

Polymorphism of legumin subunits from field bean (*Vicia faba* L. var. *minor*) and its relation to the corresponding multigene family

C. Horstmann¹, B. Schlesier¹, A. Otto², S. Kostka², K. Müntz¹

¹ Institut für Pflanzengenetik und Kulturpflanzenforschung, Corrensstrasse 3, D-06466 Gatersleben, Germany
² Zentralinstitut für Molekularbiologie, D-13122 Berlin-Buch, Germany

Received: 24 April 1992 / Accepted: 22 December 1992

Abstract. Legumin, which amounts to approximately 55% of the seed protein in field beans (Vicia faba L. var. *minor*), is a representative of the 12S storage globulin family. The 12S storage globulins are hexameric holoprotein molecules composed of different types of polymorphic subunits encoded by a multigene family. 'Type-A' legumin subunits contain methionine whereas 'type-B' are methionine-free subunits. Sequencing of two different type A-specific cDNAs, as well as an FPLC/HPLC-based improvement of subunit fractionation and peptide mapping with subsequent partial amino-acid sequencing, permit the assignment of some of the polymorphic legumin subunits to members of the multigene family. Two different type A subunits (A1 and A2) correspond to the two different cDNA clones pVfLa129 (A2) and 165 (A1), but microheterogeneity in the amino-acid sequences indicates that polymorphic variants of both representatives of this type may exist. Two groups of published type B-specific gene sequences (LeB7, and LeB2, LeB4, LeB6, respectively) are represented by two polymorphic subunit fractions (B3I, B3II, and B4I, B4II). A seventh clone, LeB3, encodes one of the large legumin subunits that is only a minor component of the legumin seed protein complex.

Key words: cDNA – Legumin subunits – Polymorphism – Gene assignment – *Vicia faba*

Communicated by G. Wenzel Correspondence to: K. Müntz

Introduction

Legumin-like 12S globulins represent one of the major storage protein classes in the plant kingdom and have been demonstrated in the seeds of many mono- and di-cotyledonous plants and of the gymnosperm *Ginkgo biloba* (Jensen and Berthold 1989), as well as in the spores of the fern *Matteucia* (Templeman et al. 1987).

The 12S globulins are hexameric proteins with molecular masses between 300 kDa and 400 kDa. Each monomer (or subunit) consists of two polypeptide chains which are linked by at least one disulphide bond (for reviews see Casey and Domoney 1984; Nielsen 1984: Casey et al. 1985). The constituent polypeptide chains are referred to as A- and B-chains in the case of glycinin from soybean, and α - and β -chains in case of 12S globulins of most other plants, or - because of their isoelectric points - as acidic and basic chains. The molecular masses are about 35-40 kDa for the α -, and 18–20 kDa for the β -chains. Biosynthetic studies (Croy et al. 1980; Spencer and Higgins 1980) and DNA sequencing (Croy et al. 1982) indicate that 12S globulin subunits are synthesised as precursor molecules comprising both α and β sequences covalently linked by a peptide bond. The precursors form trimers which are proposed to represent the form in which the leguminlike globulins are transported to the protein bodies. where the final proteolytic processing generates the 'mature' α - and β -chains which assemble into the hexameric 12S form (Chrispeels et al. 1982; Akazawa and Hara-Nishimura 1985). This mechanism seems to be universal for 12S globulins from seeds of different plant families.

The structure, formation, and deposition of field bean legumin correspond to the general principles described above for 12S storage globulins (Müntz et al.

1986). The subunits of field bean legumin exhibit heterogeneity in electrophoresis (Matta et al. 1981; Tucci et al. 1991) and ion-exchange chromatography (Horstmann 1983, 1984). Four major subunits (nos. 1-4) with relative molecular masses of about 60 k have been isolated from the legumin of Vicia faba. The occurrence of two minor legumin subunits with relative molecular masses of 75 k (no. 5), and 80 k (no. 6) was also demonstrated (Horstmann 1983, 1984). According to amino-acid analysis, peptide mapping (Horstmann 1983), and partial sequence analysis (Otto et al. 1984), subunits 1 and 2 belong to a methionine-containing ('type A') and subunits 3 and 4 to a methionine-free ('type B') variant of legumin subunits, encoded by a corresponding multigene family. The question arises of how the polymorphic subunits of field bean legumin correspond to members of its 12S multigene family. The present paper extends our knowledge on the polymorphism of legumin subunits and assigns several of these subunits to members of the 12S globulin-specific multigene family of field beans (V. faba L. var. minor).

Materials and methods

Synthesis and sequencing of legumin A-specific cDNA

cDNA was synthesised on poly $(A)^+$ RNA using the Amersham cDNA-Synthesis-System according to the supplier's instructions. The C-tailed cDNA was annealed to the plasmid pUC18 which was G-tailed at the *PstI* site. Competent HB101 cells (BRL) were transformed with the plasmid.

Legumin A-specific clones were identified by colony hybridisation with the legumin A-type clones pVfc30, pVfc53, and pVfc77 representing the C-terminal end of the α -chain and the nearly complete β -chain coding regions (Wobus et al. 1986).

Nucleotide sequences were determined by a solid phase variant of the chemical degradation method (Rosenthal et al. 1986) and partially by the dideoxy sequencing procedure (Sanger and Coulson 1978).

Chromatographic methods

HPLC and FPLC analyses were performed using a HPLC system with two analytical pumps, a 50 B programmer and a variable wavelength spectral photometer (all parts from Dr. Herbert Knauer KG, Bad Homburg, Germany).

Separation of legumin subunits

Fifty Milligrams of legumin, isolated from mature seeds of *V. faba* L. var. *minor* cv 'Fribo' as described by Horstmann (1983), were dissolved in 5 ml of buffer A [0.015 mol/1 Tris (hydroxy)-methylaminomethan (Tris), 6 mol/1 urea, adjusted to pH 7.5 with diluted sulphuric acid] and applied to a MonoQ HR 10/10 anion exchange column (Pharmacia, Uppsala, Sweden), equilibrated with buffer A. The subunits were eluted with a linear gradient (72 min) from buffer A to buffer B (0.15 mol/l sodium sulphate in buffer A) at a flow rate of 2 ml/min. Appropriate fractions (Fig. 1) were dialysed against distilled water at 4 °C and freeze-dried.



Fig. 1. Separation of legumin subunits by anion exchange chromatography on MonoQ HR 10/10. SDS-PAGE of the main fractions is shown in the upper part of the figure

Separation of α - and β -chains

Between 10 and 50 mg of the appropriate legumin subunit were dissolved in 0.5 to 3 ml of buffer A. For reduction of the disulphide bonds, dithiothreitol (3/4 of the subunit weight) was added and after standing 1 h at room temperature the mixture was injected into a MonoQ HR 10/10 column. The β -polypeptides (together with dithiothreitol and its oxidation product) were eluted with buffer A, the α -polypeptides were eluted by applying a short (10 min) gradient from buffer A to buffer B, flow rate 2 ml/min. The fractions containing the polypeptides were dialysed against distilled water at 4 °C and freeze-dried.

Peptide mapping of α - and β -polypeptides

Digestion of the polypeptides with trypsin. One milligram of polypeptide was dissolved or suspended in 1 ml 0.2 mol/l NH₄ HCO₃, pH 8.5. Ten microliters of a trypsin solution (1 mg trypsin, TPCK-treated, Serva, Heidelberg, Germany) in 1 ml 0.001 N HCl were added after 0, 1, and 2 h to the polypeptide solution, the mixture was incubated for a total time of 5 h at 37 °C; thereafter the reaction mixture was lyophilised in a vacuum centrifuge.

HPLC of the tryptic peptides. Buffer 1: $0.05 \text{ mol/l } \text{KH}_2 \text{ PO}_4$, adjusted to pH 2.9 with diluted phosphoric acid. Buffer 2: 40 ml buffer 1 + 60 ml acetonitrile (special grade for UV-spectroscopy, PCK Schwedt, East Germany).

HPLC was done on a column, $250 \text{ mm} \times 4 \text{ mm}$, of LiChrosorb RP18, $5 \mu \text{m}$ [(E. Merck, Darmstadt, Germany) prepacked by Dr. Herbert Knauer KG (Bad Homburg, Germany)], equilibrated with buffer 1. Tryptic peptides, obtained from the step above, were dissolved in 0.3 ml 0.1% trifluoro-acetic acid, and 0.15 ml was injected into the column.

Elution scheme: 0 min to 2 min from 100% buffer 1/0% buffer 2 to 93% buffer 1/7% buffer 2, 2 min to 144 min from 93% buffer 1/7% buffer 2 to 35% buffer 1/65% buffer 2; flow rate was 1 ml/min, detection wavelength 215 nm.

For the semi-preparative isolation of selected peptides the appropriate fractions of several runs, corresponding to about 10 mg α -polypeptide of subunit 3-I (α -3-I) or subunit 4-II (α -4-II), respectively, of the tryptic peptides, were collected and lyophilised. The residues were dissolved in 0.1% trifluoro-acetic acid, and, for desalting, applied to the LiChrospher RP18 column equilibrated with 0.1% trifluoro-acetic acid. After elution of the salts the peptides were recovered by applying a 5-min gradient from 0.1% trifluoro-acetic acid to 0.1% trifluoro-acetic acid/60% acetonitrile and lyophilization.

Sequence analysis of the tryptic peptides

Sequence analysis was performed on a gas-liquid-phase protein sequencer (Model 477A, Applied Biosystems) equipped with the Model 120A on-line HPLC system for the separation of the phenylthiohydantoin (PTH) amino-acid derivatives of each Edman degradation cycle. Approximately 0.2–3 nmol-aliquots of the isolated and desalted tryptic peptides were loaded onto a polybrene-coated precycled glass-fibre filter located on the sequencer cartridge. All operations, including those related to the identification of the PTH amino acids, were carried out as described in the manual for the sequencer.

Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE was carried out at room temperature in a slab gel $(15 \text{ cm} \times 15 \text{ cm} \times 0.8 \text{ mm})$ with an 18% running gel according to Casey (1979) and the Tris/borate buffer system after Neville (1971). The electrophoresis was run at a current of 24 mA and a maximal voltage of 250 V.

Results

Amino-acid sequences of legumin subunits from V. faba var. minor

The first complete sequence of a genomic DNA coding for a legumin subunit of V. faba var. minor was published by Bäumlein et al. (1986). Four more apparently functional genes and two pseudogenes, all coding for B-type subunits have been sequenced (Heim et al. 1989). No genomic clones for type A subunits could be isolated up to now, but several cDNA sequences for subunits of this type were elucidated: a partial cDNA sequence comprising the coding region of the C-terminal part of the α -chain and the almost complete β -chain of a type A legumin subunit (Wobus et al. 1986), one cDNA sequence for a complete type A subunit with signal sequence, and another cDNA sequence lacking the information for the signal sequence and the first nine amino acids of the α -chain (Schlesier et al. 1990).

Figure 2 shows the type A amino-acid sequences derived from the cDNA clones pVfLA129 and pVfLA165 (Schlesier et al. 1990). It is obvious from this figure that these sequences are very similar to each other, the main difference being the insertion of 23 amino acids into LA129 leading to a double repeat of a stretch of 11 amino acids which is present in both subunits at positions 163-174 downstream from the propolypeptide N-terminus. The other differences are mainly conservative amino-acid exchanges. A comparison of the corresponding parts of these type A sequences with the partial type A sequence published by Wobus et al. (1986) reveals that this sequence differs in a few positions from LA129 as well as from LA165, probably reflecting additional polymorphism of the type A subunits.

Separation of legumin subunits

The use of FPLC strongly improved the separation of the type B legumin subunits (Lambert et al. 1987). Subunit no. 3, as well as no. 4, could be further fractionated on a MonoQ anion exchanger into at least two main fractions each, named 3-I, 3-II and 4-I, 4-II, respectively (see Fig. 1). Shoulders in this 'B' part of the chromatograms indicate that some minor members are additionally present in this group. All these subunits showed patterns and positions similar to α - and β polypeptides in SDS-PAGE after reduction of the disulphide bonds, only the α -chain of 4-I showed two bands. A much better separation was also achieved for the high-molecular-mass group of subunits 5 and 6 as compared to the chromatographic method previously described (Horstmann 1983). A further subunit of this group was also detected (Fig. 1). On the other hand, the separation of the type A subunits 1 and 2 was comparable to the one obtained on DEAE-Sepharose CL-6B (Horstmann 1983); changing the chromatographic conditions did not lead to an improvement of the separation of these two subunits.

Peptide mapping of type B legumin subunits and the sequences of selected tryptic peptides

The α - and β -chains of the main type B subunits, 3-I, 3-II, 4-I, and 4-II, were fractionated by ion exchange chromatography, digested with trypsin, and the tryptic peptides mapped by means of HPLC. The peptide maps of all the β -chains closely resembled each other. Among the digestion products of the α -chains two peaks for each chain were found which had slightly

870				
-----	--	--	--	--

A1	< No	t determine	d>_			
A2	MAKLLALSLSL	CFLLFSSCFA	LREQSQQNEC	QLERLDALEP	DNRIESEGGL	IETWNPNNRQ
	-20 Signal		*	0 2	0 3	• 40 ×
	<	———→				
A1	R					
A2	FRCASVALSR	ATLQRNALRR	PYYSNAPQEI	YIQQGNGYFG	MVFPSCPETF	EEPOOSEOGE
	* 5	* 0 6	0 7	* 0 8	0 9	0 100
A1	_R	Q		I	T	
A2	GGRYRDSHQK *	VNRFREGDII	AVPTGIVFWM *	YNDQDTPVIA *	ISLTDIGSSN *	NQLDQMPRRF *
	11	0 12	0 13	0 14	0 150	0 160
A1	GNQEQEF	LRYQHQQGVK	EEQDND	R		N
A2	YLA		GNQE	QEFLRYQHQQ	GGKEEQDNDG	NNIFSGFKRD
				* 170	* 180	* 190
			_			
Al	F	<u> </u>	R			
A2	FLEDALNVNR *	HIVDRLQGRN *	EDEEKGAIVK *	VKGGLSIITP *	PERQARHPRG *	SRQEEDEDED
	200	210	220	230	240	250
			C	x-chain ← -	→ ß-chain	
A1	_DE		<u> </u>		·	
A2	EKEERQPS *	HHKSRRGEDE	DDKEKRHSQK	GESRRHGDN/C	G LEETVCTAKI	RLNIGSSSSP
	260	270	280	290	300	310
A1		v				
A2	DIYNPQAGRI	KTVTSLDLPV	LRWLKLSAEH	GSLRKNAMFV	PHYNLNANSI	LYALKGRARL
	* 320	* 330	* 340	* 350	* 360	* 370
A1		G	V			
A2	QVVNCNGNTV *	FDEELEAGRA	LTVPQNYAVA *	AKSLSDRFTY	VAFKTNDRAG	IARLAGTSSV
	380	390	400	410	420	430
A1	LL.		F SR	EK	ISSSTTRV	
A2	INDMPVDVVA	ATFNLERNEA	RQLKSNNPFK	FLVPPRESQ-	K F	ASA
	* 440	* 450	* 460	* ~ 470		
				• • •		

Fig. 2. Amino-acid sequences of A-type subunits derived from cDNA clones LA165 (A1) and LA129 (A2). For A1 only the amino acids differing from A2 are printed. Broken line. gaps. A horizontal single and double line above A1 indicates the sequence repeat in A1. The processing site between the α - and β -chains is indicated

different retention times in the group 3 and in the group 4 peptide maps, respectively; within each group no differences were observed in the maps. These peptides were isolated and sequenced; for each peptide the complete amino-acid sequence could be identified in spite of difficult clusters of glutamic acid. The peptides with shorter retention times had 18, and the ones with longer retention times 20, amino acids. The first 18 amino acids of the longer peptides were identical to the 18 amino acids of the shorter ones within each group; the two additional amino acids at the C-terminus of the longer peptides resulted from an incomplete tryptic cleavage at a Lys, which forms the C-terminal amino acid of the shorter peptide. The sequences of the longer peptides are shown in Fig. 3, arrows indicate the points where only partial cleavage was observed. The comparable peptides of the two groups differ in only one amino acid (underlined in Fig. 3): amino acid no. 13 is a glutamine in the peptides of group 3, but a glutamic acid in the peptides of group 4.

Discussion

Amino-acid sequences of legumin subunits from V. faba var. minor

Figure 3 shows a comparison of the sequences derived from the genomic DNA clones LeB4 (Bäumlein et al. 1986), LeB6, and LeB7 (Heim et al. 1989). The sequences of LeB2, LeB6, and LeB7 are not complete, each lacking the signal sequence and the first 127 amino acids of the α -chain (as compared to LeB4). The amino-acid sequence derived from the genomic DNA



Fig. 3. Schematic representation of the type B subunits derived from genomic clones LeB4, LeB6, and LeB7. The *triangle* indicates the deletion of six amino acids in LeB6. *Thick vertical line*, processing site between the α - and β -chains; *thin vertical line*, different amino acids at this position for group 3 and group 4 subunits. *Broken line*, sequence parts not determined for LeB6 and LeB7. The sequences of the differing tryptic peptides are shown; *arrows*, site of partial tryptic cleavage (see text)

clone LeB2 (Heim et al. 1989) is not included in Fig. 3 since it is equivalent to the comparable part of LeB4; the few nucleotide exchanges at the DNA level do not result in amino-acid exchanges. Comparison of the N-terminal amino-acid sequence determined for the β -chain of the type B subunit 3 (Otto et al. 1984) with the respective structures of the DNA-derived LeB amino-acid sequences shows that these sequence parts are in all cases completely identical and therefore, that, these clones are specific for type B sequences. The similarity within the B group in the comparable parts of the four DNA-derived type B sequences is much higher than within the A group. LeB4 and LeB7 differ at only two amino-acid positions, one close to the C-terminus of the α - (sequence part printed in Fig. 3), and the other near the C-terminus of the β -chain (thin vertical line, Fig. 3). A deletion of 18 nucleotides in LeB6 leads to a lack of six amino acids (nos. 204-209) in the corresponding α part.

The sequencing of the cDNA clone pVfLa129 and the derived type A amino-acid sequence for the first time present a comparison of type A and type B subunits (Fig. 4). The N-terminal amino acids of isolated α -chains of both complete types are known from amino-acid sequencing (Horstmann 1984), therefore the first 21 amino acids of LA129, and the first 22 amino acids of LeB4, could be assigned as signal sequences. It is remarkable that the type A sequences differ to a considerable degree (more than 50%) from the type B sequences, a fact that we had already noticed after partial amino-acid sequencing of isolated α - and β -chains of V. faba legumin (Horstmann 1984; Otto et al. 1984) and which was confirmed by the results of previous partial cDNA sequencing (Wobus et al. 1986).

Heim et al. (1989) identified a further genomic DNA clone, LeB3, that apparently possesses a sequence coding for one of the high-molecular-mass legumin subunits (molecular mass 75-80 kDa) we found in the legumin of *V. faba* (Horstmann 1983, 1984). A cDNA sequence for an 80-kDa subunit of *Pisum* legumin has already been described by Domoney and Casey (1984).

Assignment of some of the polymorphic legumin subunits to members of the multigene family

The number of genes coding for subunits of legumin from *Vicia faba* was estimated to be approximately 12 for the B-, and "less than 10" for the A-type (Heim et al. 1989). These numbers may comprise genes that, in spite of different nucleotide sequences, code for subunits with identical amino-acid sequences (i.e., genes LeB2 and LeB4), as well as non-functional pseudogenes which cannot code for polypeptides. Nevertheless, this high number of genes coding for the two main subunit types indicates that the heterogeneity of the isolated legumin subunits might not be due to further processing steps or to artifacts arising during their isolation, but rather reflects the heterogeneity of the legumin gene family.

(1) A-type sequences. Cyanogen bromide-cleavage of the isolated polypeptide chains from subunits 1 and 2 indicated that β -chain 1 contains only one Met, but β -chain 2 contains two Met, whereas for the α -chains the CNBr cleavage patterns reveal no differences in Met content (Horstmann 1984). Inspection of the type A sequences LA129 and LA165 shows that LA165 contains only one Met residue in the β -chain region. LA129 possesses an additional Met at position 41 in front of the C-terminus. The sizes of the cyanogen bromide-cleavage products of the β -chains (Horstmann 1984) correspond to the sizes that can be calculated from the amino-acid sequences. Therefore, subunit 1 could be related to LA165, and subunit 2 to LA129, though the sequences are not fully identical with the results of amino-acid sequencing: amino acid no. 331 is Leu in the β chain of LA129, and Val at the corresponding position of LA165, while for the respective position of subunit 1 of the isolated β -chain a Leu is found (Otto et al. 1984). This points to a relationship with LA129. The absence or presence of an additional Met can be regarded as a marker for the relationship of LA165 to subunit 1, and of LA129 to subunit 2. These differences, as well as those found by a comparison of

872

	-20	-10		10	20	30	
A2	MA KLLALSL	SLC FLLFSS-	CFA -LREQSQ	QNE COLERLD.	ALE PDNRIES	EGG LIETWNPNNR	
в4	MS KPFLSLL	SLS LLLFTST	CLA TSSEFDR	LNQ CRLDNIN	ALE PDHRVES	EAG LTETWNPNHP	
	-20	-10		* 10	20	30 4 0	
Signal sequences							
	40	50	60	70	80	90	
A2	QFRCASVALS	RATLQRNALR	RPYYSNAPQE	IYIQQGNGYF	* GMVFPSCPET	* FEEPQQSEQG	
В4	ELRCAGVSLI	RRTIDPNGLH	LPSYSPSPQL	IYIIQGKGVI	+ + ++ + GLTLPGCPQT	YQEPRSSQSR	
	50	0 6	ד כ זי	о 8	D 90	0 100	
	100	110	120	130	140	150	
22	* EGGRYRDS	HOKUNDEDEC	* DTTAVPTCTV	* FWMVNDODTP			
B/		+++ +++ +	++++ + ++	+ ++ + + +	++++ +		
04	2037220FD3 *	10001000000000000000000000000000000000	DIIAIPBGIF * 1 12		LVAISLUUIS	*	
	110	J 120	J 13(J 14	J 150	0 100	
	160 *	170			180 *	190 *	
A2	RRFYLAGNQE	QEFLRY		QHQ	QGGKEEQDND	GNNIFSGFKR	
В4	RVFYLGGNPE	VEFPETQEEQ	QERHQQKHSL	PVGRRGGQHQ	QEEESEEQKD	GNSVLSGFSS	
	170	D 180	D 190	200	210	0 220	
	200	210	220	230	240	250	
A2	* DFLEDALNVN	* RHIVDRLQGR	* NEDEEKGAIV	* KVKGGLSIIT	* PPERQARHPR	* GSRQEEDEDE	
в4	++ + EFLAQTFNTE	++ + EDTAKRL-~R	SPRDKRNQIV	+ +++ ++ RVEGGLRIIN	+ + PEGQQ	++ + + EEEEEEE	
	230	C	* 240	250	260	270	
	260	270	290	200	200	210	
	200	270 *	280	290 *	*	*	
A2	THE	HKSRRGEDE I	JDKEKRHSQKG	ESRRHGDNGL + +++	EETVCTAKLR	LNIGSSSSPD ++ +	
84	EEKQRSE	Q ·		GRNGL	SETICSLKIR	ENIAQPARAD	
				280	290	300	
	320	330	340	350	360	370	
A2	IYNPQAGRIK	TVTSLDLPVL	RWLKLSAEHG	SLRKNAMFVP	HYNLNANSIL	YALKGRARLQ	
В4	LYNPRAGSIS	TANSLTLPIL	RYLRLSAEYV	RLYRNGIYAP	HWNINANSLL	YVIRGEGRVR	
	310	320	330	340	350	360	
	380	390	400	410	420	430	
A2	* VVNCNGNTVF	* DEELEAGRAL	* TVPQNYAVAA	* KSLS~DRFTY	* VAFKTNDRAG	* IARLAGTSSV	
B4	++ ++ ++ IVNSQGNAVF	+ + DNKVRKGQLV	++++ ++ VVPQNFVVAE	+ QAGEEEGLEY	+++++++ LVFKTNDRAA	VSHVQQVFRA	
	* 370	380	390	* 400	410	420	
	440	450	460	170			
	440 *	45U *	400 *	4/U *	DACA		
A2	INDMPVDVVA + ++ +	ATFNLERNEA + +	RQLKSNNPFK ++	TLVPPRESQK	KASA		
B4	TPADVLA	NAFGLKQKQV	TELKLSGNRG	FLAHLÖ-2Ö2	100 *		
	430 440 450 460						

pVfLa129- and pVfLa165-derived amino-acid sequences with previously published partial cDNA sequences (Wobus et al. 1986), can be taken as indications of microheterogeneities in groups of polymorphic subunits corresponding to A1 and A2, respectively.

(2) B-type sequences. Comparison of the N-terminal peptide sequences from β -chains of subunits 3 and 4 (Otto et al. 1984) and thin-layer chromatograms of tryptic peptides from the α -chains of subunits 3 and 4

Fig. 4. Comparison of the amino-acid sequences derived from cDNA clone LA129 (A2) (Schlesier et al. 1990) and genomic DNA clone LeB4 (B4) (Bäumlein et al. 1986)

(Horstmann 1983) do not reveal any differences. At least two peptides differing in retention times for the α -chains of subunits 3 and 4, respectively, were found among their tryptic peptides by means of reversed phase HPLC. Comparable peptides belong to a part of the sequences close to the C-terminus of the α -chain, the longer peptides end only six amino acids in front of the cleavage site between α and β . They differ only in amino acid no. 13, with a glutamine in the peptide of group 3, and a glutamic acid in the peptide of group 4.



Fig. 5. Relationships between DNA sequences and isolated subunits of legumin from *Vicia faba*

Only the DNA-derived amino-acid sequences for LeB7 would give a corresponding tryptic peptide containing Gln. All the other sequences – LeB2, LeB4, and LeB6 – possess Glu at this site, as shown for LeB4 and LeB6 in Fig. 3. Therefore, such an exchange of Glu for Gln is not due to a deamidation of this Gln during the isolation or purification procedures, and these subunits can be assigned to the genomic DNA clones LeB2, LeB4, and LeB6.

A summary of the relationships between DNA sequences and isolated subunits of V. faba legumin, showing the different groups of the legumin subunits, is illustrated in Fig. 5. These results demonstrate that the diversity of the legumin gene family is also found in the gene products, the legumin subunits. A similar differentiation in several classes is also known for the subunits of 12S globulins from the seeds of other legumes, e.g., in Pisum legumin A (Lycett et al. 1984) (type A), legumins K and J (Gatehouse et al. 1988) (type B), and the high-molecular-mass legumin polypeptide for which a cDNA has been isolated (Domoney and Casey 1984). or the group I (type A) and group II (type B) glycinins from soybean (Nielsen 1984; Nielsen et al. 1989). It is interesting to note that in the case of pea legumin the most abundant subunit is of the A-type, the B-type subunits K and J are present in only minor amounts, though the DNA-derived amino-acid sequences of both types are very similar to the respective sequences of V. faba legumin subunits.

Sequence data for 12S globulin families from other plant species are still rare. Two different amino-acid sequences derived from two cDNA clones (clones 94 and 134) specific for the 12S globulins of cottonseed (Chlan et al. 1986) seem also to belong to two different 12S globulin families, but an unambiguous assignment of the sequences to one of the two legumin families is not possible. Each of the two sequences possesses about the same similarity score for either type A or type B sequences. Different cDNA-derived amino-acid sequences have also been described for the legumin-like glutelin from *Oryza sativa* (Masumura et al. 1989; Takaiwa et al. 1989). However, in this case the differences are small as compared to those found for the two main types of legumin subunits in *V. faba*.

References

- Akazawa T, Hara-Nishimura I (1985) Topographic aspects of biosynthesis, extracellular secretion, and intracellular storage of proteins in plant cells. Annu Rev Plant Physiol 36:441–472
- Bäumlein H, Wobus U, Pustell J, Kafatos FC (1986) The legumin gene family: structure of a B-type gene of Vicia faba and a possible legumin gene-specific regulatory element. Nucleic Acids Res 6:2707-2720
- Casey R (1979) Immunoaffinity chromatography as a means of purifying legumin from *Pisum* (pea) seeds. Biochem J 177:509-520
- Casey R, Domoney C (1984) The genetics of legume storage proteins. Phil Trans R Soc Lond B304:349-358
- Casey R, Domoney C, Ellis N (1985) Legume storage proteins and their genes. Oxford Surveys Plant Mol Cell Biol 3:1-95
- Chlan CA, Pyle JB, Legocki AB, Dure III L (1986) Developmental biochemistry of cottonseed embryogenesis and germination. XVII. cDNA and amino-acid sequences of members of the storage protein families. Plant Mol Biol 7:475-489
- Chrispeels MJ, Higgins TJV, Spencer D (1982) Assembly of storage protein oligomers in the endoplasmic reticulum and processing of the polypeptides in the protein bodies of developing pea cotyledons. J Cell Biol 93:306–313
- Croy RRD, Gatehouse JA, Evans IM, Boulter D (1980) Characterisation of the storage protein subunits synthesized in vitro by polyribosomes and RNA from developing pea (*Pisum* sativum L.). I. Legumin. Planta 148:49-56

- Croy RRD, Lycett GW, Gatehouse JA, Yarwood JN, Boulter D (1982) Cloning and analysis of cDNAs encoding plant storage protein precursors. Nature 295:76–79
- Domoney C, Casey R (1984) Storage protein precursor polypeptides in cotyledons of *Pisum sativum* L. Identification, and isolation of a cDNA clone for an 80000-Mr legumin-related polypeptide. Eur J Biochem 139:321-327
- Gatehouse JA, Bown D, Gilroy J, Levasseur M, Castleton J, Ellis THN (1988) Two genes encoding 'minor' legumin polypeptides in pea (*Pisum sativum L.*). Characterization and complete sequence of the LegJ gene. Biochem J 250:15-24
- Heim U, Schubert R, Bäumlein H, Wobus U (1989) The legumin gene family: structure and evolutionary implications of *Vicia* faba B-type genes and pseudogenes. Plant Mol Biol 13:653–663
- Horstmann C (1983) Specific subunit pairs of legumin from Vicia faba. Phytochemistry 22:1861–1866
- Horstmann C (1984) New aspects of the subunit structure of Vicia faba legumin. Kulturpflanze 32:S109-S116
- Jensen U, Berthold H (1989) Legumin-like proteins in gymnosperms. Phytochemistry 28:1389–1394
- Lambert N, Plumb GW, Wright DJ (1987) Application of highperformance liquid chromatography to the assessment of subunit heterogeneity in plant 11s storage globulins. J Chromatogr 402:159–172
- Lycett GW, Croy RRD, Shirsat A, Boulter D (1984) The complete nucleotide sequence of a legumin gene from pea (*Pisum sativum* L.). Nucleic Acids Res 12:4493–4506
- Masumura T, Kidzu K, Sugiyama Y, Mirsukawa N, Hibino T, Tanaka K, Fujii S (1989) Nucleotide sequence of a DNA encoding a major rice glutelin. Plant Mol Biol 12:723-725
- Matta N, Gatehouse JA, Boulter D (1981) The structure of legumin of Vicia faba L. – a reappraisal. J Exp Bot 32: 183–197
- Müntz K, Horstmann C, Schlesier B (1986) Seed proteins and their genetics in *Vicia faba* L. Biol Zblatt 105:107–120

- Neville JR DM (1971) Molecular weight determination of protein-dodecyl sulfate complexes by gel electrophoresis in a discontinuous buffer system. J Biol Chem 246:6328-6334
- Nielsen NC (1984) The chemistry of legume storage proteins. Phil Trans R Soc Lond B304:287-296
- Nielsen NC, Dickinson CD, Cho TJ, Thanh VH, Scallon BJ, Fischer RL, Sims TL, Drews GN, Goldberg RB (1989) Characterization of the glycinin gene family from soybean. Plant Cell 1:313-328
- Otto A, Kraft R, Etzold G (1984) N-terminal sequence analysis of basic subunits of legumin from *Vicia faba* by solid-phase sequencing. Kulturpflanze 32: S219–S221
- Rosenthal A, Jung R, Hunger HD (1986) Optimized conditions for solid phase sequencing: simultaneous chemical cleavage of a series of long DNA fragments immobilized on CCS anion-exchange paper. Gene 42:1–9
- Sanger F, Coulson AR (1978) The use of thin acrylamide gels for DNA sequencing. FEBS Lett 87:107-110
- Schlesier B, Bassüner R, Hai NV, Müntz K (1990) The cDNAderived primary structure of two distinct legumin A-subunit precursors from field bean (*Vicia-Faba* L.). Nucleic Acids Res 18:7146
- Spencer D, Higgins TJV (1980) The biosynthesis of legumin in maturing pea seeds. Biochem Int 1:502-509
- Takaiwa F, Kikuchi S, Oono K (1989) The complete nucleotide sequence of new type cDNA coding for rice storage protein glutelin. Nucleic Acids Res 17:3289
- Templeman TS, Demaggio AE, Stetler DA (1987) Biochemistry of fern spore germination: globulin storage proteins in *Matteucia struthiopteris* L. Plant Physiol 85:343–349
- Tucci M, Capparelli R, Costa A, Rao R (1991) Molecular heterogeneity and genetics of Vicia-Faba seed storage proteins. Theor Appl Genet 81:50–58
- Wobus U, Bäumlein H, Bassüner R, Heim U, Jung R, Müntz K, Saalbach G, Weschke W (1986) Characteristics of two types of legumin genes in the field bean (*Vicia faba L. var. minor*) genome as revealed by cDNA analysis. FEBS Lett 201:74–80