CHARACTERIZATION OF SOYBEAN (HYDROXY)PROLINE-RICH EARLY NODULINS.

Henk J. Franssen, Ben Scheres, Clemens van der Wiel and Ton Bisseling.

Department of Molecular Biology, Agricultural University, De Dreijen 11, 6703 BC Wageningen, The Netherlands.

INTRODUCTION

Legume root nodule formation can be divided into three major stages denoted as "pre-infection", "infection and nodule formation" and "nodule function" (1). During the formation of the root nodule the nodulin genes (2) are differentially expressed (3,4) and dependent on the time point of the induction of expression they are divided into early and late nodulin genes. The onset of expression of this latter group of genes coincides with the start of nitrogen fixation. The leghemoglobin genes are the best characterized genes resorting in this group. Early nodulin genes are expressed markedly before the nitrogen fixation process starts and hence are likely to be involved in nodule formation and infection. In order to understand these processes our research program is focussed on early nodulin genes. Previously we described the isolation and characterization of a soybean early nodulin cDNA clone, pGmENOD2, coding for N-75, a proline-rich protein, involved in root nodule morphogenesis (5). Here we report the analysis of two other soybean cDNA clones representing early nodulin genes. Both encode proline-rich proteins as well.

RESULTS AND DISCUSSION

From a cDNA library prepared against poly A(+) RNA of soybean root nodules, several clones were isolated that represent early nodulin genes. Here we will describe the characteristics of two of these clones, pGmENOD13 an pGmENOD55 and we will compare them with the previously studied pGmENOD2. Both pGmENOD13 and pGmENOD55 hybridize to nodule RNA and not to RNA from uninoculated roots on a Northern blot. The genes represented by these clones are expressed at least 4 days prior to the leghemoglobine genes. Therefore these two genes are expressed when the plant is infected and the nodule structure is formed. So the ENOD13 and ENOD55 early nodulins might be involved in one of these processes. To discriminate between these two possibilities we analyzed nodules formed by Bradyrhizobium japonicum 3160, a mutant that induces the formation of nodules essentially devoid of infection threads and intracellular bacteria although occasionally a few bacteria can be observed (6). While the ENOD13 mRNA is present at wild type level in nodules formed by 3160, expression of the ENOD55 gene was hardly detected in these nodu-

les. Therefore it is likely that the early nodulin ENOD13 is involved in nodule morphogenesis, like N-75 and not in the infection process. Furthermore we can conclude that the early nodulin ENOD55 is not essential for the formation of the nodule structure. Therefore it is likely that its expression is correlated either to infection thread formation or bacterial release from the infection threads or processes occurring as soon as the bacteria are released. In which of these processes ENOD55 is involved remains to be elucidated.

The nucleotide sequence of the inserts of pGmENOD13 and pGmENOD55 was determined. The amino acid sequence derived from the nucleotide sequence revealed that both the ENOD13 and ENOD55 polypeptides are rich in proline. In the ENOD55 polypeptide the prolines are confined to an internal domain of 32 amino acids in which proline and serine occur alternating. The ENOD13 protein is built up of repeating pentapeptides containing 2 prolines, like the ENOD2 protein and therefore this protein shows structural homology to N-75.

Besides pGmENOD2, pGmENOD13 and pGmENOD55 we have isolated other cDNA clones representing genes that also encode for proline-rich early nodulins. About half of the early nodulin cDNA clones we isolated from cDNA libraries against poly A(+) RNA from pea or soybean nodules encode proline-rich proteins. This striking result might indicate that the root nodule is an unusual organ because so many different proline-rich proteins appear to be present. Before we will discuss this we will first compare these proline-rich early nodulins with other proline-rich plant proteins. However, no proline-rich proteins have been described in plants sofar. All these proteins in fact contain hydroxyproline that is posttranslationally formed. In most cases these proteins are highly glycolysated and are named hydroxyproline-rich glycoproteins (HRGPs). Soybean nodule tissue is extremely hydroxyproline-rich (7). Therefore it is possible that the proline-rich early nodulins ENOD2, ENOD13 and ENOD55 are in fact hydroxyproline-rich. To date four classes of HRGPs in plants have been described (8): A) the extensins, which are structural cell wall proteins (9,10); B) the arabinogalactan proteins (AGP, 11); C) the solanaceae lectins (12) and D) the hydroxyproline-rich agglutinins (13).

The amino acid composition of a representative of each class and the ENOD2 (= N-75), the ENOD13 and the ENOD55 proteins are shown in table 1. The amino acid composition of ENOD13 and N-75 are distinguished from the HRGPs by a remarkably low content of alanine (high in AGPs), glycine and cysteine (high in lectin) and serine (high in extensin) and a high content of glutamic acid (Table 1). The proline residues within the ENOD55 protein are restricted to an internal domain. In the regions flanking this proline-rich domain the amount of glycine, leucine, methionine and valine is high, while alanine and cysteine are low, an amino acid composition not found in any of the HRGPs in table 1. Thus by comparing the amino and composition of the known HRGPs and the ENOD55, ENOD2 and ENOD13 early nodulins, it can be concluded that the amino acid composition of none of the described early nodu-

Table 1. Comparison of the amino acid composition, expressed as mol. %, of typical representatives of plant hydroxyproline-rich glycoproteins and proline-rich proteins, obtained from either amino acid analysis or derived from the DNA sequence. All amino acids not mentioned comprise less than 2 mol. % each in all cases. Putative signal peptides are not included in the amino acid composition.

		Amino acid												
protein	plant	Pro + Hyp	Ser	Glu	Glx	Gln	His	Lys	Tyr	Ala	Val	Gly	Leu	Cys	Reference
extensin	carrot	42.2	10.9		2.6		9.5	11.7	12.0	1.4	4.0	0.0	0.0	0.0	(10)
AGP	French bean	29.6	18.2		4.5		0.6	2.6	0.6	16.2	3.2	6.5	3.2	0.6	(11)
lectin	potato	21.9	12.6		6.9		0.0	3.7	3.3	4.1	0.4	12.2	1.2	10.6	(12)
agglutinin	potato	50.9	9.4		1.1		5.1	15.9	6.2	0.9	3.8	1.1	0.2	0.1	(13)
ENOD2	soybean	44.4	0.4	16.6		5.4	9.5	9.5	7.5	0.0	2.1	0.4	1.2	0.4	(5)
ENOD13	soybean	42.3	2	7.7		-	4.0	11.5	9.6	0.0	4.0	2.0	6.0	0.0	this report
ENOD55	soybean	19.5	26.8	1.2		3.7	3.6	6.1	-	1.2	7.3	8.5	9.8	11*	this report

+derived from the DNA sequence.

*this mol % is constituted of 1.2% cys and 9.8% meth.

lins resembles the amino acid composition of the previously described HRGPs. The structural cell wall protein, extensin, is characterized by the pentapeptide Ser(HyP)4, a repeat that does not occur in any of the three early nodulins. However, the fact that both the ENOD2 and the ENOD13 protein are built up of proline-rich repeating elements strongly suggests that these early nodulins are also structural cell wall proteins.

HRGP mRNA has been shown to accumulate in response to wounding and infection by pathogens and therefore the HRGPs are thought to be involved in the defence response (14,15). Several authors (16,17) have described the legume-Rhizobium symbiosis as a controled pathogenic interaction. In this view one might hypothesize that the ENOD2 and ENOD13 gene expression is part of a defence response. On the other hand, recently two soybean cDNA clones representing genes encoding two different proteins, built up of a repeating pro-pro-val-tyr-lys unit have been isolated. One cDNA clone representing a gene which is differentially expressed during growth of the hypocotyl and it encodes SbPRP1 (18). The other cDNA clone 1A10 (19) represents a gene which is expressed in axis of germinating seedlings and it is proven that this gene encodes a hydroxyproline-rich cell wall protein (19). In addition it has been shown that also the extensin genes are developmentally regulated for example during soybean seed formation (20). Whether the same extensin genes are induced during pathogenic interactions and plant development is unknown.

From these data it can be concluded that several genes encoding (hydroxy) proline-rich proteins exist in plants and these genes are developmentally regulated. For some genes it has

been proven that their expression is related to a defence response. To study whether the expression of the proline-rich early nodulin genes reflect a defence response we followed the expression of the extensin genes during nodule development. By Northern blot analysis using a sunflower extensin cDNA clone as a probe (a kind gift of Terry Thomas, Texas A & M Univ.) we showed that the expression of the extensin genes is at comparable low levels in both nodules and roots (our unpublished results). In addition Werner has shown that phytoalexins do not accumulate in soybean nodules (21). These results show that it is unlikely that nodule formation can be regarded as a defence response from the plant. Hence it is probable that the proline-rich nodulins encoded by ENOD2 and ENOD13 are related to the HRGPs that are involved in developmental processes. The occurrence of a set of plant genes that encode different (hydroxy) proline-rich cell wall proteins which are involved in different developmental programs in the plant raises the question why a plant needs this variety of cell wall proteins. Do these proteins have specific functions in the cell walls of the different tissues and if so, what is the correlation between function and the amino acid composition of the HRGPs?

The group of HRGPs that are built up of repeating units e.g. extensin, SbPRP1, 1A10 and the ENOD2 and ENOD13 early nodulins are rich in proline as well as tyrosine (table 1, 18, 19). For extensin it has been shown that intramolecular isodityrosine cross-linkages and intermolecular dityrosine cross-linkages are formed (22, 23). Extensive intra- and intermolecular cross-linking results in the insolubilization of extensin in the cell wall and the insoluble extensin polymers are thought to harden the cell wall (24). The extent of isodityrosine cross-linking decreases markedly when the pH is reduced from 7 to 4 (25), showing that the formation of these linkages depends on the pH. Cooper and Varner (25) have suggested two mechanisms by which cross-linking could be regulated; (i) wall pH might control the level of activity of the enzyme that catalyses cross-linking; (ii) at a certain pH the extent of cross-linking is determined by the ratio of acidic and basic amino acids within the (hydroxy) proline-rich protein. Since the second possibility involves just the characteristics of the protein it will be interesting to compare the amino acid compositions of the HRGPs. Extensin, SbPRP1, 1A10 and the early nodulins ENOD2 and ENOD13 are all rich in the basic amino acids, histidine and lysine (up to 20%). However, only the proline-rich early nodulins contain also a high level of the acidic amino acid glutamic acid. Thus, these (hydroxy) proline-rich proteins have different ratios of basic and acidic amino acids and hence the extent of tyrosine cross-linking will vary between the different proteins. This of course, will lead to different cell wall properties.

The ratio of acidic and basic amino acid residues present in the proline-rich proteins determines the charge of the polypeptide. Depending on this charge these polypeptides will interact with other structural components of the cell wall (20) and in that way the pKI of the protein will be involved in determining

the cell wall structure. It is conceivable that different tissues in plants have other requirements for their cell wall structure and hence it is understandable that plants have a set of genes encoding (hydroxy) proline-rich proteins available to fulfil these different demands. A good example of a tissue that has a very specific cell wall is the nodule inner cortex. It has been shown that this tissue forms an O_2 barrier in the nodule (26) and it has been proposed that this barrier is formed by a dense packing of the cells. Recently we have shown by in situ hybridization that the ENOD2 gene is exclusively expressed in the inner cortex of the nodule. Since it is very likely that the cell wall structure is involved in the packing of the cells we hypothesize that the ENOD2 early nodulin might infact be involved in creating this oxygen barrier.

LITERATURE CITED

1. Vincent, J.M. 1980. Factors controlling the legume-Rhizobium symbiosis. In: Nitrogen Fixation, eds. Newton, W.E. and Orme-Johnson, W.H. University Park Press, Baltimore, Vol. 2. pp 103-131.
2. Van Kammen, A. 1984. Suggested nomenclature for plant genes involved in nodulation and symbiosis. Plant Mol. Biol. Rep. 2, 43-45.
3. Govers, F., Gloudemans, T., Moerman, M., Van Kammen, A. and Bisseling, T. 1985. Expression of plant genes during the development of pea root nodules. EMBO J. 4, 861-867.
4. Goudemans, T., de Vries, S.C., Bussink, H.J., Malik, N.S.A., Franssen, H.J., Louwerse, J. and Bisseling, T. 1987. Nodulin gene expression during soybean (Glycine max) nodule development. Plant Molec. Biol. 8, 395-403.
5. Franssen, H.J., Nap, J.P., Gloudemans, T., Stiekema, W., van Dam, H., Govers, F., Louwerse, J., van Kammen, A. and Bisseling, T. 1987. Characterization of cDNA for nodulin-75 of soybean: a gene product involved in early stages of root nodule development. Proc. Natl. Acad. Sci. USA 84, 4495-4499.
6. Regensburger, B., Meyer, L., Filser, M., Weber, J., Studer, D., Lamb, J.W., Fischer, H-M., Hahn, M. and Hennecke, H. 1986. Bradyrhizobium japonicum mutants defective in root nodule bacteroid development and nitrogen fixation. Arch. Microbiol. 144, 355-366.
7. Cassab, G.I. 1986. Arabinogalactan proteins during the development of soybean root nodules. Planta 168, 441-446.
8. McNeil, M., Darvil, A.G., Fry, S.C. and Albersheim, P. 1984. Structure and function of the primary cell walls of plants. Ann. Rev. Biochem. 53, 635-663.
9. Van Holst, G.J. and Varner, J.E. 1984. Reinforced polyproline II conformation in an hydroxyproline-rich cell wall glycoprotein from carrot root. Plant Physiol. 74, 247-251.
10. Chen, J. and Varner, J.E. 1985. An extracellular matrix protein in plants: characterization of a genomic clone for carrot extensin. EMBO J. 4, 2145-2151.
11. Van Holst, G.J., Klis, F.M., de Wildt, P.J.M., Hazenberg, C.A.M., Buijs, J. and Stegwee D. 1981. Arabinogalactan pro-

tein from a crude cell organelle fraction of Phaseolus vulgaris L. Plant Physiol. 68, 910-913.
12. Allen, A.K., Desai, N.N., Neuberger, A. and Creeth, J.M. 1978. Properties of potato lectin and the nature of its glycoprotein linkages. Biochem. J. 171, 665-674.
13. Leach, J.E., Cantrell, M.A. and Sequira, L. 1982. A hydroxyproline-rich bacterial agglutinin from potato: extraction, purification and characterization. Plant Physiol. 70, 1353-1358.
14. Showalter, A.M., Bell, J.N., Cramer, C.L., Bailen, J.A., Varner, J.E. and Lamb, C.J. 1985. Accumulation of hydroxyproline-rich glycoprotein mRNAs in response to fungal elicitor and infection. Proc. Natl. Acad. Sci. USA 82, 6551-6555.
15. Corbin, D.R., Sauer, N. and Lamb, C.J. 1987. Differential regulation of a hydroxyproline-rich glycoprotein gene family in wounded and infected plants. Molec. and Cell. Biol. 7, 4337-4344.
16. Vance, C.P. 1983. Rhizobium infection and nodulation: a beneficial plant disease? Ann. Rev. Microbiol. 37, 394-424.
17. Djordjevic, M.A., Gabnel, D.W. and Rolfe, B.G. 1987. Rhizobium: the refined parasite of legumes. Ann. Rev. Phytopathol. 25, 145.
18. Hong, J.C., Nagao, R.T. and Key, J.L. 1987. Characterization and sequence analysis of a developmentally regulated putative cell wall protein gene isolated from soybean. J. Biol. Chem. 262, 8367-8376.
19. Averyhart-Fullard, V., Datta, K. and Marcus, A. 1988. A hydroxyprolinerich protein in the soybean cell wall. Proc. Natl. Acad. Sci. USA 85, 1083-1085.
20. Cassab, G.I., Nieto-Sotelo, J., Cooper, J.B., Van Holst, G.J. and Varner, J.E. 1985. A developmentally regulated hydroxyproline-rich glycoprotein from the cell walls of soybean seed coats. Plant Physiol. 77, 532-535.
21. Werner, D., Mellor, R.B., Hahn, M.G. and Grisebach, H. 1985. Soybean root response to symbiotic infection. Glyceolin I accumulation in an ineffective type of soybean nodules with an early loss of the peribacteroid membrane. Z. naturforsch. 40C, 171-181.
22. Epstein, L. and Lamport, D.T.A. 1984. An intramolecular linkage involving isodityrosine in extensin. Phytochemistry 23, 1241-1246.
23. Fry, S.C. 1986. Cross-linking of matrix polymers in the growing cell walls of Angiosperms. Ann. Rev. Plant Physiol. 37, 165-186.
24. Cooper, J.B. and Varner, J.E. 1983. Insolubilization of hydroxyproline-rich glycoprotein in aerated carrot root slices. Biochem. Biophys. Res. Commun. 112, 161-167.
25. Cooper, J.B. and Varner, J.E. 1984. Cross-linking of soluble extensin in isolated cell walls. Plant Physiol. 76, 414-417.
26. Witty, J.F., Minchin, F.R., Skøt, L. and Sheehy, J.E. 1986. Nitrogen fixation and oxygen in legume root nodules. Oxford Surveys of Plant Molecular and Cellular Biology 3, 275-315.