Phytochemistry, Vol. 26, No. 2, pp. 335–343, 1987.
Printed in Great Britain.

0031–9422/87 $3.00 + 0.00

PURIFICATION OF A SEED GLYCOPROTEIN: *N*-TERMINAL AND DEGLYCOSYLATION ANALYSIS OF PHASEOLIN

HERBERT E. PAAREN, JERRY L. SLIGHTOM,* TIMOTHY C. HALL,† ADAM S. INGLIS‡ and ROBERT J. BLAGROVE‡

Agrigenetics Corporation, Advanced Research Division, 5649 East Buckeye Road, Madison, Wisconsin 53716, U.S.A.; ‡CSIRO 343 Royal Parade, Parkville, 3052, Victoria, Australia

(*Revised received* 29 *April* 1986)

Key Word Index—*Phaseolus vulgaris*; Leguminosae; G-1 globulin; purification; HPLC; *N*-terminal sequencing; deglycosylation.

Abstract—Phaseolin, the major storage protein of the French bean *Phaseolus vulgaris* cv. Tendergreen, has been isolated and purified by either ion-exchange chromatography or reversed-phase HPLC. These purification procedures were used to fractionate the native protein aggregate into its characteristic subunit components. Amino-terminal sequence analysis was performed on the intact peptide subunits. Native phaseolin was chemically cleaved at a unique tryptophan residue which is proximal to the *N*-terminal region of the protein with BNPS–skatole and the resulting peptide fragments were isolated *via* reversed-phase HPLC. Chemical and enzymatic sequence results obtained from these peptide fragments are in full agreement with the results obtained for the full length peptide subunits. These *N*-terminal analyses show that the signal peptide cleavage process is somewhat random resulting in the phaseolin polypeptides having possibly three or four different *N*-termini. Native phaseolin and purified subunits were chemically deglycosylated with trifluoromethanesulphonic acid in the presence of an anisole scavenger. One-dimensional SDS-PAGE analysis of the deglycosylated products show that differential glycosylation is largely responsible for much of the observed molecular weight heterogeneity found among phaseolin polypeptides.

INTRODUCTION

Seed storage proteins have been the subject of much biochemical and genetic research because of their major role as a food source. In addition to the prospect of increasing the nutritional value of a number of food crops, the study of seed storage proteins also offers insights into the molecular mechanics of tissue-specific developmental regulation, post-translational processing, transport and accumulation of proteins. Molecular cloning of a number of seed protein genes has now made it possible to investigate these biochemical phenomena at the nucleotide sequence level with the ultimate goal being the identification of those regions of the gene responsible for regulation and processing of the expressed product. Information of this nature will ultimately provide the basis for the rapid production of crop species possessing desirable traits (e.g. increased nutritional value, disease and pest resistance) introduced by recombinant DNA technology.

Phaseolin is the major group of seed storage glycoproteins from French bean (*Phaseolus vulgaris* L.) and represents 30–50% of the total protein in mature seeds [1]. In the seeds, phaseolin is present as an aggregate of three polypeptide subunits forming a protomer with M_r of approximately 150 000 [2, 3]. One-dimensional SDS-polyacrylamide gel electrophoresis (SDS-PAGE) of the dissociated subunits, isolated from the cultivar Tendergreen, reveals three polypeptides termed α, β and γ having molecular weights of 51–53 000, 47–48 000, and 43–46 000, respectively [2, 4]. Two-dimensional SDS-PAGE shows that these polypeptides contain charge isomers as they resolve into at least five polypeptide forms [4]. Peptide mapping of phaseolin polypeptides after proteolytic and chemical cleavages indicate that all of the phaseolin polypeptides are similar in amino acid composition and organization [5, 6]. These phaseolin polypeptides are all encoded in 16S *m*RNA species [7], as shown by *in vitro* translation experiments [8], and they accumulate rapidly in developing seed cotyledons, beginning when the cotyledons are about 7 mm in length and continuing until they reach 17–19 mm in length [9]. Crosses between *P. vulgaris* cultivars, which show different SDS-PAGE phaseolin profiles, demonstrate that these polypeptide patterns are inherited as single units which are allelic and codominantly expressed [10]. Thus, phaseolin genes may be organized in closely linked gene families.

One particular phaseolin gene has been cloned, λPVPh177.4, and also its *m*RNA counterpart, *c*DNA clone pPVPh31 [11, 12]. Nucleotide sequence comparison of these clones revealed a 1990 bp phaseolin gene consisting of 80 bp of 5′ and 135 bp of 3′ non-coding DNA and 515 bp of intron DNA which splits the 1263 bp of coding DNA into six exons. The derived amino acid sequence indicates that this phaseolin gene encodes for a β-type phaseolin polypeptide. The presence of a highly hydrophobic *N*-terminus suggested that phaseolin may have a signal peptide of about 21 amino acid residues.

*Present address: The Upjohn Company, 7242-267-5, Kalamazoo, MI 49001 U.S.A.

†Present address: Texas A & M University, Department of Biology, College Station, TX 77843 U.S.A.

Agrigenetics Advanced Science Company Publication No. 53.

However, the precise length was uncertain because the *N*-terminus of the native β-type phaseolin polypeptide was unknown until this report (see below).

Two *N*-glycosyl carbohydrate–peptide linkage sites (Asn–X–Thr or Ser) were also identified from the derived amino acid sequences [12]. Heterogeneity of phaseolin peptides can be accounted for to a small degree by the variation in amino acid composition, as shown by the nucleotide sequence of six distinct phaseolin *c*DNA clones [13, 14]. However, the contribution that post-translational factors, such as glycosylation and signal peptide processing, have on polypeptide variation have not been directly addressed. The studies reported here confirm the presence of a signal peptide, define the *N*-termini of mature phaseolin polypeptides, and reveal the role that glycosylation plays in the observed heterogeneity of the native phaseolin seed storage proteins.

RESULTS AND DISCUSSION

Separation of phaseolin polypeptides by ion exchange chromatography

Classical purifications of G1 globulin fractions from legume seeds have exploited the ionic strength–pH dependent aggregation properties of these proteins. Phaseolin has been shown by ultracentrifugation studies to undergo a reversible pH regulated aggregation. At high pH ($\sim$ 11), native phaseolin is dissociated into its peptide components with 3S sedimentation coefficients; at neutral pH, a 7S protomeric trimer of peptides is formed while at pH 4–5 an 18S tetramer of trimers is found to exist [15, 16]. By adjusting the pH and ionic strength of a crude bean flour extract of *Phaseolus vulgaris* L. cv. Tendergreen seeds, the G1 phaseolin fraction can be precipitated away from other soluble seed storage proteins, principally the G2 lectin. This procedure provides a purified preparation which has been used in further experimental studies (see Fig. 1).

The isoelectric points of the three major molecular weight classes of phaseolin subunits (α, β and γ) incubated under denaturing conditions are known to be in the range of pH 5.5–5.8 from 2-D IEF-SDS gel analysis [4]. Consequently, anion-exchange chromatography is the method of choice for fractionation. When phaseolin was chromatographed under denaturing conditions on DEAE–Sephacel, fractionation of the major peptide components was achieved using a shallow NaCl gradient (0–0.15 M). The results are shown in Fig. 2 along with SDS-PAGE of the major peak fractions. Three major peaks were separated: the first peak (A) consists mainly of the β and γ subunits, the second peak (B) is a mixture of α and β subunits and the third peak (C) resembles native phaseolin with α, β and γ subunits being present.

Two aspects of these results are of interest. Firstly, the 2-D IEF-SDS gel analysis of denatured phaseolin suggests that such a large separation of components would not be possible given the rather narrow range of phaseolin subunit isoelectric points (pH 5.5–5.8). However, separation on the basis of charge was confirmed by electrophoresis under dissociating, non-sieving conditions (8 M urea–15 mM phosphate buffer, pH 7.6) on cellulose acetate membranes (data not shown). Fractions A and B possess approximately the same mobility while fraction C is much more acidic in character. The most likely explanation is that fraction C contains bound phytic acid. High levels of phytic acid have been found associated with

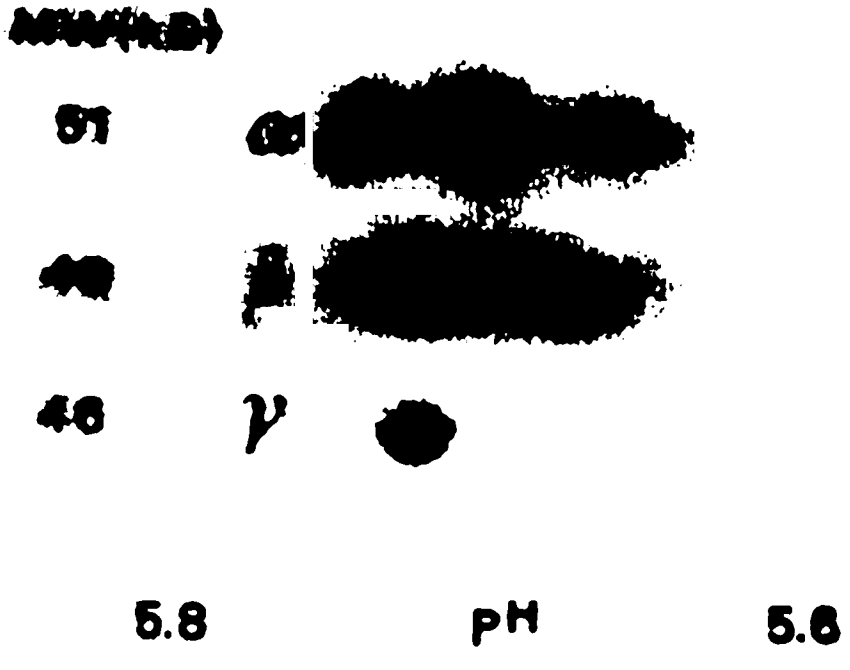

Fig. 1. Two-dimensional IEF-SDS gel of acid precipitated native phaseolin.

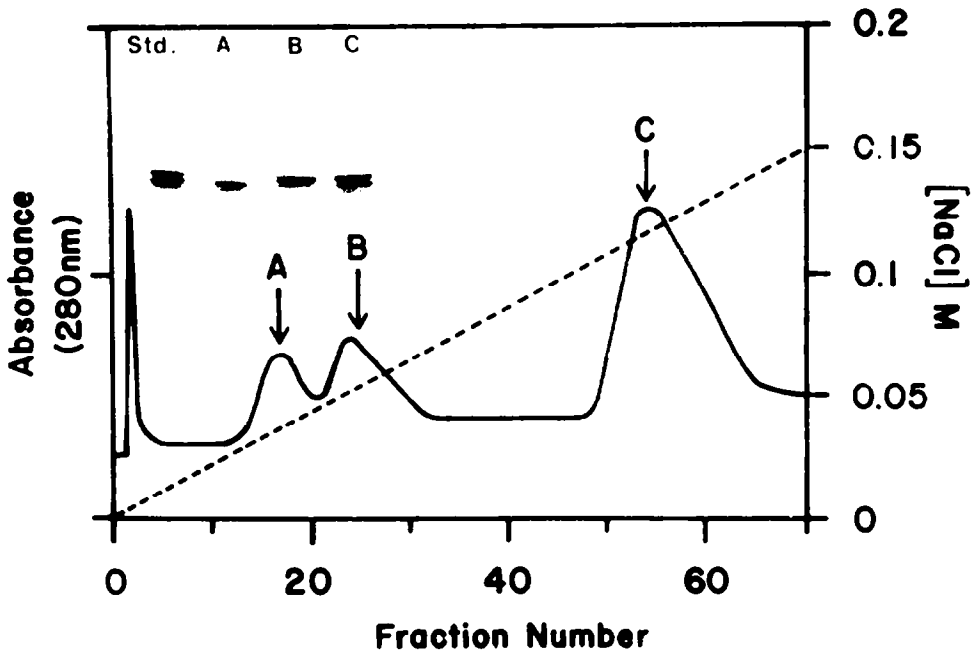

Fig. 2. DEAE-Sephacel ion-exchange chromatography of phaseolin (8 M urea, pH 7.6). One-dimensional SDS-PAGE analysis of peak fractions indicated by arrows is shown also.

phaseolin [17]. Presumably the bound phytic acid is removed when phaseolin samples are heated in urea prior to IEF analysis.

Secondly, a β subunit ($M_r \sim 48\,000$) species is apparently present in both fractions A and B. This could be accounted for by peak tailing and overlap; however, the fractions appear to be baseline separated. Alternatively, two β subunit species differing in net charge might be present. The sequence analyses of phaseolin *c*DNA clones indicate that phaseolin polypeptides are encoded by two general gene families which differ by only 14 amino acid residues [13, 14]. The α-type phaseolin polypeptides derived from the larger transcripts have a net charge of minus four when compared with the β-type polypeptides at neutral pH.

Both α- and β-type gene family members contain two *N*-glycosyl recognition sites and differential glycosylation appears to be responsible for much of the size heterogeneity found among these polypeptides [14, 18, 19], thus complicating the results shown in Fig. 2. Deglycosylation of the individual phaseolin polypeptide bands (isolated from a one-dimensional gel, see below) and SDS-PAGE analysis revealed that the β-type phaseolin band derived from native phaseolin by SDS-PAGE contains both α- and β-type polypeptides that co-migrate because they contain different amounts of carbohydrate ([18], see below). Thus, it would seem that DEAE-Sephacel chromatography is simply separating α- and β-type phaseolin polypeptides with peak A containing two β-type polypeptides glycosylated to different extents, whereas peak B contains two α-type polypeptides with different amounts of carbohydrate attached. Moreover, it follows that γ phaseolin is really a β-type polypeptide of low carbohydrate content. This interpretation is supported by *in vitro* and *in vivo* analysis of oligosaccharide chain additions to phaseolin polypeptides [18].

Separation of phaseolin polypeptides by HPLC

High-performance liquid chromatography (HPLC) purification of peptides and proteins is now a well established technique [20]. The introduction of wide-pore (300 Å) silica-based supports with a variety of alkyl bonded phases and the development of ion-pairing solvent systems permits rapid, high resolution fractionation of complex protein/peptide mixtures in a reversed-phase mode. When native phaseolin was solubilized in 0.1% TFA and chromatographed on a Baker wide-pore diphenyl column with a 20–60% linear gradient of 0.1% TFA–acetonitrile into 0.1% TFA–H_2O two peaks were obtained. One-dimensional SDS-PAGE analysis showed the earliest eluting peak to be mostly α subunit and the later eluting peak to be a mixture of β- and γ-type subunits (Fig. 3). Because macromolecules are separated on the basis of their overall hydrophobicity, elution of the more ionic glycosylated α-type subunits before the less ionic, glycosylated β-type subunits is expected. When the gradient was modified to include near isocratic elution steps, a different separation profile was obtained (Fig. 4). Analysis of fractions collected across the major peaks by SDS-PAGE showed a unique subunit composition for the three major peaks. The earliest eluting peak (fractions 1–3) consisted of α- and β-type subunits with the leading edge of the peak being enriched in the α-component. The following peak (fractions 4–6) was composed of β and γ subunits, again indicating two distinct β-type species, while the third peak showed the pattern of native phaseolin and was probably due to undissociated protein. Phaseolin has also been chromatographed on a variety of other bonded alkyl phase, wide-pore columns with similar results (data not shown), thus confirming reversed-phase chromatography as a viable means of fractionating glycosylated plant proteins.

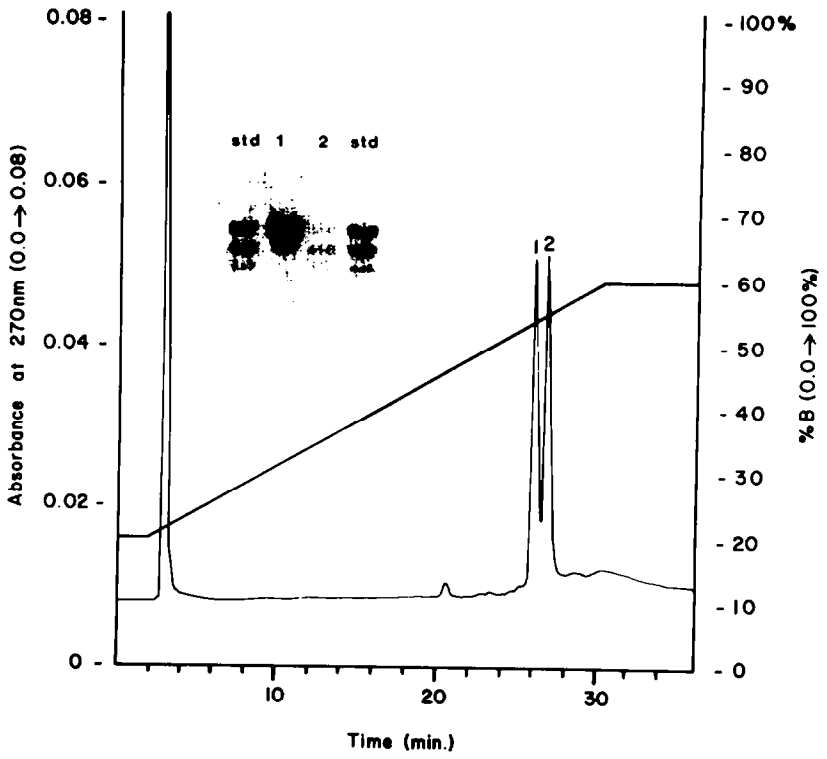

Fig. 3. Reversed-phase HPLC of 100 μg of phaseolin on a Baker 4.0 × 250 mm, 10 μm diphenyl bonded phase 300 Å pore column and SDS-PAGE analysis of fractions. Solvent A: 0.1% TFA–H_2O, solvent B: 0.1% TFA–acetonitrile; flow 1.0 ml/min; gradient as shown.

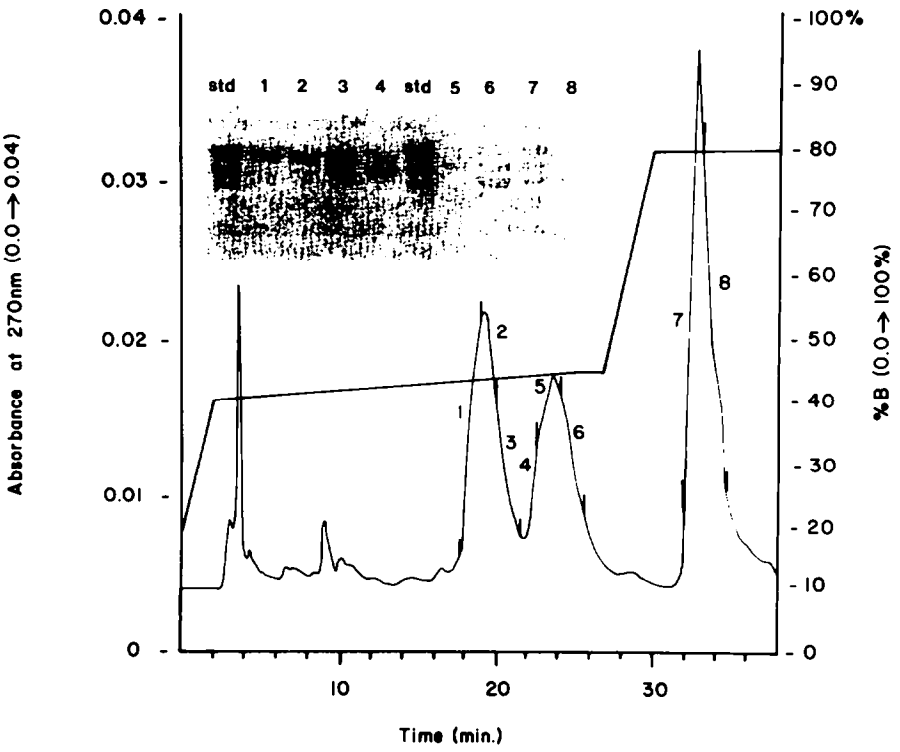

Fig. 4. Reversed-phase HPLC of 160 μg of phaseolin and SDS-PAGE analysis of fractions as described in Fig. 3. Gradient as shown.

Identification of phaseolin N-*terminal amino acid residues*

Throughout our analytical studies on phaseolin, constant reference was made to the nucleotide sequence data and the corresponding amino acid sequence. This information has greatly aided the design of experiments and the analysis of results. This was particularly the case in the *N*-terminal sequence analyses of native phaseolin subunits.

The encoded phaseolin protein possesses an amino-terminal region which is very characteristic of a signal peptide, i.e. two positively charged arginyl residues at positions 2 and 4 followed by a long stretch of hydrophobic amino acids [12]. The fact that phaseolin synthesis and subsequent accumulation in protein body structures requires membrane translocation also strengthens the case for signal peptide processing. The amino terminal sequences of phaseolin were found to be identical for the first 120 amino acids in all the *c*DNA clones studies [13, 14]. This region also contained the single tryptophan residue of the phaseolin protein at position 45 and thus represented a unique site for chemical cleavage and production of a small *N*-terminal peptide fragment which might facilitate sequence analysis.

The tryptophan specific reagent BNPS–skatole was used to cleave native phaseolin [21]. Under mildly denaturing conditions in the presence of a large (70-fold) excess of reagent, a group of peptides was isolated by reversed-phase HPLC (Fig. 5). The largest peak from this peptide fraction was rechromatographed and the major component isolated. This fraction was hydrolysed in constant boiling HCl–1.0% thioglycolic acid and amino acid analysis was performed on the hydrolysate by OPA-precolumn techniques [22]. The results of the amino acid analysis are given in Table 1 and are consistent with the amino acid composition expected from the *c*DNA nucleotide sequence [12]. The presence of alanine in this analysis suggests that it must be an *N*-terminal residue because from *c*DNA sequence analysis it is uniquely positioned farthest from the BNPS–skatole cleavage site. However, the fractional amount of alanine obtained from the hydrolysis of the skatole derived peptides (0.3 equivalents) may imply some heterogeneity at the *N*-terminus. Alanine contamination was judged not to be a factor because none was observed in control hydrolysis reactions. Partial equivalents of threonine and serine are expected in amino

Table 1. Amino acid composition of skatole peptide

Amino acid	HCl hydrolysate*	From DNA sequence
Asp†	5.1	5
Glu‡	4.9	5
Ser	3.6	4
Thr	0.6	1
Arg	0.9	1
Ala	0.3	1
Tyr	0.9	1
Phe	1.7	2
Leu	1.3	1
Pro§	0.8	1

*200 pmoles peptide hydrolysed.
†(Asp + Asn).
‡(Glu + Gln).
§Determined by Dns-Cl modification.

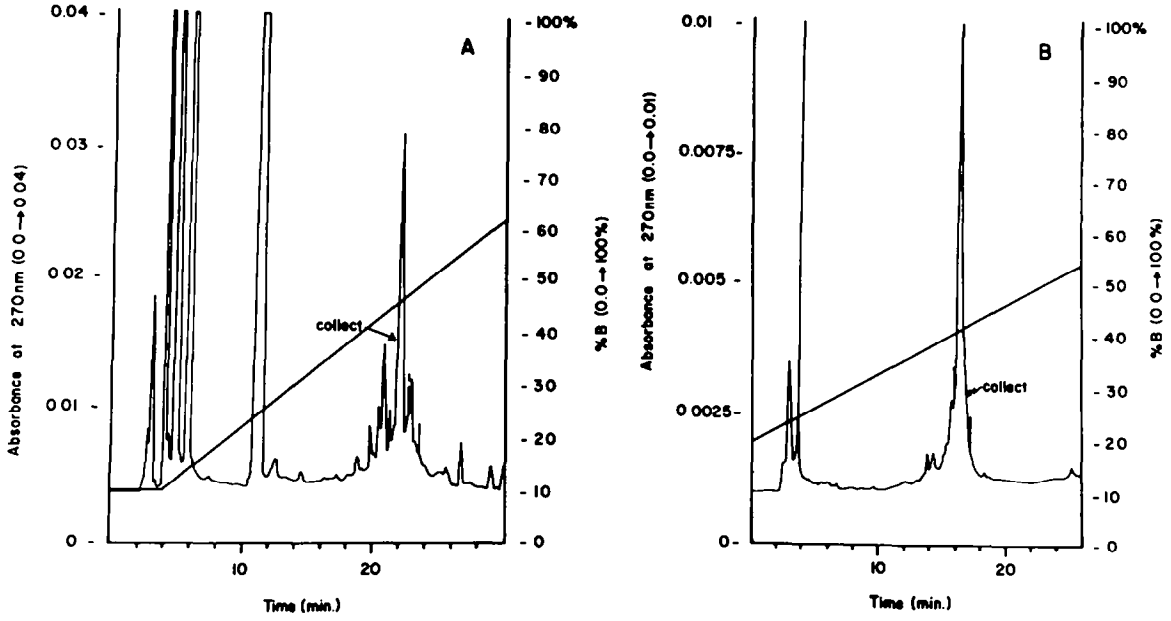

Fig. 5. Reversed-phase HPLC isolation of the peptide fragments produced from BNPS–skatole treatment of phaseolin separated on a Waters Associates 3.9 × 300 mm, 10 μm cyanopropyl bonded phase, 60 Å pore size column. Solvent A: 0.1% TFA–H_2O, solvent B: 0.1% TFA–acetonitrile, flow 1.0 ml/min, gradient as shown. A: initial isolate from crude reaction mixture; B: rechromatography of peak fraction from A.

PHOTO 26/2-B

acid analysis due to acid catalysed decomposition of these residues.

The BNPS–skatole derived peptide was also subjected to *N*-terminal sequence analysis with aminopeptidase M [23]. The time-course shown in Fig. 6 reveals rapid liberation of serine, threonine, leucine and arginine. At 180 min, the amounts of leucine and arginine are approaching 1 equivalent while the amount of threonine did not appreciably change from 0.4 equivalents after 90 min. This fractional amount of threonine represents evidence for a heterogeneous population of *N*-terminal phaseolin peptides because 1 equivalent would be expected from the *c*DNA sequence. Serine amounts continue to increase beyond 1 equivalent after 90 min because of the presence of another serine residue within the *N*-terminal region. Glutamic acid values were omitted from the graph for the sake of clarity. No alanine was observed in this experiment and since only a fractional amount of the skatole peptide pool (~ 30%) would be expected to possess an *N*-terminal alanine residue from the amino acid analysis, its absence in the aminopeptidase digestion suggests the possibility of an *N*-blocked alanine, while useful sequence information is obtained from those peptides having different *N*-termini.

Edman degradation has long been known to be the most definitive approach to peptide sequencing. Using this automated methodology, 500 pmoles of the BNPS–skatole-derived phaseolin peptide were sequenced for 16 degradative cycles with the results shown in Table 2. The presence of leucine in the first three cycles suggests *N*-terminal heterogeneity of phaseolin peptides with the ratio of T:S:L being 6:3:1. Approximately 70% (~ 350 pmoles) of the skatole-derived peptides were involved in the initial degradative cycle with no alanine residues being detected. An *N*-blocked alanine is a plausible explanation for the 30% of non-reactive peptide; this amount correlates quite well with the fractional amount of alanine observed in the amino acid hydrolysis of the skatole peptides. However, it is not uncommon to experience a low yield at the first step when sequencing.

The DEAE–Sephacel-separated phaseolin subunits were also sequenced by automated Edman techniques. These results, summarized in Table 3, are in agreement with those obtained for the skatole-derived peptides. Again, the 'ragged' nature of *N*-terminal phaseolin processing was evident; both DEAE-fractions showed two *N*-termini, threonine (major) and serine (minor), with the mixture of α and β peptides showing relatively more terminal serine. The initial yield for both sequence runs was of the order of 65%.

Two factors can be seen as potential contributors to the non-specific post-translational processing of the phaseolin subunits. Phaseolin is encoded by a closely linked family of genes which differ slightly in nucleotide sequence [13, 14]. The amino acid substitutions and deletions corresponding to the nucleotide changes may influence the signal peptide processing by changing the secondary and tertiary structure of the precursor molecule. This hypothesis can now be tested by analysing the product encoded by a single cloned phaseolin gene in yeast [24] where signal peptide processing is known to occur. The second factor is the rather unique methionine doublet at the initiation site. If signal peptide processing is affected by distance from the initiation site, then each methionine can independently act as an initiator [25] and this would certainly result in heterogeneity in signal peptide processing.

Deglycosylation of phaseolin polypeptides

Seed storage proteins have been extensively studied by gel electrophoresis techniques with banding patterns of specific cultivars providing a basis for establishing the pattern of inheritance of G1 polypeptides [4]. Because G1-phaseolin proteins are glycosylated, gene family analysis by electrophoretic techniques is subject to potential charge and molecular weight contributions introduced by glycosylation and oligosaccharide processing.

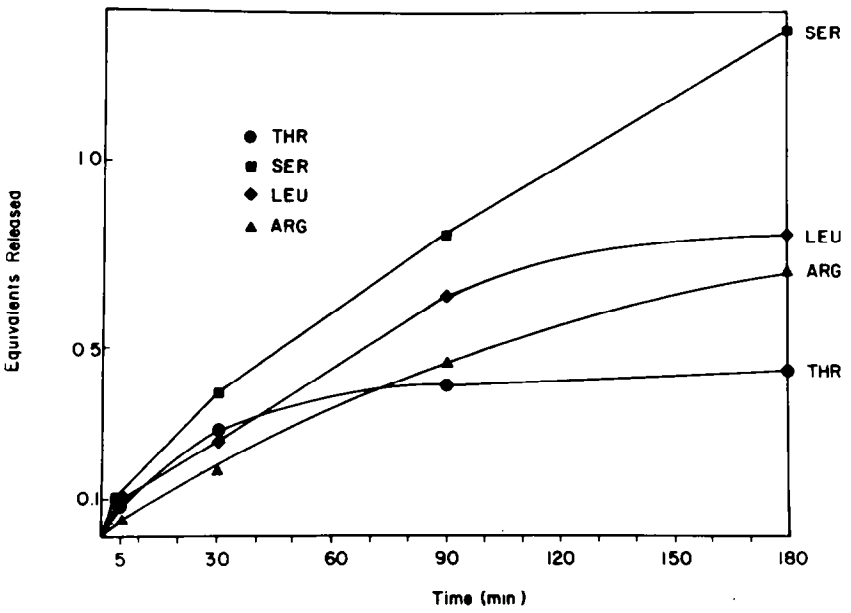

Fig. 6. Aminopeptidase M time course digestion of BNPS–skatole produced peptides.

Table 2. Amino acid sequence of skatole peptides

Cycle number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Expected from DNA	T	S	L	R	E	E	E	E	S	Q	D	N	P	F	Y	Z
Found	T	ND*	L	R	E	E	E	E	ND	Q	D	N	P	F	Y	F
	Ratio of T:S:L, 6:3:1 in first cycle															

*ND—not detected.

Table 3. Phaseolin subunit *N*-terminal amino acid sequences

Cycle number	1	2	3	4	5	6	7	8	9	10	11	12	13	14
α/β Subunits	T	S	L	R	E	E	E	E	S	Q	D/N	N	P	F
β/γ Subunits	T	S	L	R	E	E	E	E	S	Q	D	N	P	F
	Ratio of T:S, ~ 2:1 in first cycle													

N-Glycosylation is a well-known process in which the initial asparagine-linked oligosaccharide is capable of undergoing a number of processing steps [26]. Processing can produce molecular weight variants in addition to charge variants *via* the introduction of sialic acid residues, glucosamine derivatives, and mannose-6-linked phosphates and sulphates. Post-translational modifications of this type make genetic analysis of glycoprotein gene families by protein gel techniques difficult and subject to misinterpretation.

To directly analyse the gene products from multigene glycoprotein families, removal of the carbohydrate residues would be advantageous. A number of enzymes are available which accomplish deglycosylation; however, substrate specificity and steric factors have been shown to limit their overall utility [27, 28]. Chemical deglycosylation is an attractive alternative to the enzymatic approach. Under anhydrous conditions, hydrogen fluoride–pyridine [29] and trifluoromethanesulphonic acid [30, 31], both in the presence of an anisole scavenger, have proven effective on a variety of substrates.

When native phaseolin, dried over P_2O_5 under high vacuum, was treated with these reagents for 3.5 hr at 25°, only the TFMSA-treated sample produced any change in the one-dimensional gel pattern. HF–pyridine was totally ineffective while within 0.5 hr TFMSA produced two peptide components which migrated slightly faster on SDS-PAGE than the β and γ components of the native phaseolin triplet (Fig. 7). Glycoprotein staining of the gel yielded only faint bands corresponding to the position of the deglycosylated peptide components (results not shown). This residual staining may be due to the fact that TFMSA removes all of the *N*-linked oligosaccharide except for the terminal *N*-acetylglucosamine [30].

These results are very similar to those obtained from the *in vitro* translation of bean seed 16S *m*RNA which produced two polypeptides with slightly greater mobility than the β and γ subunits of native phaseolin [8]. Sequence analysis of *c*DNA clones of seed *m*RNA [13, 14] also indicates the presence of only two gene families which is in agreement with the chemical deglycosylation and *in vitro* translation experiments. Trifluoromethanesulphonic acid–anisole deglycosylation of native phaseolin therefore establishes that each of the native subunits (α, β and γ) is glycosylated to some extent since co-migration of the native and deglycosylated polypeptides is not observed. This is consistent with the results of Bollini *et al.* [18]. Glycoprotein staining of SDS-PAGE-separated native phaseolin substantiates this conclusion, however, high loadings are necessary for the γ subunit to react with the fuchsin–sulphite reagent [32].

Fig. 7. SDS-PAGE (16%) of TFMSA deglycosylated phaseolin. A, Deglycosylated phaseolin; B, native phaseolin.

Native phaseolin was fractionated into its α, β and γ subunit components by preparative SDS-PAGE techniques [33] and each individual subunit was chemically deglycosylated in order to determine the contribution of bound carbohydrate to the observed size classes. Figure 8 shows that deglycosylation of the purified α and γ subunits produce single polypeptide products; however, the relative decrease in molecular weight for the α subunit is about twice that obtained for the γ subunit. Thus, the γ subunit contains less carbohydrate. This result is consistent with the observation that glycosylation can occur at either one or both of the available sites [18]. Deglycosylation of the purified β-phaseolin subunit yields both α- and β-type polypeptides which indicates that differential glycosylation of both α- and β-polypeptides occurs. These results clearly show the effect that glycosylation has on these polypeptides and supplies further evidence supporting the claims that these phaseolin polypeptides are encoded by two unique gene families [13, 14, 18].

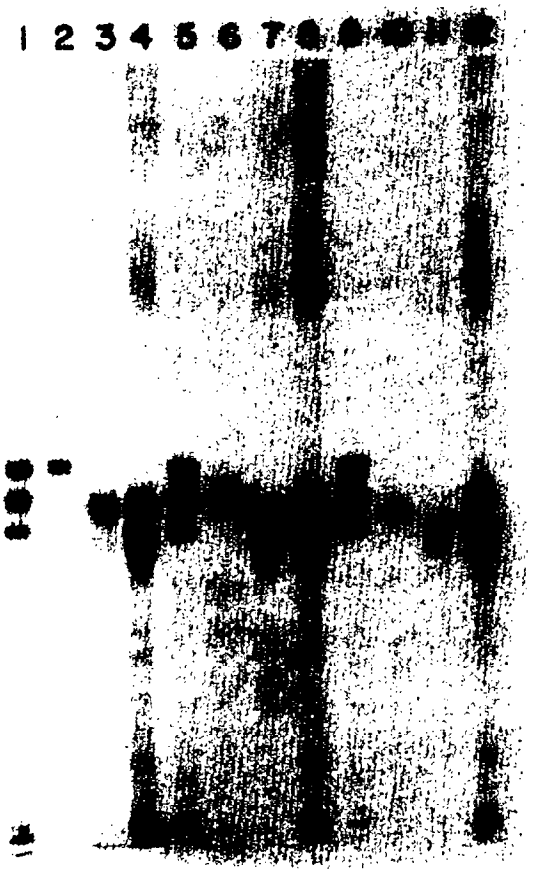

Fig. 8. SDS-PAGE (13%) of TFMSA deglycosylated phaseolin subunits. (1) native phaseolin, (2) purified α subunit, (3) deglycosylated α subunit, (4) deglycosylated native phaseolin, (5) native phaseolin, (6) purified β subunit, (7) deglycosylated β subunit, (8) deglycosylated native phaseolin, (9) native phaseolin, (10) purified γ subunit, (11) deglycosylated γ subunit, (12) deglycosylated native phaseolin.

EXPERIMENTAL

Materials. Aminopeptidase M was purchased from Boehringer-Mannheim. BNPS–skatole, trifluoroacetic acid (TFA), fluoraldehyde, amino acid standard H and PTH–amino acid standards were obtained from Pierce Chemical Co., DEAE–Sephacel from Pharmacia, and trifluoromethanesulphonic acid (TFMSA) and anisole from Aldrich Chemical Co. HPLC grade MeCN, MeOH and H_2O were purchased from Fisher Scientific.

Protein/peptide purification. Native phaseolin was extracted from dry seeds of *Phaseolus vulgaris* L. cv. Tendergreen purchased from Olds Seed Co., Madison, WI, as described [15]. Phaseolin subunits were fractionated on DEAE–Sephacel (8 M urea, pH 7.6) with a linear NaCl gradient (0–0.15 M). HPLC of phaseolin was performed on a Varian Associates Series 5000 liquid chromatograph using a Baker Wide Pore 5 μm, 4.6 × 250 mm diphenyl column with gradients of 0.1% TFA–MeCN into 0.1% TFA–H_2O. Phaseolin peptide fragments produced by BNPS–skatole cleavage were purified on a Waters μ-Bondapak 10 μm, 3.9 × 300 mm CN column with gradients of 0.1% TFA–MeCN into 0.1% TFA–H_2O. SDS-PAGE was performed as described [34] and protein bands were visualized by the silver staining procedure of Merril *et al.* [35]. SDS-PAGE separated phaseolin subunits were electroeluted and isolated as described by Hunkapiller *et al.* [33].

BNPS–skatole cleavage of phaseolin [21]. To a soln of 500 μg of native phaseolin (10 nmoles) in 500 μl of 80% HOAc was added 250 μg of BNPS–skatole (70-fold excess). The reaction was brought to 2 M guanidine–HCl by the addition of 100 mg of the denaturant and allowed to react overnight at room temp. The soln was taken to dryness, redissolved in 1.0 ml of H_2O, extracted with 2 × 300 μl $CHCl_3$ and 1 × 300 μl EtOAc. The resulting crude mixture was centrifuged briefly to sediment the insoluble material which formed during the organic extraction. The supernatant was removed and the pellet resuspended in 500 μl of 6 M guanidine–HCl.

Trifluoromethanesulphonic acid (TFMSA) deglycosylation of native phaseolin [30, 31]. A mixture of 600 μl TFMSA and 300 μl anisole was cooled at 0° and added to 4.2 mg native phaseolin which had been dried over P_2O_5 under high vacuum. The reaction was vortexed briefly to solubilize the protein and kept at 0° for 1.0 hr. At the end of this time, 3 ml of ice cold Et_2O was added with mixing followed by the careful addition of 3.0 ml of ice cold pyridine–H_2O 1:1. The Et_2O layer was separated after vortexing and the crude reaction was extracted twice with 3.0 ml portions of ether. The final soln was dialysed against 1 M NaCl followed by H_2O. The precipitated protein was lyophilized to dryness.

The purified native subunits (5–10 μg) were dried over P_2O_5 under high vacuum overnight, then treated with 50 μl TFMSA–anisole 2:1 at 0° for 1.0 hr. The reaction was quenched with 300 μl ice cold Et_2O followed by 300 μl of pyridine–H_2O 1:1. The resulting mixture was extracted × 4 with 300 μl of cold Et_2O and the aq. phase concd to 150 μl of volume. The protein was precipitated by the addition of 600 μl of MeOH, 150 μl $CHCl_3$ and 450 μl of H_2O. The soln was vortexed and centrifuged and the upper layer discarded. Addition of 450 μl of MeOH to the lower phase and centrifugation resulted in a protein pellet suitable for gel analysis [36].

Amino acid analysis. HPLC purified peptides from the BNPS–skatole digestion of native phaseolin were hydrolysed in constant boiling HCl, 1% thioglycolic acid for 22 hr at 110°. Aliquots of the hydrolysate were derivatized for 1.0 min at room temp. with an equal volume of fluoraldehyde to produce the highly fluorogenic OPA–amino acid analogues which were separated on a Varian Associates 5000 series liquid chromatograph

utilizing an Ultrasphere (5 μm, 4.6 × 250 mm ODS) column and a gradient of MeOH into 50 mM NaOAc, pH 6.8, tetrahydrofuran 99:1 as described [22]. Detection was accomplished with a Kratos-Schoeffel FS-970 fluorometer (ex 330 nm, em 419 nm) and peak areas were integrated on a Varian Associates CDS 401 data processor. Analyses were done in triplicate.

Protein/peptide sequencing. Native phaseolin (200 μg, ~ 4 nmoles) was dissolved in 98% HCO_2H and degradations were carried out on an Edman–Begg type sequenator with a stationary stainless steel reaction module as described [37]. Amino acid thiazolinones were converted in 1 N HCl at 80° for 7 min to the corresponding phenylthiohydantoins which were identified by HPLC on a Zorbax ODS (Dupont) column at 45° using an NH_4OAc–MeCN gradient. The Laboratory Data Control system was equipped with a chromatography control module and dual absorbance monitors set at 269 and 320 nm. The initial yield of *N*-terminal amino acids was about 65%. BNPS–skatole peptides (1.5 μg, ~ 500 pmoles) were dissolved in 20 μl 0.1% HOAc and sequenced on an Applied Biosystems 470A gas-phase sequencer. The isolated PTH–amino acids were identified by HPLC on an Ultrasphere (5 μm, 4.6 × 250 mm ODS) column at 40°, employing a NaOAc–MeCN gradient. A Varian Associates 5000 series chromatograph was interfaced with a WISP 710B (Waters Associates) autoinjector and samples were detected at 269 nm on a Hitachi variable wavelength UV monitor. Approximately 70% of the applied peptide underwent initial *N*-terminal cleavage.

REFERENCES

1. Ma, Y. and Bliss, F. A. (1978) *Crop. Sci.* **17**, 431.
2. Hall, T. C., McLeester, R. C. and Bliss, F. A. (1977) *Plant Physiol.* **59**, 1122.
3. Blagrove, R. J., Lilley, G. G., Van Donkelaar, A., Sun, S. M. and Hall, T. C. (1984) *Int. J. Biol. Macromol.* **6**, 137.
4. Brown, J. W. S., Ma, Y., Bliss, F. A. and Hall, T. C. (1981) *Theor. Appl. Genet.* **59**, 83.
5. Ma, Y., Bliss, F. A. and Hall, T. C. (1980) *Plant Physiol.* **66**, 897.
6. Bollini, R. and Vitale, A. (1981) *Physiol. Plant.* **52**, 96.
7. Hall, T. C., Sun, S. M., Buchbinder, B. U., Pyne, J. W., Bliss, F. A. and Kemp, J. D. (1980) in *Genome Organization and Expression in Plants* (Leaver, C. J., ed.), p. 259. Plenum, New York.
8. Hall, T. C., Ma, Y., Buchbinder, B. U., Pyne, J. W., Sun, S. M. and Bliss, F. A. (1978) *Proc. Natn. Acad. Sci. U.S.A.* **75**, 3196.
9. Sun, S. M., Mutschler, M. A., Bliss, F. A. and Hall, T. C. (1978) *Plant Physiol.* **61**, 918.
10. Brown, J. W. S., Bliss, F. A. and Hall, T. C. (1981) *Theor. Appl. Genet.* **60**, 251.
11. Sun, S. M., Slightom, J. L. and Hall, T. C. (1981) *Nature* **289**, 37.
12. Slightom, J. L., Sun, S. M. and Hall, T. C. (1983) *Proc. Natn. Acad. Sci. U.S.A.* **80**, 1987.
13. Hall, T. C., Slightom, J. L., Ersland, D. R., Murray, M. G., Hoffman, L. M., Adang, M. J., Brown, J. W. S., Ma, Y., Matthews, J. A., Cramer, J. H., Barker, R. F., Sutton, D. W. and Kemp, J. D. (1983) in *Structure and Function of Plant Genomes* (Ciferi, O. and Dure, L., III, eds), p. 123. Plenum, New York.
14. Slightom, J. L., Drong, R. F., Klassy, R. C. and Hoffman, L. M. (1985) *Nucl. Acids Res.* **13**, 6483.
15. Sun, S. M. and Hall, T. C. (1975) *J. Agric. Food Chem.* **23**, 184.
16. Sun, S. M., McLeester, R. C., Bliss, F. A. and Hall, T. C. (1974) *J. Biol. Chem.* **249**, 2118.
17. Bourdillon, J. (1951) *J. Biol. Chem.* **189**, 65.
18. Bollini, R., Vitale, A. and Chrispeels, M. J. (1983) *J. Cell Biol.* **96**, 999.
19. Lioi, L. and Bollini, R. (1984) *Plant Mol. Biol.* **3**, 345.
20. Regnier, F. E. (1983) *Methods Enzymol.* **91**, 137.
21. Fontana, A. (1972) *Methods Enzymol.* **25**, 419.
22. Umagot, H. and Kucera, P. (1982) *J. Chromatogr.* **239**, 463.
23. Jones, B. N., Pääbo, S. and Stein, S. (1981) *J. Liq. Chromatogr.* **4**, 565.
24. Cramer, J. H., Lea, K. and Slightom, J. L. (1985) *Proc. Natn. Acad. Sci. U.S.A.* **82**, 334.
25. Kozak, M. (1984) *Nucl. Acids Res.* **12**, 3873.
26. Hubbard, S. C. and Ivatt, R. J. (1981) *Ann. Rev. Biochem.* **50**, 555.
27. Plummer, T. H., Jr. and Tarentino, A. L. (1981) *J. Biol. Chem.* **256**, 10243.
28. Tarentino, A. L. and Plummer, T. H., Jr. (1982) *J. Biol. Chem.* **257**, 10776.
29. Coudron, C., Ellis, K., Philipson, L. and Schwartz, N. B. (1980) *Biochem. Biophys. Res. Commun.* **92**, 618.
30. Edge, A. S. B., Faltynek, C. R., Hof, L., Reichert, L. E., Jr. and Weber, P. (1981) *Analyt. Biochem.* **118**, 131.
31. Desai, N. N., Allen, A. K. and Neuberger, A. (1983) *Biochem. J.* **211**, 273.
32. Fairbanks, G., Steck, T. L. and Wallach, D. F. H. (1971) *Biochemistry* **10**, 2606.
33. Hunkapiller, M. W., Lujan, E., Ostrander, F. and Hood, L. E. (1983) *Methods Enzymol.* **93**, 227.
34. Laemmli, U. K. (1970) *Nature* **277**, 680.
35. Merril, C. R., Goldman, D., Sedman, S. A. and Ebert, M. H. (1981) *Science* **211**, 1437.
36. Wessel, D. and Flügge, U. I. (1984) *Analyt. Biochem.* **138**, 141.
37. Inglis, A. S., Sutherland, W. J. and Woods, E. F. (1983) *Proc. Aust. Biochem. Soc.* **15**, 31.