Phytochemistry, Vol. 25, No. 9, pp. 2109–2112, 1986.
Printed in Great Britain.

0031–9422/86 $3.00 + 0.00
Pergamon Journals Ltd.

THE AMINO ACID SEQUENCES OF THE α SUBUNITS OF THE LECTINS FROM THE SEEDS OF *LATHYRUS HIRSUTUS* AND *LATHYRUS TINGITANUS*

A. YARWOOD, M. RICHARDSON, B. SOUSA-CAVADA*, D. PÈRE† and P. ROUGÉ†

Department of Botany, University of Durham, Durham, DH1 3LE, U.K.; *Department of Biochemistry, Federal University of Pelotas, Pelotas (RS), Brazil; †Laboratoire de Botanique et Biologie Cellulaire, Faculté des Sciences Pharmaceutiques, Université Paul Sabatier, 34 chemin des Maraîchers, 31062 Toulouse, France

(*Received* 6 *January* 1986)

Key Word Index—*Lathyrus hirsutus*; *L. tingitanus*; Leguminosae; Vicieae; lectins; amino acid sequences.

Abstract—The amino acid sequences of the α_1 and α_2 subunits of the isolectins (LhL1 and LhL2) from seeds of *Lathyrus hirsutus* and the α subunit of the lectin from *L. tingitanus* were determined by analysis of peptides derived from the proteins by separate digestions with chymotrypsin and the protease from *S. aureus* V8. The α_1 subunit of the *L. hirsutus* LhL1 isolectin differed from the α_2 form in LhL2 only in having an extra lysine inserted near the C-terminus. The *L. tingitanus* α subunit differed from the *L. hirsutus* α_1 in three positions and from *L. hirsutus* α_2 in four.

INTRODUCTION

We have recently reported the amino acid sequences of the α subunits (light chains) and β subunits (heavy chains) of a number of lectins and isolectins isolated from seeds of the genus *Lathyrus* [1–3]. These sequences and the secondary structures predicted from them [4] reveal that these two-chain legume lectins are highly conserved three domain proteins which afford an opportunity of studying speciation not only in the genus *Lathyrus* but also within the tribe Vicieae [5]. We now wish to report the amino acid sequences of the α subunit from the *Lathyrus tingitanus* lectin and the α_1 and α_2 subunits from the isolectins (LhL1 and LhL2) from seeds of *L. hirsutus*.

RESULTS AND DISCUSSION

Separate digestions of the α subunits of the lectins from *Lathyrus hirsutus* and *L. tingitanus* with chymotrypsin and the protease from *Staphylococcus aureus* V8 yielded suitably overlapping peptides, all of which were readily separated by reverse-phase HPLC and from which the complete sequences of the α proteins could be deduced by microsequence analysis using the DABITC/PITC double coupling method (Fig. 1).

The observed specificity of the chymotrypsin used was as expected, except that it gave rise to anomalous hydrolyses of the Arg^{21}–Ile^{22} peptide bonds in all three proteins, as was the case in the α subunits from *L. cicera* and *L. articulatus* [2]. The only example of anomalous hydrolysis by the *S. aureus* V8 protease occurred at Ser^{25}–Ala^{26} in the *L. hirsutus* α_1 subunit and again was similar to the anomalous hydrolyses of SER-X peptide bonds that we have noted previously [2, 6]. The failure of this enzyme to hydrolyse the peptide bond E^{47}–E^{48} in the *L. tingitanus* α subunit is in accord with the results of other workers [7].

The amino acid sequences are shown in Fig. 1 together with the details of the peptides from which they were deduced. Both the *L. hirsutus* α_1 and *L. tingitanus* α proteins consisted of 54 amino acids, whereas *L. hirsutus* α_2 contained only 53. The M_r calculated from the amino acid sequences were in good agreement with the estimates of the sizes (M_r 6000) of the subunits made by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE). The sequences shown in Fig. 1 were also in agreement with the amino acid compositions of the subunits except for some discrepancies in the values for aspartic acid, glycine and alanine.

The amino acid sequences of the *L. hirsutus* α_1 and α_2 subunits were identical except for the apparent deletion of Lys^{53} in the α_2 form, and both differed from the *L. tingitanus* α subunit in positions 9, 47 and 48. Figure 2 shows a comparison of the amino acid sequences of the *L. hirsutus* and *L. tingitanus* α lectins with those of the other *Lathyrus* species which have been determined [1, 3, 8, 9] and with the α chains of the other five members of the tribe Vicieae which are currently available [10–14]. The most striking feature of this comparison is that the *L. tingitanus* α subunit which has Ile^9 is the first representative of the genus *Lathyrus* and the tribe Vicieae so far to contain an amino acid other than Val in this position. It should be noted however that the homologous positions in the single-chain lectins from the tribes Diocleae [15, 16], Hedysareae [17] and Phaseoleae [18, 19] are not invariant.

This change in position 9 brings the number of invariant residues found amongst the α lectins of the genus *Lathyrus* down to 41 out of the 53–55 amino acids which are compared, and there are 28 which remain invariant throughout the tribe Vicieae. These results confirm that the *Lathyrus* and Vicieae lectins are highly conserved molecules, a feature which becomes even more apparent when their secondary structures are predicted and their hydropathic profiles computed [4]. It appears that all these legume lectins possess a very similar three-dimensional conformation whatever their minor variations in molecular structure [4]. This retention of such a highly conserved structure is almost certainly related to

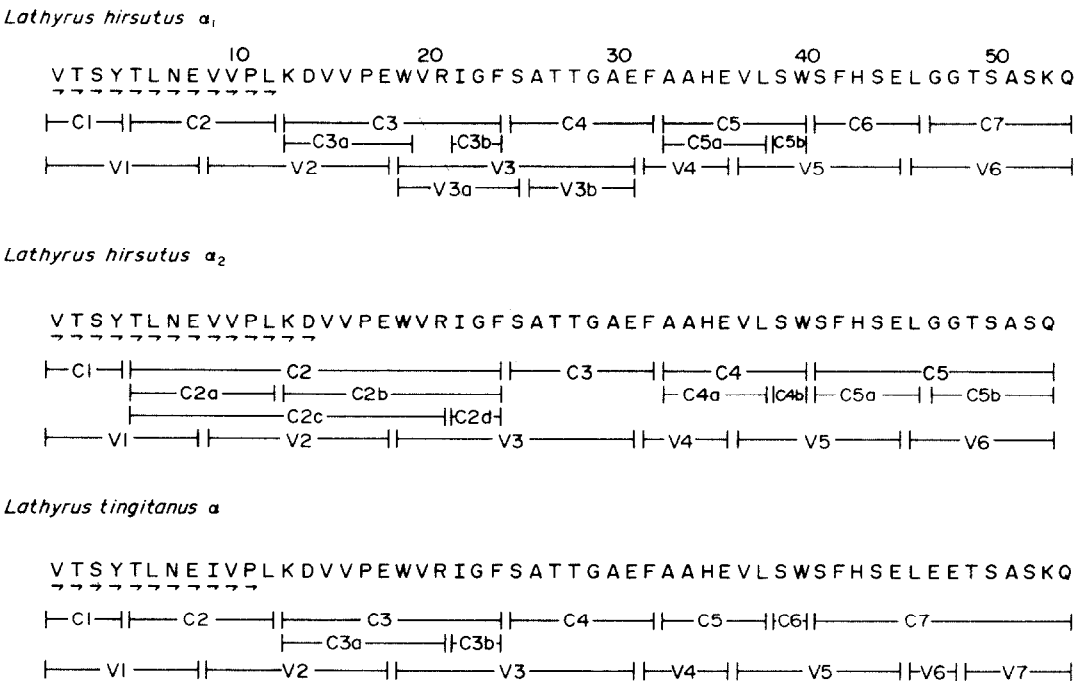

Fig. 1. The amino acid sequences of the α subunits of the lectins from *Lathyrus hirsutus* and *L. tingitanus*. C, chymotryptic peptides; V, peptides from digestion with protease from *S. aureus* V_8. (—) Amino acids identified by the DABITC microsequencing method applied to the intact subunits.

	10 20 30 40 50
Lathyrus ochrus α_1	ETSYTLNEVVPLKEFVPEWVRIGFSATTGAEFAAHEVLSWFFHSELAGTSSSN
L. ochrus α_2	----------SV-----
L. articulatus α	A----------S----------
L. aphaca α	A---------A--DV------V----------S-Q---S---G--
L. cicera α_1	V----------DV----------S-----GE--A-KQ
L. cicera α_2	V----------DV----------S-----GE--A-dQ
L. odoratus α	V----------DV----------S-----G---G-QK
L. sativus α	V----------DV----------S-----DG---A-KQS
L. hirsutus α_1	V----------DV----------S-----G---A-KQ
L. hirsutus α_2	V----------DV----------S-----G---A-dQ
L. tingitanus α	V-------I----DV----------S-----EE--A-KQ
Lens culinaris α	V----------DV----------S-N-Q-GH--K-
Pisum sativum α	V-----SD--S--DV----------Y---------S-----S----dKQ
Vicia cracca α	V-----SD-----DV----------P---Y---------S-----S----dKQ
Vicia faba α	L-G---S------DV----------Y-T------T-L---T-P-N
Vicia sativa α	S---G-SA-----DV----------D-Y--Q--H--S-----G----

Fig. 2. Homology of the α lectins from the tribe Vicieae. The sequences of the *Lathyrus ochrus* α chains are as in ref. [1], *L. articulatus*, *L. aphaca* and *L. cicera* as in ref. [2], *L. odoratus* in ref. [8], *L. sativus* in ref. [9], *Lens culinaris* in ref. [10], *Pisum sativum* in ref. [11], *Vicia cracca* in ref. [12], *V. faba* in ref. [13] and *V. sativa* in ref. [14]. Dashes indicate the same amino acid as in *L. ochrus* α_1. d, Deletions inserted to facilitate the comparison of sequences.

their physiological role(s) within the plant, which may in turn be related to the three functional features (hydrophobic cavity, cation binding sites and carbohydrate binding sites) which have so far been identified in these molecules [20, 21].

In Fig. 2 it can be seen that of the residues in homologous positions to those contributing to the hydrophobic cavity in the concanavalin A molecule [20], namely Leu^{12}, Val^{16}, Val^{20}, Ile^{22}, Phe^{42} and Ser^{44}, the only one to vary is residue 22 in *L. aphaca* which is replaced by a similarly hydrophobic Val. Also the three residues of the α chains shown in Fig. 2 which are in homologous positions to those believed to be involved in the carbohydrate binding of Con A [21], namely Gly^{29}, Ala^{30} and Glu^{31} are invariant apart from an Asp^{30} in *Vicia faba.*

Table 1 shows the minimal number of nucleotide substitutions which are required to interconvert these α lectin proteins from the *Lathyrus* species examined so far. Inspection of this matrix shows that the closest affinities occur between *hirsutus* α_1 and *cicera* α_1, and between *hirsutus* α_2 and *cicera* α_2. *Lathyrus tingitanus* is also quite closely related to *L. cicera* and *L. hirsutus*, and these three species form a group with *L. odoratus* and *L. sativus*. Another distinct group is formed by the α proteins from *L. ochrus* and *L. articulatus*, but *L. aphaca* appears to be relatively unrelated to either of these other groupings.

EXPERIMENTAL

Seeds of *Lathyrus hirsutus* L. and *L. tingitanus* L.* were harvested from plants grown under field conditions. The seeds (300 g) were ground and the resulting meal was soaked overnight at 4° with continuous stirring in 2 l. of 50 mM Tris, 150 mM NaCl buffer (pH 7.6). After centrifugation the proteins in the supernatant were subjected to a fractional precipitation with $(NH_4)_2SO_4$. The 30–60% $(NH_4)_2SO_4$ ppts were dialysed against the extraction buffer and the solns (500 ml) were filtered through columns (5 × 70 cm) of Sephadex G-100 equilibrated in the same buffer. The lectins retained by the columns were eluted by adding 0.1 M Glc to the eluting buffer, ppted with 90% (w/v) $(NH_4)_2SO_4$ and extensively dialysed against the Tris-buffered saline.

The two *L. hirsutus* isolectins were separated by chromatofocusing on a PBE 96 (Pharmacia) column (1 × 30 cm) using 25 mM triethylamine–HCl (pH 11) as the starting buffer. Elution was performed with a 45-fold dilution of Pharmalyte pH 8–10.5 (Pharmacia) adjusted to pH 7.0 with 0.2 M HCl. Two main peaks were obtained, corresponding respectively to LhL1 (pH 8.4–8.3) and LhL2 (pH 7.9–7.7). These *L. hirsutus* isolectins were precipitated with 90% (w/v) $(NH_4)_2SO_4$ and extensively dialysed against H_2O. The *L. tingitanus* lectin was also purified by chromatofocusing using the same column with 25 mM imidazole–HCl (pH 7.4) as the starting buffer and an 8-fold dilution of Polybuffer 74 (Pharmacia) adjusted to pH 4.0 with 0.2 M HCl as the eluting buffer. This procedure yielded three well separated peaks. The first to elute (between pH 5.9 and 5.5) was the major one, and contained the *L. tingitanus* lectin which was collected and treated as described above.

The light (α) and heavy (β) subunits of the *L. hirsutus* isolectins and the *L. tingitanus* lectin were separated by chromatography on a column (2.6 × 100 cm) of Biogel P-60 (Bio-Rad) equilibrated with 6 M guanidine–HCl. Two peaks corresponding respectively

Table 1. Comparison of the sequences of the α subunits of the lectins from the genus *Lathyrus* and other species of the tribe Vicieae

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1. *L. cicera* α_1	0	1	4	3	2	4	4 or 5	11	10	11	12	12	10	10	19	21
2. *L. hirsutus* α_1		0	3	4	3	3	4	10	9	10	12	12	9	9	21	20
3. *L. hirsutus* α_2			0	1	6	5	6 or 7	12	11	12	14	9	12	12	18	17
4. *L. cicera* α_2				0	5	6	7 or 8	13	12	13	14	9	13	13	19	18
5. *L. tingitanus* α					0	6	6	13	11	12	14	14	12	12	23	23
6. *L. odoratus* α						0	6 or 7	10	10	11	13	12	11	11	21	20
7. *L. sativus* α							0	13 or 14	12	13	15	15 or 16	12 or 13	12 or 13	24	23 or 24
8. *L. aphaca* α								0	7	9	10	14	15	16	20	20
9. *L. articulatus* α									0	2	5	13	15	15	19	18
10. *L. ochrus* α_1										0	4	14	16	16	20	20
11. *L. ochrus* α_2											0	13	17	17	22	22
12. *Lens culinaris* α												0	20	20	18	19
13. *Pisum sativum* α													0	2	16	16
14. *Vicia cracca* α														0	16	16
15. *Vicia faba* α															0	15
16. *Vicia sativa* α																0

The numbers show the minimal number of nucleotide substitutions required to interconvert the α proteins

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

to the heavy β (first peak) and light α (second peak) subunits were obtained in all three cases. The guanidine–HCl was eliminated by extensive dialysis against H_2O and the subunits recovered by lyophilization.

The purity of the various isolectins and their subunits was confirmed by PAGE at basic [22] and acid pH [23] and checked by analytical isolectric focusing in the presence of 8 M urea between pH 8.0 and pH 5.0 [24]. The M_rs of the isolectins were estimated by chromatography on a column (1 × 100 cm) of AcA 54 (IBF, Villeneuve la Garenne) buffered and eluted with 0.2 M Pi (pH 7.2) containing 150 mM NaCl. The M_rs of the subunits were estimated by SDS-PAGE on gradient (4–30%) acrylamide gel slabs, using a 40 mM Tris–acetate (pH 7.4) buffer containing 0.2% SDS and 0.1% Na_2 EDTA for electrophoresis. The samples in 1% SDS and 5% mercaptoethanol were boiled for 10 min and electrophoresis was conducted at 150 V until the tracking dye (bromophenol blue) had migrated to the bottom of the gel slabs.

Enzyme digestions and separation of peptides. Samples (3 mg) of the α subunits were digested separately with chymotrypsin and the protease from *S. aureus* V_8 as described in ref. [16]. The mixtures of peptides produced by these methods were resolved by reverse phase HPLC on a Micropac MCH-10 column (0.4 × 30 cm, Varian) 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 the A at 214 nm.

Sequence determination. The intact α subunits and the peptides derived from them by enzymic digestion were subjected to microsequence analysis using the DABITC/PITC double coupling method [25]. The presence of tryptophan in certain peptides was confirmed by staining on paper with *p*-dimethylaminobenzaldehyde. The *C*-terminal residues of certain peptides and the intact α subunits were determined by digestion with carboxypeptidase A as in ref. [1] in combination with the dansyl technique as in ref. [26]. For quantitative amino acid analyses samples were hydrolysed in twice glass-distilled 6.0 M HCl at 110° for 18 hr and analysed on a Beckman model 119 BL amino acid analyser.

Acknowledgements—We are grateful to the Science and Engineering Research Council, the Smith Kline Foundation and the CNPq (Brazilian National Council for Scientific and Technological Development) for financial support. We also thank J. Gilroy and D. Bown for valuable technical assistance and Professor D. Boulter for the provision of certain facilities.

REFERENCES

1. Richardson, M., Rougé, P., Sousa-Cavada, B. and Yarwood, A. (1984) *FEBS Letters* **175**, 76.
2. Sousa-Cavada, B., Richardson, M., Yarwood, A., Père, D. and Rougé, P. (1985) *Phytochemistry* **25**, 115.
3. Yarwood, A., Richardson, M., Sousa-Cavada, B. and Rougé, P. (1985) *FEBS Letters* **184**, 104.
4. Rougé, P., Richardson, M., Yarwood, A., Chatelain, C., Sousa-Cavada, B. and Père, D. (1986) in *Lectins, Biology, Biochemistry, Clinical Biochemistry* (Bøg-Hansen, T. C. and Van Driesche, E., eds) Vol. 5. W. de Gruyter, Berlin (in press).
5. Rougé, P., Richardson, M., Yarwood, A., Ranfaing, P., Sousa-Cavada, B. and Père, D. (1986) Proc. Colloque Lathyrus, Pau, 9–13 Sept 1985 (in press).
6. Campos, F. D. A. P. and Richardson, M. (1984) *FEBS Letters* **167**, 221.
7. Hitz, H., Schafer, D. and Wittmann-Liebold, B. (1977) *Eur. J. Biochem.* **75**, 497.
8. Sletten, K., Kolberg, J. and Michaelsen, T. E. (1983) *FEBS Letters* **156**, 253.
9. Sletten, K. and Kolberg, J. (1983) *Hoppe-Seyler's Z. Physiol. Chem.* **364**, 1047.
10. Foriers, A., de Neve, R., Kanarek, L. and Strosberg, A. D. (1978) *Proc. Natl. Acad. Sci. U.S.A.* **75**, 1136.
11. Richardson, C., Behnke, W. D., Freisheim, J. H. and Blumenthal, K. M. (1978) *Biochim. Biophys. Acta* **537**, 310.
12. Baumann, C. M., Strosberg, A. D. and Rüdiger, H. (1982) *Eur. J. Biochem.* **122**, 105.
13. Hopp, T. P., Hemperly, J. J. and Cunningham, B. A. (1982) *J. Biol. Chem.* **257**, 4473.
14. Gebauer, G., Schiltz, E. and Rüdiger, H. (1981) *Eur. J. Biochem.* **16**, 319.
15. Cunningham, B. A., Wang, J. L., Waxdal, M. J. and Edelman, G. M. (1975) *J. Biol. Chem.* **250**, 1503.
16. 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.
17. Kouchalakos, R. N., Bates, O. J., Bradshaw, R. A. and Hapner, K. D. (1984) *Biochemistry* **23**, 1824.
18. Vodkin, L. O., Rhodes, P. R. and Goldberg, R. B. (1983) *Cell* **34**, 1023.
19. Hoffman, L. M., Ma, Y. and Barker, R. F. (1982) *Nucleic Acids Res.* **10**, 7819.
20. Becker, J. W., Reeke, G. N., Wang, J. L., Cunningham, B. A. and Edelman, G. M. (1975) *J. Biol. Chem.* **250**, 1513.
21. Hardman, K. D. and Ainsworth, C. F. (1976) *Biochemistry* **15**, 1120.
22. Davis, B. J. (1964) *Ann. N.Y. Acad. Sci.* **121**, 404.
23. Reisfeld, R. A., Lewis, U. J. and Williams, D. E. (1962) *Nature* **195**, 281.
24. Wrigley, C. (1968) *Sci. Tools.* **15**, 17.
25. Chang, T. Y., Brauer, D. and Wittman-Liebold, B. (1978) *FEBS Letters* **93**, 205.
26. Richardson, M. (1974) *Biochem. J.* **137**, 101.