Aust. J. Biol. Sci., 1977, **30**, 33–45

Amino Acid Sequence of the Smaller Subunit of Conglutin γ, a Storage Globulin of *Lupinus angustifolius*

T. C. Elleman

Division of Protein Chemistry, CSIRO, Parkville, Vic. 3052.

Abstract

The amino acid sequence of the smaller subunit of conglutin γ, the simplest of the three globulins from the seeds of *Lupinus angustifolius* cv. Uniwhite, has been determined. The subunit was homogeneous and contained 154 amino acid residues, including five sulphur-containing amino acids—a considerably higher content than is found in most other legume storage proteins. There was no indication of the complexity experienced in studies of many other legume storage proteins. This is perhaps the first sequence of a subunit of a legume storage protein to be determined.

Introduction

Seed proteins form an important component of animal and human nutrition. The seeds of legumes are of particular interest since the storage globulins which make up the bulk of the protein of these seeds contain adequate levels of most essential amino acids and somewhat complement the deficiencies of the cereals (Bressani 1973). The lower quantities of nitrogenous fertilizer required for growth by legumes, because of the symbiotic nitrogen fixing bacteria in the root nodules, also renders legumes particularly important in the ever-increasing demand for protein.

Unfortunately most legume seeds have a deficiency in sulphur-containing amino acids, and an improvement in this respect is highly desirable. Since storage proteins are restricted to seeds, manipulation of the type and composition of protein produced should be possible without deleterious effects on the parent plant, and to this end a study of the storage proteins of legumes and the controls over their biosynthesis is desirable.

Lupinus angustifolius cv. Uniwhite constitutes an important crop in Western Australia for animal and potentially for human nutrition (Gladstones 1970). In addition to the two storage globulins, conglutin α and β, equivalent to the legumin-like and vicillin-like globulins found in many legumes (Danielsson 1949), a third storage globulin, conglutin γ, is also present. Conglutin γ has been shown to be significantly richer in its content of sulphur-containing amino acids, and of threonine and valine than are either of the other lupin storage proteins or legume proteins generally. It has been suggested that an increase in the synthesis of this protein would result in a significant improvement in the nutritional value of this legume (Gillespie and Blagrove 1974, 1975; Blagrove and Gillespie 1975).

It has been possible to purify conglutin γ so that, by the criteria of sodium dodecyl sulphate (SDS)–polyacrylamide gel electrophoresis and cellulose acetate electrophoresis in 8·0 M urea, it appears to be a homogeneous protein (Blagrove and Gillespie

1975). Impurity of the storage proteins of legumes, and the large number of similar subunits (Millerd 1975; Derbyshire *et al.* 1976) have previously hampered studies on these proteins. The discovery of conglutin γ with its homogeneity and simple two-subunit structure offers hope for studying the regulation of seed protein synthesis, and the possible manipulation of the proportions of the storage proteins. To this end a study of the sequence of conglutin γ has been commenced. The present paper reports the sequence of the smaller of the two subunits—subunit 2.

Materials and Methods

Enzymes

Trypsin [L-(1-tosylamido-2-phenyl)ethylchloromethylketone-treated], chymotrypsin and pepsin were obtained from Worthington Biochemical Corp., Freehold, N.J., U.S.A. Thermolysin was from Calbiochem., Los Angeles, Calif., U.S.A. Cyanogen bromide was from Koch–Light Laboratories, Colnbrook, Bucks, England.

Preparation of Proteins

Conglutin γ was prepared from lupin seeds (*L. angustifolius* cv. Uniwhite) by the method of Blagrove and Gillespie (1975). 200 mg (5 μmol) of conglutin γ was dissolved in 40 ml of a buffer at pH 8·1, 8·0 M urea, 0·1 M in tris, 0·002 M in EDTA. 180 mg dithiothreitol (1·2 mmol) was added and the solution was stirred under nitrogen for 2 h at room temperature and for a further 1 h at 50°C. After cooling to room temperature the pH was adjusted to 8·4, and 2·2 g iodoacetic acid (12 mmol) was added in small aliquots; the pH of 8·4 was maintained to within $\pm$0·2 pH units with 5 M sodium hydroxide solution, and light was excluded. After a further 15 min at room temperature 1·2 ml β-mercaptoethanol was added, and the pH was maintained at 8·4. The solution was dialysed against deionized water and freeze-dried.

The reduced and carboxymethylated protein (190 mg) was dissolved in 7 ml of a potassium phosphate buffer of ionic strength 0·3, pH 7·0, 8·0 M urea, and loaded onto a column (200 by 2·5 cm diameter) of Sephadex G100 equilibrated with a potassium phosphate buffer of ionic strength 0·3, pH 7·0, 6·0 M urea, 1% β-mercaptoethanol. A flow rate of 15 ml/h was used, and 9·3-ml fractions were collected. The column effluent was monitored at 278 nm. Two subunits were separated by this procedure (Fig. 1) and each was re-run after dialysis and freeze-drying. Throughout the preparation the purity of the protein and of its constituent subunits was monitored on cellulose acetate strips in 0·05 M sodium phosphate, 8·0 M urea, pH 7·0, using a Beckman Microzone electrophoresis system.

Amino Acid Analysis

Protein and peptides (about 10 nmol) were routinely hydrolysed with 5·7 M HCl (0·4 ml + 0·5 μl β-mercaptoethanol) in sealed evacuated tubes at 105°C for 24 h, or longer if necessary. Amino acid compositions were determined using half of each hydrolysate with a modified Beckman model 120 C amino acid analyser using a single column technique. Conglutin γ subunit 2 was also hydrolysed with 3 M *p*-toluenesulphonic acid and the tryptophan content was determined using an 18-cm column (Liu and Chang 1971).

Oxidation of the protein, to convert methionine and methionine sulphoxide to methionine sulphone prior to acid hydrolysis, was performed with a 50-fold molar excess of performic acid over total thioether (Neumann *et al.* 1962). This procedure permitted a quantitative estimate of methionine residues which are unreactive with cyanogen bromide because of oxidation by air to the sulphoxide.

Homoserine lactone was converted to homoserine prior to analysis by dissolving the dried hydrolysate in 100 μl of 0·1 M potassium carbonate and leaving it in a dessicator overnight under very light vacuum. No β-mercaptoethanol was added to analyses for which this treatment was proposed, since this practice was found to result in extensive loss of proline.

Preparation of Peptides and Sequence Determination

Conglutin γ subunit 2 (60 mg) was dissolved in 6 ml of formic acid containing 200 mg cyanogen bromide. 2·6 ml of water was added, the containing vessel was flushed with nitrogen and the solution was left in the dark for 16 h. The formic acid was removed directly using a glass freeze-drier without

prior dilution. The glassy material obtained was dissolved in 0·5 ml formic acid, diluted with an equal volume of water, and applied to a column (240 by 1·5 cm diameter) of Sephadex G50 Superfine equilibrated with 50% (v/v) formic acid. A flow rate of 4·0 ml/h was used, and 3-ml fractions were collected. The column effluent was monitored at 280 nm (Fig. 2), and pooled fractions were again directly freeze-dried.

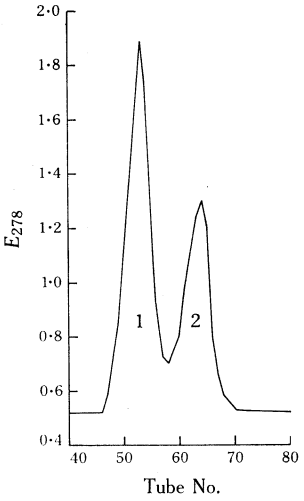


Fig. 1. Separation of the subunits of reduced carboxymethylated conglutin γ on Sephadex G100. The details appear in the text.

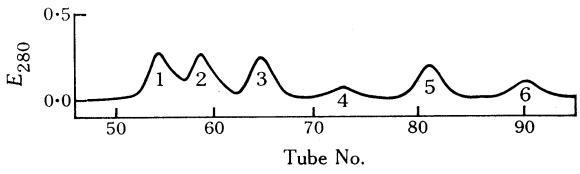


Fig. 2. Separation of the fragments from cleavage of the *S*-carboxymethylated derivative of conglutin γ subunit 2 with cyanogen bromide on Sephadex G50 in 50% (v/v) formic acid. The details appear in the text.

Peptide sequences were determined from the *N*-terminus by the dansyl–Edman procedure (Gray 1967). The sequences of peptides containing 15 or fewer residues, excepting 4T2c2, were determined completely by this procedure, identifying the final residues as the free amino acid or dansyl amino acid without prior acid hydrolysis. 1-*N*,*N*-Dimethylaminonaphthalene-5-sulphonyl amino acids were identified by the method of Woods and Wang (1967). When a peptide contained a larger number of residues it was further cleaved by secondary enzymic digestion and all data on the sequence studies of the smaller fragments have been presented (Tables 3, 4, 5 and 6). Enzymic digestions, excepting peptic digestions, were performed in 0·2 M *N*-ethylmorpholine acetate, pH 8·4, at 37°C for 4 h using an enzyme-substrate ratio of 1 : 100 (w/w). Fragments from cyanogen bromide cleavage were first dissolved or suspended in 100 μl of an aqueous solution of 20% (v/v) triethylamine below 18°C and then 1·0 ml of 0·2 M *N*-ethyl morpholine buffer, pH 8·4, was added. Peptic digestions were performed in 5% (v/v) formic acid, the fragments being dissolved in 98% (w/w) formic acid and the solution diluted with water.

The resultant peptide mixtures were separated either by high-voltage paper electrophoresis at pH 6·5, 3·5, 1·9 and 9·0 (Ambler 1963) or, if a complicated pattern arose, by a prior fractionation on a column (140 by 0·8 cm diameter) of Sephadex G25 Superfine in 0·1 M ammonium bicarbonate at pH 7·8 using a flow rate of 6 ml/h and collecting 3·0-ml fractions. The column effluent was monitored at 280 and 230 nm. Peptides on paper electrophoretograms were routinely located by guide-strips using 0·25% (w/w) ninhydrin, 1% (v/v) collidine in ethanol, and then by the chlorination method of Reindel and Hoppe (1954) applied to the same strips. Peptides containing tryptophan were located by overstaining the ninhydrin-stained guide-strips with 1% (w/w) *p*-dimethylaminobenzaldehyde in concentrated HCl–acetone (1 : 9 v/v), and histidine-containing peptides were located with Pauly stain (Easley 1965) after fading the ninhydrin colour with concentrated HCl–acetone (1 : 100

v/v). Elution of peptide bands from paper was performed with aqueous 1% (w/v) NH_3 solution. Amide groups were assigned to peptides on the basis of electrophoretic mobility (m) at pH 6·5 (Offord 1966) relative to aspartic acid ($m = -1·0$).

Nomenclature

All initial fragments used for sequence studies were obtained from cyanogen bromide cleavage, and the numerical prefix for each peptide indicates the fragments from which that peptide was derived. These fragments are numbered in the order in which they were eluted from the Sephadex G50 column. The peptides obtained by further degradation of these fragments by trypsin, chymotrypsin, thermolysin or pepsin are designated by the letters, T, C, Th, or P respectively, whilst other peptides obtained by secondary digestion of these peptides are designated by the additional suffix t, c, th or p depending on the secondary enzyme used.

Table 1. Chemical compositions of conglutin γ subunit 2 and fragments from cyanogen bromide cleavage

Results are expressed as residues/mole. Values in parentheses are inferred from sequence studies. Hexosamine was absent from hydrolysis. Abbreviations: RCM, reduced carboxymethylated; SCMCys, *S*-carboxymethyl cysteine; SCMHcy, *S*-carboxymethyl homocysteine; Hse, homoserine

Amino acid	RCM conglutin γ subunit 2		Fragments from CnBr cleavage						Sequence studies
	24 h	48 h	1	2	3	4	5	6	
Lys	5·7	6·0	5·8	5·6	3·6	2·1 (2)	3·4 (4)	0·0	6
His	5·7	5·7	5·6	2·8	2·8	3·0 (3)	0·0	2·8 (3)	6
Arg	4·5	4·9	4·7	4·8	2·0	3·0 (3)	2·0 (2)	0·0	5
SCMCys	3·0	2·9	2·9	2·8	1·0	2·0 (2)	1·0 (1)	0·0	3
Asp	16·7	17·6	17·0	16·0	8·0	10·0 (4)	6·0 (4)	2·0 (0)	8
Asn	—	—	—	—	—	(6)	(2)	(2)	10
Thr	6·9	6·7	6·8	1·4	5·4	1·2 (1)	0·1 (0)	5·5 (6)	7
Ser	13·4	12·0	13·8	9·9	8·0	5·5 (6)	5·0 (5)	4·6 (5)	16
Glu	11·0	11·0	11·0	8·3	6·0	5·2 (3)	3·1 (1)	3·0 (2)	6
Gln	—	—	—	—	—	(2)	(2)	(1)	5
Pro	5·4	5·4	5·2	4·0	4·0	1·2 (1)	2·7 (3)	1·0 (1)	5
Gly	15·3	15·8	15·6	11·8	7·6	8·0 (8)	4·0 (4)	4·0 (4)	16
Ala	10·5	10·8	10·6	8·6	5·7	5·0 (5)	4·0 (4)	2·0 (2)	11
Val	11·6	13·6[C]	12·8[C]	10·3	6·4	6·4 (7)	3·7 (4)	3·0 (3)	14
Ile	6·7	6·9[D]	6·7[D]	4·1	5·5	1·3 (1)	2·7 (3)	2·8 (3)	7
Leu	12·7	12·9	13·0	10·9	4·9	8·0 (8)	3·2 (3)	2·0 (2)	13
Tyr	3·8	3·8	3·7	2·0	2·6	1·0 (1)	1·0 (1)	1·7 (2)	4
Phe	8·2	8·9	8·6	5·6	4·8	4·3 (4)	2·0 (2)	2·9 (3)	9
Met	1·3[A]	1·3	0·4	0·2	0·0	0·0	0·0	0·0	2
Hse	0·1	0·2	1·0	0·5	1·7	0·0	1·0 (1)	1·0 (1)	—
Trp	1·0[B]	—	—	—	—	—	— (1)	—	1
SCMHcy	0·1	0·1	0·4	0·2	0·2	0·0	0·0	0·0	—
N-terminal residue	Ser		Ser	Pro	Ser	Val	Pro	Ser	

[A] $MetSO_2$: 1·3 residues/mole after performic acid oxidation.
[B] Trp : 1·0 residue/mole by *p*-toluenesulphonic acid hydrolysis (24 h).
[C] Val : 14 residues/mole after 72 h hydrolysis.
[D] Ile : 7 residues/mole after 72 h hydrolysis.

Results

The amino acid analyses of conglutin γ subunit 2 and the cyanogen bromide fragments are shown in Table 1. The molecular weight based on the sequence study is 16691 for the *S*-carboxymethyl derivative, confirming the estimate of 17000 by Blagrove and Gillespie (1975) from SDS–polyacrylamide gel electrophoresis. Serine

was found as the *N*-terminal residue by dansylation in urea (Gray 1967), and the sequence of the first 17 residues was confirmed using a protein sequenator constructed in these laboratories. No *C*-terminal residue could be detected by selective tritiation with ^3H_2O in pyridine–acetic anhydride (Holcomb *et al.* 1968). The presence of proline found by the sequence study accounts for this, since *C*-terminal proline cannot form the intermediary oxazalone and so cannot incorporate tritium.

Neither hexosamine nor hexose is present in this subunit. The hexose found by the orcinol–H_2SO_4 method (Winzler 1955) amounted to only 0·5 residue/mole and traces might be expected since the subunit was prepared on Sephadex.

Treatment of the subunit with cyanogen bromide produced six fragments instead of the expected three (Table 1), due to only partial splitting occurring at each methionine residue. Since performic acid oxidation prior to hydrolysis showed low amounts of methionine sulphone, this inability to cleave completely was not due to partial oxidation of the methionine to the sulphoxide. Traces of the characteristic degradation products of the carboxymethyl sulphonium salt of methionine, namely carboxymethyl homocysteine and homoserine, found in acid hydrolysates of the conglutin γ subunit 2, as well as methionine suggested that the methionine residues had been partially carboxymethylated.

Cyanogen bromide fragments 1, 2 and 3 represent material which contained the derivatized methionine, and so were resistant to cyanogen bromide cleavage. End group studies and amino acid analyses showed that fragment 1 was uncleaved reduced carboxymethylated subunit 2, fragment 2 represented the *C*-terminal portion of the molecule (equivalent to fragments 5 and 4), and fragment 3 represented the overlapping *N*-terminal portion of the molecule (equivalent to fragments 6 and 5).

Table 2. Amino acid sequence of peptides from regions overlapping the methionine residues

Results are expressed as residues/mole

2T1	
Composition	$His_{0.8}$ $Arg_{1.0}$ $SCMCys_{0.7}$ $Asp_{3.2}$ $Ser_{2.6}$ $Glu_{3.0}$ $Gly_{4.0}$ $Ala_{2.0}$ $Val_{3.6}$ $Ile_{1.0}$ $Leu_{1.3}$ $Phe_{1.8}$ $Hse_{0.5}$ $Met_{0.2}$
Sequence	Ile-Ser-Ser-Glx-Asx-Phe-Hse-Val-Glx-
	→ → → → → → → → →
3T1	
Composition ($m = +0{\cdot}42$)	$Lys_{1.0}$ $Asp_{2.0}$ $Pro_{1.0}$ $Ala_{1.0}$ $Hse_{0.5}$ $Met_{0.1}$ $SCMHcy_{0.3}$
Sequence	Ala-Asn-Asn-Hse-Pro-Lys
	→ → → → →

When the sequences of fragments 4, 5 and 6 had been determined their order was apparent from a consideration of the end groups and compositions of the larger fragments. Peptides containing overlapping regions between fragments 6 and 5, and 5 and 4 were isolated from fragments 3 and 2 respectively by overnight tryptic hydrolyses followed by paper electrophoresis for the former and, for the latter, chromatography on Sephadex G50 in 0·1 M ammonium bicarbonate, pH 7·8, then paper electrophoresis (Table 2).

For each of the smaller cyanogen bromide fragments a complete set of peptides was isolated from one type of enzymic digestion. For fragments 4 and 5 the enzyme was trypsin, whilst for fragment 6 chymotrypsin was used (Tables 3, 4, 5 and 6). The

Table 3. Amino acid analyses of peptides 4T1 and 4T2 and peptides derived from them

Results are expressed as residues/mole

Amino acid	4T1	4T1c1	4T1c2	4T2	4T2c1	4T2c2	4T2c3	4T2th1	4T2th2	4T2th3	4T2th4	4T2th5	4T2p1
Lys													
His	1·0		0·9	1·6		1·5		2·0			1·7		
Arg	1·0		1·0	1·0			1·0						
SCMCys	0·8	0·6											
Asp	2·0	1·0	1·0	2·1		1·1	1·0		1·0	1·0	1·0	1·0	1·0
Thr													
Ser	1·0	0·7											
Glu	2·0	2·0		3·0		1·9	1·0		2·0		2·0		
Pro													
Gly	3·6	1·0	2·8	2·0	1·0	1·0		1·0					
Ala	1·9	1·0	1·0	2·8	1·8	1·0		1·0			0·7		
Val	3·4	2·0	2·0	1·8		1·7				1·5		1·5	2·0
Ile				0·9	0·8								
Leu	1·2	1·0		3·6	1·0	2·0	1·0	0·9	2·0		2·0	1·0	1·0
Tyr													
Phe	1·0		0·9	1·1		0·9				1·0		1·0	1·0
m	−0·20	−0·50	+0·24	*i*[A]	0·0	*i*[A]	−0·39	+0·75	−0·68	−0·32	−0·29	−0·28	−0·30

[A] *i* = insoluble.

order of peptides within each set was determined by the isolation of overlapping peptides from other enzymic digests (Fig. 3).

When the sequences of large peptides from primary enzymic digests could not be determined completely by the dansyl–Edman procedure (namely peptides 4T1 and 4T2), or mobilities indicated the presence of both acidic and amidated residues, further enzymic degradation was performed. Peptide 4T1c1 had two acidic residues; removal of four residues left a peptide of mobility −0·69, which still had two acidic residues. Removal of the fifth residue decreased the mobility to −0·44, indicating an aspartic acid had been removed. Peptide 4T1c2 contained only a single residue of asparic acid or asparagine. If the histidine is assumed to carry a complete positive charge at pH 6·5, then aspartic acid is indicated, and this was confirmed by amino-peptidase (microsomal) digestion. Dansyl–Edman degradation on peptide 4T1 to the extent of 16 residues established the continuity of peptides 4T1c1 and 4T1c2. Peptide 4T2 was insoluble in dilute aqueous buffers below pH 9, and the fragment obtained from chymotryptic digestion of this peptide containing the central region (4T2c2) was also insoluble. Soluble overlapping peptides were obtained from thermo-lytic and peptic digestion, so enabling the sequence to be determined. The amidated residue of peptide 4T2th2 was determined by measurements of mobility after the successive removal of each glutamic acid residue.

Table 4. Amino acid analyses of peptides from fragment 4

Results are expressed as residues/mole

Amino acid	4T3	4T4	4T5	4T6	4Th1	4Th2	4Th3	4Th4	4P1	4P2
Lys		1·0	1·0				1·0	0·8		1·9
His										
Arg	1·0				1·0	2·0			1·7	
SCMCys				0·9				0·8		0·9
Asp		1·6		4·0				1·0	1·1	2·9
Thr				1·0				0·8		1·0
Ser	0·9	1·9	0·9	1·0		0·9	0·9	0·9	1·0	3·6
Glu						1·0			1·0	
Pro				1·1						
Gly		1·3	1·0		1·0	1·0		1·1	1·1	1·2
Ala					2·0					
Val		1·0				1·0			0·9	
Ile										
Leu		1·0		2·0		1·0	1·0		0·9	2·0
Tyr			1·0					1·0		1·0
Phe		0·9		1·0					0·9	
m	+0·60	+0·30	+0·45	−0·45	+0·58	+0·36	+0·58	0·0	0·0	+0·21

Two tryptic peptides from fragment 5 contained both acid and amidated residues. Peptide 5T6 remained neutral at pH 6·5 after removal of the first residue, showing this to be an asparagine; amino acid analysis confirmed that a residue had been removed. A positive Ehrlich reaction showed this peptide to contain tryptophan, the position of which was inferred from the absence of a dansyl amino acid at the penultimate position. Peptide 5T7 was fragmented with thermolysin to produce a fragment 5T7th1 devoid of any homoserine which might lactonize during the Edman

Table 5. Amino acid analyses of peptides from fragment 5

Results are expressed as residues/mole

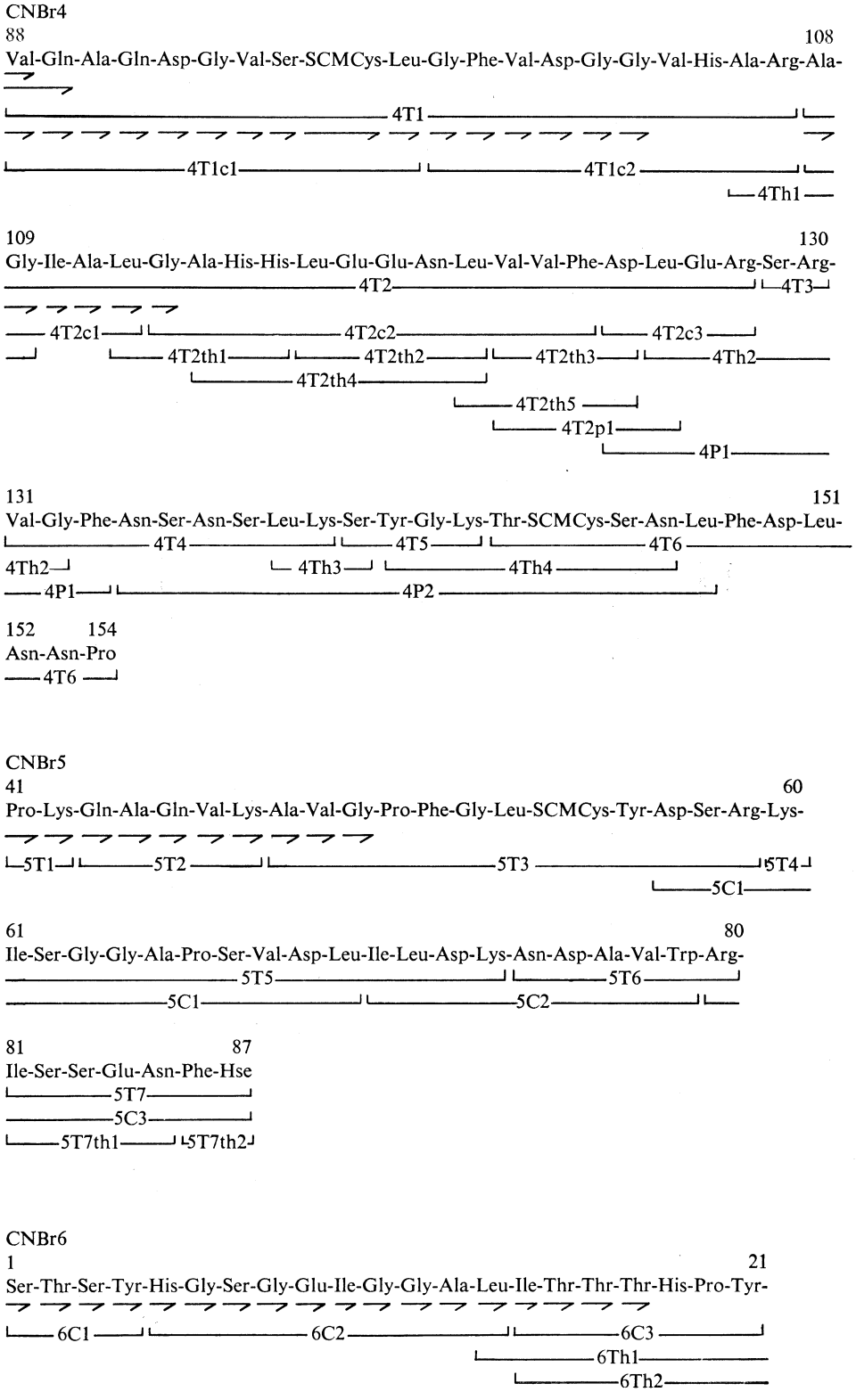
Amino acid	5T1	5T2	5T3	5T4	5T5	5T6	5T7	5T7th1	5T7th2	5C1	5C2	5C3
Lys	0·9	0·6		1·0	0·7					1·0	1·0	
His												
Arg			0·8			1·0				1·1		0·9
SCMCys			0·8									
Asp			1·1		2·1	1·9	1·0	1·0		2·0	3·0	1·0
Ser			1·0		2·1		1·7	1·9		2·8		1·9
Thr												
Glu		2·0					1·0	1·0				1·0
Pro	1·0		1·0		1·1					0·9		
Gly			1·8		2·0					2·0		
Ala		1·0	0·8		1·0	1·0				1·1	1·0	
Val		0·6	0·6		0·7	0·9				1·1	1·0	
Ile					1·4		0·9	0·9		1·0	0·9	1·0
Leu			0·7		1·5					1·1	1·0	
Tyr			0·6									
Phe			0·6				0·6		1·0			1·0
Hse							1·0		1·0			1·0
Trp						+						
m	+0·69	+0·41	−0·21	+0·90	−0·19	+0·0	−0·30	−0·37	0·0	0·0	−0·20	0·0

Table 6. Amino acid analyses of peptides from fragment 6

Results are expressed as residues/mole

Amino acid	6C1	6C2	6C3	6C4	6C5	6C6	6C7	6C8	6Th1	6Th2	6Th3	6Th4	6Th5	6Th6	6Th7
Lys															
His		0·8	1·0		1·0				1·1	1·0	1·0				
Arg															
SCMCys															
Asp								2·0							2·0
Thr	0·8		2·8	0·9			0·9		4·0	3·8				1·0	
Ser	1·9	0·9			1·7						2·0				
Glu		1·0				1·0	1·0					1·0	1·0	1·0	
Pro			1·1						1·0	1·0					
Gly		4·0													
Ala		1·0						1·0							1·0
Val				1·0		1·0	1·0		0·7		1·0		0·7	1·0	0·9
Ile		1·0	1·0		1·0				1·0	1·0		1·0	0·9		
Leu		1·0		1·0					0·3		1·0				
Tyr	1·0		0·9						1·0	1·0					
Phe					1·0	1·0	1·0					1·0	1·0	1·0	0·9
Hse								1·0							0·9
m	0·0	−0·20	+0·30	0·0	n.d.[A]	−0·46	0·0	0·0	+0·23	+0·27	+0·38	−0·38	−0·33	0·0	0·0

[A] Not determined.


CNBr4
88
108
Val-Gln-Ala-Gln-Asp-Gly-Val-Ser-SCMCys-Leu-Gly-Phe-Val-Asp-Gly-Gly-Val-His-Ala-Arg-Ala-
4T1
4T1c1
4T1c2
4Th1
109
130
Gly-Ile-Ala-Leu-Gly-Ala-His-His-Leu-Glu-Glu-Asn-Leu-Val-Val-Phe-Asp-Leu-Glu-Arg-Ser-Arg-
4T2
4T3
4T2c1
4T2c2
4T2c3
4T2th1
4T2th2
4T2th3
4Th2
4T2th4
4T2th5
4T2p1
4P1
131
151
Val-Gly-Phe-Asn-Ser-Asn-Ser-Leu-Lys-Ser-Tyr-Gly-Lys-Thr-SCMCys-Ser-Asn-Leu-Phe-Asp-Leu-
4T4
4T5
4T6
4Th2
4Th3
4Th4
4P1
4P2
152
154
Asn-Asn-Pro
4T6
CNBr5
41
60
Pro-Lys-Gln-Ala-Gln-Val-Lys-Ala-Val-Gly-Pro-Phe-Gly-Leu-SCMCys-Tyr-Asp-Ser-Arg-Lys-
5T1
5T2
5T3
5T4
5C1
61
80
Ile-Ser-Gly-Gly-Ala-Pro-Ser-Val-Asp-Leu-Ile-Leu-Asp-Lys-Asn-Asp-Ala-Val-Trp-Arg-
5T5
5T6
5C1
5C2
81
87
Ile-Ser-Ser-Glu-Asn-Phe-Hse
5T7
5C3
5T7th1
5T7th2
CNBr6
1
21
Ser-Thr-Ser-Tyr-His-Gly-Ser-Gly-Glu-Ile-Gly-Gly-Ala-Leu-Ile-Thr-Thr-Thr-His-Pro-Tyr-
6C1
6C2
6C3
6Th1
6Th2

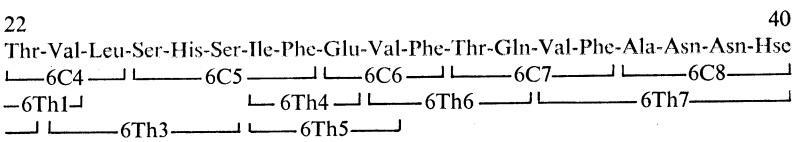


Fig. 3. Amino acid sequences of fragments 4, 5 and 6 from cleavage of *S*-carboxymethylated derivative of conglutin γ subunit 2 with cyanogen bromide. The disposition of these fragments within the subunit is indicated above the terminal residues.

degradation so rendering mobility measurements spurious. Free asparagine was found as the *C*-terminal residue of peptide 5T7th1 after four Edman degradations.

Fragment 6 produced two chymotryptic peptides which required some further investigation of their mobilities. Peptide 6C2 was negatively charged at pH 6·5 even though the glutamic acid residue was balanced by an *N*-terminal histidine residue. The proximity of the positively charged amino group presumably lowers the pK of the imidazole nitrogen. Removal of this histidine residue increased the mobility to −0·29. Peptide 6C8 was neutral at pH 6·5 which indicated either two asparagine residues or one aspartic acid residue and a lactonized homoserine. Treatment of this peptide with anhydrous trifluoroacetic acid for 1 h at 37°C produced a fragment which was basic at pH 6·5 (m = +0·54), indicating the presence of two residues of asparagine, and a homoserine residue which had been lactonized by the acid treatment.

The complete amino acid sequence of the smaller subunit of conglutin γ is shown in Fig. 4.

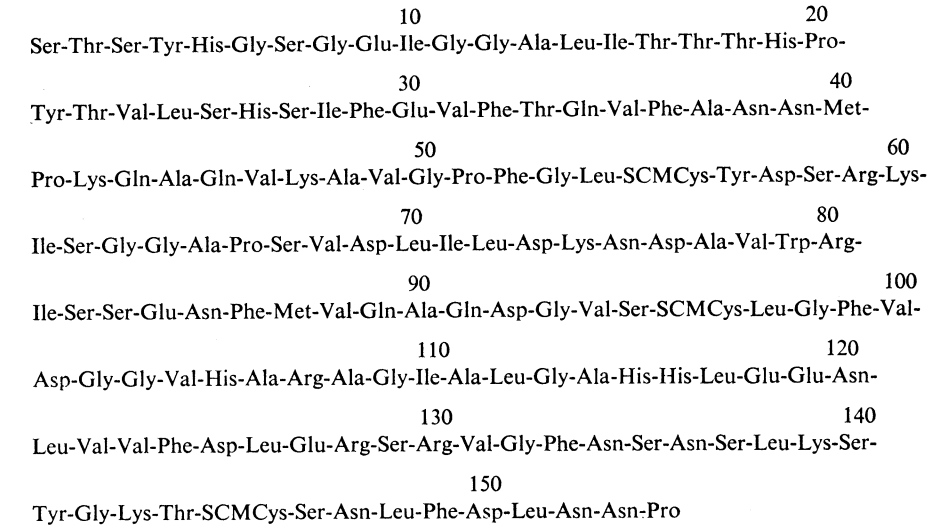


Fig. 4. Complete amino acid sequence of the *S*-carboxymethylated derivative of conglutin γ subunit 2. Residue numbers are shown above residues.

Discussion

The extensive carboxymethylation of the methionine residues observed was unexpected in view of the short reaction time and must be attributable to the large excess of iodoacetic acid. An amount equimolar to the total thiol rather than the five-fold

excess used would probably have been more selective. However, peptides containing the carboxymethylsulphonium salt of methionine were quite beneficial in the determination of the amino acid sequence. No attempt was made to cleave at the peptide bonds formed by these derivatives of methionine since, unlike the *S*-carbamidomethyl derivative and the *S*-cyano derivatives, the *S*-carboxymethyl derivative is quite stable, so permitting the sequence determination by conventional means of fragments containing this residue. On acid hydrolysis of peptides this derivative was degraded largely to homoserine and its lactone with traces of *S*-carboxymethyl homocysteine and methionine, whilst acid hydrolysis of the dansyl *S*-carboxymethyl methionine peptides produced dansyl homoserine and its lactone and traces of dansyl *S*-carboxymethyl homocysteine.

All fragments from cyanogen bromide cleavage, except fragment 5, were insoluble in the normal non-disaggregating buffers generally used in sequence studies, so 50% (v/v) aqueous formic acid was used for dissolving these peptides during fractionation. However, enzymes were able to digest finely suspended insoluble fragments, producing smaller peptides which were soluble in aqueous buffers at near neutrality, with the notable exception of peptide 4T2. The amino acids of peptide 4T2 have an extremely high helix-forming capability (Chou and Fasman 1974), and one side of the proposed helix would be hydrophobic whilst the other side would contain basic and acidic residues in juxtaposition at i and $i+3$, possibly permitting internal compensation of charge and so aggregation, and insolubility. Peptide 4T2 on chymotryptic digestion produced an insoluble peptide from the middle section, even though this chymotryptic peptide contained two glutamic acid residues and two histidine residues in only 12 residues.

The variability of the relative proportions of the three conglutins in lupin seeds has been examined already with regard to species and cultivar differences (Gillespie and Blagrove 1975), and the effect of differences in levels of applied sulphur has also been determined (Blagrove *et al.* 1976). The homogeneity and simple two-subunit structure of conglutin γ—a protein of the highest nutritional value so far reported in legume seeds—is in contrast to the complexity of other legume storage proteins (Blagrove and Gillespie 1975; Millerd 1975; Derbyshire *et al.* 1976), and should inspire further studies on the sequences and controls of lupin storage proteins.

Acknowledgments

I am indebted to my colleagues Dr C. McC. Roxburgh for performing the amino acid analysis, Mr A. S. Inglis for the sequenator results, and Mr R. E. Guthrie for the preparation of conglutin γ.

References

Ambler, R. P. (1963). The amino acid sequence of *Pseudomonas* cytochrome c-551. *Biochem. J.* **89**, 349–78.

Blagrove, R. J., and Gillespie, J. M. (1975). Isolation, purification and characterisation of the seed globulins of *Lupinus angustifolius*. *Aust. J. Plant Physiol.* **2**, 13–27.

Blagrove, R. J., Gillespie, J. M., and Randall, P. J. (1976). Effect of sulphur supply on the seed globulin composition of *Lupinus angustifolius*. *Aust. J. Plant Physiol.* **3**, 173–84.

Bressani, R. (1973). In 'Nutritional Improvement of Food Legumes by Breeding'. pp. 15–42. (Protein Advisory Group of the United Nations System: New York.)

Chou, P. Y., and Fasman, G. D. (1974). Prediction of protein conformation. *Biochemistry* **13**, 222–45.

Danielsson, C. E. (1949). Seed globulins of the *Gramineae* and *Leguminosae*. *Biochem. J.* **44**, 387–400.

Derbyshire, E., Wright, D. J., and Boulter, D. (1976). Legumin and vicilin, storage proteins of legume seeds. *Phytochemistry* **15**, 3–24.

Easley, C. W. (1965). Combination of specific colour reactions useful in the peptide mapping technique. *Biochim. Biophys. Acta* **107**, 386–8.

Gillespie, J. M., and Blagrove, R. J. (1974). Improving the content of essential amino acids in lupin seed. *Search* **6**, 600–1.

Gillespie, J. M., and Blagrove, R. J. (1975). Variability in the proportion and type of subunits in lupin storage globulins. *Aust. J. Plant Physiol.* **2**, 29–39.

Gladstones, J. S. (1970). Lupins as crop plants. *Field Crop Abstr.* **23**, 123–48.

Gray, W. (1967). Sequential degradation plus dansylation. In 'Methods in Enzymology'. (Ed. C. H. W. Hirs.) Vol. XI, pp. 469–75. (Academic Press: New York.)

Holcomb, G. N., James, S. A., and Ward, D. N. (1968). A critical evaluation of the selective tritiation method of determining *C*-terminal amino acids and its application to luteinizing hormone. *Biochemistry* **7**, 1291–6.

Liu, T. Y., and Chang, Y. H. (1971). Hydrolysis of proteins with *p*-toluenesulphonic acid. *J. Biol. Chem.* **246**, 2842–8.

Millerd, A. (1975). Biochemistry of legume seed proteins (peas, broadbean). *Annu. Rev. Plant Physiol.* **26**, 53–72.

Neumann, N. P., Moore, A., and Stein, W. H. (1962). Modification of the methionine residues in ribonuclease. *Biochemistry* **1**, 68–75.

Offord, R. E. (1966). Electrophoretic mobilities of peptides on paper and their use in the determination of amide groups. *Nature (London)* **211**, 591–3.

Reindel, F., and Hoppe, W. (1954). Über eine Färbemethode zum Anfärben von Aminosäuren, Peptiden und Proteinen auf Papierchromatogrammen und Papierelektropherogrammen. *Chem. Berl.* **87**, 1103–7.

Winzler, R. J. (1955). Determination of serum glycoproteins. In 'Methods of Biochemical Analysis'. (Ed. D. Glick.) Vol. 2, pp. 279–311. (Interscience Pub. Inc.: New York.)

Woods, K. R., and Wang, K. T. (1967). Separation of dansyl-amino acids by polyamide layer chromatography. *Biochim. Biophys. Acta* **133**, 369–70.

Manuscript received 20 September 1976