough

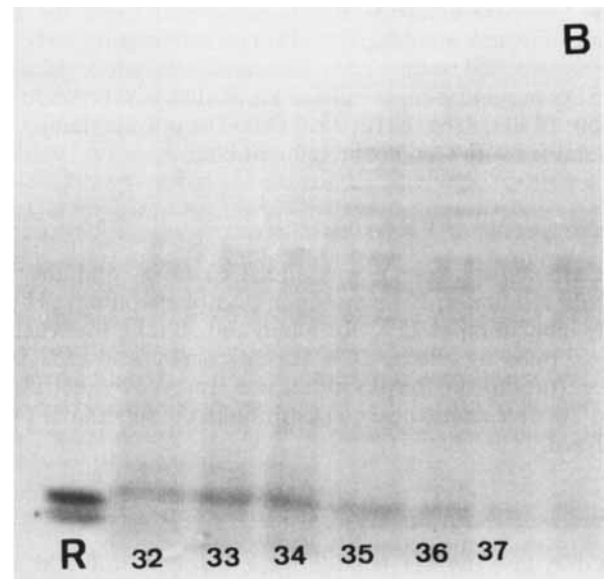
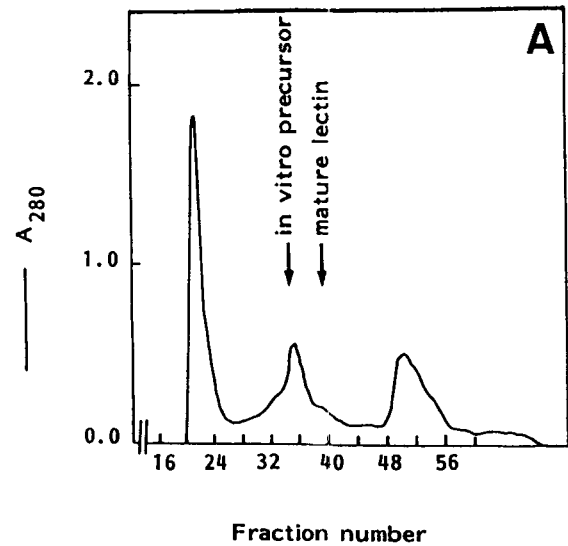


Fig. 2. Gel filtration on a Superose 12 column of a total wheat-germ cell-free system after translation of poly(A)-rich RNA from snowdrop (A). Fractions of 300 μ l were collected and the lectin immunoprecipitated. The elution position of the lectin was determined in a separate run. NaCl/Pi containing 0.2 M mannose was used as running buffer. The flow rate was 20 ml/h. Lectin polypeptides were immunoprecipitated and analyzed by SDS/PAGE and subsequently by fluorography (B). The numbers under each lane correspond to the fraction number shown in A. In lane R *in vivo* labeled precursor and lectin were run

endoplasmic reticulum) and used in *in vitro* translation experiments, two polypeptide bands with molecular masses of 17 kDa and 15 kDa appeared on the autofluorogram indicating that the lectin is partially processed (results not shown). Since these experiments were highly dependent on the batch of embryos used, we tried to confirm the existence of a signal peptide by synthesizing the lectin in *Xenopus oocytes*.

Lectin polypeptides synthesized in *Xenopus oocytes*

In order to determine whether *X. laevis* oocytes are able to synthesize the snowdrop lectin they were injected with

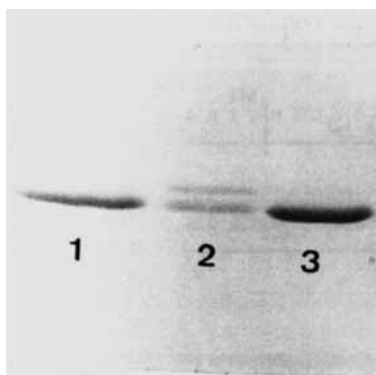


Fig. 3. SDS/PAGE of the *in vivo* synthesized lectin in snowdrop ovaries. Lane 1, lectin isolated from the soluble fraction; lane 2, lectin isolated from the organelle fraction; lane 3, snowdrop lectin

poly(A)-rich RNA isolated from snowdrop ovaries. After 26 h of incubation, in the presence of [3 H]leucine, oocytes were homogenized and lectin isolated by affinity chromatography on mannose–Sephacrose. As shown in Fig. 1 *Xenopus* oocytes synthesized a lectin polypeptide with a molecular mass of 15 kDa. Since it is generally accepted that oocytes can process plant preprotein polypeptides (Larkins et al., 1979; Boston et al., 1982) we assume that the signal sequence had been cleaved off. *Xenopus* oocytes have also been shown to secrete both animal secretory proteins (Lane, 1981) and intracellularly deposited plant proteins (Bassuner et al., 1983) when injected with the corresponding mRNA. However, different proteins are exported at intrinsically different rates (Colman and Morser, 1979). Although it has been reported before that the snowdrop lectin is synthesized on the rough endoplasmic reticulum (Van Damme and Peumans, 1988) and is probably located in protein bodies as is the case for many other lectins and storage proteins, we found no evidence for secretion of the lectin at least not within the incubation period of 26 h (results not shown).

Isolation of the *in vivo* lectin precursor

A lectin fraction enriched in precursor was isolated from the organelle fraction of snowdrop ovaries by affinity chromatography on mannose–Sephacrose. As shown on SDS/PAGE (Fig. 3) the lectin isolated from this fraction contains both the mature lectin and the lectin precursor. The soluble fraction, however, only contains the mature lectin polypeptide.

N-terminal amino acid sequence analysis of the lectin preparation containing both the lectin precursor and the mature lectin revealed only one sequence. Since this amino acid sequence is identical to that of an independently sequenced single isolectin of snowdrop (results not shown) we can conclude that both the mature lectin and the *in vivo* lectin precursor contain the same N-terminal sequence unless the precursor polypeptide is blocked. Since no evidence could be obtained for glycosylation of any of the precursor forms (results not shown) we assume that the *in vivo* synthesized precursor will have an additional sequence at the C terminus which can account for the 2-kDa difference in molecular mass found between the mature lectin polypeptide and both the *in vivo* precursor and the *Xenopus* translation products.

Structural analysis of the mature protein

Using chemical cleavage and modification, and both trypsin and carboxypeptidase Y digestion, the complete amino acid sequence of the snowdrop lectin was determined. As shown in Fig. 4, the complete amino acid sequence of GNA consists of 105 amino acid residues/subunit. The snowdrop lectin is a pure tetrameric protein rich in asparagine, glycine and leucine, and contains three cysteine residues. Based on the amino acid sequence composition, the molecular mass of GNA was calculated to be 11 724 Da. This GNA protein was readily and directly sequenced up to residue 57. However, four amino acid residues between residue 51 and 54 were ambiguous in the first sequencing.

GNA contains two methionine residues/subunit located at positions 25 and 56. Therefore the cleavage of methionine residues of GNA using cyanogen bromide is a convenient way for separating the large GNA C-terminal peptide from the two N-terminal small peptides (Fig. 4). The elution profile of GNA on a Bio-gel P-6 gel-filtration column following treatment with cyanogen bromide reagent is shown in Fig. 5. Two peptide fractions were obtained: M1, obtained by combining fractions 65–70 and M2, obtained from fractions 82–95. The molecular size of M1 and M2, calculated from SDS/urea (12.5%) PAGE appeared to be approximately 4.2 kDa and 2.0 kDa, respectively. However, M2 contained two peptides (M2-1 and M2-2), which consist of position 1–25 (M2-1) and 26–56 (M2-2). We were unable to separate these peptides; hence, they were sequenced simultaneously. The results indicate that they were present in equimolar amounts and correspond to the N-terminal sequence of 1–56.

M1, which represents the C-terminal peptide isolated from methionine cleavage of GNA by cyanogen bromide incubation, was sequenced from the glutamine residue at position 57 to the arginine residue at position 101. Moreover, digestion of the intact protein and peptide M1 with carboxypeptidase Y yielded mostly two amino acids, suggesting that the C-terminal sequence was Thr-Gly.

Acetylation of M1 blocked two lysyl residues at positions 70 and 90, protecting them from tryptic digestion, while blocking the N-terminal residue at position 57 of the M1 peptide. Trypsin cleaved acetylated M1 at arginines 92 and 101, yielding an N-terminally blocked peptide encoding residues 57–92 and two small peptides which were not separated. Sequencing the mixture confirmed the 93–101 sequence (MA-1 fragment) already determined from the intact peptide M1. A new sequence which contained the two C-terminal amino acids of M1 (and of the intact lectin) formed fragment MA-2 (residues 102–105).

Two cysteines were assigned to peptide M2-2. If the peptide was alkylated with vinylpyridine in the absence of reductant, only background was found at residues 29 and 52, while Pth-Pro-Glu-Cys (PTH, phenylthiohydantoin) was detected at both positions after reduction and alkylation, indicating that a disulfide bond existed between the two cysteines. We must note that this disulfide bond is very difficult to reduce in the intact protein. However, when M1 was alkylated without prior reduction, Pth-Pro-Glu-Cys was found at residue 86, indicating the presence of free cysteine. The disulfide bond does not appear to be involved in sugar binding inasmuch as GNA containing 10% dithiothreitol did not lose its precipitating activity with *Saccharomyces cerevisiae* yeast α -mannan.

Several instances of micro-heterogeneity of the GNA amino acid sequence were found at positions 51 (S or H), 53 (F or Y) and 86 (S or C). This is not surprising since the lectin

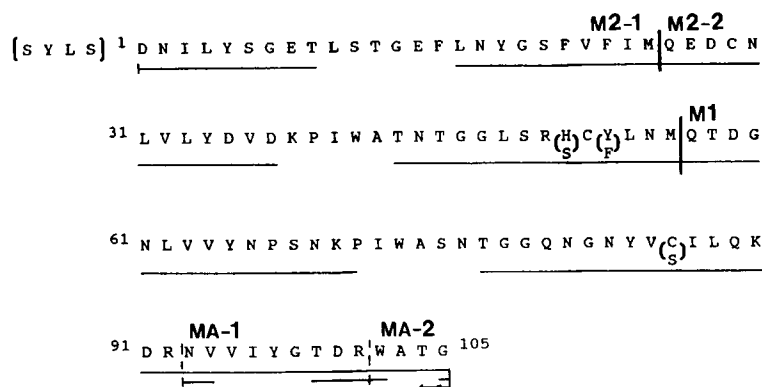


Fig. 4. Complete amino acid sequence of the snowdrop lectin. (—) Extent of the fragments used to construct the sequence, assignment of residues by amino acid sequencing; (←) results of carboxypeptidase Y digestion; open parentheses (), possession of heterogeneity of amino acid residue. Solid vertical lines (|) show points of cleavage by cyanogen bromide. Dotted vertical lines (:) show points of cleavage by trypsin on peptide M1

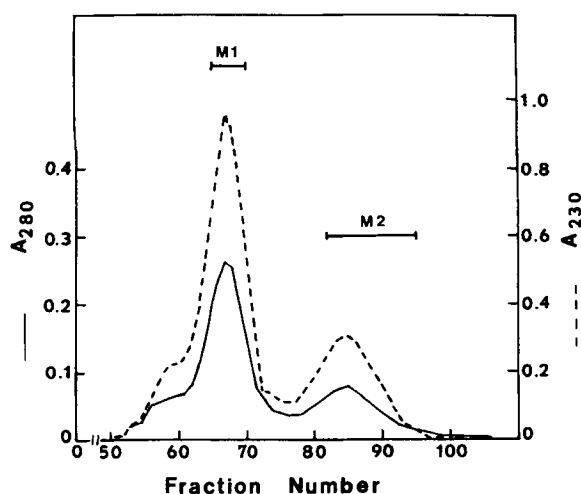


Fig. 5. Elution pattern of cyanogen bromide fragments of snowdrop lectin on a column of Bio-Gel P-6 (2.5×105 cm) in 0.2 M acetic acid

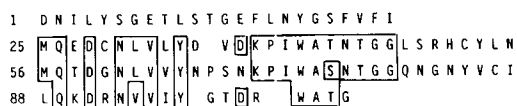


Fig. 6. Internal homology of segments 25–47, 56–79 and 89–104 of the snowdrop lectin sequence

is a very complex mixture of isolectins as was shown upon analysis of the purified protein by ion-exchange chromatography or isoelectric focusing (Van Damme et al., 1988). Moreover we now have evidence for the occurrence of multiple isoforms at the molecular level since different lectin cDNA clones which definitely differ in their amino acid sequence have been isolated (E. Van Damme, unpublished results). Moreover, 30% of GNA possessed four more amino acids (Ser-Tyr-Leu-Ser) at the N-terminal amino acid residue (Fig. 4). It is suggested that these two GNA amino acid sequences undergo differential enzyme processing during their differentiation.

Interestingly a portion of amino acid sequence of Met25–Gly47 has strong identity between Met56–Gly79 and Gln89–Thr104, showing 70% and 35% identity, respectively (Fig. 6).

Characterization of the snowdrop lectin cDNA

In order to confirm the existence of a pre-proform of the lectin, a cDNA library was constructed in the multifunctional phagemid pT₃T₇ 18U using poly(A)-rich RNA isolated from

snowdrop ovaries. The nucleotide sequence of the cDNA lectin clone selected by colony hybridization and the deduced amino acid sequence are depicted in Fig. 7. The lectin clone contains a 570-bp open-reading frame with one possible initiation codon at position 18. Translation starting with this codon generates a 157 amino acid polypeptide with a calculated molecular mass of 16917 Da, corresponding in size to the *in vitro* translation product for the snowdrop lectin. The 3' untranslated region contains six in-frame termination codons and one potential polyadenylation signal at position 532. Although the cDNA library was constructed using an oligo-dT primer for first strand synthesis, the lectin clone has no poly(A) tail.

Using the rules for protein processing of Von Heijne (1986) three possible cleavage sites were identified between residues 18 and 19, residues 19 and 20, and residues 23 and 24 of the amino acid sequence deduced from the cDNA clone. Processing of the signal peptide at the most probable cleavage site between residues 23 and 24 (following the rules of Von Heijne) generates a lectin polypeptide with a calculated molecular mass of 14602 Da starting with the first amino acid of the protein and will account for a 2315-Da decrease in molecular mass upon transport of the polypeptide across the endoplasmic reticulum. Cleavage of the signal peptide between residues 19 and 20 will generate a lectin polypeptide possessing four amino acids more at the N terminal, as found in protein sequence determinations.

Comparison of the C-terminal-deduced amino acid sequence of both the lectin clone and the protein reveals that the first stop codon occurs 29 amino acids further downstream of Gly105. Therefore we assume that a C-terminal extension (2944 Da) is removed during post-translational processing of the protein. The hydrophobic character of this C-terminal peptide is consistent with the possibility that it is removed post-translationally. The loss of C-terminal extensions has

CAACTACAAGTTACAAA ATG GCT AAG GCA AGT CTC CTC ATT TTG GCC GCC ATC TTC CTT 59
Met Ala Lys Ala Ser Leu Leu Ile Leu Ala Ala Ile Phe Leu

GGT GTC ATC ACA CCA TCT TGG CTG AGT[▼]GAC AAT ATT TTG TAC TCC GGT GAG ACT CTC TCT 119
Gly Val Ile Thr Pro Ser Cys Leu Ser Asp Asn Ile Leu Tyr Ser Gly Glu Thr Leu Ser

ACA GGG GAA TTT CTC AAC TAC GGA AGT TTC GTT TTT ATC ATG CAA GAG GAC TGC AAT CTG 179
Thr Gly Glu Phe Leu Asn Tyr Gly Ser Phe Val Phe Ile Met Gln Glu Asp Cys Asn Leu

GTC TTG TAC GAC GTG GAC AAG CCA ATC TGG GCA ACA AAC ACA GGT GGT CTC TCC CGT AGC 239
Val Leu Tyr Asp Val Asp Lys Pro Ile Trp Ala Thr Asn Thr Gly Gly Leu Ser Arg Ser

TGC TTC CTC AGC ATG CAG ACT GAT GGG AAC CTC GTG GTG TAC AAC CCA TCG AAC AAA CCG 299
Cys Phe Leu Ser Met Gln Thr Asp Gly Asn Leu Val Val Tyr Asn Pro Ser Asn Lys Pro

ATT TGG GCA AGC AAC ACT GGA GGC CAA AAT GGG AAT TAC GTG TGC ATC CTA CAG AAG GAT 359
Ile Trp Ala Ser Asn Thr Gly Gly Gln Asn Gly Asn Tyr Val Cys Ile Leu Gln Lys Asp

AGG AAT GTT GTG ATC TAC GGA ACT GAT CGT TGG GCT ACT GGA[▼]ACT CAC ACC GGA CTT GTT 419
Arg Asn Val Val Ile Tyr Gly Thr Asp Arg Trp Ala Thr Gly Thr His Thr Gly Leu Val

GGA ATT CCC GCA TCG CCA CCC TCA GAG AAA TAT CCT ACT GCT GGA AAG ATA AAG CTT GTG 479
Gly Ile Pro Ala Ser Pro Pro Ser Glu Lys Tyr Pro Thr Ala Gly Lys Ile Lys Leu Val

ACG GCA AAG TAA TGA CCGGTGATCTTTAACTTGCATGTATGTGGGAAGAGTAATAAAAATAAGTCATTGGAGA 553
Thr Ala Lys ■ ■

TAATCGACCTCGTCGGC 570

Fig. 7. Nucleotide sequence and deduced amino acid sequence of snowdrop lectin cDNA clone LECGNA 2. The putative processing sites for the signal sequence and the C-terminal extension, are indicated by arrowheads. Two stop codons at the end of the coding region are indicated with squares. The potential polyadenylation site is underlined

been reported for several plant lectins [wheat germ agglutinin B, Raikhel and Wilkins (1987); rice lectin, Wilkins and Raikhel (1989); barley lectin, Lerner and Raikhel (1989); concanavalin A, Carrington et al. (1985); pea lectin, Higgins et al. (1983)] as well as plant storage proteins such as thaumatin (Edens et al., 1982), napin (Ericson et al., 1986) and 2S albumin (Krebbes et al., 1988), and β -1,3-glucanase (Shinshi et al., 1988). However, the possible function of these C-terminal peptides remains unclear.

Computer searches of the EMBL DNA sequence databases have not revealed any significant homology of the snowdrop lectin sequence with any other protein. Comparison of the protein sequence of the snowdrop lectin with several mannose-binding lectins isolated from plants [e.g. *Canavalia*, Wang et al. (1975); pea, Higgins et al. (1983); lentin, Foriers et al. (1981)] or rat serum and liver (Ikeda et al., 1987; Drickamer et al., 1986) or *E. coli* (Klemm, 1984) indicated only about 20% identity. However, three mannose-specific bulb lectins isolated from snowdrop, daffodil and amaryllis belonging to the same family Amaryllidaceae showed high similarity among each other (approximately 80–89%) (I. J. Goldstein, unpublished results).

This work was supported in part by grant GM 29470 from the National Institute of Health and grants of the Nationale Bank and the National Fund for Scientific Research (Belgium, FGWO grant

2005989 N). W. P. is Research Director and E. V. D. Research Assistant of this fund. We would like to thank Prof G. Préaux and Dr C. De Greef for their help with the translation experiments in *Xenopus* oocytes. Furthermore, we are most grateful to Drs N. De Clercq, Dr F. Claessens and Dr K. Hemschoote for their advice in molecular cloning experiments and to V. Feytons and R. De Greef for their expert technical assistance, and Ms L. Loesel and K. Mountjoy for assistance in protein sequencing.

REFERENCES

- Balzarini, J., Schols, D., Neyts, J., Van Damme, E., Peumans, W. & De Clercq, E. (1990) *Antimicrob. Agents Chemother.* 35, 410–416.
- Bassuner, R., Huth, A., Manteuffel, R. & Rapoport, T. (1983) *Eur. J. Biochem.* 133, 321–326.
- Boston, R. S., Miller, T. J., Mertz, J. E. & Burgess, R. R. (1982) *Plant Physiol.* 69, 150–154.
- Carrington, D. M., Auffret, A. & Hanke, D. E. (1985) *Nature* 313, 64–67.
- Colman, A. & Morser, J. (1979) *Cell* 17, 517–526.
- Drickamer, K., Dordal, M. S. & Reynolds, L. (1986) *J. Biol. Chem.* 261, 6878–6887.
- Edens, L., Heslinga, L., Klok, R., Ledebouer, A. M., Maat, J., Toonen, M. Y., Visser, C. & Verrips, C. T. (1982) *Gene* 18, 1–12.
- Ericson, M. L., Rodin, J., Lenman, M., Glimelius, K., Josefsson, L.-G. & Rask, L. (1986) *J. Biol. Chem.* 261, 14576–14581.
- Etzler, M. E. (1986) in *The lectins: properties, functions and applications in biology and medicine* (Liener, I. E., Sharon, N. & Goldstein, I. J., eds) pp. 371–435, Academic Press Inc.
- Finkelstein, R. R. & Crouch, M. L. (1986) *Plant Physiol.* 81, 907–912.
- Foriers, A., Lebrun, E., Van Rapenbusch, R., de Neve, R. & Strosberg, A. D. (1981) *J. Biol. Chem.* 256, 5550–5560.
- Fujiwara, M., Ishida, Y., Nimura, N., Toyama, A. & Kinoshita, T. (1987) *Anal. Biochem.* 166, 72–78.
- Higgins, T. J. V., Chandler, P. M., Zurawski, G., Button, S. C. & Spencer, D. (1983) *J. Biol. Chem.* 258, 9544–9549.
- Ikeda, K., Sannoh, T., Kawasaki, N., Kawasaki, T. & Yamashina, I. (1987) *J. Biol. Chem.* 262, 7451–7454.
- Klemm, P. (1984) *Eur. J. Biochem.* 143, 395–399.
- Krebbes, E., Herdies, L., De Clercq, A., Seurinck, J., Leeman, J., Van Damme, J., Segura, M., Gheysen, G., Van Montagu, M. & Vandekerckhove, J. (1988) *Plant Physiol.* 87, 859–866.
- Laemmli, U. K. (1970) *Nature* 227, 680–685.
- Lane, C. D. (1981) *Cell* 24, 281–282.
- Larkins, B. A., Pedersen, K., Handa, A. K., Hurkman, W. J. & Smith, L. D. (1979) *Proc. Natl Acad. Sci. USA* 76, 6448–6452.
- Lerner, D. R. & Raikhel, N. V. (1989) *Plant Physiol.* 91, 124–129.
- Mierendorf, R. C. & Pfeffer, D. (1987) *Methods Enzymol.* 152, 556–562.
- Peumans, W. J., Carlier, A. R. & Delaey, B. M. (1980) *Plant Physiol.* 66, 584–587.
- Peumans, W. J., Stinissen, H. M., Tierens, M. & Carlier, A. R. (1982) *Plant Cell Rep.* 1, 212–214.
- Raikhel, N. V. & Wilkins, T. A. (1987) *Proc. Natl Acad. Sci. USA* 84, 6745–6749.
- Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl Acad. Sci. USA* 74, 5463–5467.
- Shibuya, N., Berry, J. E. & Goldstein, I. J. (1988) *Arch. Biochem. Biophys.* 267, 676–680.
- Shields, D. & Blobel, G. (1978) *J. Biol. Chem.* 253, 3753–3756.
- Shinshi, H., Wenzler, H., Neuhaus, J.-M., Felix, G., Hofsteenge, J. & Meins, F. Jr (1988) *Proc. Natl Acad. Sci. USA* 85, 5541–5545.
- Siflow, C. D., Hammett, J. R. & Key, J. L. (1979) *Biochemistry* 18, 2725–2731.
- Swank, R. T. & Munkres, K. D. (1971) *Anal. Biochem.* 39, 462–477.
- Van Damme, E. J. M., Allen, A. K. & Peumans, W. J. (1987) *FEBS Lett.* 215, 140–144.
- Van Damme, E. J. M. & Peumans, W. J. (1988) *Plant Physiol.* 86, 922–926.

- Van Damme, E. J. M., Allen, A. K. & Peumans, W. J. (1988) *Physiol. Plant.* 73, 52–57.
- Van Damme, E. J. M. & Peumans, W. J. (1990) *Planta* 182, 605–609.
- Van Der Wilden, W., Gilkes, N. R. & Chrispeels, M. J. (1980) *Plant Physiol.* 66, 390–394.
- Von Heijne, G. (1986) *Nucleic Acids Res.* 14, 4683–4690.
- Wang, J. L., Cunningham, B. A., Waxdal, M. J. & Edelman, G. M. (1975) *J. Biol. Chem.* 250, 1490–1502.
- Wilkins, T. A. & Raikhel, N. V. (1989) *The Plant Cell* 1, 541–549.