Phytochemistry, 1972, Vol 11, pp 59 to 64 Pergamon Press Printed in England

THE STRUCTURE OF VICILIN OF *VICIA FABA*

C. J. BAILEY* and D BOULTER

Department of Botany, University of Durham, Durham City, England

(*Received* 28 *May* 1971)

Abstract—Four sub-units of vicilin were separated by sodium dodecyl sulphate gel electrophoresis and their molecular weights determined by calibration The amino acid composition of vicilin was determined by ion-exchange chromatography and by other methods for half-cystine and tryptophan Tryptic peptide maps and *N*-terminal analysis following cyanogen bromide cleavage, were used to deduce a chemical molecular weight of vicilin of 100–130 $\times$ 10^3 These structural results on vicilin are discussed in relationship to its biological role

INTRODUCTION

THE SEED storage globulins of legumes have been classically divided into two fractions, legumin and vicilin [1,2] In a previous paper, legumin from broad beans has been shown to be a homogeneous protein with a complex structure [3] In this paper we report the results of a similar investigation of vicilin

RESULTS

The purity of preparations of vicilin from *Vicia faba* was examined by polyacrylamide gel electrophoresis. With acrylamide concentrations between 4·2 and 7·5%, vicilin migrated in the gel as a single band (Fig 1) However, the width of the band was so great that it was unlikely that only one component was present, nevertheless, no fractionation of this band was achieved using a variety of experimental conditions. When *S*-carboxymethyl vicilin (CM-vicilin) or vicilin preparations containing 1% (v/v) 2-mercaptoethanol were subjected to sodium dodecyl sulphate (SDS) gel electrophoresis, four main bands (sub-units) were observed (Fig. 1). The molecular weights of these sub-units, determined by calibration, were 66,000, 60,000, 56,000 and 36,000, and dye-binding measurements indicated that they were present in about equimolar proportions, traces of other polypeptides were also observed on the original gels Correspondingly, vicilin contained four different *N*-terminal amino acids in major yields together with several others in much lower yields (Table 1)

The amino acid composition of vicilin was determined by ion-exchange chromatography of acid hydrolysates of CM-vicilin (Table 2), and by other methods for half-cystine and tryptophan. A very large discrepancy in the values for the total half-cystine content was found between the cysteic acid and the *S*-carboxymethyl cysteine methods (Table 3). Examination of hydrolysates of performic-oxidized vicilin by high-voltage paper electrophoresis at pH = 2 showed the presence of three substances, which moved towards the anode One of these was cysteic acid, the others had mobilities the same as those of the

* Present address Department of Biochemistry, Trinity College, Dublin, Republic of Ireland.

[1] T B OSBORNE, *The Vegetable Proteins*, Longmans-Green, New York (1924).
[2] C E DANIELSSON, *Biochem J* **44**, 387 (1949)
[3] C J BAILEY and D BOULTER, *Eur J Biochem* **17**, 460 (1970)

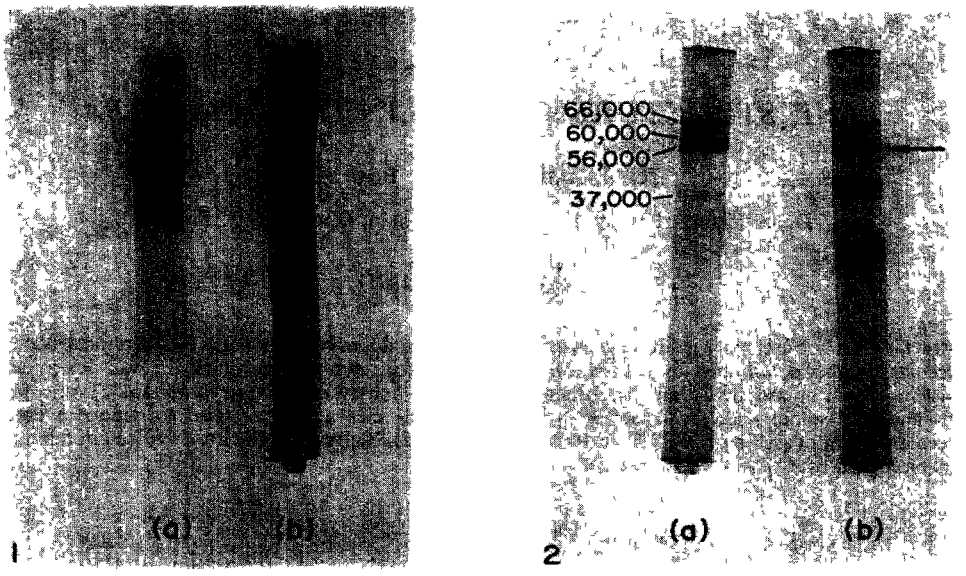

FIG 1 ACRYLAMIDE GEL ELECTROPHORESIS OF NATIVE VICILIN
(a) On 7 5% gel (run with a spacer gel) (b) On 4 2% gel (no spacer)

FIG 2 SDS GEL ELECTROPHORESIS OF CM-VICILIN AND CM-GLOBULIN
(a) CM-vicilin (b) CM-globulin, the arrow indicates the band which is common to both legumin[3] and vicilin

TABLE 1 *N*-TERMINAL ANALYSES OF VICILIN PREPARATIONS

Preparation	Dansyl amino acids found	*N*-terminal amino acids deduced
Vicilin	Leu 3+, Thr 3+, Ser 2+, bis-Lys 2+, Asp+, Glu +, Val s tr , Gly s tr , Ala s tr , *o*-Tyr v g , ϵ-Lys v g	Leu, Thr, Ser, Lys
CNBr-cleaved, succinyl CM-vicilin	Leu 3+, Val +, Ser 2+, Thr 2+, ϵ-Lys 2+, *o*-Tyr 3+	Leu, Val
Succinyl CM-vicilin	ϵ-Lys+	
Succinyl CM-vicilin (CNBr control) treated with formic acid	Ser +, Thr +, ϵ-Lys 2+, *o*-Tyr 2+	

The fluorescent intensities of the spots are classified on the following scale v g (very great), 3+, 2+, +, s tr (strong trace), tr (trace) Intensities below + are not usually considered significant

TABLE 2 AMINO ACID ANALYSIS OF HYDROLYSATES OF CM-VICILIN

Amino acid	Recovery (g amino acid residue/100 g protein)	
	20 hr	70 hr
Aspartic acid	10 1 ± 1 25	12 3 ± 0 7
Threonine	2 62 ± 0 14	2 58 ± 0 03
Serine	4 32 ± 0 21	3 06 ± 0 17
Glutamic acid	17 05 ± 0 25	17 1 ± 0 2
Proline	6 25 ± 0 25	5 67 ± 0 44
Glycine	2 64 ± 0 04	2 98 ± 0 03
Alanine	2 40 ± 0 17	2 65 ± 0 04
Valine	4 59 ± 0 49	5 35 ± 0 03
Methionine	0 19 ± 0 01	0 20 ± 0 02
Isoleucine	4 77 ± 0 21	5 73 ± 0 1
Leucine	8 74 ± 0 08	9 71 ± 0 02
Tyrosine	3 73 ± 0 10	2 67 ± 0 03
Phenylalanine	5 65 ± 0 08	6 83 ± 0 02
Lysine	7 38 ± 0 00	8 25 ± 0 22
Histidine	2 35 ± 0 07	2 72 ± 0 00
Arginine	8 33 ± 0 03	9 21 ± 0 09
	91 11	97 01

The figures are the mean values of duplicate analyses after hydrolysis in 6 N HCl at 108° *in vacuo* for the stated times

products obtained by sulphate esterification of threonine and serine [4] As all these compounds are eluted at the same position on sulphonic acid resins,[4] the 'cysteic acid' content of performic-oxidized vicilin must be an over-estimate of the half-cystine; for this reason the total half-cystine value obtained by determination of half-cystine as *S*-carboxymethylcysteine has been preferred (Table 4).

Carbohydrate analysis showed that vicilin contained small but significant quantities of neutral sugars (0·5% w/w) and insignificant quantities of hexosamines (≤0·2% w/w).

TABLE 3 CYSTEINE AND TOTAL HALF-CYSTINE CONTENT OF VICILIN

Method	Amount measured (g amino acid residue/ 100 g protein)
Total half-cystine as *S*-carboxymethyl cysteine*	0 13 ± 0 05
Total half-cystine as cysteic acid	0 83 ± 0 05
Cysteine by DTNB† titration	0 00
Cysteine by DTNB titration in the presence of 1% sodium dodecyl sulphate	0 00

* Mean value of duplicate analyses for 70 hr and a single analysis of 20 hr hydrolyses

† DTNB Di-thiobisnitrobenzoic acid

[4] K MURRAY and C MILSTEIN, *Biochem J.* **105**, 491 (1967)

TABLE 4 THE AMINO ACID COMPOSITION OF VICILIN

Amino acid	Content* (g/100 g protein)	Content* (mole/10^5 g protein)
Aspartic acid	12 3	107
Threonine†	2 62	25 9
Serine†	4 77	54 8
Glutamic acid	17 1	133
Proline	6 25	64 4
Glycine	2 98	52 3
Alanine	2 65	37 2
Valine	5 35	54 1
Methionine	0 20	1 5
Isoleucine	5 73	50 7
Leucine	9 71	85 9
Tyrosine†	4 10	25 2
Phenylalanine	6 83	46 5
Lysine	8 25	64 4
Histidine	2 72	19 8
Arginine	9 21	59 2
Cystine/2	0 13	1 3
Tryptophan‡	0 0	0
	100 90	

* The highest values from Table 2 have been taken unless otherwise stated

† Extrapolated to zero time

‡ Tryptophan was determined by spectrophotometric titration with N-bromosuccinimide,[5] UV analysis by the Goodwin-Morton method,[6] and colorimetrically,[7] the values obtained were 0($\leq$0 04%), 0 17% and 0 07% In view of the difficulties of tryptophan analysis in the presence of high molar ratios of tyrosine and of the occasional false positive result recorded by the colorimetric method (e g ref 8), we have preferred the result from N-bromosuccinimide titration

About 125 ninhydrin-positive peptides were resolved in peptide maps of tryptic digests of CM-vicilin Autoradiographs of the peptide maps prepared from similar digests of ^{14}C-carboxymethylated vicilin and ^{12}C-carboxymethylated-^{35}S-vicilin were complex although similar, about 30–40 peptides had some radioactivity in both the ^{14}C and ^{35}S preparations, but none were sufficiently active to be major peptides

When the *N*-terminal amino acids of CM-vicilin were first blocked by succinylation and the preparation digested by the methionine-specific CNBr method, four *N*-terminal amino acids were found (Table 1) Succinyl CM-vicilin contained only blocked *N*-termini, but after treatment of similar preparations with formic acid as a control of the CNBr reaction, *N*-terminal serine and threonine were present It appeared therefore, that these amino acids were released by the action of formic acid, and that CNBr treatment itself releases only *N*-terminal leucine and valine as a consequence of methionyl peptide bond cleavage.

[5] A PATCHORNIK, W B LAWSON and B WITKOP, *J Am Chem Soc* **80**, 4747 (1958)
[6] T W GOODWIN and R A MORTON, *Biochem J.* **40**, 628 (1946)
[7] J R SPIES and D C CHAMBERS, *Analyt Chem* **20**, 30 (1948)
[8] E A NOLTMANN, T A MAHOWALD and S A KUBY, *J Biol Chem* **237**, 1146 (1962)

DISCUSSION

The very wide diffuse band obtained on gel electrophoresis of vicilin is not typical of a pure protein. However, since it was found impossible to fractionate the protein of the band by variation in the experimental conditions, it is unlikely that protein impurities were the reasons for the phenomenon Rather more likely are the possibilities that vicilin is either a complex mixture of related proteins, or that during electrophoresis complex association/dissociation reactions took place between vicilin molecules or sub-units. Examination of the dissociated vicilin on SDS gels and *N*-terminal amino acid analyses showed that four sub-units were present in significant quantities, and that other polypeptides were present in such smaller quantities These latter are impurities in the sense that they could not be present in other than trace molar quantities in a native molecule having a molecular weight of 186,000 daltons.

Furthermore, the large number of cystine peptides observed on autoradiography of ^{14}C and ^{35}S-tryptic peptide maps is not consistent with a simple protein containing only 1·3 mole of half-cystine per 10^5 g (Table 4) The observed pattern could only arise from a complex mixture of cystine-containing impurities, each present in a small quantity. It seems likely that the approx. 2·4 residues per mole of half-cystine indicated by the analysis of vicilin (Table 3), are not present in the major sub-unit polypeptides. In considering autoradiographs of the ^{14}C- and ^{35}S-tryptic digest maps, these should be similar with respect to the cystine peptides but the ^{35}S map should also contain methionine peptides known to be present from the CNBr cleavage experiment However, no differences corresponding to methionine peptides on the ^{35}S map were observed, probably because the specific activity of methionyl-^{35}S was only 0 04 relative to cystinyl-^{35}S after labelling *in vivo* using the cotyledon slice system.[9] It is of interest to note that both *in vitro* and *in vivo* labelling gave the same results.

The most likely impurity to be present in the vicilin preparation is the similar storage protein, legumin. This however, was not present, since its characteristic 23,000 dalton polypeptide[3] was absent from the SDS gel of vicilin (Fig 1) Comparison of this gel with that of Fig. 1 of Ref. 3 shows that both proteins have a sub-unit of molecular weight 56,000. This band was not resolved into more than one component when total globulin samples (i e. mixtures of vicilin and legumin) were submitted to electrophoresis Several workers[10,11] have suggested that legumin and vicilin are similar, and the present results give a possible molecular explanation of the similarity, although it remains to be shown whether or not the 56,000 component is identical in both proteins

A number of workers have measured the molecular weight of vicilin as 186,000 [2,12] The molecular weight sum of the four sub-units observed on SDS gels in our experiments is 219,000 Clearly the physical unit observed in the ultracentrifuge cannot contain stoichimetric quantities of all four components Of the possible explanations for this discrepancy, simple heterozygosity seems unlikely in view of differing molecular weights of the observed sub-units, but the evidence is consistent with the view that vicilin is polymorphic Complex polymorphism might explain the diffuse band found on gels.

The observed numbers of tryptic peptides (approximately 125) and of new end groups following CNBr cleavage (two), both suggest a chemical molecular weight of about 100–130,000. As this is much less than the value of 186,000 from physical methods, it indicates

[9] C J Bailey and D Boulter, *Planta* **95**, 103 (1970)
[10] J Klotz and V Turkova, *Biologia Pl* **5**, 29 (1963)
[11] P. Jackson, D Boulter and D A Thurman, *New Phytol* **68**, 25 (1969).
[12] A D Shutov and I A. Vaintraub, *Biokhimiya* **31**, 726 (1966).

that the four major polypeptides contain extensive regions of identical amino acid sequence In view of this chemical evidence, the four cistrons which code for vicilin are probably related The observed complexity of the protein could be explained at the genetic level by duplication(s) of existing gene(s) with subsequent independent development. This process of gene duplication could have a selective advantage for the plant Thus, the developing cotyledon mainly synthesises storage protein which represents a large proportion of the total protein synthesized and the possession of multiple genes coding for storage protein could be useful in maximizing the rate of seed development Duplication of genes and parts of genes seems to have occurred frequently during evolution [13]

Other workers have also reported evidence of vicilin heterogeneity. Most notably Vaintraub, Shutov and Klimenko,[14] have fractionated vicilin from *Vicia sativa* and *V. ervilia*, the fractionation of vicilin from *Pisum sativum* is also reported [15] Further studies are needed to produce a satisfactory biochemical description of the storage globulin known as vicilin

EXPERIMENTAL

Seeds of *Vicia faba* L (var Triple White) were obtained from the Tyneside Seed Stores Plants were grown from this seed in the University Botanic Garden, Durham, during the summer of 1969

Purification of vicilin This started from the purified globulin fraction described previously [3] The total precipitate at 70% $(NH_4)_2SO_4$ saturation from 300 g of seed, was suspended in 0 2 M NaCl (1 l.) and adjusted to pH 4 7 with acetic acid After clarification at 23,000 *g* for 1 hr, the solution was dialysed overnight against tap H_2O, then for 2 days with 2 changes of distilled H_2O The globulin precipitate was taken up in 0 2 M NaCl (1 l) at pH 4 7, centrifuged and dialysed as before This isoelectric precipitation and dialysis cycle was repeated a further two times, and the final vicilin precipitate was freeze-dried

Analytical methods The techniques used for amino acid analysis,[3] *N*-terminal analysis,[3] acrylamide gel electrophoresis,[3,16] peptide mapping,[16] radioisotope labelling *in vivo* with ^{35}S-sulphate,[3,9] carbohydrate analysis,[3] and carboxymethylation of protein with iodoacetate,[3] are described in previous papers as indicated

Acknowledgements—This work was supported by grants from the Agricultural Research Council

[13] R L WATTS, in *Phytochemical Phylogeny* (edited by J B HARBORNE), p 145, Academic Press, New York (1970)
[14] I A VAINTRAUB, A D SHUTOV and V G KLIMENKO, *Biokhimiya* **27**, 349 (1962)
[15] V GHETIE and L BUZILA, *Rev Roum Biochim* **5**, 271 (1968)
[16] C J BAILEY and D BOULTER, *Biochem J* **113**, 669 (1969)

Key Word Index—Vicia faba, Leguminosae, broad bean, vicilin, protein, amino acid composition, subunits