THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 262, No. 28, Issue of October 5, pp. 13392-13403, 1987
Printed in U.S.A.

Structure, Position, and Biosynthesis of the High Mannose and the Complex Oligosaccharide Side Chains of the Bean Storage Protein Phaseolin*

(Received for publication, December 22, 1986)

Arnd Sturm‡§, J. Albert Van Kuik¶, J. F. G. Vliegenthart¶, and Maarten J. Chrispeels‡||

From the ‡Department of Biology C-016, University of California, San Diego, La Jolla, California 92093 and the ¶Department of Bio-Organic Chemistry, Transitorium III, Utrecht University, The Netherlands

Phaseolin, the major storage protein of the common bean (*Phaseolus vulgaris*), is a glycoprotein which is synthesized during seed development and accumulates in protein storage vacuoles or protein bodies. The protein has three different *N*-linked oligosaccharide side chains: $Man_9(GlcNAc)_2$, $Man_7(GlcNAc)_2$, and Xyl-$Man_3(GlcNAc)_2$ (where Xyl represents xylose). The structures of these glycans were determined by ^1H NMR spectroscopy. The $Man_9(GlcNAc)_2$ glycan has the typical structure found in plant and animal glycoproteins. The structures of the two other glycans are shown below.

```
                B
              Manα1→6
                        4′
                      Manα1→6
                A                 3          2            1
              Manα1→3           Manβ1→4GlcNAcβ1→4GlcNAcβ1→N-Asn
   D1         C          4
 Manα1→2Manα1→2Manα1→3

                4′
              Manα1→6
                        3          2            1
                      Manβ1→4GlcNAcβ1→4GlcNAcβ1→N-Asn
                4
              Manα1→3

              Xylβ1→2
```

Phaseolin was separated by electrophoresis on denaturing gels into four size classes of polypeptides. The two abundant ones have two oligosaccharides each, whereas the less abundant ones have only one oligosaccharide each. Polypeptides with two glycans have $Man_7(GlcNAc)_2$ attached to Asn^{252} and $Man_9(GlcNAc)_2$ attached to Asn^{341}. Polypeptides with only one glycan have Xyl-$Man_3(GlcNAc)_2$ attached to Asn^{252}. Both these asparagine residues are in canonical glycosylation sites; the numbering starts with the N-terminal methionine of the signal peptide of phaseolin. The presence of the $Man_7(GlcNAc)_2$ and of Xyl-$Man_3(GlcNAc)_2$ at the same asparagine residue (position 252) of different polypeptides seems to be controlled by the glycosylation status of Asn^{341}. When Asp^{341} is unoccupied, the glycan at Asn^{252} is complex. When Asn^{341} is occupied, the glycan at Asn^{252} is only modified to the extent that 2 mannosyl residues are removed.

The processing of the glycans, after the removal of the glucose residues, involves enzymes in the Golgi apparatus as well as in the protein bodies. Formation of the Xyl-$Man_3(GlcNAc)_2$ glycan is a multistep process that involves the Golgi apparatus-mediated removal of 6 mannose residues and the addition of 2 *N*-acetylglucosamine residues and 1 xylose. The terminal *N*-acetylglucosamine residues are later removed in the protein bodies. The conversion of $Man_9(GlcNAc)_2$ to $Man_7(GlcNAc)_2$ is a late processing event which occurs in the protein bodies. Experiments in which [^3H]glucosamine-labeled phaseolin obtained from the endoplasmic reticulum (*i.e.* precursor phaseolin) is incubated with jack bean α-mannosidase show that the high mannose glycan on Asn^{252}, but not the one on Asn^{341}, is susceptible to enzyme degradation. Incubation of [^3H]

* This work was supported by grants (to M. J. C.) from the National Science Foundation (Metabolic Biology) and United States Department of Energy Grant DE-FG03-86ER13497. The work in J. F. G. V.'s laboratory was supported by the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organization for the Advancement of Pure Research (Zuiver Wetenschappelijk Onderzoek). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ Supported by a fellowship from the Deutsche Forschungsgemeinschaft.

|| To whom correspondence should be addressed.

glucosamine-labeled phaseolin obtained from the Golgi apparatus with jack bean β-*N*-acetylglucosaminidase results in the removal of the terminal *N*-acetylglucosamine residues from the complex chain. These observations are discussed in relation to the hypothesis that the control of glycan modification is determined largely by the accessibility of the glycan chains to the glycosidases and glycosyltransferases in the Golgi apparatus.

The asparagine-linked oligosaccharides found on plant glycoproteins, like those of other eukaryotes, fall into two general categories: high mannose and complex oligosaccharides. The high mannose oligosaccharides generally have the formula $Man_{5-9}(GlcNAc)_2$, with a branching pattern of mannosyl residues similar to that of the high mannose oligosaccharides of animal and yeast cells (1, 2). The complex oligosaccharides have a $Man_3(GlcNAc)_2$ core to which one or more of the following sugars may be attached: xylose, fucose, *N*-acetylglucosamine, and/or galactose. The fucose is attached to the proximal *N*-acetylglucosamine of the chitobiose core, as in lima bean agglutinin (3) and laccase of *Acer pseudoplatanus* (4); whereas the xylose is attached to the β-linked mannose residue, as in stem bromelain (5), the protease inhibitor of *Caesalpinia pulcherrima* (6), and laccase (4). The structures of only a few *N*-linked glycans of plant glycoproteins have been determined. Because of our interest in the biosynthesis and transport of storage proteins and lectins in developing legume seeds (7), we have made a detailed study of the structure and biosynthesis of the oligosaccharide side chains of phaseolin, the major storage protein of the common bean. This protein accumulates during seed development in special protein storage vacuoles (protein bodies).

When phaseolin isolated from cotyledons of beans (*Phaseolus vulgaris* cv. Greensleeves) is subjected to SDS-PAGE,[1] it can be separated into four polypeptides, more accurately called size classes (8). These four polypeptide classes, referred to as A, B, C, and D, range in M_r from 52,000 to 45,000. Each polypeptide band can be further resolved by two-dimensional electrophoresis into two or more polypeptides with different isoelectric points near pH 5 (9, 10).

Previous reports from our and other laboratories detailed some important events during the biosynthesis and processing of phaseolin (11). When cotyledon mRNA is translated *in vitro*, two phaseolin polypeptide size classes of M_r 48,000 (α) and 45,000 (β) are made (12). *In vivo*, phaseolin is synthesized on polysomes bound to the rough ER (13). The polypeptides are cotranslationally glycosylated with either one or two oligosaccharide chains, converting the α-polypeptides into polypeptides A and B, with A having two oligosaccharide chains and B having one (11). Analogously, the β-polypeptides become the glycosylated polypeptides C and D, with C having two oligosaccharide chains and D having one. Transport of these glycosylated polypeptides through the Golgi complex (14) and their ultimate deposition in the protein bodies (15) are accompanied by further processing steps.

The *N*-linked oligosaccharides of plant glycoproteins are synthesized initially as $Glc_3Man_9(GlcNAc)_2$ groups (16, 17) which are cotranslationally transferred from dolichol pyrophosphate to nascent polypeptide chains. Processing of the oligosaccharides starts in the ER, continues in the Golgi apparatus, and involves numerous glycosidases and glycosyltransferases. There is a considerable amount of information concerning the processing of glycoproteins in animal cells (18, 19), but very little is known about these events in plant cells. The processing of *N*-linked glycans in plant cells also begins with the loss of 3 glucose residues, followed by the removal of up to 7 mannose residues and the possible addition of fucose, xylose, *N*-acetylglucosamine, and galactose residues. The result of these processing events is a variety of glycans that range in size up to $Hex_{15-17}(GlcNAc)_2$ (where Hex represents hexose) when assayed on Bio-Gel P-4 columns (20). Developing cotyledons of leguminous seeds represent an excellent system to study the processing of *N*-linked glycans (7, 22).

Phaseolin is encoded by a small multigene family (23). Gene copy number analysis indicates that there are approximately seven phaseolin genes/haploid genome (24). The phaseolin gene family can be divided into two main gene types, α and β, which encode the two polypeptide size classes α and β detectable after translation of mRNA *in vitro*. The nine longest of the published nucleotide sequences of phaseolin cDNA clones show 98% homology between the α- and β-type genes (24). All amino acid sequences derived from these cDNA sequences have two canonical glycosylation sites. We refer to these sites as Asn^{252} and Asn^{341}, counting amino acid residues from the initiating methionine of the signal sequence. Asn^{252} is located in a sequence Gly-Asn-Leu-Thr-Glu in a hydrophilic protein domain, and Asn^{341} is in a Val-Asn-Phe-Thr-Gly sequence in a hydrophobic protein domain.

In this paper, we report the structures of the three different phaseolin oligosaccharides as determined by ^1H NMR spectroscopy. We show their distribution among the various phaeolin polypeptides and the specific glycosylation sites (asparagine residues). In addition, we provide evidence for specific processing steps in the Golgi complex and the protein bodies.

MATERIALS AND METHODS AND RESULTS[2]

The structures of the three principal glycans present in Pronase digests of purified phaseolin were determined by ^1H NMR and are shown in the Miniprint (Figs. 1–5). The principal polypeptides A and C each have $Man_9(GlcNAc)_2$ and $Man_7(GlcNAc)_2$ in equal proportions. Polypeptide D has mainly a small complex glycan: $Xyl\text{-}Man_3(GlcNAc)_2$. We have no analytical data on polypeptide B because we were unable to purify sufficient amounts. However, indirect evidence indicates that its glycan resembles the one from polypeptide D. Both polypeptides B and D have a single glycan (11), the polypeptides stain poorly with the Schiff stain for glycoproteins (42) (data not shown), and the glycan is endo-β-*N*-acetylglucosaminidase H-resistant (Fig. 2).

Identification of the Glycosylation Site for the Three Different Phaseolin Oligosaccharides—The amino acid sequences of

[1] The abbreviations used are: SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; ConA, concanavalin A; Xyl, xylose; ER, endoplasmic reticulum; HPLC, high pressure liquid chromatography; endo H, endo-β-*N*-acetylglucosaminidase; PBS, phosphate-buffered saline; FAB-MS, fast atom bombardment-mass spectrometry; TPCK, L-1-tosylamido-2-phenylethyl chloromethyl ketone; Ph 1 and 2, phaseolin 1 and 2.

[2] Portions of this paper (including "Materials and Methods," part of "Results," Figs. 1–5 and 9–11, and Table I) are presented in miniprint at the end of this paper. Miniprint is easily read with the aid of a standard magnifying glass. Full size photocopies are available from the Journal of Biological Chemistry, 9650 Rockville Pike, Bethesda, MD 20814. Request Document No. 86-4358, cite the authors, and include a check or money order for $6.40 per set of photocopies. Full size photocopies are also included in the microfilm edition of the Journal that is available from Waverly Press.

the phaseolin polypeptides derived from the nucleotide sequences of nine different cDNAs show two possible glycosylation sites/polypeptide: Asn^{252} and Asn^{341} (16, 46). To find out which oligosaccharide is attached to which asparagine residue, affinity-purified phaseolin from cotyledons labeled for 24 h with [^{3}H]glucosamine was digested with trypsin. The resulting mixture of peptides and glycopeptides was fractionated by HPLC on a C_{18} column (Fig. 6*A*). Only one major radioactive peak (fraction 81) was observed (Fig. 6*B*). An aliquot of the peak fraction was digested with Pronase, and gel filtration of the resulting glycopeptides resolved three peaks corresponding in size to $Man_9(GlcNAc)_2Asn$, $Man_7(GlcNAc)_2(Gly)Asn$, and $Xyl\text{-}Man_3(GlcNAc)_2(Gly)Asn$ (data not shown). Next, the tryptic glycopeptides from HPLC fraction 81 were fractionated by ConA chromatography. When the radioactive material which did not bind to ConA was rerun on the reverse-phase C_{18} column, a major peak at fraction 81 appeared again (Fig. 6*C*). Amino acid sequencing of the first five amino acids of this tryptic glycopeptide (peak I) revealed the sequence Gln-Asp-Asn-Thr-Ile (QDNTI), which is identical with the amino acid sequence of the predicted tryptic glycopeptide that includes Asn^{252} (Fig. 7). Sizing of the glycan on Bio-Gel P-4 after exhaustive Pronase diges-

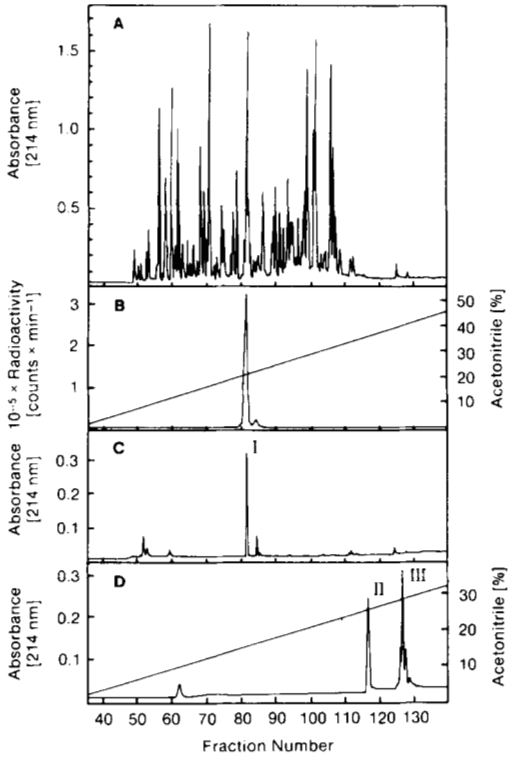

FIG. 6. **Reverse-phase HPLC separation of tryptic peptides from phaseolin.** *N*-[^{3}H]Acetylglucosamine-labeled phaseolin, prepared as described for Fig. 1, was digested with L-1-tosylamido-2-phenylethyl chloromethyl ketone-treated trypsin. *A–C*, tryptic peptides separated on a Vydac C_{18} column. The column was eluted with 0.1 M sodium phosphate, pH 2.2, for 30 min, followed by a linear gradient of 0–50% acetonitrile for 120 min. *D*, tryptic peptides were separated on an Alltech C_8 column. This column was eluted with 0.1% (v/v) trifluoroacetic acid in water for 30 min, followed by a linear gradient of 0–35% acetonitrile in 0.1% trifluoroacetic acid for 150 min. *A*, elution profile of the tryptic peptides as measured by absorbance at 214 nm. *B*, elution profile as measured by radioactivity. The radioactive peak (fraction 81) was applied to a ConA column with results similar to those shown in Fig. 3*A*. *C*, elution profile of the glycopeptide fraction which did not bind to ConA (peak I in Fig. 3*A*). *D*, elution profile on the C_8 column of the glycopeptide fraction that bound tightly to ConA (peak II in Fig. 3*A*). Peaks I (*C*), II, and III (*D*) were submitted to amino acid sequence analysis.

```
                240
Glycopeptide      QDNTI
Phaseolin       SKQDNTIGNEFGNLTERTDNSLNVLISSIEMEEGALFVPH
                  ▲           ▲

                280
Phaseolin       YYSKAIVILVVNEGEAHVELVGPKGNKETLEYESYRAELS

                320
Glycopeptide                      ATSNV
Phaseolin       KDDVFVIPAAYPVAIKATSNVNFTGFGINANNNNRNLLAGK
                                  ▲                    ▲
```

FIG. 7. **Comparison of the amino acid sequences of the tryptic glycopeptides obtained by HPLC with the derived amino acid sequence from phaseolin cDNA (36).** The beginning and end of the two tryptic glycopeptides (predicted) are marked with *arrowheads*. The two glycosylation sites, Asn-*X*-Thr, are *underlined*.

tion showed it to be the complex phaseolin glycan Ph 1. When the fraction 81 glycopeptides that bound ConA were fractionated by HPLC on a C_8 column, we observed two major radioactive peaks (Fig. 6*D*, peaks II and III). These glycopeptides were subjected to amino acid sequencing, and the oligosaccharides were analyzed by gel filtration after exhaustive digestion with Pronase. Analysis of peak II identified it as $Man_9(GlcNAc)_2$ attached to a peptide starting with the sequence Ala-Thr-Ser-Asn-Val (ATSNV). Similarly, peak III was identified as $Man_7(GlcNAc)_2$ attached to a peptide starting with the sequence Gln-Asp-Asn-Thr-Ile (QDNTI). Comparison of these amino acid sequences with the two predicted tryptic glycopeptides indicates that $Man_9(GlcNAc)_2$ is attached to Asn^{341} and $Man_7(GlcNAc)_2$ is attached to Asn^{252} (Fig. 7).

Biochemical Characterization of the Phaseolin Oligosaccharide Intermediates—To study the biosynthesis and processing of the phaseolin oligosaccharides, we labeled cotyledons for 3 and 24 h with [^{3}H]glucosamine, isolated phaseolin, and separated the polypeptides by preparative SDS-PAGE. Polypeptides A, C, and D were obtained by electroelution and digested with Pronase, and the glycopeptides were fractionated on ConA (Fig. 8). The polypeptide size class B, which was not clearly separated from peptides A and C on the gels, remained as a contaminant in the size classes A and C. The glycopeptides prepared from polypeptides A and C after a 3-h labeling period all bind tightly to ConA-agarose. After 24-h labeling, 88% of this glycopeptide fraction still binds to ConA. In contrast, polypeptides of the phaseolin size class D isolated after 3-h labeling contained a mixture of glycans which either did not bind to the ConA column or were slightly retarded (Fig. 8, *lower left panel*, peaks A and B, respectively) or bound tightly and were eluted with α-methylmannoside. After 24-h labeling, the proportion of the complex glycans was about 70% of the total and now eluted as one major peak from the ConA column (Fig. 8, *lower right panel*).

The glycopeptides resolved by ConA chromatography were analyzed by sizing on a long column (1 × 100 cm) of Bio-gel P-4 before and after treatment with various glycosidases. The results of these experiments are presented in Figs. 9–11 and can be summarized as follows. Initially (3-h labeling period), both polypeptides A and C have two $Man_9(GlcNAc)_2$ glycans. One glycan remains as $Man_9(GlcNAc)_2$, and the other is slowly processed to $Man_7(GlcNAc)_2$. Polypeptide D has mainly a complex glycan which, in short-time labeling experiments, has terminal *N*-acetylglucosamine residues. After 24-h labeling, these terminal *N*-acetylglucosamine residues are no longer present.

Localization of the Phaseolin Oligosaccharide Processing Events—To determine whether the different oligosaccharide

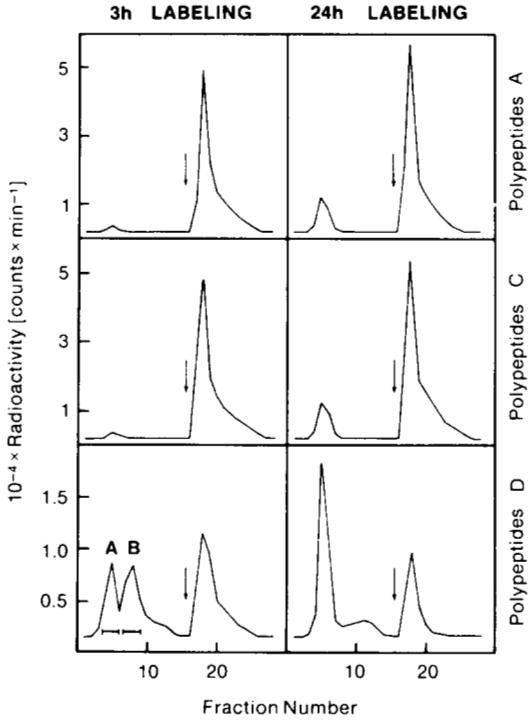

FIG. 8. **Analysis of the glycopeptides from phaseolin size classes A, C, and D by ConA affinity chromatography.** Each of 20 cotyledons was labeled with 6 μCi of [^{3}H]glucosamine for 3 and 24 h, and phaseolin was isolated by immunoaffinity chromatography. Purified phaseolin was subjected to preparative SDS-PAGE, and the individual polypeptides were recovered by electroelution (polypeptide B remained as a contaminant in polypeptides A and C). Glycopeptides were obtained by exhaustive digestion with Pronase, followed by gel filtration on a small column (2 × 18 cm) of Bio-Gel P-4 to remove small peptides and amino acids. The purified glycopeptides were then subjected to chromatography on ConA-agarose. The *arrows* indicate the starting point of elution with 200 mM α-methylmannoside.

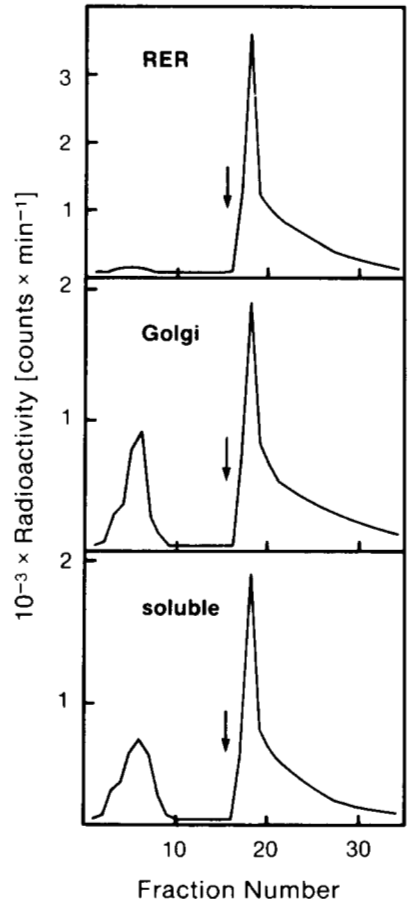

FIG. 12. **ConA chromatography of glycopeptides from phaseolin isolated from rough ER, Golgi, and soluble fractions.** Each of 20 cotyledons was labeled with 6 μCi of [^{3}H]glucosamine for 1 h. The radioactive tissue was homogenized in buffered sucrose containing 2 mM $MgCl_2$ and then separated into an organelle and soluble fraction on a column of Sepharose 4B. The organelles were fractionated on linear gradients of 16–54% (w/w) sucrose, and the rough ER (*RER*) and Golgi apparatus were identified by marker enzymes as described (14). Fractions containing the rough ER and Golgi apparatus were diluted with an equal volume of phosphate-buffered saline containing 1% Tween, and the phaseolin was isolated using anti-phaseolin-Sepharose. Phaseolin from the soluble fraction was isolated by immunoaffinity chromatography. Each fraction of phaseolin was digested by Pronase, purified by passage through the short Bio-Gel P-4 column, and then fractionated on ConA-agarose columns as described for Fig. 1. The *arrows* indicate the starting point of elution with 200 mM α-methylmannoside.

processing events take place in the cell, cotyledons were labeled for 1 h with [^{3}H]glucosamine, and the organelles were fractionated on sucrose gradients. The homogenization procedure disrupts the large protein bodies, and their proteins become admixed with cytosolic proteins (28). Rough ER and the Golgi apparatus were collected according to their sucrose densities of 1.13 and 1.18 g·cm^{-3}, respectively. Radiolabeled phaseolin, isolated from the membrane preparations with anti-phaseolin-Sepharose in the presence of Tween 20 and from the soluble fraction by the affinity method of Stockman *et al.* (27), was digested with Pronase; and the glycopeptides obtained were submitted to ConA affinity chromatography. The glycopeptides of total phaseolin isolated from the ER carry high mannose-type oligosaccharides exclusively as they all bind to ConA-agarose (Fig. 12, *upper panel*). On the other hand, about 25% of the glycopeptides of phaseolin isolated from the Golgi apparatus and the soluble fractions do not bind to ConA (Fig. 12, *middle* and *lower panels*). This is consistent with the known role of the Golgi apparatus in the conversion of high mannose to complex side chains on glycoproteins (see Ref. 19).

The various glycopeptides obtained by ConA affinity chromatography were further analyzed by gel filtration (Fig. 13). In addition to the glycopeptides described above, those obtained from soluble phaseolin labeled for 3 and 24 h with *N*-[^{3}H]acetylglucosamine were analyzed. Gel filtration of the ConA-binding (ConA+) glycopeptides from the ER-localized phaseolin shows that they co-migrated with $Man_9(GlcNAc)_2Asn$. They are not distinguishable from the ConA+ glycopeptides found in the Golgi apparatus or the soluble fraction after 1-h labeling. However, the ConA+ glycopeptides of soluble phaseolin labeled for 3 h can be resolved into a major peak corresponding to $Man_9(GlcNAc)_2Asn$ and a minor peak that co-migrated with $Man_7(GlcNAc)_2(Gly)Asn$. After a 24-h labeling period, the $Man_9(GlcNAc)_2Asn$ and $Man_7(GlcNAc)_2(Gly)Asn$ peaks are present in a ratio of 1:1, as found for the mature protein. Thus, the processing from $Man_9(GlcNAc)_2$ to $Man_7(GlcNAc)_2$ is only apparent after a long labeling period and is therefore presumed to take place after the transport of phaseolin to the protein bodies.

Gel filtration of the Golgi apparatus-derived phaseolin glycopeptides that did not bind ConA (ConA−) yielded two peaks that eluted in positions corresponding to the complex glycopeptides obtained from 3-h-labeled polypeptide D (compare the *Golgi panel* in Fig. 13 to the *CONTROL panel* in Fig. 10). Treatment of these Golgi apparatus-derived glycopeptides with β-*N*-acetylglucosaminidase generated a product which

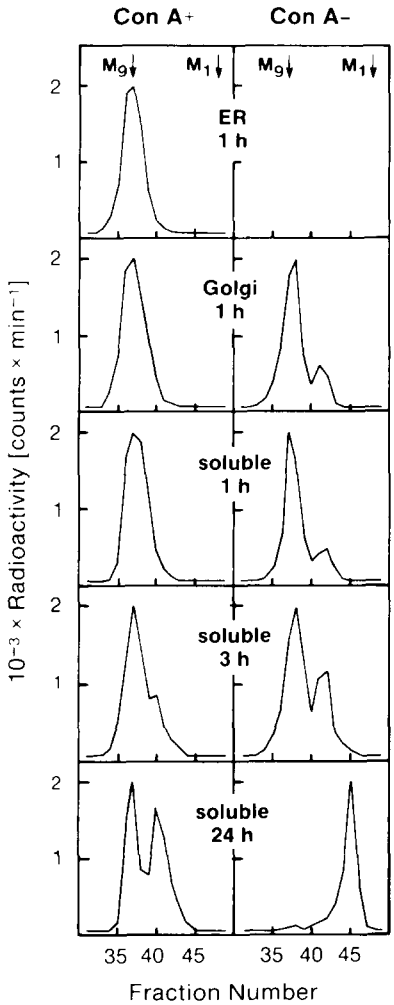

FIG. 13. **Gel filtration of glycopeptides from phaseolin isolated from rough ER, Golgi,and soluble fractions.** Phaseolin glycopeptides obtained by ConA affinity chromatography as described for Fig. 12 were analyzed by gel filtration as described for Fig. 9. Glycopeptides of phaseolin from the soluble fraction labeled for 3 and 24 h with [^{3}H]glucosamine were included. Glycopeptides which were not retained by ConA-agarose columns are labeled as *ConA*−, and those which required α-methylmannoside elution are designed *ConA*+. *Arrows* M_9 and M_1 indicate the elution positions of $Man_9(GlcNAc)_2Asn$ and $Man_1(GlcNAc)_2Asn$, respectively.

co-migrated with Xyl-$Man_3(GlcNAc)_2$(Gly)Asn, and simultaneous treatment with β-*N*-acetylglucosaminidase and α-mannosidase resulted in a product that co-migrated with Xyl-$Man_1(GlcNAc)_2$(Gly)Asn (data not shown). The complex (ConA−) glycopeptides obtained from the soluble phaseolin fraction after 1-h labeling showed the same size distribution as was found for the same in the Golgi apparatus. After a 3-h labeling period, the glycopeptide corresponding to the peak with less terminal *N*-acetylglucosamine residues increased in level. After 24-h labeling, all the ConA− glycopeptides comigrated with Xyl-$Man_3(GlcNAc)_2$(Gly)Asn. Thus, the processing of a high mannose glycan to a complex glycan occurs in the Golgi complex where the complex glycan obtains terminal *N*-acetylglucosamine residues. This is followed by the slow removal of terminal *N*-acetylglucosamine residues after the protein arrives in the protein bodies.

Accessibility of Phaseolin High Mannose Oligosaccharides to α-Mannosidase Digestion in Vitro—Phaseolin isolated from the rough ER carries only high mannose-type oligosaccharides with 9 mannose residues (Fig. 13, *upper left panel*). Digestion of undenatured ER-derived phaseolin with jack bean α-mannosidase converts approximately half of the $Man_9(GlcNAc)_2$ groups into an oligosaccharide shortened by 4 mannose residues (data not shown). To determine if this accessibility to α-mannosidase *in vitro* related to the attachment site of the oligosaccharide (Asn^{252} or Asn^{341}), we analyzed tryptic glycopeptides of the α-mannosidase-treated phaseolin.

The mixture of tryptic peptides was separated by C_8 reverse-phase HPLC (see Fig. 6*D*). The glycopeptides were identified by measuring the radioactivity in each fraction, and the N termini of the glycopeptides were determined by amino acid sequence analysis. The oligosaccharides attached to these glycopeptides were analyzed by sizing them on a calibrated column of Bio-Gel P-4 after exhaustive digestion with Pronase (Fig. 14, *middle panel*). Two controls are included in Fig. 14. Tryptic peptides, obtained from phaseolin which was not incubated with α-mannosidase, were separated on the C_8 column. One-half of each glycopeptide fraction was digested with Pronase, and the other half with α-mannosidase followed by Pronase. The reaction products were analyzed by sizing them on a calibrated column of Bio-Gel P-4 (Fig. 14, *upper* and *lower panels*, respectively). The results can be summarized as follows. The N-terminal region of the glycopeptide in HPLC fraction 118 (Fig. 6, HPLC peak II) is H_2N-Ala-Thr-Ser-Asn-Val (ATSNV) and carries $Man_9(GlcNAc)_2$. Comparison with the amino acid sequences of the two predicted glycopeptides (Fig. 7) shows that this high mannose oligosaccharide is attached to Asn^{341} in the hydrophobic protein domain. When present in the undenatured ER-derived glycoprotein, this oligosaccharide is not accessible to jack bean

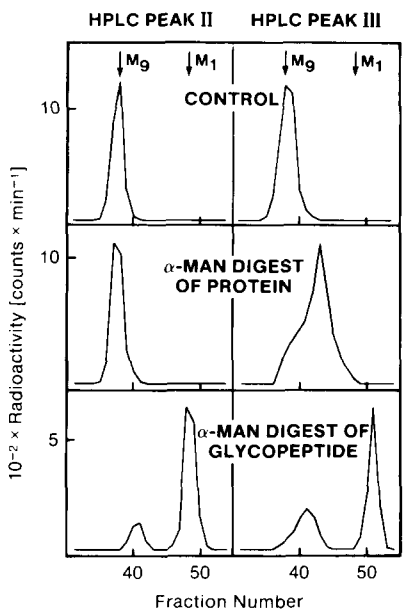

FIG. 14. **Gel filtration of glycopeptides from rough ER-derived phaseolin.** The procedures for labeling and isolating phaseolin from rough ER were as described for Fig. 12 with the exception that the cotyledons were labeled for 2 h. An aliquot of phaseolin (25,000 cpm) was treated for 24 h with α-mannosidase. Both the treated sample and an untreated sample were then digested with trypsin, and the tryptic peptides were separated on a C_8 HPLC column as described for Fig. 6. The tryptic glycopeptides so obtained were then digested with Pronase, and the reaction products were analyzed by gel filtration on the long Bio-Gel P-4 column (*upper* and *middle panels*). In addition, samples of phaseolin not treated with α-mannosidase were subjected to trypsin digestion, HPLC fractionation, and finally α-mannosidase treatment before Pronase digestion and gel filtration (*lower panel*). *Arrows* M_9 and M_1 indicate the elution positions of $Man_9(GlcNAc)_2Asn$ and $Man_1(GlcNAc)_2Asn$, respectively.

α-mannosidase. It is readily degraded by α-mannosidase in the corresponding tryptic glycopeptide; however, the main digestion product co-migrates on the Bio-Gel P-4 column with $Man_1(GlcNAc)_2Asn$ (fraction 48), although a significant proportion is still somehow protected by the attached peptide and only shortened by 2–4 mannose residues.

The N-terminal region of the glycopeptide in HPLC fraction 128 (Fig. 6, HPLC peak III) is H_2N-Gln-Asp-Asn-Thr-Ile (QDNTI) and also carries $Man_9(GlcNAc)_2$. Comparison with the two predicted glycopeptides (Fig. 7) indicates that this oligosaccharide is attached to Asn^{252} in the hydrophilic protein domain. When ER-derived phaseolin (undenatured) is treated with jack bean α-mannosidase, this high mannose oligosaccharide is shortened by 2–4 mannose residues. Most of the corresponding tryptic glycopeptide is totally accessible to α-mannosidase, which becomes trimmed to $Man_1(GlcNAc)_2(Gly)Asn$ (fraction 51). However, as seen for the HPLC peak II tryptic glycopeptide, some of the material is partially resistant to α-mannosidase and only shortened by 2–4 mannose residues. It is not entirely clear why glycopeptides obtained by trypsin digestion are only partially degradable by α-mannosidase, as the corresponding glycopeptides from Pronase digests are completely degraded by α-mannosidase to $Man_1(GlcNAc)_2(Gly)Asn$ (data not shown). We favor the explanation that the peptide portion of the glycopeptide causes steric hindrance of the enzyme. The experiments discussed here show that the oligosaccharides of ER-derived phaseolin, which are accessible to jack bean α-mannosidase *in vitro*, are the same ones which are normally modified: $Man_9(GlcNAc)_2$ to $Man_7(GlcNAc)_2$ in the phaseolin polypeptides A and C in the protein bodies and $Man_9(GlcNAc)_2$ to Xyl-$Man_3(GlcNAc)_2$ in polypeptide D in the Golgi apparatus.

Accessibility of the Terminal N-Acetylglucosamine Residues of the Complex Phaseolin Oligosaccharide to β-N-Acetylglucosaminidase Digestion in Vitro—Phaseolin isolated from the Golgi fraction of cotyledons labeled for 1 h with [^{3}H]glucosamine was digested with Pronase, and the resulting glycopeptides were separated by ConA chromatography. When the glycopeptides that did not bind ConA were treated with β-*N*-acetylglucosaminidase, more than 90% of the label appeared as free *N*-acetylglucosamine (data not shown). Thus, more than 90% of the *N*-[^{3}H]acetylglucosamine incorporated into the complex side chains of 1-h-labeled phaseolin resides in terminal (β-*N*-acetylglucosaminidase-accessible) residues.

The sensitivity of the terminal *N*-acetylglucosamine residues of the complex oligosaccharides in native phaseolin to β-*N*-acetylglucosaminidase *in vitro* was studied by treating *N*-[^{3}H]acetylglucosamine-labeled phaseolin with the enzyme. The phaseolin was obtained from cotyledons which had been labeled for 1 h with [^{3}H]glucosamine. After this short labeling time, most of the *N*-[^{3}H]acetylglucosamine is in terminal *N*-acetylglucosamine residues rather than in the chitobiose core. After inactivation of β-*N*-acetylglucosaminidase, phaseolin was digested by Pronase, and the glycopeptides obtained were separated by ConA affinity chromatography. Analysis of the nonretarded glycopeptide fraction of the enzyme-treated and control phaseolin showed that the terminal *N*-acetylglucosamine residues can be removed by treatment of native phaseolin with β-*N*-acetylglucosaminidase. The amount of radioactivity in *N*-[^{3}H]acetylglucosamine in *in vitro* treated phaseolin was less than 10% compared to the controls. Gel filtration analysis of the same glycopeptides from *in vitro* treated phaseolin showed that the glycopeptides co-migrated with the mature glycopeptides obtained from mature phaseolin in the protein bodies (data not shown).

DISCUSSION

The results reported in this paper confirm and extend the limited amount of information available on the structure of *N*-linked glycans of plant glycoproteins in general and seed storage glycoproteins in particular. Plant glycoproteins have been shown to contain both high mannose and complex *N*-linked glycans, and it is known that the latter derive from the former (1–7, 14, 23, 47, 48). We isolated three different abundant glycans from affinity-purified phaseolin and determined their structures. Two are typical high mannose oligosaccharides, and one is complex. The two high mannose oligosaccharides have 9 and 7 mannose residues and share the unique branching pattern of mannosyl residues reported for high mannose glycans of yeast and animal cells (1, 2). In addition to these, phaseolin contains small amounts of Man_6 and Man_8 glycans (48). By collecting the peak fractions of the glycopeptides separated on Bio-Gel P-4, we obtained samples for ^{1}H NMR analysis which were uncontaminated by these minor glycan species. The complex glycan of phaseolin with 3 mannose residues and 1 xylose residue linked β1→2 to the β-linked mannose seems to be a common component of complex *N*-linked plant oligosaccharides (4–6). A β1→2-linked xylose residue has been found in only one animal glycan (44). In addition to this core, most complex glycans of plant glycoproteins have an α1→3-linked fucose residue on the proximal *N*-acetylglucosamine, and some have *N*-acetylglucosamine, galactose, and fucose residues emanating from the core.

By sequencing tryptic phaseolin glycopeptides which were separated by reverse-phase HPLC and identified by their specific oligosaccharide, we have shown that $Man_9(GlcNAc)_2$ is always attached to Asn^{341}, whereas $Man_7(GlcNAc)_2$ and Xyl-$Man_3(GlcNAc)_2$ are always attached to Asn^{252}. Thus, the $Man_9(GlcNAc)_2$ glycan is located in a hydrophobic protein domain, and the glycans which undergo processing are both in a hydrophilic part of the protein (24).

Polypeptides A and C have only high mannose glycans with $Man_7(GlcNAc)_2$ attached to Asn^{252} and $Man_9(GlcNAc)_2$ attached to Asn^{341}. Polypeptide D has only one glycan: mostly Xyl-Man_3(Glc-NAc)$_2$ and a smaller amount of $Man_7(GlcNAc)_2$. It is likely that polypeptide D is a mixture of two slightly different polypeptides: one with $Man_7(GlcNAc)_2$ and one with Xyl-$Man_3(GlcNAc)_2$. Xyl-$Man_3(GlcNAc)_2$ is always attached to Asn^{252}, and $Man_7(GlcNAc)_2$ also occupies this site on the related polypeptide, as it does on polypeptides A and C. The presence of a glycine residue at the N-terminal side of asparagine clearly identifies this glycosylation site when glycopeptides are analyzed. Whether polypeptide D lacks a second glycosylation site or has an unoccupied glycosylation site is not clear. All nine cDNAs which have been sequenced give rise to polypeptides with two glycosylation sites. This means that it is more likely that the second site is unoccupied rather than nonexistent. We interpret the presence of the two types of glycans in polypeptide D as evidence for incomplete processing when the protein passes through the Golgi apparatus. We speculate that a portion of the D polypeptides reach the protein bodies with $Man_9(GlcNAc)_2$ glycans and that these are slowly processed to $Man_7(GlcNAc)_2$ (see below). Whereas we have no data on the glycan of polypeptide B, we postulate that it is the same as polypeptide D. The method used to purify the polypeptides (SDS-PAGE followed by electroelution) did not allow us to isolate polypeptide B away from polypeptides A and C because the three polypeptides are not clearly separated on the overloaded gels.

Biochemical Characterization and Localization of the Phaseolin Oligosaccharide Intermediates—By labeling cotyledons for different intervals with [^{3}H]glucosamine, we were able to

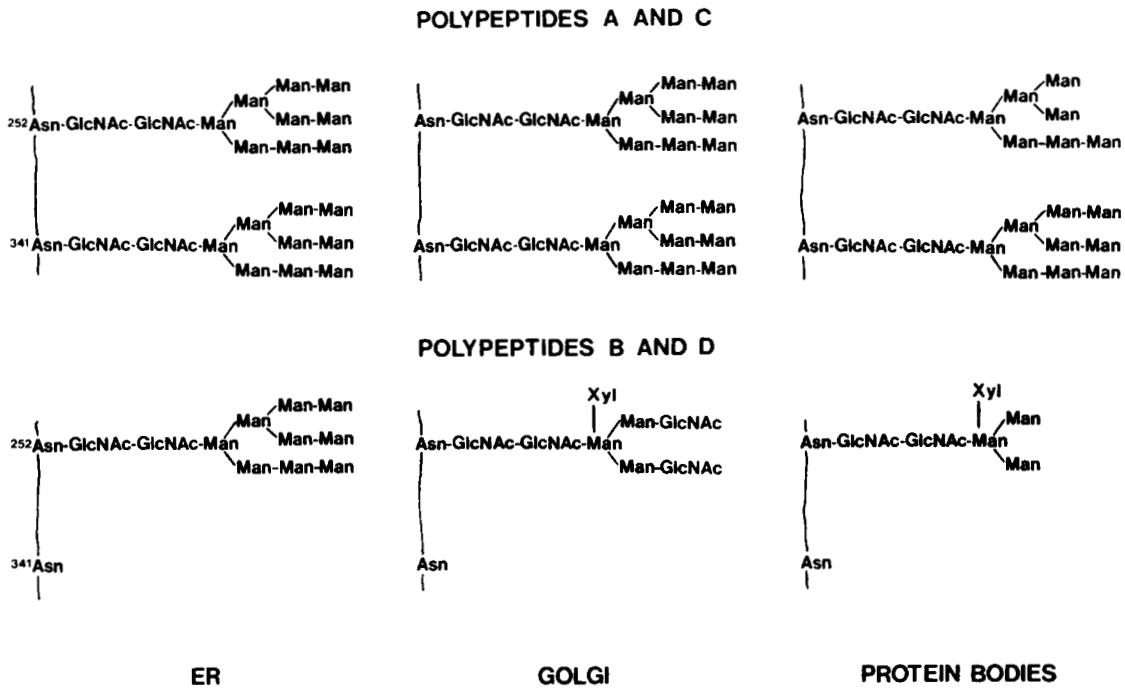

FIG. 15. **Proposed pathway for the processing of phaseolin oligosaccharides.**

identify early and late glycan processing events, and subcellular fractionation experiments permitted us to place the early processing events in the Golgi apparatus and late events in the protein bodies. The various processing steps are summarized in Fig. 15. In the ER, we find only the $Man_9(GlcNAc)_2$ form of glycans, indicating that the removal of terminal glucose residues occurs very quickly. In the Golgi apparatus, we find unprocessed high mannose glycans and fully processed complex glycans with terminal *N*-acetylglucosamine residues. This indicates that the intermediate processing steps occur rapidly in the Golgi apparatus. Transport of the glycoproteins to the protein bodies occurs with terminal *N*-acetylglucosamine residues attached to the complex glycans. Our findings confirm that in plants the transformation of high mannose oligosaccharides into complex glycans occurs in the Golgi complex (14). Plant glycan processing enzymes such as mannosidases I (49) and II[3] have recently been characterized but not yet localized, whereas some of the glycosyltransferases such as fucosyltransferase (50) and *N*-acetylglucosamine transferase (51) have been identified and localized in the Golgi apparatus. Our finding that processing (removing of *N*-acetylglucosamine and mannose) occurs in the protein bodies is consistent with the observation that protein bodies contain α-mannosidase and β-*N*-acetylglucosaminidase (52).

Accessibility of Glycans to Modifying Enzymes—Experiments with glycoproteins from plants (53), yeast (54), and virus envelopes (55) show that the control of glycan modification is, in large part, determined by the accessibility of glycans to the processing enzymes in the Golgi complex. This hypothesis derives from studies in which native glycoproteins are incubated in the absence or presence of glycosidases (endo-β-*N*-acetylglucosaminidase H, α-mannosidase) and then run on SDS gels. Any change in M_r caused by glycosidase treatment can yield an estimate of the extent of glycan accessibility to modifying enzymes. The results obtained here are in agreement with the hypothesis stated above. We observed that the $Man_9(GlcNAc)_2$ on Asn^{252} of ER-derived phaseolin is partially susceptible to α-mannosidase (2–4 mannose residues are removed), whereas the $Man_9(GlcNAc)_2$ on Asn^{341} is not altered by exposure to the enzyme. Thus, only the glycan which can be modified in the Golgi apparatus is accessible to the glycosidase *in vitro*. As discussed above, the glycan on Asn^{252} only becomes complex when the second glycosylation site is unoccupied. In addition, we found that the terminal *N*-acetylglucosamine residue(s) on the complex glycan of Golgi apparatus-derived phaseolin can be removed by incubation of the native protein with β-*N*-acetylglucosaminidase. It is of particular interest in this regard that the glycan which is accessible to glycosidases (and presumably to glycosyltransferases as well) is in a hydrophilic region of the polypeptide (Asn^{252}), whereas the glycan which is not accessible (Asn^{341}) is in a hydrophobic region of the polypeptide (24). It may well be that high mannose chains in hydrophobic pockets are sequestered and thus remain unmodified. However, the converse is not necessarily true. When analyzing the glycans of phytohemagglutinin, we found an unmodified chain, which is not accessible to α-mannosidase *in vitro*, attached to an asparagine residue in a hydrophilic region (56).

Of particular interest with respect to the processing of glycans is our finding that the presence of a glycan on Asn^{341} may determine whether the glycan at Asn^{252} is processed or not. We interpret our results as showing that the absence of a glycan from Asn^{341} results in the more extensive processing of the glycan on Asn^{252}. If glycans are not accessible, they remain unmodified. The absence of a glycan on Asn^{341} may result in greater accessibility to the Golgi enzymes of the glycan attached to Asn^{252}.

Role of Terminal N-Acetylglucosamine Residues—Our results show that the complex side chain of phaseolin polypeptide D acquires terminal *N*-acetylglucosamine residues in the Golgi apparatus, which are later removed in the protein bodies. Earlier (57), we made a similar observation for the complex side chain of phytohemagglutinin, which, like phaseolin, moves via the ER-Golgi apparatus-protein body transport pathway in beans. Since protein bodies are similar to lysosomes (both contain a full complement of acid hydrolases (see Ref. 7), it is tempting to speculate that these terminal *N*-

[3] A. Elbein, personal communication.

acetylglucosamine residues may play a role in targeting in the same manner that terminal phosphate residues target lysosomal hydrolases to lysosomes (58). Such a role for the terminal *N*-acetylglucosamine residues seems unlikely in view of the finding that tunicamycin does not inhibit the transport of unglycosylated phytohemagglutinin to the protein bodies (59). Furthermore, many protein body proteins, such as soybean glycinin and pea legumin, are not glycoproteins.

Our laboratory has recently investigated (60) the structural requirements of oligosaccharides that can accept xylose and/or fucose residues during the formation of complex glycans in the plant Golgi apparatus. Only glycans containing at least 1 terminal *N*-acetylglucosamine residue can serve as acceptors of xylose and fucose. As both phytohemagglutinin and phaseolin contain glycans with transient *N*-acetylglucosamine residues and these same glycans contain xylose and/or fucose, then the terminal *N*-acetylglucosamine residues may simply constitute recognition markers for later Golgi apparatus-mediated processing events.

Acknowledgments—We thank D. P. Delmer, who initiated the work on the glycans of phaseolin in her own laboratory, for sending us the results of her preliminary investigations. We gratefully acknowledge C. Versluis (Department of Analytical Chemistry, University of Utrecht, The Netherlands) for recording the mass spectra and Dr. Kenneth D. Johnson for useful discussions and constructive criticism of this manuscript.

REFERENCES

1. Vliegenthart, J. F. G., Dorland, L., and van Halbeek, H. (1983) *Adv. Carbohydr. Chem.* **41,** 209–374
2. Dorland, L., van Halbeek, H., Vliegenthart, J. F. G., Lis, H., and Sharon, N. (1981) *J. Biol. Chem.* **256,** 7708–7711
3. Misaka, A., and Goldstein, I. J. (1977) *J. Biol. Chem.* **252,** 6995–6999
4. Takahashi, N., Hotta, T., Ishihara, H., Mori, M., Tejima, S., Bligny, R., Akazawa, T., Endo, S., and Arata, Y. (1986) *Biochemistry* **25,** 388–395
5. Ishihara, H., Takahashi, N., Oguri, S., and Tejima, S. (1979) *J. Biol. Chem.* **254,** 10715–10719
6. Hase, S., Koyama, S., Daiyasu, H., Takemoto, H., Hara, S., Kobayashi, Y., Kyoyoku, Y., and Ikenaka, T. (1986) *J. Biochem. (Tokyo)* **100,** 1–10
7. Chrispeels, M. J. (1984) *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **304,** 309–322
8. Bollini, R., Van der Wilden, W., and Chrispeels, M. J. (1982) *Physiol. Plant.* **55,** 82–92
9. Bollini, R., and Vitale, A. (1981) *Physiol. Plant.* **52,** 96–100
10. Brown, T. W. S., Ma, F., Bliss, F. A., and Hall, T. C. (1981) *Theor. Appl. Genet.* **59,** 83–88
11. Bollini, R., Vitale, A., and Chrispeels, M. J. (1983) *J. Cell Biol.* **96,** 999–1007
12. Hall, T. C., Ma, Y., Buchbinder, B. U., Pyne, J. W., Sun, S. M., and Bliss, F. A. (1978) *Proc. Natl. Acad. Sci. U. S. A.* **75,** 3196–3200
13. Bollini, R., and Chrispeels, M. J. (1979) *Plant (Berl.)* **146,** 487–501
14. Chrispeels, M. J. (1983) *Planta (Berl.)* **158,** 140–151
15. Bollini, R., and Chrispeels, M. J. (1978) *Planta (Berl.)* **142,** 291–298
16. Lehle, L. (1981) *FEBS Lett.* **123,** 63–66
17. Hori, H., James, D. W., and Elbein, A. D. (1982) *Arch. Biochem. Biophys.* **215,** 12–21
18. Hubbard, S. C., and Ivatt, R. J. (1981) *Annu. Rev. Biochem.* **50,** 555–583
19. Kornfeld, R., and Kornfeld, S. (1985) *Annu. Rev. Biochem.* **54,** 631–664
20. Hori, H., and Elbein, A. D. (1983) *Arch. Biochem. Biophys.* **220,** 415–425
21. Deleted in proof
22. Chrispeels, M. J. (1985) *Oxf. Surv. Plant Mol. Cell Biol.* **2,** 43–68
23. Hall, T. C., Slightom, J. L., Ersland, D. R., Murray, M. G., Hoffman, L. M., Adang, M. J., Brown, J. W. S., Ma, Y., Matthews, J. A., Cramer, J. H., Barker, R. F., Sutton, D. W., and Kemp, J. D. (1983) in *Structure and Function of Plant Genomes* (Ciferri, O., and Dure, L. S., eds) pp. 123–142, Plenum Publishing Corp., New York
24. Slightom, J. L., Drong, R. F., Klassy, R. C., and Hoffman, L. M. (1985) *Nucleic Acids Res.* **13,** 6483–6498
25. Sun, S. M., Muttschler, M. A., Bliss, F. A., and Hall, T. C. (1978) *Plant Physiol.* **61,** 918–923
26. Spencer, D., Higgins, T. J. V., Button, S. C., and Davey, R. A. (1980) *Plant Physiol.* **66,** 510–515
27. Stockman, D. R., Hall, T. C., and Ryan, D. S. (1976) *Plant Physiol.* **58,** 272–275
28. Chrispeels, M. J., Higgins, T. J. V., Craig, S., and Spencer, D. (1982) *J. Cell Biol.* **93,** 5–14
29. Deleted in proof
30. Smith, J. A., Hurrell, J. G. R., and Leach, S. J. (1978) *Anal. Biochem.* **87,** 299–305
31. Laemmli, U. K., and Favre, M. (1973) *J. Mol. Biol.* **80,** 575–599
32. Bonner, W. M., and Laskey, R. A. (1974) *Eur. J. Biochem.* **46,** 83–88
33. Hodge, J. E., and Hofreiter, B. T. (1962) *Methods Carbohydr. Chem.* **1,** 288–294
34. Hsieh, P., Rosner, M. R., and Robbins, P. W. (1983) *J. Biol. Chem.* **258,** 2548–2554
35. Eastoe, J. E. (1966) in *Glycoproteins* (Gottschalk, A., ed) pp. 112–150, Elsevier/North-Holland Biomedical Press, New York
36. Gray, W. R. (1972) *Methods Enzymzol.* **25,** 121–138
37. Ernst, R. R. (1966) *Adv. Magn. Res.* **2,** 1–135
38. Deleted in proof
39. Hewick, R. M., Hunkapiller, M. W., Hood, L. E., and Dreyer, W. J. (1981) *J. Biol. Chem.* **256,** 7990–7997
40. Hunkapiller, M. W., and Hood, L. E. (1983) *Methods Enzymol.* **51,** 486–493
41. Trimble, R. B., and Maley, F. (1984) *Anal. Biochem.* **141,** 515–522
42. Zacharius, R. M., Zell, T. E., Morrison, J. M., and Woodlock, J. J. (1969) *Anal. Biochem.* **30,** 148–152
43. Roberts, G. C. K., and Jardetzky, O. (1970) *Adv. Protein Chem.* **24,** 447–545
44. van Kuik, J. A., van Halbeek, H., Kamerling, J. P., and Vliegenthart, J. F. G. (1985) *J. Biol. Chem.* **260,** 13984–13988
45. Paz-Parente, J., Strecker, G., Leroy, Y., Montreuil, J., Fournet, B., van Halbeek, H., Dorland, L., and Vliegenthart, J. F. G. (1983) *FEBS Lett.* **152,** 145–152
46. Slightom, J. L., Sun, S. M., and Hall, T. (1983) *Proc. Natl. Acad. Sci. U. S. A.* **80,** 1897–1901
47. Vitale, A., Warner, T. G., and Chrispeels, M. J. (1984) *Planta (Berl.)* **160,** 256–263
48. Neeser, J.-R., del Vedovo, S., Mutsaers, J. H. G. M., and Vliegenthart, J. F. G. (1985) *Glycoconjugate J.* **2,** 355–364
49. Szumilo, T., Kaushal, G. P., Hori, H., and Elbein, A. D. (1986) *Plant Physiol.* **81,** 383–389
50. Sturm, A., and Kindl, H. (1983) *FEBS Lett.* **160,** 165–168
51. Chrispeels, M. J. (1985) *Plant Physiol.* **78,** 835–838
52. Harris, N., and Chrispeels, M. J. (1975) *Plant Physiol.* **56,** 292–299
53. Faye, L., Johnson, K. D., and Chrispeels, M. J. (1986) *Plant Physiol.* **80,** 206–211
54. Trimble, R. B., Maley, F., and Chu, F. K. (1983) *J. Biol. Chem.* **258,** 2562–2567
55. Hsieh, P., Rosner, M. R., and Robbins, P. W. (1983) *J. Biol. Chem.* **258,** 2555–2561
56. Sturm, A., and Chrispeels, M. J. (1986) *Plant Physiol.* **81,** 320–322
57. Vitale, A., and Chrispeels, M. J. (1984) *J. Cell Biol.* **99,** 133–140
58. Sly, W. S., and Fischer, H. D. (1982) *J. Cell. Biochem.* **18,** 67–85
59. Bollini, R., Ceriotti, A., Daminati, M. G., and Vitale, A. (1985) *Physiol. Plant.* **65,** 15–22
60. Johnson, K. D., and Chrispeels, M. J. (1986) *Plant Physiol.* **80,** Suppl. 84 (abstr.)

Continued on next page.

Supplementary material to:

Structure and position of the high mannose and the complex oligosaccharide sidechains of the bean storage protein phaseolin

Arnd Sturm, J. Albert Van Kuik, J.F.G. Vliegenthart and Maarten J. Chrispeels

MATERIALS AND METHODS

Plant Material: Plants of *Phaseolus vulgaris* L. cv. Greensleeves were grown in a greenhouse as described (13). Beans were harvested for the isolation of mature phaseolin when pods turned yellow. All other experiments were carried out with developing cotyledons weighing 125-175 mg, when accumulation of phaseolin is quite rapid (25).

Isolation of Mature Phaseolin: Phaseolin was isolated and purified with slight modifications as described (15). Cotyledons (40g) of mature beans were homogenized in a cold mortar in extraction buffer (100 mM Tris/HCl, pH 7.8, containing 1 mM EDTA and 12% sucrose) and the homogenate was centrifuged for 90 min at 95,000xg. The supernatant was passed in aliquots (20 ml) through a Sephadex G50 (Pharmacia Fine Chemicals Inc.) column (2.2 x 16 cm), equilibrated with phosphate buffer (25 mM Na-phosphate, pH 7.0 containing 25 mM NaCl). The cloudy fractions containing the proteins in the excluded volume were combined and constitute the protein extract. Aliquots of 40 ml of the protein extract were loaded on a DEAE-cellulose (Whatman Chemical Separation Ltd.) column (2.2 x 7 cm), equilibrated with phosphate buffer. The column was rinsed with the same buffer for several hours and then eluted with 200 ml of a NaCl gradient (25-400 mM) in 25 mM Na-phosphate buffer, pH 7.0. Fractions of 4 ml were collected and analyzed for protein by reading the absorption at 280 nm. The main protein peak was combined (fractions 13-27) and dialyzed against water overnight, followed by lyophilization. The yield of phaseolin was 2g from 40g of mature beans.

Isolation of Radiolabeled Phaseolin: Radioactive labeling of excised cotyledons was done as described by Spencer et al. (26) using D-[6-^{3}H]glucosamine HCl (36 Ci/mmol, Amersham Co.) or L-[^{35}S]methionine (1086 Ci/mmol, New England Nuclear Corp.). Each of 20 cotyledons were labeled with 6µCi of precursor for 24 h, and the radioactive tissue was collected by cutting a thin slice from the cotyledon with a razor blade. The remainder of the cotyledon was discarded.

The radioactive tissue was homogenized in 6 ml of buffer (30 mM Na-phosphate, pH 7.4 with 450 mM NaCl containing 1% (w/v) Tween 20) and the homogenate centrifuged for 10 min at 1,000xg to remove debris. The affinity procedure of Stockman et al. (27) was used to isolate phaseolin from the clear supernatant. Sepharose 4B covalently coupled to phaseolin (9 mg/ml of gel) was added to the homogenate and the slurry was adjusted to pH 4.6 with 1M acetic acid. After allowing the mixture to stand at pH 4.6 for 20 min, the phaseolin-Sepharose was separated from the homogenate by filtration in a small plastic column and the gel rinsed with 0.1M Na-acetate, pH 4.6, containing 0.4M NaCl. The radiolabeled phaseolin was eluted with 50 mM Tris/HCl, pH 7.8 and 1 mM EDTA. The homogenate was adjusted back to pH 7.4 and the isolation procedure was repeated twice. The combined phaseolin samples were dialyzed against distilled water and lyophilized.

Isolation of Radiolabeled Phaseolin from ER, Golgi, and Protein Bodies: The tissue was homogenized in 100 mM Tris-Cl, pH 7.8, containing 2 mM $MgCl_2$ and 12% (w/w) sucrose using a mortar and pestle. This homogenization procedure disrupts the fragile protein bodies and releases the storage proteins into the soluble fraction (28). The cell walls and debris were removed by centrifugation at 1,000xg for 5 min. Organelles were separated from the soluble proteins by passage through a Sepharose 4B column (1.5 x 20 cm, Pharmacia Fine Chemicals Inc.). The column was loaded with 3 ml supernatant and eluted with the homogenization medium. Fractions of 1.5 ml were collected and the two fractions containing the most light-scattering material (organelles) were used for the isolation of rough ER and Golgi. Late fractions with a high absorbance at 280 nm (soluble proteins) were used for the isolation of phaseolin originating from the protein bodies. The organelle fraction was layered onto a linear sucrose gradient (16-54%) containing 100 mM Tris/HCl, pH 7.8, and 2 mM $MgCl_2$. The gradients were centrifuged at 150,000xg for 2 h at 4°C. Golgi and rough ER were identified as described by Chrispeels (1^M) at sucrose densities of 1.13g.cm^{-3} and 1.18g.cm^{-3}, respectively. The membrane fractions were diluted with an equal volume of PBS containing 1% Tween 20, and subjected to anti-phaseolin Sepharose affinity chromatography as described below. Phaseolin was isolated from the soluble protein fraction by the affinity method of Stockman et al. (27).

Immunological Techniques: 100 µg of purified phaseolin in 400 µl PBS (10 mM Na-phosphate, pH 7.4, containing 150 mM NaCl) were emulsified with 0.4 ml Freund's complete adjuvant (Behring Diagnostics) and injected subcutaneously in 5-6 portions into a New Zealand white rabbit obtained from a local breeder. Three more boosts were given at intervals of 2 weeks. After 2 additional weeks, blood was taken from the ear vein. For the preparation of serum, blood was allowed to clot for 2 h at room temperature and then overnight at 4°C. The clot was removed by centrifugation in a Sorvall SS-34 rotor for 10 min. The supernatant was collected and recentrifuged. IgG specific for phaseolin was prepared by affinity chromatography using phaseolin-Sepharose as an affinity gel as described by Smith et al. (30). Coupling of Sepharose 4B (Pharmacia Fine Chemicals Inc.) to phaseolin and phaseolin-specific IgG was performed according to the manufacturer's specifications. Isolation of radioactive phaseolin polypeptides with anti-phaseolin Sepharose was carried out in PBS containing 1% Tween 20, 0.2M Glycine/HCl, pH 2.2, containing 0.5M NaCl was used to elute the bound phaseolin. The polypeptides obtained were dialyzed against water and lyophilized.

Separation of Phaseolin by Polyacrylamide Gel Electrophoresis: Separation of polypeptides denatured with sodium dodecyl sulfate (SDS) was carried out according to the method of Laemmli (31). Fluorography of polyacrylamide gels was done as described by Bonner and Laskey (32). The lyophilized phaseolin was dissolved in denaturing buffer (67.5 mM Tris/HCl, pH 6.8, 4% SDS, 5% β-mercaptoethanol, 10% glycerol and 0.002% bromphenol blue; 10 mg phaseolin/ml) and denatured by boiling for 3 min. Purified phaseolin (3 mg) was fractionated by SDS-PAGE on a slab gel (0.2 x 11 x 14 cm). After staining the gel with Coomassie brilliant blue R, the phaseolin bands were cut out and electroeluted overnight at 350V (electroelution buffer: 2.5 mM Tris, 19.2 mM glycine and 0.1% SDS). The electroeluted polypeptides were precipitated with acetone, recovered by centrifugation, and dried under vacuum.

Digestion of Phaseolin with Pronase and Large-Scale Isolation of Phaseolin Glycopeptides: Lyophilized phaseolin (750 mg) was dissolved in 60 ml of 100 mM Tris/HCl, pH 8.0, and 60 mg Pronase (Behring Diagnostics) and some drops of toluene were added. The solution was incubated at 37°C. After 24 h and 48 h, the same amount of Pronase was added and the incubation was continued. After 62 h, the reaction was stopped by adding 2 ml of glacial acetic acid and the precipitated material was removed by centrifugation. The clear supernatant was lyophilized and the residue was dissolved in 25 ml of water. To separate the glycopeptides from salt, amino acids, and small peptides, the solution was loaded in 2 aliquots on a Sephadex G15 column (2.2 x 22 cm) which was equilibrated with water. Fractions (1 ml) were collected and analyzed for neutral sugars by the phenol/sulfuric acid method (33). The glycopeptide peak was lyophilized and subjected to further fractionation by ConA-chromatography.

Trypsin Digestion of Radiolabeled Phaseolin Polypeptides: Trypsin digestion was done as described (34). [^{3}H]GlcN labeled-phaseolin (1.5 mg) was dissolved in 1 ml of 50 mM ammonium carbonate, pH 8.0 and heated at 100°C for 3 min. 150µg of trypsin (TPCK-treated) were added and the solution was incubated at 37°C. After 2 h the same amount of trypsin was added and the incubation continued for another 2 h. The reaction was stopped by heating at 100°C for 3 min. The solution of tryptic peptides was lyophilized and the residue prior to HPLC analysis dissolved in 0.5 ml of 0.1M sodium phosphate, pH 2.2, centrifuged to remove particulate matter and filtered through a 0.2 µm Nylon Filter (Western Analytical Prod.).

ConA-Chromatography: The glycopeptide mixture was dissolved in 5 ml of ConA buffer (50 mM Tris/HCl, pH 7.0, containing 200 mM NaCl, 1 mM each of $CaCl_2$, $MgCl_2$ and $MnCl_2$, and 0.05% Na-azide) and loaded at room temperature on a column (2 x 16 cm) of ConA-agarose (Sigma Chemical Company), equilibrated and eluted with ConA buffer. Fractions (4 ml) were collected and glycopeptides were detected by the neutral sugar assay. Glycopeptides which did not bind ConA (ConA-) were pooled and lyophilized, and bound glycopeptides (ConA+) were eluted with 200 mM α-methylmannoside (aMM, Sigma Chemical Company) in ConA buffer. The first 10 aMM-containing fractions were combined and lyophilized. The ConA- and ConA+ samples were dissolved in a small amount of water and desalted over a Sephadex G15 column (2.2 x 22 cm). The glycopeptide-containing fractions were lyophilized and further fractionated by gel-filtration chromatography.

Gel-Filtration Chromatography: Glycopeptide fractions (4 mg) were dissolved in 0.5 ml 0.1M acetic acid and then run on a column (1.0 x 100 cm) of Biogel P-4 (minus 400 mesh, Bio-Rad Laboratories) equilibrated with the same solvent. Fractions of 1 ml were collected and analyzed for neutral sugars (33) or radioactivity. The Biogel P-4 column was calibrated with various standards, including $Man_9GlcNAc$, ManGlcNAc, $Man_6(GlcNAc)_2Asn$, $Man_3(GlcNAc)_2Asn$, and GlcNAcAsn (2-acetamido-1-(L-β-aspartamido)-1,2-deoxy-β-D-glucose).

The void volume of the long Biogel P-4 column (1.0 x 100 cm) is 35 ml. $Man_9(GlcNAc)_2Asn$ elutes in fraction 37 and $Man_3(GlcNAc)_2Asn$ in fraction 48. A size difference of 1 mannose residue corresponds to a shift of 1.4 fractions. Different sugars have a different influence on the elution position of glycopeptides on this Biogel column: 1 GlcNAc corresponds to a shift of 3 fractions, whereas Xyl or Fuc cause a shift of only 0.5 fraction. In the low molecular range of the column, small size differences cause larger fraction shifts, i.e. free Man elutes in fraction 63, whereas free GlcNAc elutes in fraction 58.

Treatment of Phaseolin Glycopeptides with Glycosidases: Endo H (*Streptomyces plicatus*, Miles Laboratories Inc.) treatment of glycopeptides was performed by incubation in 500 µl of 100 mM Na-acetate, pH 5.8, containing 10 mU of enzyme. Treatment with α-mannosidase (from jack-bean, Sigma Chemical Co.) was carried out in 500 µl of 50 mM Na-acetate, pH 5.8, 5 mM $ZnSO_4$ with 5 U of enzyme. Incubation with β-N-acetylglucosaminidase (from jack-bean, Sigma Chemical Co.) was performed in 500 µl of 50 mM Na-citrate, pH 4.6, with 5 U of enzyme. All glycosidase incubations were carried out at 37°C for 24 h under a toluene atmosphere.

Endo H-Digest of Bulk Phaseolin Glycopeptides: 2.8 mg of each glycopeptide were dissolved in 0.5 ml 0.1M sodium acetate buffer, pH 5.8. 15mU of endo H (Miles Laboratories Inc.) and a drop of toluene were added and the solution incubated at 37°C. After 24 h and 48 h, 10mU of endo H were added and the incubation continued. The reaction was stopped after 62 h by addition of 10µl of acetic acid, and the mixture fractionated on a Biogel P-4 column (1.0 x 100 cm), which was equilibrated with 0.1M acetic acid. 1 ml fractions were collected and 20µl analyzed for neutral sugars (33) and for glycopeptides by the ninhydrin method (35).

Amino-Terminal End-Group Analysis: Determination of the N-terminal amino acid using dansyl-chloride was performed according to the method of Gray (36).

FAB Mass Spectrometry: Positive-ion mass spectra were recorded on a VG Analytical ZAB-2f reversed-geometry mass spectrometer (Department of Analytical Chemistry, Utrecht University, The Netherlands). The sample, dispersed in glycerol, was bombarded with xenon atoms having a kinetic energy of approximately 7 keV. The sputtered ions were extracted and accelerated with a potential of 8 kV.

500-MHz ^{1}H NMR Spectroscopy: Glycopeptides were repeatedly exchanged in ^2H_2O (99.96 atom % ^{2}H, Aldrich) with intermediate lyophilization. ^{1}H NMR spectra were recorded on a Bruker WM-500 spectrometer (SON hf NMR facility, Department of Biophysical Chemistry, University of Nijmegen, The Netherlands) operating at 500 MHz in the Fourier transform mode at a probe temperature of 27°C. Resolution enhancement of the spectra was achieved by Lorentzian-to-Gaussian transformation (37). Chemical shifts are given relative to sodium 4,4-dimethyl-4-silapentane-1-sulfonate but were actually measured indirectly to acetone in ^2H_2O (δ = 2.225 ppm) (2).

High-Performance Liquid Chromatography: Separation of the tryptic peptides was carried out by reverse phase HPLC on a Beckman 342 Gradient Liquid Chromatograph equipped with a Vydac C-18 column (5 microns, 4.6 x 250 mm). For further purification of the glycopeptides, an Alltech C-8 column (5 microns, 4.6 x 250 mm) was used. In all cases, the peptide separation was carried out at room temperature at a flowrate of 1 ml/min. The gradients are described in detail in the legends of the figures. Fractions of 1 ml were collected. Radioactivity in the fractions was measured by liquid scintillation counting.

Automatic Protein Sequencing: Samples of glycopeptides from sodium phosphate HPLC gradients were adjusted to neutral pH with 5N NaOH, lyophilized and desalted on a small G-15 column prior to amino acid sequencing. Samples from trifluoroacetic acid containing HPLC gradients were lyophilized without further treatments. After lyophilization, the glycopeptides were dissolved in 1% trifluoroacetic acid and subjected to automatic Edman degradations using an Applied Biosystems Model 470A Gas-Phase-Sequencer employing the standard "No Vac" program supplied by the manufacturer (39). Phenylthiohydantoin derivatives were identified by HPLC on a Brownlee C-18 column using procedures described previously (40). The HPLC system used with this column consisted of a Perkin Elmer Series 4 liquid chromatograph, an LC-85B spectrophotometric detector equipped with a 1.4µl flow cell, an ISS-100 automatic sample injector, an LCI 100 computing integrator, and a Model 7500 computer employing Chrom III software.

Endo H Treatment of Phaseolin Polypeptides: The endo H treatment of denatured phaseolin polypeptides was carried out with slight modifications as described by Trimble and Maley (41). 50µg of protein in 50µl of phosphate-buffered saline (10 mM Na-phosphate, pH 7.4, containing 150 mM NaCl) were added to 60µg of SDS in 50µl of water (1.2 fold weight excess of SDS) and heated for 2 min in a boiling water bath. The clear solution was brought to a final concentration of 0.1% Triton X-100 and 0.25% β-mercaptoethanol and heated in a boiling water bath for another 2 min. 20µl of endo H buffer (1M Na-acetate, pH 5.8) were added. The digestion was carried out for 48 h at 37°C under a toluene atmosphere with 10mU of enzyme. Another 10 mU of endo H was added after 24 h. The protein was precipitated with acetone, recovered by centrifugation, and dried under vacuum.

Treatment of Phaseolin with α-Mannosidase and Analysis of the Oligosaccharides by HPLC and Gel Filtration Chromatography: [^{3}H]glucosamine labeled phaseolin isolated with anti-phaseolin Sepharose was dialyzed against water. About 25000 cpm in 0.5 ml of water were treated with jack bean α-mannosidase as described above. The samples were then boiled for 2 min to inactivate the α-mannosidase, and the protein was precipitated overnight with 5 vol. of acetone. The precipitated protein was now digested with trypsin (treated with tosylamido-2-phenylethyl chloromethyl ketone) and the peptides obtained were separated by reverse phase HPLC as described above. The glycopeptides were further digested with Pronase and the reaction products analyzed by gel filtration chromatography on a Biogel P-4 column (1.0 x 100 cm).

Treatment of Golgi Phaseolin with β-N-Acetylglucosaminidase: Affinity-purified phaseolin isolated from the Golgi after a 1-h labeling period with [^{3}H]glucosamine was dialyzed against water. About 50000 cpm in 0.9 ml of water were treated with β-N-acetylglucosaminidase as described above. The samples were then boiled for 2 min to inactivate β-N-acetylglucosaminidase, and the protein was precipitated overnight with acetone. The precipitated protein was now digested with Pronase, and the glycopeptides obtained were separated by ConA-affinity chromatography. The unretarded glycopeptide fractions were desalted and analyzed by gel filtration chromatography.

RESULTS

Biochemical Characterization of the Oligosaccharides of Mature Phaseolin: When purified phaseolin is separated by one-dimensional SDS-PAGE, four polypeptides (size classes of polypeptides) labeled A, B, C and D (Fig. 1). Polypeptides B and C are difficult to resolve when mature phaseolin is the starting material, but are easily resolved when phaseolin obtained from the ER is used (11). The relative abundance of the four polypeptides based on Coomassie blue staining is 1:0.1:1:0.3 (A:B:C:D) (data not shown). By comparing the Coomassie Blue stain with the Schiff Stain for glycoproteins (42), it became obvious that B and D stain poorly for carbohydrate (data not shown).

The number of glycans per polypeptide was determined by Bollini et al. (11) who showed that A and C have two oligosaccharide chains while B and D have only one. The nature of these oligosaccharide chains was partially elucidated by treating the isolated polypeptides with endo H (Fig. 2). The M_r of polypeptides A and C decreased by 4000 while the M_r of polypeptide D was unaffected. Because of its low abundance and difficulty to isolate, we did not examine polypeptide B.

To elucidate the nature of the glycans, phaseolin was isolated from mid-maturation bean cotyledons labeled for 24 h with [^{3}H]glucosamine. The protein was digested with Pronase and the glycopeptides obtained were subjected to ConA-affinity chromatography (Fig. 3A). About 18% of the labeled material eluted as a sharp peak with the ConA-buffer (Fig. 3A, peak I). The remaining 82% of the radioactivity could be eluted with 200 mM α-methylmannoside (Fig. 3A, peak II). The sizes of the glycopeptides were estimated by gel filtration on a 1 x 100 cm column of Biogel P-4 (Figs. 3B and C). The sample which was not retained by ConA (peak I) coeluted with $Man_3(GlcNAc)_2Asn$ (fraction 45) and is hereafter identified as Ph 1 (Fig. 3B). The glycopeptide Ph 1 corresponds to $Man_3Xyl(GlcNAc)_2(Gly)Asn$ (see below). The tightly bound glycopeptides (peak II) separated into two size classes corresponding to $Man_6(GlcNAc)_2Asn$ (fraction 40, called Ph 2) and $Man_9(GlcNAc)_2Asn$ (fraction 37, called Ph 3 in Fig. 3C). The position of $Man_6(GlcNAc)_2Asn$ corresponds in size to a $Man_3(GlcNAc)_2(Gly)Asn$ glycopeptide (see below).

To analyze the distribution of these three kinds of glycopeptides (Ph 1, Ph 2 and Ph 3) among the different polypeptides, the analysis described above was applied to the individual polypeptides electroeluted after SDS-PAGE of phaseolin labeled for 24 h with [^{3}H]glucosamine (Fig. 8, right panels and Figs. 9 and 11). Polypeptide B remained as a contaminant of polypeptides A and C. The results indicate that polypeptides A and C each have Ph 2 and Ph 3 in an approximate ratio of 1:1. These glycans are endo H sensitive and correspond respectively to $Man_6(GlcNAc)_2$ and $Man_9(GlcNAc)_2$ in size. The glycan of polypeptide D is a mixture of Ph 1 (70%) and Ph 2 (30%). Ph 1 is Endo H resistant and coelutes with $Man_3(GlcNAc)_2Asn$ on Biogel P-4.

Characterization of the Oligosaccharides of Mature Phaseolin by ^{1}H NMR: To elucidate the primary structure of the three phaseolin glycopeptides, 500-MHz ^{1}H-NMR spectra of the compounds in ^2H_2O were recorded. The analysis was performed on the peak fraction from the Biogel P-4 column for each glycopeptide (Ph 1, Ph 2 and Ph 3). Relevant chemical shifts for these compounds are listed in Table I. The corresponding data for the reference compound (R = XylMan_3GlcNAcGlcNAc-ol) from *Helix pomatia* α-hemocyanin (44) have been included.

The resolution enhanced, structural-reporter-group regions of the spectrum of Ph 1 are presented in Fig. 4A. The equal intensity of the anomeric proton signals in the spectrum point to the presence of a single compound. The spectral features of Ph 1 revealed a structure that is the glycopeptide analogue of compound R. This can be concluded from the spectrum of Ph 1 by the presence of R-1 (δ = 5.048 ppm) and NAc (δ = 2.010 ppm) of GlcNAc-$\underline{1}$ linked to Asn. Furthermore, GlcNAc-$\underline{2}$ R-1 is shifted upfield (Δδ = -0.026 ppm) going from R to Ph 1.

which is in agreement with the presence of GlcNAcAsn (45). The chemical shift of the other signals of Ph 1 are essentially identical with those of R. Therefore, the structure of the glycan of Ph 1 is:

```
                4'
            Manα(1→6)
                     \ 3          2          1
                      Manβ(1→4)GlcNAcβ(1→4)GlcNAcβ(1→N)Asn
                4     //
            Manα(1→3)/
                     /
            Xylβ(1→2)
```

The equal intensity of the anomeric proton signals in the spectrum of Ph 2 (Fig. 4B) as well as of Ph 3 (not shown) indicates that the fractions consist of a single compound each. From the spectral features of Ph 2, which are essentially identical with those of $Man_7(GlcNAc)_2Asn$ from an immunoglobulin M obtained from blood plasma of a patient with Waldenstrom's macroglobulinemia (1), it can be concluded that the structure of the glycan of Ph 2 is:

```
                B
            Manα(1→6)
                     \ 4'
                      Manα(1→6)
            A        /          \ 3          2          1
            Manα(1→3)            Manβ(1→4)GlcNAcβ(1→4)GlcNAcβ(1→N)Asn
            C                4   /
        D1
    Manα(1→2)Manα(1→2)Manα(1→3)
```

The ^{1}H-NMR data of Ph 3 are essentially the same as those for $Man_9(GlcNAc)_2Asn$ (1) from bovine lactotransferrin which has the following structure:

```
        D3          B
    Manα(1→2)Manα(1→6)
                     \ 4'
                      Manα(1→6)
        D2      A    /          \ 3          2          1
    Manα(1→2)Manα(1→3)            Manβ(1→4)GlcNAcβ(1→4)GlcNAcβ(1→N)Asn
        D1      C            4   /
    Manα(1→2)Manα(1→2)Manα(1→3)
```

Characterization of the Proximal GlcNAc Residue: The two high-mannose glycopeptides from polypeptides A and C were treated with endo H and the products sized on the long Biogel P-4 column (Fig. 5). Endo H treatment of the larger glycopeptide yielded two products which coeluted with $Man_8GlcNAc$ and GlcNAcAsn (compound I in Fig. 5, middle panel), as expected. The smaller glycopeptide yielded two products, one of which coeluted with $Man_7GlcNAc$. However, the small digestion product (compound II) eluted with a longer retention time than noted for GlcNAcAsn. Prior digestion of both glycopeptides with proteinase K did not alter the elution profiles of the endo H-generated products in either case (data not shown).

To investigate the peptide part of compound II, end-group analysis using dansyl chloride was carried out. This revealed the presence of Gly at the amino terminus of compound II. FAB mass spectrometry of compounds I and II confirmed this in showing molecular ions at m/z 336 $(M+H)^+$ and m/z 393 $(M+H)^+$, respectively. This supports the conclusion that compound I is GlcNAcAsn and compound II is Gly(GlcNAc)Asn.

To check the identification made by end-group analysis and FAB-MS, compounds I and II (Fig. 5, middle and lower panel) were subjected to 500-MHz ^{1}H-NMR spectroscopy. The spectral parameters of compound I were essentially identical with those of GlcNAcAsn, obtained at 360-MHz (37). Therefore, the structure of compound I could be confirmed as GlcNAcAsn. In the spectrum of compound II, all signals arising from GlcNAcAsn are present, together with a multiplet at δ = 3.664 ppm, which is attributed to the two unequal Gly H_α atoms (43). The spectral features of GlcNAc are not affected by the presence of Gly, whereas the H-α signal of Asn is shifted from δ = 3.965 ppm to δ = 4.523 ppm going from I to II. The structure of compound II could be confirmed as Gly(GlcNAc)Asn.

The presence of the additional Gly residue in compound II, but not in compound I, accounts for their different chromatographic behavior on Biogel P-4. It is not clear why the presence of a glycine would increase the retention time. The ^{1}H-NMR signals of the Asn protons in the spectrum of compound II, which are strongly influenced by the presence of Gly, are found at the same positions in the spectrum of Ph 1 and Ph 2 (data not shown). So it is likely that Ph 1 like Ph 2 has Gly attached to Asn. The presence of a Gly on the amino-side of Asn is in accordance with the known amino acid sequence of phaseolin (Fig. 7).

The analysis of the proximal GlcNAc residues of the high mannose oligosaccharides from the polypeptides A and C was extended to the $Man_7(GlcNAc)_2$ glycan of the polypeptide D. This glycopeptide has a Gly residue attached at the N-terminal side of Asn, indicating that the same Asn252 is involved in glycosylation.

The observation, that certain glycopeptides whose structures were determined by ^{1}H-NMR, have a Gly residue, allowed us to redefine the elution positions on Biogel P-4: the $Man_7(GlcNAc)_2(Gly)Asn$ glycopeptide elutes in the same position as a $Man_6(GlcNAc)_2Asn$ glycopeptide, and the $XylMan_3(GlcNAc)_2(Gly)Asn$ glycopeptide elutes in the same position as a $Man_3(GlcNAc)_2Asn$ glycopeptide. In these glycopeptides, the presence of the Gly residue paradoxically decreases their apparent size, while the additional sugar increases their size.

Analysis of the Phaseolin Oligosaccharide Intermediates: The various glycopeptides obtained by ConA chromatography (see Fig. 8) were analyzed by glycosidase treatments followed by gel filtration. When the ConA-binding glycopeptides from 3-h-labeled polypeptide A were analyzed by gel filtration chromatography, most of the radioactivity coeluted with $Man_9(GlcNAc)_2Asn$, but a shoulder corresponding to a slightly smaller glycopeptide was visible (Fig. 9, upper left panel). After a 24-h-labeling period, approximately half of the radioactivity now appears in this smaller glycopeptide (Fig. 9, upper right panel). Similar results were obtained for the ConA-binding glycopeptides derived from the polypeptide size class C (data not shown). By comparing these results with those obtained for the characterization of the phaseolin glycopeptides from the mature protein, we identified the smaller glycopeptide as $Man_7(GlcNAc)_2(Gly)Asn$. The comparison also revealed that the high mannose oligosaccharides of phaseolin labeled for 24 h with [^{3}H]glucosamine do not differ in size or proportion from those of the mature protein. Therefore, we consider, that the 24-h-labeled glycopeptides, which carry high mannose sidechains, represent the mature glycans.

Treatment of the high mannose glycopeptides from polypeptide with endo H yielded fragments that coeluted with $Man_8GlcNAc$ (fraction 41), $Man_7GlcNAc$ (fraction 44), GlcNAcAsn (fraction 56/57), and GlcNAc(Gly)Asn (fraction 60) (Fig. 9, lower panels). After 24 h of labeling, the apparent levels of $Man_8GlcNAc$ and $Man_7GlcNAc$ had reached equality. Treatment of the same glycopeptides with α-mannosidase produced products with sizes corresponding to $Man(GlcNAc)_2Asn$ (fraction 48) and $Man(GlcNAc)_2(Gly)Asn$ (fraction 51) (Fig. 9, middle panels).

Analysis of the glycopeptides from polypeptide D by ConA chromatography indicated that a sizable proportion of complex oligosaccharides were present (Fig. 8, lower panels). After 3 h of labeling the tissue, there were at least 2 fractions of glycopeptides which did not bind tightly to ConA (peaks A and B in lower left panel of Fig. 8). Fig. 10 shows the gel filtration profiles of these glycopeptide fractions before and after treatment with β-N-acetylglucosaminidase and/or α-mannosidase. In comparing the elution positions of peaks A and B from the Biogel column before and after β-N-acetylglucosaminidase treatment, it is apparent that peak A glycopeptide has one terminal GlcNAc and peak B has 2 terminal GlcNAc residues. The fact that β-N-acetylglucosaminidase treatment of peak B generated twice the amount of free-GlcNAc/glycopeptide radioactivity than it did for peak A is consistent with this interpretation. Neither glycopeptide was susceptible to α-mannosidase treatment unless this glycosidase was administered simultaneously with β-N-acetylglucosaminidase. Under these circumstances, α-mannosidase appears to remove 2 Man residues from both A and B glycopeptide fragments.

Glycosidase/gel filtration analysis of the glycopeptide fraction from 24-h-labeled polypeptide D that was not retained on the ConA column (Fig. 8, lower right panel) is shown in Fig. 11. The glycopeptide was not susceptible to endo H or β-N-acetylglucosaminidase, but treatment with α-mannosidase removed approximately 2 mannose residues, rendering a fragment that coeluted with $Man_5(GlcNAc)_2Asn$ (fraction 48). The same glycopeptide labeled for 24 h with [^{3}H]mannose, and digested with α-mannosidase, liberated radiolabelled $Man_5(GlcNAc)_2Asn$ and free mannose in a ratio 1:2 (data not shown). These results are consistent with the identification of this glycopeptide as $XylMan_3(GlcNAc)_2(Gly)Asn$, the mature glycan on polypeptide D.

A similar analysis was performed on the ConA-binding glycopeptide fraction from polypeptide D. After a 3-h labeling period with [^{3}H]GlcNAc, the major radioactive glycopeptide was identified as $Man_9(GlcNAc)_2(Gly)Asn$. After 24 h of labeling, most of the radioactivity appeared as $Man_7(GlcNAc)_2(Gly)Asn$ (data not shown).

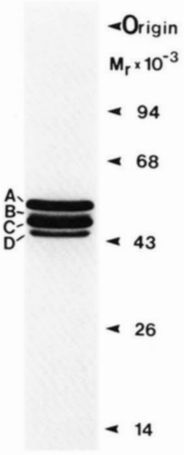

Fig. 1. Flurograph of purified phaseolin labeled with [^{3}H]glucosamine. Each of 20 cotyledons were labeled with 6μCi of [^{3}H]glucosamine for 24 h. Phaseolin was isolated and purified by affinity chromatography then subjected to SDS-PAGE. Polypeptide size class B is difficult to resolve from size class C.

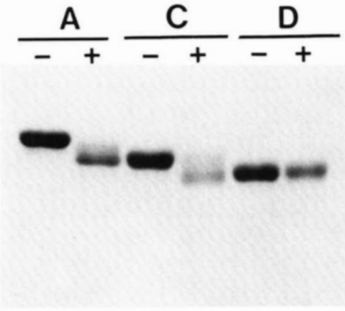

Fig. 2. Effect of endo H treatment on the mobility of phaseolin polypeptide classes A, C, and D. Each of 20 cotyledons were labeled with 6μCi of [^{35}S]methionine for 24 h. Phaseolin was isolated and purified by affinity chromatography then subjected to SDS-PAGE. The individual polypeptides were recovered by electro-elution. (Polypeptide B remained as a contaminant of A and C.) Each radioactive polypeptide was then incubated with (+) or without (−) endo H for 48 h and subjected to SDS-PAGE and fluorography.

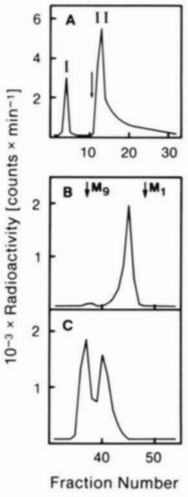

Fig. 3. ConA and gel filtration chromatography of the phaseolin glycopeptides. [^{3}H]GlcNAc-labeled phaseolin (see Fig. 1) was digested with Pronase for 48 h and the resultant glycopeptides were purified by passage through a short (2 x 18 cm) Biogel P-4 column. The purified glycopeptides were subjected to ConA-affinity chromatography (panel A). The arrow indicates the starting point of elution with 200 mM α-methylmannoside. Peaks I and II from the ConA-affinity column were further analyzed by gel filtration on a long (1 x 100 cm) Biogel P-4 column (panels B and C, respectively). In panel B, the arrows labeled M_9 and M_1 indicate the elution positions of $Man_9(GlcNAc)_2Asn$ and $Man_1(GlcNAc)_2Asn$, respectively.

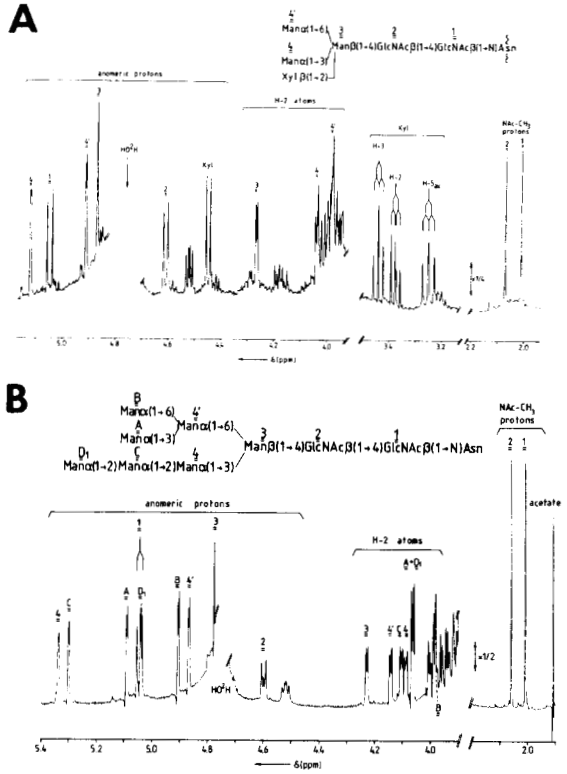

Fig. 4A,B. Structural-reporter-group regions of the resolution-enhanced 500-MHz ^{1}H-NMR spectrum of phaseolin glycopeptides Ph 1 (A) and Ph 2 (B) recorded in ^2H_2O at 27°C. The numbers in the spectrum refer to the corresponding residues in the structure. The relative intensity scale of the N-acetyl region differs from that of other parts of the spectrum as indicated.

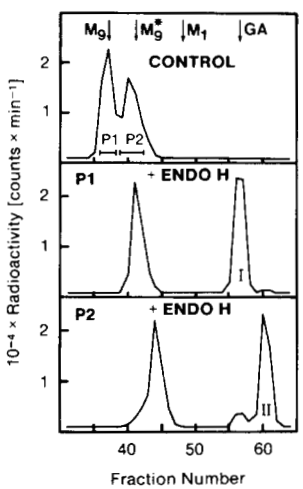

Fig. 5. Gel filtration of the ConA-binding glycopeptides from [^{3}H]GlcNAc-labeled phaseolin after endo H treatment. The glycopeptides which bound to ConA (peak b in Fig. 3A) were fractionated on Biogel P-4 yielding Ph 2 and Ph 3 (upper panel). Each peak was treated separately with endo H, and the resulting products run on the long Biogel column (middle and lower panels). The arrows in the upper panel labeled M_9, M_9*, M_1 and GA indicate the elution positions of $Man_9(GlcNAc)_2Asn$, $Man_9GlcNAc$, $Man(GlcNAc)_2Asn$, and GlcNAcAsn, respectively. The peak labeled I represents GlcNAc-Asn, while the peak labeled II is GlcNAc(Gly)Asn. It is not clear why the presence of a Gly causes retardation of this compound on Sepharose 4B.

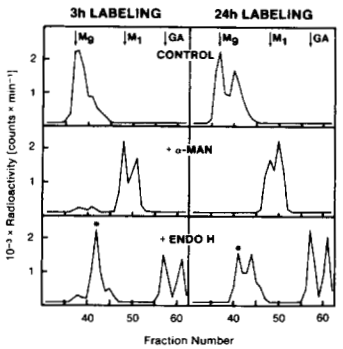

Fig. 9. Gel filtration of 3-h- and 24-h-labeled polypeptide class A glycopeptides before and after treatment with α-mannosidase or endo H. An aliquot of the glycopeptides which were eluted with α-methylmannoside from the ConA column (Fig. 1, upper 2 panels) were applied to the long (1 x 100 cm) Biogel P-4 column. The arrows in the upper panels labeled M_9, M_1 and GA indicate the elution positions of $Man_9(GlcNAc)_2Asn$, $M_1(GlcNAc)_2Asn$, and GlcNAcAsn, respectively. The asterisks in the lower 2 panels indicate the elution positions of $Man_9GlcNAc$.

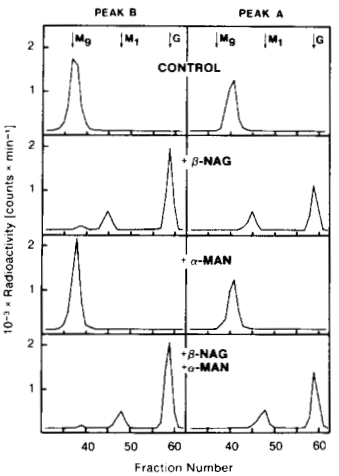

Fig. 10. Gel filtration of 3-h-labeled polypeptide class D glycopeptides before and after treatment with β-N-acetylglucosaminidase and/or α-mannosidase. Aliquots of the glycopeptides that were not retained by ConA-Agarose (Fig. 1, lower left panel, marked as A and B) were applied to the long Biogel P-4 column. The arrows in the upper 2 panels labeled M_9, M_1 and G indicate the elution positions of $Man_9(GlcNAc)_2Asn$, $Man_1(GlcNAc)_2Asn$ and free glucosamine, respectively.

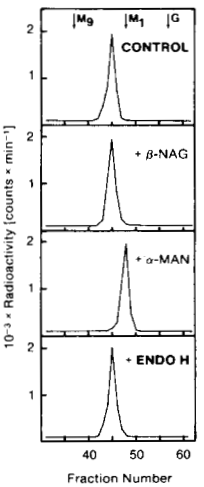

Fig. 11. Gel filtration of 24-h-labeled polypeptide class D glycopeptides before and after treatment with β-N-acetylglucosaminidase, α-mannosidase and endo H. An aliquot of the glycopeptides that were not retained by ConA-Agarose (Fig. 1, lower right panel) was further separated by gel filtration on a calibrated 100 cm column of Biogel P-4. See Fig. 10 for explanation of arrows.

Table I: ^{1}H-Chemical shifts of structural-reporter-group protons of the constituent monosaccharides for the glycopeptides derived from phaseolin (Ph 1-3), together with those for reference compound R (44).

		4'\ 3 - 2 - 1 - ol / 4 X	4'\ 3 - 2 - 1 - Asn / 4 X	B\ 4' A' 3 - 2 - 1 - Asn / D1 - C - 4	D3 - B\ 4' D2 - A' 3 - 2 - 1 - Asn / D1 - C - 4
		R	Ph 1	Ph 2	Ph 3
NAc of	1	2.057	2.010	2.009	2.012
	2	2.073	2.070	2.060	2.066
H-1 of	1		5.048	5.045	5.069
	2	4.634	4.608	4.599	4.606
	3	4.883	4.869	4.779	4.780
	4	5.122	5.121	5.337	5.331
	4'	4.913	4.911	4.869	4.866
	A			5.091	5.400
	B			4.907	5.140
	C			5.301	5.306
	D_1			5.042	5.046
	D_2				5.058
	D_3				5.038
	Xyl	4.449	4.447		
H-2 of	1	4.239			
	3	4.270	4.624	4.228	4.226
	4	4.039	4.037	4.089	4.096
	4'	3.983	3.981	4.143	4.152
	A			4.063	4.107
	B			3.986	4.020
	C			4.104	4.107
	D_1			4.063	4.064[a]
	D_2				4.072[a]
	D_3				4.072[a]
	Xyl	3.377	3.373		
H-3 of	Xyl	3.437	3.438		
H-5ax of	Xyl	3.250	3.248		

[a]Assignments may have to be interchanged.