Connolly, M. L., Kuntz, I. D., Ferrin, T. E., & Langridge, R. (1982) *J. Mol. Biol.* (in press).
Dayhoff, M. O. (1972) *Atlas of Protein Sequence and Structure*, Vol. 5, National Biomedical Research Foundation, Silver Spring, MD.
Douzou, P. (1977) *Cryobiochemistry*, pp 109–115, Academic Press, London.
Frauenfelder, H., Petsko, G. A., & Tsernoglou, D. (1979) *Nature (London) 280*, 558–563.
Fuchsman, W. H., & Appleby, C. A. (1979) *Biochemistry 18*, 1309–1321.
Geraci, G., Parkhurst, L., & Gibson, Q. H. (1969) *J. Biol. Chem. 244*, 4664–4667.
Giacometti, G. M., DiIorio, E. E., Antonini, E., Brunori, M., & Winterhalter, K. H. (1977) *Eur. J. Biochem. 75*, 267–273.
Giacometti, G. M., Brunori, M., Antonini, E., DiIorio, E. E., & Winterhalter, K. H. (1980) *J. Biol. Chem. 255*, 6160–6165.
Hayashi, Y., Yamada, H., & Yamazaki, I. (1976) *Biochim. Biophys. Acta 427*, 608–616.
Heidner, E. J., Ladner, R. C., & Perutz, M. F. (1976) *J. Mol. Biol. 104*, 707–722.
Ikeda-Saito, M., Iizuka, T., Yamamoto, H., Kayne, I. J., & Yonetani, T. (1977) *J. Biol. Chem. 14*, 4882–4887.
Imai, K., Ikeda-Saito, M., & Yonetani, T. (1980) *J. Mol. Biol. 144*, 551–565.
LaMar, G. N., Budd, D. L., Sick, H., & Gersonde, K. (1978) *Biochim. Biophys. Acta 537*, 270–283.
McKinnie, R. E., & Olson, J. S. (1981) *J. Biol. Chem. 256*, 8928–8932.
Morris, R. J., & Gibson, Q. H. (1980) *J. Biol. Chem. 255*, 8050–8053.
Norvell, J. C., Nunes, A. C., & Schoenborn, B. P. (1975) *Science (Washington, D.C.) 190*, 568–569.
Phillips, S. E. V. (1980) *J. Mol. Biol. 142*, 531.
Phillips, S. E. V., & Schoenborn, B. (1981) *Nature (London) 292*, 81–82.
Phillips, S. E. V., Hall, D., & Perutz, M. F. (1981) *J. Mol. Biol. 150*, 137–141.
Provencher, S. W., & Dovi, V. G. (1979) *J. Biochem. Biophys. Methods 1*, 313–318.
Reisberg, P. I., & Olson, J. S. (1980) *J. Biol. Chem. 255*, 4151–4158.
Satterlee, J. D., Teintze, M., & Richards, J. (1978) *Biochemistry 17*, 1456–1462.
Sharma, V. S., Geibel, J. F., & Ranney, H. M. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 3747–3750.
Stryer, C., Kendrew, J. C., & Watson, H. C. (1964) *J. Mol. Biol. 8*, 96–104.
Szabo, A. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 2108–2111.
Traylor, T. G. (1981) *Acc. Chem. Res. 14*, 102–107.
Warshel, A. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 1789–1793.
Winterhalter, K. H., Anderson, N. M., Amiconi, G., Antonini, E., & Brunori, M. (1969) *Eur. J. Biochem. 11*, 435–440.
Yip, Y. K., Waks, M., & Beychok, S. (1972) *J. Biol. Chem. 247*, 7237–7244.

Biochemical and Crystallographic Data for Phaseolin, the Storage Protein from *Phaseolus vulgaris*†

Stephen Johnson, Guy Grayson, Lindsay Robinson, Rima Chahade, and Alexander McPherson*

ABSTRACT: We have isolated and biochemically characterized a major seed protein from *Phaseolus vulgaris* (common garden green bean) that appears to be identical with a form of the storage protein of that plant seed known as phaseolin. We have further shown that it appears very similar in most properties to the storage protein from *Canavalia ensiformis* (jack bean) which we have solved to 3.0-Å resolution using X-ray diffraction techniques. [McPherson, A., & Rich, A. (1973) *J. Biochem. (Tokyo) 74*, 155–160; McPherson, A., & Spencer, R. (1975) *Arch. Biochem. Biophys. 169*, 650–661; McPherson, A. (1980) *J. Biol. Chem. 255*, 10472–10480]. We have crystallized phaseolin and conducted a preliminary X-ray diffraction analysis on it as well. The data show the crystals to be of pseudo cubic space group *P*432, although the symmetry can be, strictly speaking, only *P*1. The unit cell has one trimeric molecule of 150 000 daltons in a unit cell of dimensions $a = b = c = 66.6$ and $\alpha = \beta = \gamma = 90°$. The crystals apparently possess some form of disorder that makes reconciliation of the unit cell contents with the observed crystallographic properties difficult, although they do diffract strongly to better than 2.8-Å resolution. No evidence of twinning has been observed.

As an extension of our X-ray structural studies on canavalin, from jack beans (*Canavalia ensiformis*), we undertook a search for similar crystalline proteins from other plant species. We have isolated such a protein from *Phaseolus vulgaris* and have obtained it in crystalline form. The protein has those properties, so far as we can determine, that identifies it as phaseolin, one of the major storage proteins of the seed. This protein has recently gained some additional attention by virtue of its being the first plant protein cloned into bacteria (Sun et al., 1980). The protein demonstrates a number of interesting biochemical and physical properties which indicate that it is closely similar in a structural sense to the protein canavalin which is a major storage protein in jack beans. We have previously studied canavalin by X-ray diffraction and other biochemical techniques and have recently reported the solution

† From the Department of Biochemistry, University of California, Riverside, California 92521. *Received October 30, 1981; revised manuscript received June 15, 1982.* This research was supported by Grant GM 21398 from the National Institutes of Health.

0006-2960/82/0421-4839$01.25/0

of its structure to 3.0-Å resolution (McPherson, 1980). Application of X-ray crystallographic techniques to the *Phaseolus* globin has provided evidence that these two proteins are likely homologous to one another, at least in broad molecular structure, and could be structurally representative of the entire class of 7S seed storage proteins.

Materials and Methods

The amino acid analyses were performed on 3 times recrystallized protein that had been thoroughly dialyzed vs. distilled water and hydrolyzed in 12 N HCl for 24 h before lyophilization. The instrument used was a Beckman Model 120C amino acid analyzer, and the result is an average of three different analyses. Cysteine content was determined in a like manner after performic acid oxidation.

The number of exposed sulfhydryl groups on the native protein was determined (Ellman, 1959) by the reaction of the protein with 5,5′-dithiobis(2-nitrobenzoic acid) (Ellman's reagent). The protein sample consisted of 1.0 mL of 0.05 M phosphate buffer at pH 7.0 with 1.0 mL of a 3.5 mg/mL protein solution. The Ellman solution was composed of 39.6 mg of Ellman's reagent dissolved in 15 mL of the same buffer. To the protein solution was added 610 μL of Ellman solution and the absorbance of duplicate samples compared against a blank containing 2 mL of phosphate buffer. This was monitored at 412 nm for a period of 24 h.

An analysis of the protein for covalently bound neutral sugar was conducted by using the phenol–sulfuric acid procedure of Dubois et al. (1956). The optical density increase at 490 nm as compared with a standard curve produced by parallel reactions with known amounts of sucrose (5–50 μg) was taken as a measure of carbohydrate content. Amino sugars were sought with the amino acid analyzer.

The native molecular weight of the protein was determined by comparing its migration velocity through the column of a Waters high-pressure liquid chromatography apparatus with that of a series of standard proteins of known molecular weight. Recrystallized protein dissolved in 0.05 M phosphate buffer at pH 7.0 was run several times to demonstrate reproducibility.

The subunit molecular weight of the protein was determined numerous times by electrophoresis on sodium dodecyl sulfate ($NaDodSO_4$)–polyacrylamide gels according to Laemmli (1970) in parallel with a series of proteins of known polypeptide molecular weight. The gels were of the slab variety and were gradients of 5–15% polyacrylamide composition topped with 5% stacking gels. The apparatus employed was that produced by the LKB Co., and the electrophoresis was conducted at 25 mA and 150 V. The molecular weight of the individual peptide chains was also determined in a similar manner but on gels containing 8 M urea in addition to the $NaDodSO_4$ and other components. For trypsin digest experiments, $NaDodSO_4$ gradient gels without urea were employed.

Cross-reactivity with antibodies to the protein canavalin (from *Canavalia ensiformis*) was determined on agar diffusion plates (Ouchterlony, 1958). The antibodies to canavalin, previously certified by the same procedure to have a high titer for that protein, were produced in white rabbits over a 6-month period by standard procedures.

For preparation of the protein, approximately 225 g of seeds (cultivars used were Burpee Green Crop or Top Crop) was ground to meal with a conventional electric seed mill, stirred into acetone (2.0 mL of acetone/g of seed flour), and allowed to stand for 24 h with occasional stirring. The mixture was poured through paper filters and the extract dried overnight. The remainder of the preparation was carried out at 4 °C to minimize microbial activity. The seed meal was extracted with a solution (8.0 mL of solution/g of seed meal) of 0.5% NaCl and 0.05 M tris(hydroxymethyl)aminomethane (Tris-HCl) at pH 7.5 that was stirred vigorously for 5 h. This mixture was then filtered through cheesecloth and glass wool overnight and the supernatant spun for 30 min at 6000 rpm to remove insoluble material.

The supernatant was brought to 40% saturation with ammonium sulfate and stirred overnight. The resulting precipitate was removed by centrifugation at 8000 rpm for 40 min. Approximately 100 mL of 0.025 Tris-HCl, pH 7.5, was used to wash the tubes and dissolve the precipitate. The seed extract was centrifuged at 12 000 rpm for 20 min and the supernatant dialyzed against several changes of 0.025 M Tris-HCl, pH 7.5, over a 24-h period.

For determination of the effects of trypsin digestion, 60 μL of an appproximately 8 mg/mL solution of dissolved crystals and 60 μL of a 1.0 mg/mL solution of trypsin were mixed and incubated at room temperature. After 2, 8, 15, 30, and 60 min and 12 h, a 20-μL sample of the incubation mixture was mixed with 15 μL of the $NaDodSO_4$ sample preparation solution (700 μL of 2% bromophenol, 1.0 mL of glycerol, 1.0 g of $NaDodSO_4$, 0.735 g of Tris base, and 2.5 mL of β-mercaptoethanol adjusted to pH 7.3 with HCl and brought to 10.0 mL total volume) to stop the reaction. Aliquots (20 μL of each) were run on a 5–15% gradient acrylamide gel in parallel with a sample containing undigested protein.

The protein was initially crystallized by using the back-extraction method of Jacoby (1968). The protein was fully precipitated with 100% saturated ammonium sulfate and then sequentially resuspended and extracted with ammonium sulfate solutions of decreasing concentration. Crystals of small size but large yield grew in solutions between 75 and 90% ammonium sulfate saturation. These crystals were harvested by centrifugation after 2 weeks and dissolved in water at a concentration of about 10 mg/mL.

By use of the vapor diffusion technique employing multidepression glass plates in sealed plastic boxes (McPherson, 1982), this protein was used to grow crystals of sufficient quality for high-resolution X-ray diffraction analysis. The 20-μL droplets contained initially 5 mg/mL protein and were 50% saturated with ammonium sulfate. The reservoirs yielding the best crystals were of 25 mL volume and consisted of 60–70% saturated ammonium sulfate. No buffer was included since a broad screen of pH values using many buffers indicated no effect.

X-ray diffraction data were collected on single crystals of the protein by using precession photography with a Buerger precession camera having a crystal to film distance of 75 mm. The X-rays were of the Cu Kα spectral line produced by a GX-20 Elliott rotating anode generator operated at 40 kV and 40 mA with a focal spot size of 200 μm^2. The exposure times varied from 6 to 24 h, depending on the size of the crystal and the particular zone under investigation.

X-ray diffraction data to 3.0-Å resolution were collected on native protein crystals with an automated Enraff-Nonius CAD-4 diffractometer and a 1200-W broad focus X-ray tube operated at 50 kV and 20 mA. The intensities were collected so that one entire quadrant of the sphere of reflection was measured. The intensities were corrected for absorption and Lp factors, reduced, and scaled by using conventional computational procedures. All calculations and data management were performed on a PDP 11/34 computer running under the RSX-11M operating system.

Table I: Amino Acid Analyses of 7S Seed Storage Proteins[a]

	crystalline *Phaseolus* protein	7S *Phaseolus* fraction protein		canavalin
		Derbyshire[c]	Hall[d]	
Lys	5.8	5.6	6.5	6.0
His	1.5	2.6	3.0	2.4
Arg	5.2	5.0	5.3	7.6
Trp	*b*	0.8	1.6	0.0
Asp	13.2	12.4	13.6	13.1
Thr	7.1	3.4	2.8	4.0
Ser	8.3	6.7	6.4	5.6
Glu	16.0	15.1	19.0	16.8
Pro	3.4	2.9	3.5	4.8
Gly	4.5	2.7	3.2	3.7
Ala	4.4	3.0	3.3	4.2
Val	5.8	5.2	5.4	4.2
$^1/_2$-Cys	0.2	0.3	0.4	0.2
Met	0.5	1.1	0.7	0.7
Ile	5.2	5.6	5.4	4.7
Leu	9.1	9.1	9.4	10.6
Tyr	3.6	3.5	3.8	5.0
Phe	6.6	6.6	6.6	5.9

[a] Expressed as percentage by weight of total protein. [b] Not determined. [c] From Derbyshire et al. (1976). [d] T. C. Hall, personal communication.

Results

From high-pressure liquid chromatography in parallel with standards, the native molecular weight of the *Phaseolus* protein is 145 000 ± 5000 daltons. In addition, the protein traversed the column as a single, sharp band which indicated it to be a single, homogeneous protein specie. It does not, by this method, appear to be heterogeneous in terms of polypeptide or subunit composition or aggregation state.

$NaDodSO_4$–polyacrylamide gel electrophoresis both in the presence and in the absence of 8 M urea yielded three closely spaced bands having molecular weights in the range 48 000–53 000 daltons in a pattern consistent with that reported for phaseolin (Bollini & Vitale, 1981; Ma et al., 1980). The protein then exists as a trimer of three polypeptides of very similar molecular weights aggregated to form an oligomer of about 150 000 daltons. The canavalin molecule also is a trimer of 147 000 daltons comprised of three identical subunits of 49 000 daltons.

The amino acid composition of the crystalline protein is shown in Table I along with the composition of the 7S vicillin fractions and phaseolin reported by other investigators (Bewley & Black, 1978; Derbyshire et al., 1976) and with that of canavalin. As is apparent by visual inspection, the amino acid compositions are very similar. All are low in sulfur-containing amino acids, lack cystine residues and tryptophan, are low in histidine, and possess a characteristic distribution of the more common entries. It is of particular significance that phaseolin shares a composition similar to the jack bean protein, because canavalin, in addition to being a member of the vicillin class, is also a structural tandem duplicate (McPherson, 1980). The carboxy-terminal half of the polypeptide chain that makes up the canavalin subunit is structurally and compositionally almost the same as its amino-terminal half. This correlation is consistent with the possibility that the phaseolin could, like canavalin, also be a structurally redundant, tandem duplicate. The rabbit antibody made to canavalin, however, was found not to cross-react with the *Phaseolus* globin. Thus the two proteins certainly differ in some important respects.

The phenol–sulfuric acid test applied to our crystalline phaseolin protein indicated approximately 10% neutral sugar to be present in the molecule. The amino acid analyses of acid hydrolysates revealed two glucosamine residues. Thus we conclude that our phaseolin is indeed a glycoprotein. The sugar content we find is in agreement with determinations by other groups studying phaseolin, as is the finding of the two glucosamine residues (Bewley & Black, 1978; Derbyshire et al., 1976). We note here that we have been unable to show the presence of any carbohydrate groups on the canavalin molecule, and this remains one apparent difference between the two proteins.

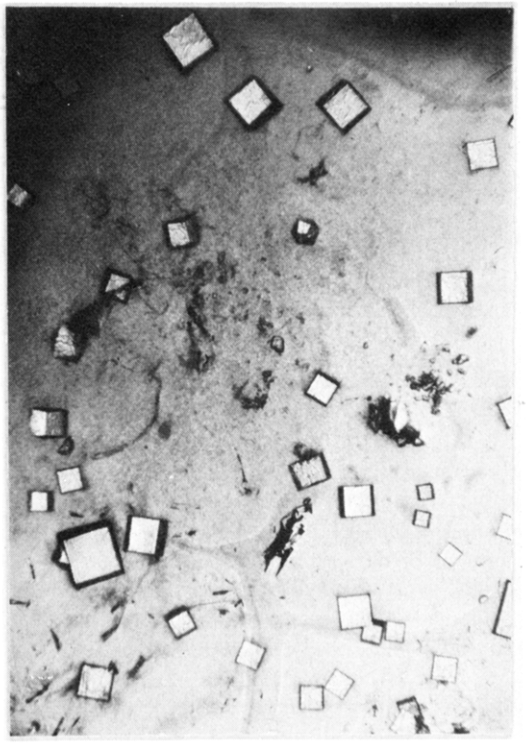

FIGURE 1: Low-power light microscope photograph of the cubic phaseolin crystals grown from 70% saturated ammonium sulfate at room temperature over 30 days time.

Determination of the exposed sulfhydryl content of the protein using Ellman's reagent gave an odd result. Very slow reaction was observed so that only after 24 h was an apparent end point reached. This corresponded to only 0.5 sulfhydryl residue per 48 000-dalton polypeptide chain. This is difficult to explain since it implies that only one out of every two, otherwise identical, polypeptide chains has a free cysteine group. This might be due to the polypeptide chain heterogeneity we observed on our $NaDodSO_4$–polyacrylamide gels. If the trimer of 144 000 daltons was a mixture of three kinds of subunits that were nearly identical but only one or two of which possessed a cysteine group and the other did not, then the result we obtained might be consistent.

The time course of digestion of the phaseolin protein by trypsin was monitored by electrophoresis of time aliquots on $NaDodSO_4$–polyacrylamide gels. The results of these experiments show that when exposed to the protease, major fragments are produced having migration rates closely corresponding to halves of the three phaseolin subunits. Over the remaining several hours there is very little breakdown of these product chains into smaller polypeptide fragments. The digestion pattern suggests that the subunit polypeptides of the protein may consist of two rather stable domains joined by a section of amino acids highly susceptible to trypsin attack. This experiment is also consistent with polypeptide chains that contain a structural redundancy. Once again, this result is virtually the same as the results obtained for canavalin from jack beans; it too is quickly cleaved to two chains of about

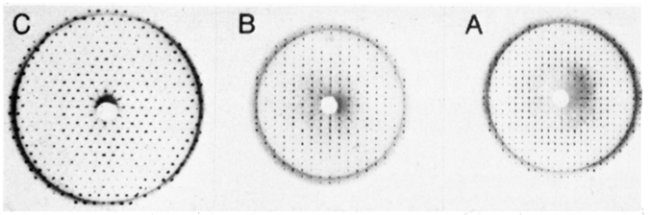

FIGURE 2: X-ray diffraction patterns of one crystal of phaseolin taken with a Buerger precession camera having a crystal to film distance of 75 mm. The precession angles were 10° in (A) and (B) and 12° in (C), and exposure times were 15 h. In (A) the X-ray beam is aligned along the 100 direction of the crystal, in (B) along the 110 direction, and in (C) along the 111 body diagonal of the crystal. In all three cases there is the appearance of pseudosymmetry that is not maintained upon close inspection.

24 000 daltons which are then relatively stable to further digestion.

The exclusive habits of the crystals, shown in Figure 1, were perfect cubes or derivatives of cubes. Under a polarizing microscope the crystals exhibited no observable birefringence and extinguished sharply when analyzer and polarizer were crossed at 90°. These observations would seem to be consistent with a cubic unit cell. When examined with X-rays, the reciprocal lattice net was seen to be perfectly cubic with equal dimensions of 66.6 Å. The photograph obtained by aligning the X-ray beam perpendicular to any major face of the crystal is shown in Figure 2A. The pattern has no systematic absences with respect to upper levels of reciprocal space, and the apparent symmetry is 4 mm. Close inspection shows, however, that the symmetry, although very close, is definitely not 4 mm and that only 4-fold symmetry is present. Alignment of the X-ray beam along the 110 direction yields the pattern seen in Figure 2B having apparent mm symmetry. Figure 2C is the photograph obtained by directing X-ray beam along the 111 direction. It shows superficial indications of the 6-mm symmetry that would be expected of a cubic unit cell, but it actually exhibits only $\bar{1}$ symmetry.

The space group of the crystals, both from the known unit cell contents and from the diffraction symmetry, is formally $P1$, but the overall symmetry of the diffraction pattern is quite close to $P432$. For further investigation of the extent of the pseudosymmetry, X-ray diffraction data were collected to 3.0-Å resolution with an automated diffractometer. The data were collected in such a way that six reflections symmetrically equivalent in space group $P432$, plus their Friedel equivalents (total of six Bijvoet pairs), were observed. After the reflections were corrected for absorption and Lorentz–polarization effects, they were merged by using linear scale factors, and a symmetry residual was computed. For the 2000 reflections observed to 3.0-Å resolution, the R factor was 0.16 of the mean structure amplitude F. Bijvoet pairs alone, however, merged with an R of 0.06. Thus we conclude that the space group of the crystals is near, but not exactly, $P432$ and that the true space group is $P1$.

The volume of the crystallographic unit cell is $V = 2.95 \times 10^5$ Å^3. Since the native molecular weight of the protein is 144 000, then by no measure or means could the unit cell accommodate any but one single 144 000 trimeric molecule of the protein, that is, three polypeptide chains of appproximately 48 000 daltons.

The crystals diffract quite well to reasonably high resolution, at least 2.8 Å, and are stable in the X-ray beam for up to 60 h without appreciable decay. They do, however, with respect to their rather large size (greater than 0.8 mm on an edge) and small unit cell, diffract rather weakly and occasionally demonstrate a high mosaic spread, a feature often indicative of some inherent internal disorder.

Discussion

This protein from *Phaseolus vulgaris* represents a large proportion of the total seed protein and has solubility properties, a molecular weight, and an amino acid composition that suggest that it is a form of phaseolin, the seed storage protein. By the same criteria as well as sulfhydryl reactivity, trypsin digest pattern, oligomer and polypeptide molecular weight, amino acid composition, and general physical properties, it appears to be closely homologous to canavalin from jack beans, which is the storage protein of that seed.

Like the *Phaseolus* protein, canavalin has been found to have a trimer molecular weight of about 144 000 and a subunit molecular weight of 48 000. It is initially cleaved by trypsin and chymotrpysin into two similar polypeptide chains of about 24 000 daltons, just as is phaseolin. The reactivity of exposed sulfhydryl groups is unusual and the same for both proteins. Canavalin has been shown, by biochemical means and independently by its structure determination using single crystal X-ray diffraction, to have a subunit that is 2-fold redundant and to have these two structural units, or domains, related by a quasi-dyad axis of symmetry. Although the same conclusion cannot yet be drawn unequivocally for the *Phaseolus* protein, the evidence suggests that it is probably true for it as well.

The interpretation of the crystallographic observations is not at all straightforward, and indeed, we must defer a full explanation until the analysis is more advanced. Placing one molecule of the protein in a triclinic unit cell of volume $V = 2.95 \times 10^5$ Å^3 presents no problem. It yields a value of $V_m = 2.05$ Å^3/dalton which is within the range of values compiled for other crystalline proteins (Matthews, 1968). The presence in the crystals of high pseudo cubic symmetry must, however, be addressed and reconciled with the composition, structure, and disposition of the protein subunits that make up the single trimer contained in the unit cell, and this we cannot do.

The final explanation will require the presence in the crystals of either some form of twinning or other inherent disorder. Since the observed symmetry does not appear to be fully consistent with any higher space group, we conclude that the formal crystallographic symmetry is $P1$. The higher symmetry might be produced by twinning, for example, unit cells of symmetry $P3$ or $R3$ [as suggested also by Drenth & Wiebanga (1955) for crystals of the seed protein excelsin]. Conceivably, it could also result from multiple orientations of a trimeric molecule having a 3-fold axis of symmetry about a 4-fold axis or alternatively from a mixture of blocks of unit cells disposed in a symmetrical mosaic pattern which is then superimposed on the true crystallographic symmetry. We observe in our diffraction patterns, however, no obvious indications as to the underlying causes. There is generally no splitting of reflections or distorted spot shapes or streaking. We often find the mosaic spread of the intensities to be quite broad, but this is not always so, and the spread is usually reasonably narrow. Our examinations of many crystals demonstrates the observed diffraction patterns to be reproducible and consistent. Because of the unusual nature of these crystals, we are proceeding cautiously and attempting to obtain a structure solution to low resolution (approximately 5.0 Å) by multiple isomorphous replacement. From this we hope a satisfactory explanation will emerge.

In investigations of the 11S proteins, it was concluded that excelsin from Brazil nuts (Drenth & Wiebanga, 1955; Schepman et al., 1972), edestin from hemp seed (Drenth & Wiebanga, 1955; Von Cleeman & Kratky, 1960; Schepman et al., 1972), globulin from tobacco seed (Drenth & Wiebanga,

1955), and globulin from sunflower seed (Plietz et al., 1978) were arranged with the subunits as trigonal antiprisms having D_3, or quasi-D_3, point group symmetry. Interestingly, when single crystals of these proteins could be obtained, they were invariably of octahedral or some other cubic habit. These, we believe not coincidentally, were the same morphologies obtained by Bailey (1942), Vickery et al. (1941), and Osborne (1892, 1902) for a large number of other varieties of seed storage proteins.

In a manner very similar to what we find for the phaseolin crystals, Drenth & Wiebanga (1955) and Astbury et al. (1935) described crystals of quasi-cubic symmetry. They concluded that their crystals of the 11S protein fraction were probably rhombohedral and of space group $R3$ or $R32$, but with $\gamma = 60°$ so that they expressed pseudo cubic symmetry $F432$. It was argued, however, for the best studied case, excelsin, that the true symmetry of the diffraction pattern at higher resolution than the 8.0 Å extent of their data was not in fact even $R3$, but of lower symmetry, $P1$. This is remarkably similar to our findings for the *Phaseolus* globin.

References

Astbury, W. T., Dickinson, S., & Bailey, K. (1935) *Biochem. J. 29*, 2351–2360.
Bailey, K. (1942) *Trans. Faraday Soc. 38*, 186–190.
Bewley, J. D., & Black, M. (1978) in *Physiology and Biochemistry of Seeds in Relation to Germination*, Vol. 1, Springer-Verlag, Berlin.
Bollini, R., & Vitale, A. (1981) *Physiol. Plant 52*, 96–100.
Derbyshire, E., Wright, D. J., & Boulter, D. (1976) *Phytochemistry 15*, 3–24.
Drenth, J., & Wiebanga, E. H. (1955) *Recl. Trav. Chim. Pays-Bas 74*, 813.
Dubois, M., Gilles, K. A., Hamilton, J. K., Rebers, P. A., & Smith, F. (1956) *Anal. Chem. 28*, 350.
Ellman, G. L. (1959) *Arch. Biochem. Biophys. 82*, 70–77.
Jacoby, W. B. (1968) *Anal. Biochem. 26*, 295, 298.
Laemmli, V. K. (1970) *Nature (London) 227*, 680.
Ma, Y., Bliss, F. A., & Hall, T. C. (1980) *Plant Physiol. 66*, 897–902.
Matthews, B. W. (1968) *J. Mol. Biol. 33*, 491–497.
McPherson, A. (1980) *J. Biol. Chem. 255*, 10472–10480.
McPherson, A. (1982) *The Preparation and Analysis of Protein Crystals*, Wiley, New York.
McPherson, A., & Rich, A. (1973) *J. Biochem. (Tokyo) 74*, 155–160.
McPherson, A., & Spencer, R. (1975) *Arch. Biochem. Biophys. 169*, 650–661.
Osborne, T. B. (1892) *J. Am. Chem. Soc. 14*, 662.
Osborne, T. B. (1902) *J. Am. Chem. Soc. 24*, 28.
Ouchterlony, O. (1958) *Prog. Allergy 5*, 1.
Plietz, P., Damaschun, H., Pamaschun, G., & Schwenke, K. D. (1978) *Acta Biol. Med. Ger. 37*, K1–K2.
Schepman, A. M. H., Wichertjes, T., & Van Bruggen, E. F. J. (1972) *Biochim. Biophys. Acta 271*, 279–285.
Sun, S. M., Slightom, J. L., & Hall, T. C. (1980) *Nature (London) 289*, 37–42.
Vickery, H. B., Smith, E. L., Hubbell, R. B., & Nolan, L. S. (1941) *J. Biol. Chem. 140*, 613–624.
Von Cleeman, J. C., & Kratky, D. (1960) *Z. Naturforsch., B: Anorg. Chem., Org. Chem., Biochem., Biophys., Biol. 156*, 525–535.

Proton Magnetic Resonance Studies of Barley and Wheat Thionins: Structural Homology with Crambin†

Juliette T. J. Lecomte, Berne L. Jones, and Miguel Llinás*

ABSTRACT: The thionins comprise a group of very basic proteins of $M_r \sim 5000$ found in the seeds of Gramineae. They each contain 45 amino acid residues arranged along a single polypeptide chain that is constrained by four disulfide bridges. Five thionins of known sequence, from barley and wheat, have been investigated and compared by ^1H NMR spectroscopy at 600 MHz. From their spectral characteristics it is concluded that the five proteins have very similar, nonrandom conformations in ^2H_2O solution. Moreover, on the basis of selective nuclear Overhauser experiments at 300 MHz, features of their secondary and tertiary structures are shown to be similar to those of crambin, a related, hydrophobic protein extracted from seeds of the crucifer *Crambe abyssinica*. The strong compositional homology of the thionins facilitates the assignment of methyl and aromatic resonances, as only a few residues are replaced and these are at known sites. The substitution of leucine for an isoleucine does not affect significantly the local magnetic environment, suggesting that those isomeric side chains easily accommodate the same spatial constraints. A fast hydrogen–deuterium exchange is observed at pH* 6.25, 25 °C. This indicates that, although of folded conformation, the thionins are structurally flexible polypeptides that efficiently expose all amides to the solvent.

In 1940, a compound of proteinic nature was isolated from a petroleum ether extract of unbleached wheat flour (Balls & Hale, 1940; Balls et al., 1942a). The substance associated with lipids, crystallized readily from an ethanol–water mixture (Balls et al., 1942a), and was toxic to bacteria (Stuart & Harris, 1942; Fernández de Caleya et al., 1972), yeasts (Stuart & Harris, 1942; Okada et al., 1970; Hernández-Lucas et al., 1974), fungi (Stuart & Harris, 1942), small laboratory animals (Coulson et al., 1942), and insect larvae (Kramer et al., 1979). Owing to the high cystine content of the protein, the name "purothionin" was coined shortly after its characterization (Balls et al., 1942b). Since then, it has been shown that the

† From the Department of Chemistry, Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213 (J.T.J.L. and M.L1.), and the Grain Marketing Research Laboratory, U.S. Department of Agriculture, Manhattan, Kansas 66502 (B.L.J.). *Received April 28, 1982.* This Research was supported by the U.S. Public Health Service, National Institutes of Health Grant GM-25213. The NMR Facilities are supported by National Institutes of Health Grants GM-27390 and RR-00292.

0006-2960/82/0421-4843$01.25/0