The complete amino acid sequences of the β_1- and β_2-subunits of the isolectins LoLl and LoLll from seeds of *Lathyrus ochrus* (L.) DC

A. Yarwood, M. Richardson†, B. Sousa-Cavada+ and P. Rougé*

Department of Botany, University of Durham, Science Laboratories, South Road Durham DH1 3LE, England, +Department of Biochemistry, Federal University of Pelotas, Pelotas (RS) Brazil, and **Laboratoire de Botanique et Biologie Cellulaire, Faculté des Sciences Pharmaceutiques, Université Paul Sabatier, 35 chemin des Maraîchers, 31062 Toulouse, France*

Received 30 January 1985; revised version received 14 March 1985

The complete amino acid sequences of the β_1- and β_2-subunits of the isolectins (LoLl and LoLll) from seeds of *Lathyrus ochrus* were determined by analysis of peptides derived from the proteins by digestion with trypsin, chymotrypsin, pepsin and the *S. aureus* V_8 protease, as well as fragments produced by cleavage with iodosobenzoic acid. Both β-subunits consisted of single polypeptide chains of 181 amino acids, which differed from one another in only 3 positions. The homology of the *Lathyrus ochrus* isolectins with the other two-chain lectins of the tribe Vicieae, and the single-chain lectins of other tribes of the Leguminosae is discussed.

Lathyrus ochrus *isolectin* *β-Subunit* *Amino acid sequence* *Homology* *Legume lectin*

1. INTRODUCTION

The lectins found in seeds of the Vicieae tribe of the Leguminosae are composed of two light (α) chains ($M_r \simeq 6000$) and two heavy (β) chains ($M_r \simeq 17000$). The amino acid sequences of the α-subunits from several genera belonging to this tribe have been determined [1–9] and have indicated the value of such data for chemotaxonomical classification [10] and studies of speciation [11]. At present, however, the potential of this method is somewhat limited by the relatively small number of complete sequences available for the large β-chains. The only sequences of β-chains which have been reported so far are those for *Vicia faba* [12,13], *Pisum sativum* [1,14] and *Lens culinaris* [3,15]. We now report the complete amino acid sequences of the β_1- and β_2-subunits of the isolectins LoL1 and LoL11 from *Lathyrus ochrus* and their homology with the corresponding two-chain lectins from the Vicieae and the single-chain lectins from other tribes of the Leguminosae.

† To whom correspondence should be addressed

2. MATERIALS AND METHODS

2.1. *Isolation of* L. ochrus *lectin (LoL)*

The lectin was isolated from extracts of a crude seed meal by fractional precipitation with $(NH_4)_2SO_4$ and affinity chromatography on Sephadex G-100 as in [8,16].

2.2. *Separation of LoL1 and LoL11 isolectins*

The isolectins LoL1 (p*I* 7.2) and LoL11 (p*I* 6.0) were separated from one another by chromafocusing on columns of PBE94 (Pharmacia AB), as in [8,16].

2.3. *Separation of isolectin and subunits*

The α- and β-subunits of the *L. ochrus* isolectins were separated by gel filtration on a column of

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/85/$3.30

Biogel P-60 equilibrated with 6 M guanidine HCl as in [8,16].

2.4. *Enzyme digestion and separation of peptides*

Samples (6–12 mg) of the β-subunits were digested separately with trypsin, chymotrypsin, pepsin and the protease from *S. aureus* V_8 as in [17]. The mixtures of peptides produced by these methods were initially fractionated on columns (1 × 200 cm) of Biogel P-6 in 0.05 M ammonium bicarbonate (pH 8.1). Further purification of peak fractions was achieved by reverse-phase HPLC on a Micro-pac MCH-10 column (0.4 × 30 cm, Varian) in a Varian Model 5000 HPLC apparatus using a linear gradient of 0–70% acetonitrile (HPLC grade S, Rathburn, Peebles) in 0.1% trifluoroacetic acid. Peptides were detected by measuring the absorbance at 214 nm.

2.5. *Cleavage with iodosobenzoic acid*

Samples (20 mg) of the β-subunits were cleaved at TRP-X peptide bonds by iodosobenzoic acid in 80% acetic acid 4 M guanidine HCl for 21 h at 20°C as in [18]. The resulting fragments were initially fractionated on columns (1 × 200 cm) of Biogel P-30 in 70% formic acid, and subsequently purified by reverse-phase HPLC on a column (0.5 × 25 cm) of Vydac C_{18} (Technicol, Stockport) using a linear gradient of 0–70% acetonitrile in 0.1% trifluoroacetic acid.

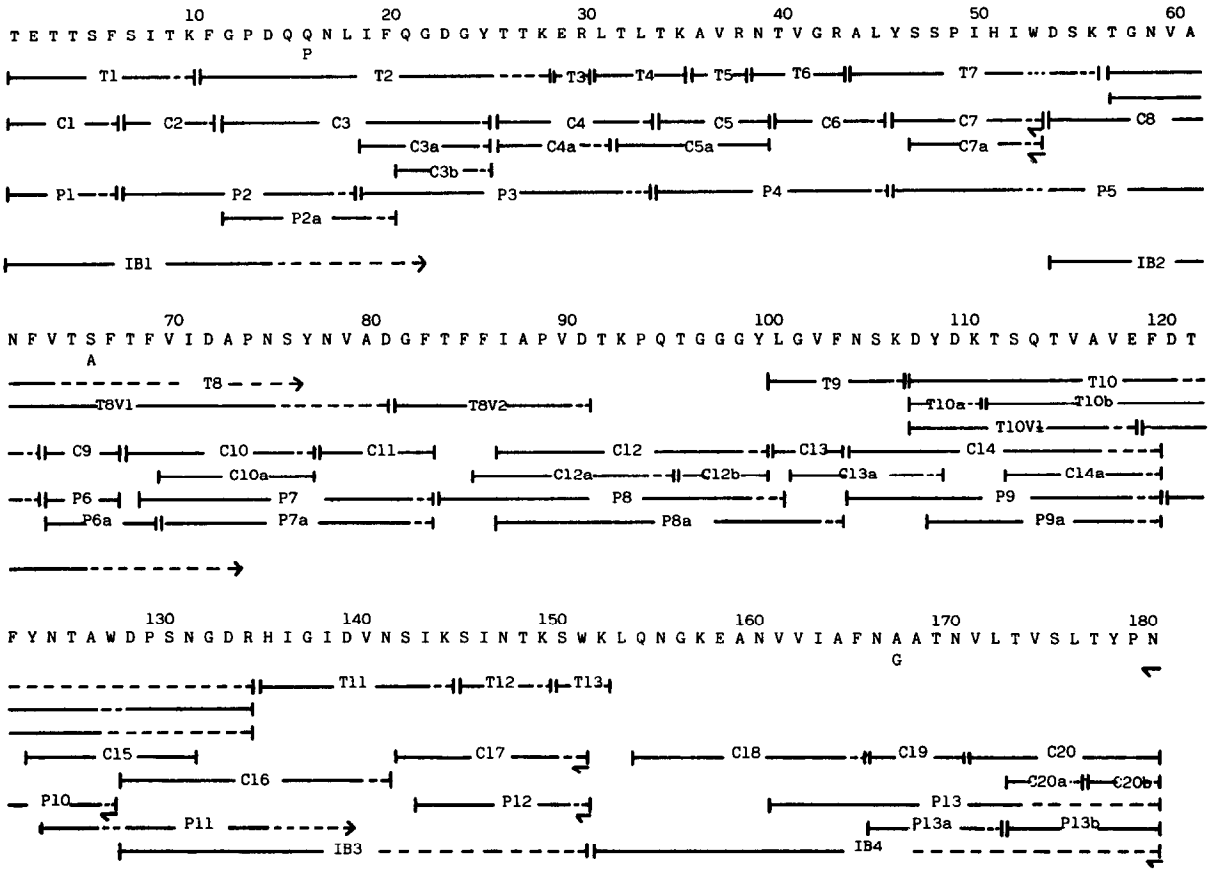

Fig.1. The amino acid sequences of the β_1- and β_2-subunits of the isolectins from seeds of *Lathyrus ochrus* (L.) DC. T_1, tryptic peptides; C, chymotryptic peptides; P, peptic peptides; V, peptides from redigestion with *S. aureus* V_8 protease; IB, fragments from cleavage with iodosobenzoic acid. (——) Regions of peptides sequenced by the DABITC method and/or the dansyl-Edman procedure; (– – –) residues which were not sequenced or yielded unsatisfactory results; (– – →) C-terminus of peptide not known; (←—) residues determined by digestion with carboxypeptidase A.

2.6. *Sequence determination*

The intact subunits, and the fragments or peptides derived from them were subjected to microsequence analysis using the 4-*N*,*N*-dimethylaminoazobenzene-4′-isothiocyanate (DABITC)/phenylisothiocyanate double coupling method, the dansyl-Edman procedure and digestion with carboxypeptidase A, all as in [17]. Amino acid compositions were determined as in [8]. The presence of tryptophan in peptides was detected by staining on paper with *p*-dimethylaminobenzaldehyde.

3. RESULTS AND DISCUSSION

The isolectins LoL1 (p*I* 7.2) and LoL11 (p*I* 6.0) which were readily separated by chromafocusing in the range pH 8.4–5.0 [8,16] each gave only one α- and one β-subunit when subjected to isoelectric focusing in the pH range 8.0–5.0 in PAGE rods containing 8 M urea [16]. The homogeneity of the β-subunits was also confirmed by SDS-PAGE [19]

Table 1

The amino acid compositions of the *Lathyrus ochrus* isolectin subunits β_1 and β_2

	L. ochrus β_1		*L. ochrus* β_2	
	Analysis	Sequence	Analysis	Sequence
Asp	24.2	28	23.5	28
Thr	22.1	23	23.3	23
Ser	14.7	14	13.9	14
Glu	8.1	10	7.2	9
Pro	7.1	7	8.5	8
Gly	13.9	13	15.2	14
Ala	14.5	12	13.2	11
Val	16.1	14	16.7	14
Ile	11.2	11	11.2	11
Leu	9.4	8	8.9	8
Tyr	5.6	7	5.2	7
Phe	14.5	13	13.9	13
Lys	12.1	11	11.9	11
His	2.0	2	2.7	2
Arg	3.8	4	4.0	4
Trp	ND	3	ND	3
Total		181		181

Residues expressed as mol/mol protein. ND, not determined. No corrections were made for decomposition or incomplete hydrolysis

and by N-terminal analysis by the DABITC/PITC double coupling method.

The digestions with trypsin, chymotrypsin and pepsin yielded suitable peptides from which most of the sequences of β_1 and β_2 could readily be deduced (fig.1). Redigestions of the long tryptic peptides T8 and T10 with the *S. aureus* V_8 protease were helpful in determining the sequences in the regions 82–91 and 121–125 where good overlaps were otherwise missing, and the fragments 184 from the cleavage with iodosobenzoic acid clarified the sequences of residues 153–162.

The amino acid sequences of the β_1- and β_2-subunits shown in fig.1 both contained 181 residues and only differed in the amino acids found in positions 16, 66 and 168. The M_r values calculated from the sequences (19 774 for β_1 and 19 713 for β_2) were in good agreement with the estimates of the sizes of the subunits made by SDS-PAGE, and the sequences were compatible with the amino acid compositions (table 1) except for the discrepancies in the values for ASP in both subunits.

No other β-chains of the lectins from the genus *Lathyrus* have been completely sequenced, although partial N-terminal sequences have been determined for both *L. odoratus* [20] and *L. sativus* [21]. On comparing the residues 1–25 of the *L. ochrus* β-chains with these other *Lathyrus* sequences only four amino acids were found to be different. These were Ser^7 (Leu in both *L. odoratus* and *L. sativus*), Ile^8 (Leu in *L. sativus*), Gly^{12} (Ser in *L. odoratus*) and β_2 Pro^{16} (Gln in both).

An alignment of all the available completed sequences of the lectins from the Leguminosae can be made which demonstrates that they are partially homologous throughout their chains (fig.2). However, to maximise the homology, appropriate deletions have to be introduced and certain other rearrangements made. The β- and α-subunits of the two-chain lectins of the Vicieae (*Lathyrus, Pisum, Vicia* and *Lens*) have to be aligned with the N- and C-terminal regions of the single-chain lectins from *Onobrychis viciifolia* [22], *Glycine max* [23] and *Phaseolus vulgaris*, respectively [24]; and the sequences of the lectins from *Canavalia ensiformis* (concanavalin A) [25] and *Dioclea grandiflora* [17] have to be rearranged as circular permutations, such that their residues, 123–237 and 1–69, align with the residues 1–181 (or 183) of the

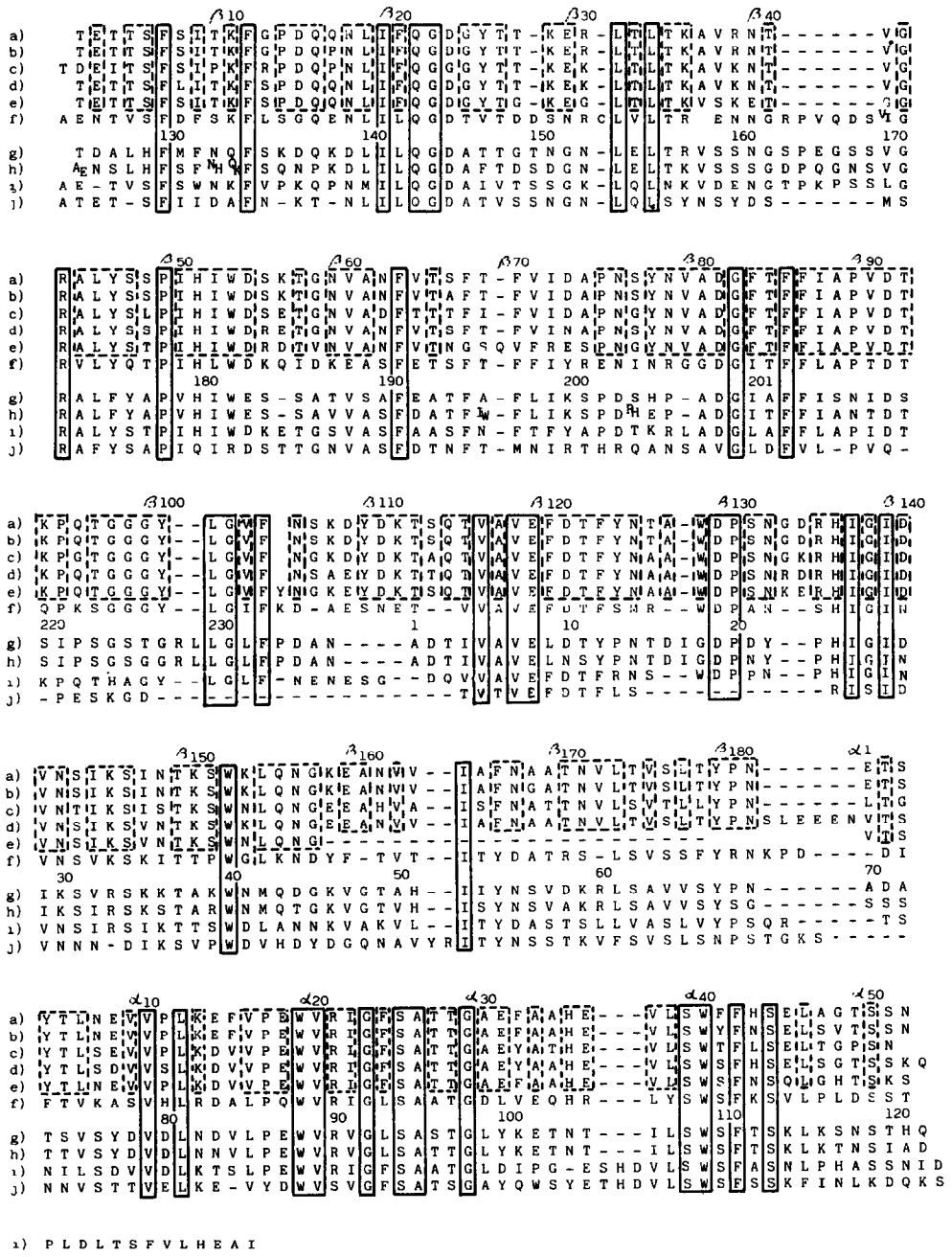

Fig.2. Homology of the amino acid sequences of the *Lathyrus ochrus* isolectins LoL1 ($\beta_1\alpha_1$) (a) and LoL11 ($\beta_2\alpha_2$) (b), with the β- and α-chains of the lectins from *Vicia faba* [12,13] (c), *Pisum sativum* [1,14] (d), Lens culinaris [3,15] (e); and the single-chain lectins from *Onobrychis viciifolia* [22] (f), *Canavalia ensiformis* [25] (g), *Dioclea grandiflora* [17] (h), *Glycine max* [23] (i) and *Phaseolus vulgaris* [24] (j). Only the *Lathyrus ochrus* (β and α) lectins and the *Canavalia ensiformis* (concanavalin A) sequences are numbered. The concanavalin A sequence is circularly permutated to facilitate the alignment. (–) Gaps (deletions) inserted in all sequences to facilitate comparisons; (×) unidentified amino acid; (——) region of sequence not determined. Solid boxes enclose residues which are invariant in all 10 proteins; dashed boxes enclose residues which are invariant in tribe Vicieae.

β-chains of the Vicieae, whilst their residues 70–122 correspond to the α-chains.

When the *L. ochrus* isolectins LoL1 and LoL11 (β- and α-subunits) are compared with the lectins of the other members of the tribe Vicieae, as shown in fig.2, and the percentage of sequence identities is calculated (table 2) there is striking evidence of the high degree of homology within the tribe. The *L. ochrus* isolectins appear to be most similar to the lectin from *Pisum sativum* (85–87% homology), with all members of the tribe showing at least 70% sequence identity. Inspection of fig.2 also reveals that the amino acids found in at least 151 positions are invariant in the tribe Vicieae. The true figure of invariant residues might be as high as 166, but for the fact that the sequence of the *Lens culinaris* β-chain appears to lack some 24 residues at its C-terminus, possibly as a result of a post-translational modification [10].

When the comparison of the lectins (fig.2, table 2) is extended beyond the two-chain lectins of the Vicieae to include the sequences of the single-chain lectins from the tribes Hedysareae (*Onobrychis*), Diocleae (*Canavalia* and *Dioclea*), Phaseoleae (*Phaseolus*) and Glycineae (*Glycine*) the values of relatedness fall to 26–51% for inter-tribe comparisons, but once again the intra-tribe homology shown between the lectins from the two members of tribe Diocleae is above 70%.

The amino acids in 35 positions were found to be conserved (invariant) in all of the lectins compared in fig.2. It is notable that these conserved amino acids include two which correspond with the residues previously identified in concanavalin A [26] as being important in the binding of Ca^{2+} and Mn^{2+}; i.e., Glu^{8} and Asp^{19}. The other important metal-binding residues also tend to be nearly invariant. For example, Asp^{10} is replaced only by an Asn in the homologous position in the *Dioclea* lectin, Asn^{14} by a Ser in *Phaseolus*, His^{24} by an Arg in *Phaseolus*, and Tyr^{12} by Phe in all except the *Dioclea* lectin. Similarly, the amino acids contributing to the three-dimensional structure of the hydrophobic binding cavity of concanavalin A [26] are very highly conserved in the homologous positions in all of the lectins. On the other hand, the residues which constitute the carbohydrate binding site in concanavalin A [26] appear to be very poorly conserved.

A further feature becomes apparent when the

Table 2

Comparison of the sequence identities among homologous legume lectins

	1	2	3	4	5	6	7	8	9	10
L. ochrus ($\alpha_1\beta_1$)		226/253	209/240	190/234	172/235	109/246 110	90/245	92/245 93	129/261	88/245
L. ochrus ($\alpha_2\beta_2$)	96.9		205/240	190/234	171/235	107/246 108	90/245	92/245 93	128/261	88/245
P. sativum ($\alpha\beta$)	87.1	85.4		190/241	174/242	112/250 113	93/251	92/251 93	123/262	87/248
V. faba ($\alpha\beta$)	81.2	81.2	78.8		164/236	105/247 106	90/246	94/246 96	131/265	82/246
L. culinaris ($\alpha\beta$)	73.2	72.8	71.9	69.5		98/245	83/246	88/246 89	112/263	74/248
O. viciifolia	44.3/44.7	43.5/43.9	44.8/45.2	42.5/42.9	40.0		91/248	94/248 97	118/260	65/247
C. ensiformis (ConA)	36.7	36.7	37.2	36.6	33.7	36.7/31.7		184/237 186	99/262	71/250
D. grandiflora	37.5/38.0	37.5/38.0	36.7/37.2	38.2/39.0	35.8/36.2	37.9/39.1	77.6/78.5		102 105/262	65/250
G. max	49.4	49.0	46.9	49.4	42.6	45.4	37.8	38.9/40.1		87/259
P. vulgaris	35.9	35.9	35.1	33.3	29.8	26.3	28.4	26.0	33.6	

Percentage identities (below diagonal) calculated from fig.2 as identities/position compared (above diagonal)

ConA and *Dioclea* lectin sequences are circularly permutated as in the alignment in fig.2. Several workers [2,17,27] have reported that these lectins contain a natural fragmentation point at the peptide bond 118–119. In this alignment the susceptible bond is situated very close to the C-termini of the single-chain lectins and α-chains, and to the N-termini of the single-chains and β-subunits. This juxtaposition is interesting in view of the very recent report that the amino acid sequence of concanavalin A derived by analysis of the DNA complementary (cDNA) to its messenger (mRNA) has neither direct nor circular homology to the other legume lectins, and that in forming mature concanavalin A there must be a transposition and ligation of two polypeptides produced from a precursor polypeptide [28].

ACKNOWLEDGEMENTS

The authors 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, P. Miller and D. Bown for valuable technical assistance and Professor D. Boulter for the provision of certain facilities.

REFERENCES

[1] Richardson, C., Behnke, W.D., Freisheim, J.H. and Blumenthal, K.M. (1978) Biochim. Biophys. Acta 537, 310–319.
[2] Hopp, T.P., Hemperly, J.J. and Cunningham, B.A. (1982) J. Biol. Chem. 257, 4473–4483.
[3] Foriers, A., De Neve, R., Kanarek, L. and Strosberg, A.D. (1978) Proc. Natl. Acad. Sci. USA 75, 1136–1139.
[4] Gebauer, G., Schiltz, E. and Rüdiger, H. (1981) Eur. J. Biochem. 113, 319–325.
[5] Baumann, C.M., Strosberg, A.D. and Rüdiger, H. (1982) Eur. J. Biochem. 122, 105–110.
[6] Sletten, K., Kolberg, J. and Michaelsen, T.E. (1983) FEBS Lett. 156, 253–256.
[7] Sletten, K. and Kolberg, J. (1983) Hoppe-Seyler's Z. Physiol. Chem. 364, 1047–1051.
[8] Richardson, M., Rougé, P., Sousa-Cavada, B. and Yarwood, A. (1984) FEBS Lett. 175, 76–81.
[9] Sousa-Cavada, B., Richardson, M., Yarwood, A., Père, D. and Rougé, P. (1985) Phytochemistry, in press.
[10] Strosberg, A.D., Lauwereys, M. and Foriers, A. (1983) in: Chemical Taxonomy, Molecular Biology and Function of Plant Lectins (Goldstein, I.J. and Etzler, M.E. eds) pp.7–20, A.R. Liss Inc., New York.
[11] Rougé, P. (1984) Biochem. Syst. Ecol. 12, 47–51.
[12] Cunningham, B.A., Hemperly, J.J., Hopp, T.P. and Edelman, G.M. (1979) Proc. Natl. Acad. Sci. USA 76, 3218–3222.
[13] Hemperly, J.J., Hopp, T.P., Becker, J.W. and Cunningham, B.A. (1979) J. Biol. Chem. 254, 6803–6810.
[14] Higgins, T.J.V., Chandler, P.M., Zurawski, G., Button, S.C. and Spencer, D. (1983) J. Biol. Chem. 258, 9544–9549.
[15] Foriers, A., Lebrun, E., Van Rapenbusch, R., De Neve, R. and Strosberg, A.D. (1981) J. Biol. Chem. 256, 5550–5560.
[16] Rougé, P. and Sousa-Cavada, B. (1985) Plant Sci. Lett., in press.
[17] Richardson, W., Campos, F.D.A.P., Moreira, R.A., Ainouz, I.L., Begbie, R., Watt, W.B. and Pusztai, A. (1984) Eur. J. Biochem. 144, 101–111.
[18] Mahoney, W.C., Smith, P.K. and Hermodson, M.A. (1981) Biochemistry 20, 443–448.
[19] Swank, R.T. and Munkres, K. (1971) Anal. Biochem. 39, 462–477.
[20] Kolberg, J., Michaelsen, T.E. and Sletten, K. (1980) FEBS Lett. 117, 281–283.
[21] Kolberg, J. and Sletten, K. (1982) Biochim. Biophys. Acta 704, 26–30.
[22] Kouchalakos, R.N., Bates, O.J., Bradshaw, R.A. and Hapner, K.D. (1984) Biochemistry 23, 1824–1830.
[23] Vodkin, L.O., Rhodes, P.R. and Goldberg, R.B. (1983) Cell 34, 1023–1031.
[24] Hoffman, L.M., Ma, Y. and Barker, R.F. (1982) Nucleic Acids Res. 10, 7819–7828.
[25] Cunningham, B.A., Wang, J.L., Waxdal, M.J. and Edelman, G.M. (1975) J. Biol. Chem. 250, 1503–1512.
[26] Becker, J.W., Reeke, G.N., Wang, J.L., Cunningham, B.A. and Edelman, G.M. (1975) J. Biol. Chem. 250, 1513–1524.
[27] Abe, Y., Iwabuchi, M. and Ishii, S.I. (1971) Biochem. Biophys. Res. Commun. 45, 1271–1277.
[28] Carrington, D.M., Auffret, A. and Hanke, D.E. (1985) Nature 313, 64–67.