*Phytochemistry*, Vol. 27, No. 6, pp. 1719–1721, 1988.
Printed in Great Britain.

0031-9422/88 $3.00 + 0.00


# AMINO ACID SEQUENCES OF TWO ATYPICAL SINGLE-CHAIN VICIEAE ISOLECTINS FROM SEEDS OF *LATHYRUS NISSOLIA*

A. YARWOOD, M. RICHARDSON, B. MORPHET, M. WESTBY, D. PÈRE* and P. ROUGÉ*

Department of Botany, University of Durham, Durham, DH1 3LE, U.K.; *Laboratoire de Botanique et Biologie Cellulaire, Faculté des Sciences Pharmaceutiques, Université Paul Sabatier, 35 chemin des Maraîchers 31062 Toulouse, France

(*Received* 30 *October* 1987)

**Key Word Index**—*Lathyrus nissolia*; Leguminosae; Vicieae; isolectins; amino acid sequences.

**Abstract**—The amino acid sequences of the isolectins (LnL1 and LnL2) from seeds of *Lathyrus nissolia* were determined by analysis of peptides derived from the proteins by separate digestions with trypsin, chymotrypsin, pepsin and the protease from *S. aureus* V8, together with chemical cleavage by iodosobenzoic acid. Both isolectins were atypical of the lectins normally found in the tribe Vicieae in that they were composed of single polypeptide chains of 235 (LnL1) and 233 (LnL2) amino acids whereas the Vicieae lectins usually consist of a heavy $\beta$ chain of *ca* 181 residues and a light $\alpha$ chain of 53–55 residues. Nevertheless the *L. nissolia* isolectins showed strong overlapping homology with both the $\beta$ and $\alpha$ chains of lectins from this tribe.

## INTRODUCTION

The amino acid sequences of lectins from numerous species belonging to the tribe Vicieae of the family Leguminosae have been reported [1–15] with the genus *Lathyrus* having received particular attention [5–9]. In nearly every case so far the lectins have been shown to have a two-chain structure $(\beta\alpha)_2$ where the heavy $\beta$ chains contain 181 amino acids, and the light $\alpha$ chains only 52–54 amino acids. The two chains of these highly conserved three domain lectins appear to have been generated by the post-translational cleavage of a larger precursor by an asparagine-specific endoproteinase [10]. We now wish to report that the two major isolectins (LnL1 and LnL2) isolated from seeds of *Lathyrus nissolia* L. are both atypical in being composed of single polypeptide chains of 235 and 233 amino acids respectively.

## RESULTS AND DISCUSSION

The two major isolectins present in mature seeds of *L. nissolia* were initially purified by the standard methods of fractional precipitation with ammonium sulphate and affinity chromatography on Sephadex G-100, as used for the isolation of lectins from other *Lathyrus* spp [5–9]. When the mixture of isolectins obtained was subjected to chromatofocusing on a column of PBE 94 two forms were recovered, isolectin 1 (LnL1) with pI of 8.4 and isolectin 2 (LnL2) with pI of 8.1. Both isolectins gave a single peak during gel filtration on a column of Biogel P-60 equilibrated and eluted with 6M guanidine HCl, and single bands on SDS-PAGE with apparent $M_r$ of 28 000.

Separate digestions of both isolectins with trypsin, chymotrypsin and pepsin and redigestions of certain peptides with the protease from *S. aureus* V8 yielded suitably overlapping peptides from which most of the sequence could be deduced by micro-sequence analysis using the DABITC/PITC double coupling method (Fig. 1). Redigestions of certain long tryptic peptides (T7 and T15) with the *S. aureus* V8 protease were helpful in determining the sequences in the regions 83–88 and 171–175 where good overlaps were otherwise missing. Cleavage at TRP-X peptide bonds with iodosobenzoic acid clarified the sequences of residues 62–69 and 158–162.

The sequences of both LnL1 and LnL2 were compatible with them being single chain polypeptides of 235 and 233 residues respectively which show strong overlapping homology (>95% identity) with the sequences of the *Lathyrus* lectins [9] starting at the N-termini of the $\beta$ chains and running through to the C-termini of the $\alpha$ chains. The sequences were in good agreement with the results of amino acid analyses of the intact isolectins and the compositions of the peptides.

The two-chain lectins found in the seeds of other species belonging to the tribe Vicieae are thought [10] to result from the post-translational cleavage of a larger precursor by an asparagine specific endoproteinase acting on the peptide bond $ASN^{181}$–$X^{182}$. Both of the *L. nissolia* isolectins contain a similar peptide bond in the exactly homologous position yet there is no apparent hydrolysis of this bond in these seeds. No traces of the shorter $\beta$ and $\alpha$ subunits were detected. The exact reasons for the absence of this cleavage during the maturation of the isolectins in *L. nissolia* are unclear. One possibility is that the appropriate enzyme may be missing in *L. nissolia*. An alternative explanation may be that the environment of the normally susceptible peptide bond in *L. nissolia* may be subtly altered. When these regions are compared (Fig. 2) it can be seen that the two-chain lectins all appear to contain a potential *N*-glycosylation site (N–X–T) at the susceptible bond whereas the single chain isolectins from *L. nissolia* do not.

Recent work has established that during the processing of pro-concanavalin A to the mature CON A [11] there is a deglycosylation involving the loss of N-

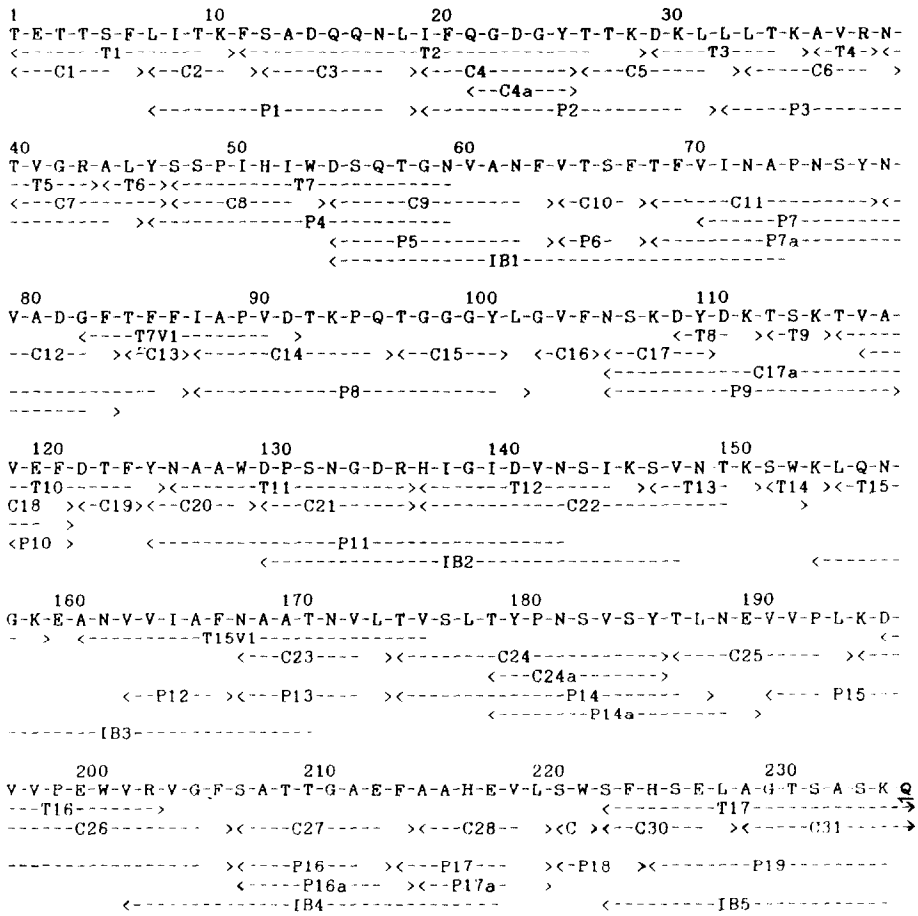

Fig. 1. The amino acid sequences of the isolectins LnL1 and LnL2 from seeds of *L. nissolia*. $T_1$ tryptic peptides; C, chymotryptic peptides; P, peptic peptides; V, peptides from redigestions with *S. aureus* V8 protease. 1B, fragments from cleavage with iodosobenzoic acid. (----) Regions of peptides sequenced by the DABITC method. The amino acid sequence of LnL1 and LnL2 only differed in five positions. The sequence of LnL2 lacked the two C-terminal residues found in LnL1, and contained an ALA (instead of PRO) in position 192, an ILE (instead of VAL) in position 203, and an ASP (instead of ALA) in position 228. These few alterations were entirely compatible with the observed small difference in the pIs of the two isolectins, LnL1 with a value of 8.4 being slightly more basic than LnL2 with a value of 8.1

| Two-chain lectins | | Ref. |
|---|---|---|
| *Lathyrus ochrus* | T-Y-P-N-E-T-S-Y | [5,6] |
| *L. cicera* | T-Y-P-N-V-T-S-Y | [7,9] |
| *L. aphaca* | T-Y-S-N-V-T-S-Y | [7,9] |
| *L. articulatus* | T-Y-P-N-A-T-S-Y | [7,9] |
| *V. faba* | L-Y-P-N-L-T-G-Y | [3,4] |
| **Single-chain lectins** | | |
| *L. nissolia* 1 and 2 | 180 185<br>T-Y-P-N-S-V-S-Y | |
| *L. sphaericus* | 188 193<br>T-Y-P-N-S-R-D-Y | [14] |

Fig. 2. Comparison of the amino acid sequences around the putative sites of post-translational cleavage which yield the $\beta$ and $\alpha$ chains of the two-chain lectins of the Vicieae with the homologous region of the single-chain isolectins from *Lathyrus nissolia*. *Indicates potential site of *n*-glycosylation. ↓ Site of post-translational cleavage.

linked high mannose oligosaccharides [12] and the cleavage of several peptide bonds by an asparagine endopeptidase together with a transpeptidation event [13]. The possibility that a glycosylation/deglycosylation event in this region is a pre-requisite for hydrolysis of the potentially susceptible peptide bond is supported by the recent report that the sequence of another atypical single chain lectin from seeds of *Lathyrus sphaericus* [14] also lacks the *N*-glycosylation sequence (Fig. 2). We are currently examining the possibility that a deglycosylation is a necessary pre-requisite for the endoproteolytic processing of the two-chain lectins in the tribe Vicieae.

## EXPERIMENTAL

Seeds of *Lathyrus nissolia* L.* were harvested from plants grown under field conditions. The seeds were ground, extracted, and the extract subjected to fractional pptn with $(NH_4)_2SO_4$ and the lectins purified by affinity chromatography on Sephadex G-100 by the methods described previously [8].

The two *L. nissolia* isolectins were separated by chromatofocusing in the pH-range between pH 9.4 and pH 6.0 on a PBE 94 (Pharmacia) column equilibrated with 25 mM ethanolamine–AcOH (pH 9.4) buffer and eluted with a 10-fold dilution of Polybuffer 96 (Pharmacia) adjusted to pH 6.0 with 1.0 M AcOH. The isolation of the isolectins subunits was carried out by filtration on a Bio-Gel P60 (Bio-Rad) column equilibrated with 6.0 M guanidine–HCl. Both isolectins gave a single peak.

The $M_r$ of the LnL was estimated by gel filtration by using an AcA 54 (IBF, Villeneuve-la-Garenne, France) column equilibrated with 0.2 M PBS (pH 7.0). The $M_r$ of the isolectin subunits were estimated by SDS-PAGE in gradient (4–30% acylamide) polyacrylamide gel slabs using 40 mM Tris–acetate (pH 7.4) buffer containing 0.1% Na EDTA and 0.2% SDS for the electrophoretic run. Samples, made 1% SDS and 5% ME, were heated at 100° for 10 min before electrophoresis.

*Enzyme digestions and separation of peptides.* Samples (3 mg) of the isolectins were digested separately with trypsin chymotrypsin and pepsin and certain of the resulting large peptides were also redigested with the protease from *S. aureus* V8 as described in ref. [15]. The mixtures of peptides produced by these methods were resolved by reverse phase HPLC on a Vydac $C_{18}$ column (25 cm × 4.6 mm, 218 TP54, Technicol Ltd, Stockport) in a Varian model 5000 HPLC apparatus using a linear gradient of 0–70% MeCN (HPLC grade S, Rathburn, Scotland) in 0.1% $CF_3CO_2H$. Peptides were detected by measuring $A_{214}$.

*Cleavage with iodosobenzoic acid.* Samples (5 mg) of the isolectins were cleaved at TRP-X peptide bonds by iodosobenzoic acid in 80% HOAc–4 M guanidine–HCl for 21 hr at 20° as in ref. [16]. The resulting fragments were initially fractionated on columns (1 × 150 cm) of Biogel P-30 in 70% $HCO_2H$ and subsequently purified by reverse phase HPLC as described above.

*Sequence determination.* The intact isolectins and the peptides derived from them by enzymic digestion or chemical cleavage were subjected to micro-sequence analysis using the DABITC-/PITC double coupling method [17]. The presence of tryptophan in certain peptides was confirmed by staining on paper with *p*-dimethylaminobenzaldehyde. Quantitative amino acid analyses of the isolectins were obtained using a Beckman model 119 BL amino acid analyser. Quantitative analyses of certain peptides were obtained using a Water PICO-TAG system. All other peptides were analysed qualitatively using the dansyl TLC method [15].

*Acknowledgements*—We are grateful to the Science and Engineering Research Council and the Smith Kline Foundation for financial support. We also thank J. Gilroy for valuable technical assistance and Professor D. Boulter for the provision of certain facilities.

## REFERENCES

1. Higgins, T. J. V., Chandler, P. M., Zurawski, G., Button, S. C. and Spencer, D. (1983) *J. Biol. Chem.* **258**, 9544.
2. Richardson, C., Behnke, W. D., Freisheim, J. H. and Blumenthal, K. M. (1978) *Biochim. Biophys. Acta* **537**, 310.
3. Cunningham, B. A., Hemperly, J. J., Hopp, T. P. and Edelman, G. M. (1979) *Proc. Natl Acad. Sci. U.S.A.* **76**, 3218.
4. Hemperly, J. J., Hopp, T. P., Becker, J. W. and Cunningham, B. A. (1979) *J. Biol. Chem.* **254**, 6803.
5. Richardson, M., Rougé, P., Sousa-Cavada, B. and Yarwood, A. (1984) *FEBS Letters* **175**, 76.
6. Yarwood, A., Richardson, M., Sousa-Cavada, B. and Rougé, P. (1985) *FEBS Letters* **184**, 104.
7. Sousa-Cavada, B., Richardson, M., Yarwood, A., Père, D. and Rougé, P. (1986) *Phytochemistry* **25**, 115.
8. Yarwood, A., Richardson, M., Sousa-Cavada, B., Père, D. and Rougé, P. (1986) *Phytochemistry* **25**, 2109.
9. Rougé, P., Richardson, M., Chatelain, C., Yarwood, A., Sousa-Cavada, B. and Père, D. (1986) in *Lectins V* (Bøg-Hansen, T. C. and Van Driessche, E., eds) p. 185. W. de Gruyter, Berlin.
10. Higgins, T. J. V. (1984) *Annu. Rev. Plant Physiol.* **35**, 191.
11. Carrington, D. M., Auffret, A. and Hanke, D. (1985) *Nature* **313**, 64.
12. Herman, E. M., Shannon, L. M. and Chrispeels, M. J. (1985) *Planta* **165**, 23.
13. Bowles, D. J., Marcus, S. E., Pappin, D. J. C., Findlay, J. B. C., Eliopoulos, E., Maycox, P. R. and Burgess, J. (1986) *J. Cell Biol.* **102**, 1284.
14. Richardson, M., Yarwood, A. and Rougé, P. (1987) *FEBS Letters* **216**, 145.
15. Richardson, M., Campos, F. D. A. P., Moreira, R. A., Ainouz, I. L., Begbie, R., Watt, W. B. and Pusztai, A. (1984) *Eur. J. Biochem.* **144**, 101.
16. Mahoney, W. C., Smith, P. K. and Hermodson, M. A. (1981) *Biochemistry* **20**, 443.
17. Chang, T. Y., Brauer, D. and Wittman-Liebold, B. (1978) *FEBS Letters* **93**, 205.

* Mature plants grown from these seeds, were authenticated by taxonomists and voucher specimens are deposited in the herbarium of the University of Toulouse.