Biochem. J. (1982) **207**, 629–632
Printed in Great Britain

The post-translational proteolysis of the subunits of vicilin from pea (*Pisum sativum* L.)

John A. GATEHOUSE, Grantley W. LYCETT, Ronald R. D. CROY and Donald BOULTER
Department of Botany, University of Durham, Science Laboratories, South Road, Durham DH 1 3LE, U.K.

(*Received 12 August 1982/August 13 September 1982*)

Tryptic-peptide profiles and amino acid sequencing of purified pea (*Pisum sativum* L.) vicilin subunits were used to show that their sequences were interrelated. Comparison with the nucleotide sequence of a cloned vicilin complementary DNA (mRNA) showed that all vicilin subunits could be derived from 50000-M_r precursors containing up to two sites for post-translational proteolytic cleavage, and allowed these subunits to be located relative to the precursor.

The storage protein vicilin is an abundant component of pea (*Pisum sativum* L.) seeds, its major subunits of M_r 50000, 33000, 19000, 16000, 13500 and 12500 being prominent components of the band pattern on SDS/polyacrylamide-gel electrophoresis of pea seed extracts. A structural model for vicilin, envisaging the molecule as a multimer, probably a trimer, of 50000-M_r subunits has been put forward (Gatehouse *et al.*, 1981), and it suggests that molecules are assembled from a range of different 50000-M_r subunits of slightly differing properties, some of which may contain 'nicks' in the polypeptide chains accounting for the polypeptides of $M_r < 50000$ observed in purified vicilin. In support of this model, only precursor subunits of apparent $M_r \simeq 50000$ (50000 and 47000) have been shown to be produced *in vivo* and by cell-free synthesis *in vitro* programmed with pea seed RNA (Croy *et al.*, 1980*b*). Also, some evidence has been presented that the subunits of $<50000\,M_r$ in vicilin are produced by post-translational cleavage of the 50000-M_r precursor subunits (Gatehouse *et al.*, 1981; Chrispeels *et al.*, 1982; Spencer *et al.*, 1980), possibly in a manner analogous to the well-established post-translational proteolytic cleavage of the legumin 60000-M_r precursor (Croy *et al.*, 1980*a*). Further evidence is presented here for the processing of vicilin polypeptides through comparisons of protein sequence and sequences predicted by cDNA species encoding vicilin precursor polypeptides, and a model showing the derivation of the polypeptides of $<50000\,M_r$ from the 50000-M_r precursors is put forward.

Abbreviations used: SDS, sodium dodecyl sulphate; cDNA, complementary DNA; h.p.l.c., high-pressure liquid chromatography.

Experimental

The isolation and characterization of the vicilin cDNA clone pRC 2.2.1, containing an insert of approx. 1000 base-pairs in pBR322, was described previously (Croy *et al.*, 1982). cDNA was sequenced by the methods of Maxam & Gilbert (1980) and Seif *et al.* (1980). Vicilin was purified from mature seeds of *Pisum sativum* var. Feltham First as described by Gatehouse *et al.* (1981) and separated into its component subunits by means of ion-exchange chromatography on DEAE-cellulose (Whatman) at pH 8.5 in 50 mM-Tris/HCl/8 M-urea, with a gradient of NaCl (0–0.4 M) for elution, followed by gel-filtration chromatography on Sephacryl S-200 (Pharmacia) in 70% (v/v) formic acid, as outlined by Hirano *et al.* (1982). The homogeneity of subunit fractions was verified by SDS/polyacrylamide-gel electrophoresis (Laemmli, 1970).

Purified vicilin subunits were digested with trypsin treated with 1-chloro-4-phenyl-3-L-tosylamidobutan-2-one ('TPCK') (Millipore) at an enzyme/protein ratio of 1:50 (w/w), at room temperature for 16 h in 0.1 M-NH_4HCO_3, pH 8.2. The mixture was then freeze-dried and the peptides were separated by h.p.l.c. on a reverse-phase column [Varian Micropak MCH-10 (C_{18}, 10 μ)]. Samples were dissolved in 0.1% trifluoroacetic acid (Mahoney & Hermodson, 1980) and injected in a volume of 100 μl. The column was run at a flow rate of 1.0 ml/min at 30°C on a Varian LC5060 chromatography system and peptides were eluted with a 0–70% acetonitrile gradient (in 0.1% trifluoroacetic acid) over 120 min. The peak elution profiles were used as peptide maps for comparative purposes. Peptides were sequenced by the manual diaminobenzoyl isothiocyanate method described by Chang *et al.* (1978).

0306-3275/82/120629-04$01.50/1 © 1982 The Biochemical Society

Results and discussion

Polypeptides homogeneous by SDS/polyacrylamide-gel electrophoresis were obtained corresponding to all the vicilin subunits except the minor components of M_r 35000 and 30000, which were only partially purified. H.p.l.c. analysis of tryptic peptides from the purified subunits (results not shown) suggested that the amino acid sequences of the various vicilin subunits were interrelated. Subunits of M_r 19000 and 13500 had tryptic peptides in common with those of the 33000-M_r subunit, but had dissimilar peptide profiles themselves. Tryptic peptides of the 16000- and 12500-M_r subunits were very similar, but different from those of the 33000-, 19000- and 13500-M_r subunit. In agreement with the suggestion that vicilin subunits of $M_r < 50000$ were derived from precursors of $M_r \simeq 50000$ (Gatehouse *et al.*, 1981), all these subunits had tryptic peptides in common with those of the 500000-M_r subunits. Subunits of essentially the same M_r but differing charge, as separated by the ion-exchange-chromatography step, had very similar tryptic peptide profiles. The minor subunits of M_r 35000 were tentatively shown to be similar to the 33000-M_r subunits, but those of 30000M_r wsre only partially similar, apparently having in addition peptides in common with the 16000- and 13500-M_r subunits.

Amino acid-sequence data from purified tryptic peptides was used to confirm the relationships between the major vicilin subunits. The close similarity of the 16000- and 12500-M_r subunits was shown by the six identical sequences of five to seven amino acids from corresponding tryptic peptides, listed in Table 1. The 16000-M_r subunit is known to be glycosylated (Badenoch-Jones *et al.*, 1981), and this may be the only significant difference between the two subunits. Corresponding peptides from the 33000- and 19000-M_r, and 33000- and 13500-M_r subunits also contained identical amino acid sequences (Table 1), but no peptides from any of the 19000-, 16000- or 13500-M_r subunits had sequence similarity. Finally, peptides from the 50000-M_r subunits were shown to be very similar in sequence to peptides from the 33000-M_r, 19000-M_r and 16000-M_r subunits (Table 1). These sequences allowed the model for the derivation of vicilin subunits shown in Fig. 1 to be deduced. If the 50000-M_r vicilin precursor polypeptides are assumed to contain up to two potential sites for proteolytic cleavage, generating three fragments, α, β and γ, the subunits may be designated thus: 50000 M_r, $\alpha+\beta+\gamma$ (i.e. no cleavage); 33000M_r, $\alpha+\beta$; 19000M_r, α; 13500M_r, β; 16000 and 12500M_r, γ. Comparisons based on tryptic peptide profiles alone tentatively designate the 35000-M_r subunits as $\alpha+\beta$ and the 30000-M_r subunits as $\beta+\gamma$.

The ordering of the subunits relative to a 50000-M_r precursor (shown in Fig. 1) was deduced

Table 1. *Amino acid sequences of identical tryptic peptides isolated from digests of different purified vicilin subunits*

Where alternative residues are given in parentheses, both residues were found in different peptides. In (*e*) the peptides given are closely similar rather than identical.

(a) 16000-M_r and 12500-M_r subunits (also predicted by cDNA sequence)

Q V Q L/I Y R
E L/I A F P G S...
Q S Y F A N A...
N F L/I A G...
D H E E E...
L/I S P G D...

(*b*) 33000-M_r and 19000-M_r subunits

F D Q R
N S F N L/I E R
L/I F E N L/I Q N Y R
N L/I L/I E A S F N...
V L/I D L/I A L/I P V...
L/I P A G T L/I A Q N...

(*c*) 33000-M_r and 13500-M_r subunits (also predicted by cDNA sequence)

E Q L/I E E E L/I S K
S V S S E S G P F N L/I R
F F E L/I T P E K
N E N Q G K
S S D P L/I Y S N N...
V S R

(*d*) 50000-M_r, 33000- and 19000-M_r subunits

N L/I L/I E A S F N...
V L/I D L/I A L/I P...
A T L/I T Y...
F Q T L/I F...

(*e*) 50000-M_r and 16000-M_r subunits (also predicted by cDNA sequence)

{ L T P G N V F V L/I P A G R 50000-M_r
{ L/I S P G D V F V L/I P A G H... 16000-M_r

{ ($\frac{D}{E}$) L/I A F P... 50000-M_r
{ E L/I A F P G S... 16000-M_r

by comparison of peptide sequences with the relevant nucleotide sequences from the vicilin cDNA clone pRC2.2.1, which has previously been shown to hybrid-select mRNA species coding for vicilin 50000-M_r precursor (Croy *et al.*, 1982). As shown in Fig. 1, the 5′-end of this cDNA contained sequences coding for peptides found in the 33000-M_r subunit. Extending this sequence towards the *N*-terminal of the 33000-M_r subunit gave overlap with 19000-M_r subunit peptides, whereas extending the sequence towards the *C*-terminal of the 33000-M_r subunits, contiguous with the middle sequence of the cDNA, gave overlap with the 13500-M_r sequence (Hirano *et al.*, 1982). Sequences towards the 3′-end of the cDNA corresponded to those of tryptic peptides from the 16000M_r-subunit (Fig. 1). The order of the subunits is thus as shown

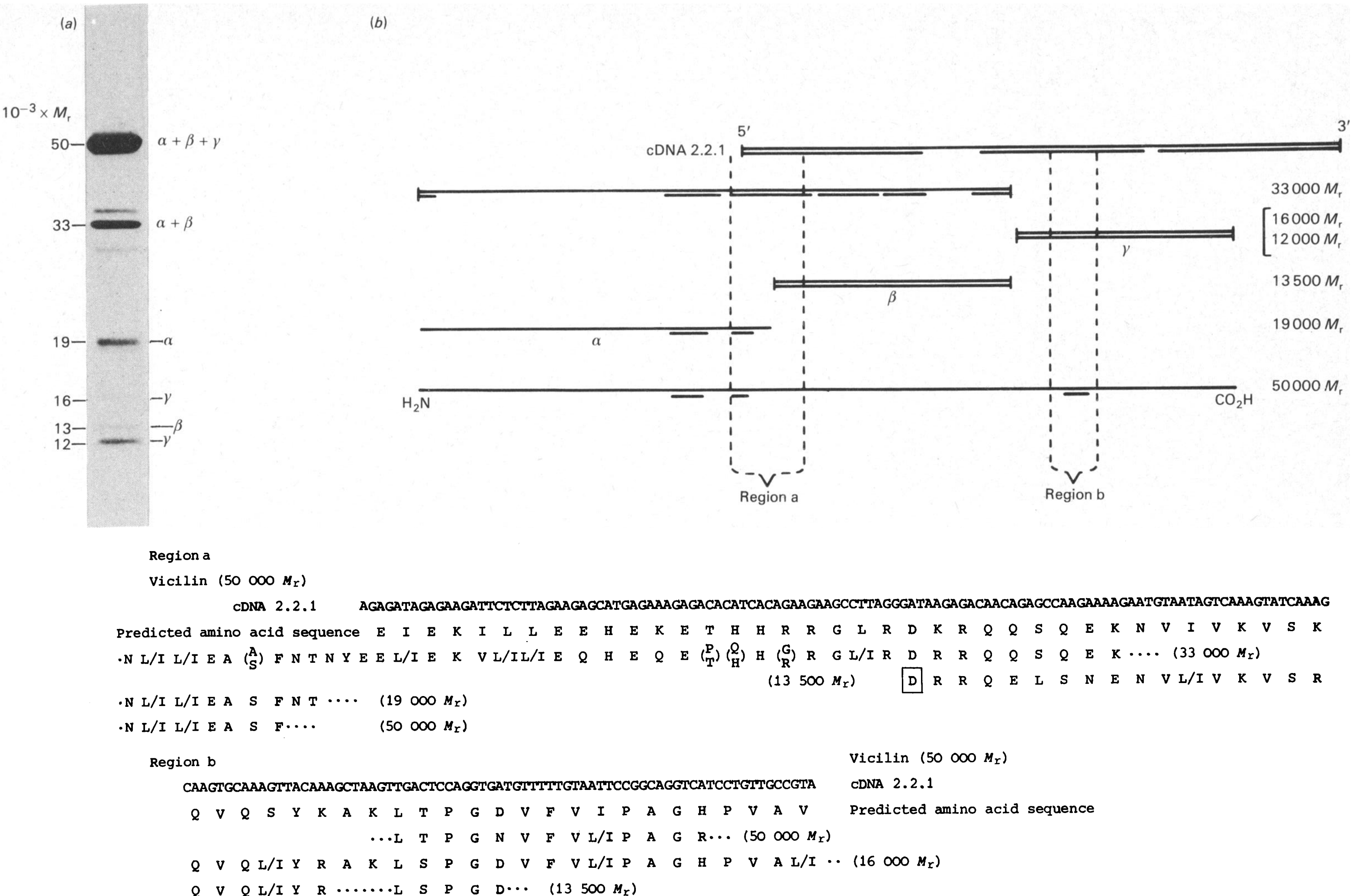

Fig. 1. *(a) SDS/polyacrylamide-gel electrophoresis of purified vicilin, showing the one-dimensional separation of the component subunits and the assignment of polypeptide sequences, and (b) proposed vicilin model, showing the origins and alignment of the major subunits of less than 50000 M_r, relative to the 50000 M_r subunits and precursors*

In (*b*), sequenced regions are represented by double lines. Regions a and b are shown as full sequences. Where alternative residues are given in parentheses, both residues were found in different peptides.

in Fig. 1: H_2N-α-β-γ-CO_2H. Although the overall agreement between the amino acid sequence predicted from the vicilin cDNA and the sequenced vicilin peptides is good, discrepancies in sequence (mostly conservative amino acid substitutions) occur, as shown in regions a and b in Fig. 1. Since at least 20 vicilin subunits may be observed on two-dimensional gel analysis (result not shown), such sequence micro-heterogeneity is expected. Complete amino acid-sequence data for the 13500-M_r (β) subunit has already been published (Hirano *et al.*, 1982); the complete sequences for the other subunits have yet to be published.

References

Badenoch-Jones, J., Spencer, D., Higgins, T. J. V. & Millerd, A. (1981) *Planta* **153**, 201–209

Chang, J. Y., Braver, D. & Wittman-Liebold, B. (1978) *FEBS Lett.* **93**, 205–214

Chrispeels, M. J., Higgins, T. J. V., Craig, S. & Spencer, D. (1982) *J. Cell Biol.* **93**, 5–14

Croy, R. R. D., Gatehouse, J. A., Evans, I. M. & Boulter, D. (1980*a*) *Planta* **148**, 48–56

Croy, R. R. D., Gatehouse, J. A., Evans, I. M. & Boulter, D. (1980*b*) *Planta* **148**, 57–63

Croy, R. R. D., Lycett, G. W., Gatehouse, J. A., Yarwood, J. N. & Boulter, D. (1982) *Nature (London)* **295**, 76–79

Gatehouse, J. A., Croy, R. R. D., Morton, H., Tyler, M. & Boulter, D. (1981) *Eur. J. Biochem.* **118**, 627–633

Hirano, H., Gatehouse, J. A. & Boulter, D. (1982) *FEBS Lett.* in the press

Laemmli, U. K. (1970) *Nature (London)* **227**, 680–685

Mahoney, W. C. & Hermodson, M. A. (1980) *J. Biol. Chem.* **255**, 11199–11203

Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 449–560

Seif, I., Khoury, G. & Dhar, R. (1980) *Nucleic Acids Res.* **8**, 2225–2240

Spencer, D., Higgins, T. J. V., Button, S. C. & Davey, R. (1980) *Plant Physiol.* **66**, 510–515