Proc. Natl. Acad. Sci. USA
Vol. 81, pp. 1406–1410, March 1984
Biophysics

Crambin in phospholipid vesicles: Circular dichroism analysis of crystal structure relevance

(membrane proteins/secondary structure/plant toxins)

B. A. WALLACE†, N. KOHL†, AND M. M. TEETER‡

†Department of Biochemistry, Columbia University, New York, NY 10032; and ‡Department of Chemistry, Boston University, Boston, MA 02215

Communicated by Elkan R. Blout, November 18, 1983

ABSTRACT **Crambin, a hydrophobic plant seed protein that exhibits sequence homology to membrane-active plant toxins, was incorporated into phospholipid vesicles. Circular dichroism spectroscopy indicates that its structure in vesicles is nearly identical to its structure in 60% ethanol solution, the solvent from which the protein was crystallized. The secondary structure predicted from the circular dichroism data of the ethanol solution closely resembles that determined by x-ray diffraction of the crystals. This agreement suggests that the x-ray structure may form a useful basis for modeling the structure and behavior of lipophilic plant toxins. Finally, because the structure of crambin has been determined in an organic solvent medium, it provides a protein standard for examination of the effect of solvent dipole moment on the circular dichroism spectra of proteins, which may be important for interpretation of data for membrane proteins.**

The hydrophobic plant protein crambin is isolated from an aqueous acetone extract of delipidated plant seed (1). The protein is insoluble in water but dissolves in ethanol. Crystals prepared from a 60% ethanol solution diffract to 0.88 Å (2), the highest resolution of any protein known to date. The amino acid sequence (3) and crystal structure (4) indicate that crambin contains amphipathic helices, which may be an important structural feature in some lipid-associated proteins. This amphipathic character—i.e., segregation of charged and hydrophobic amino acids on opposite sides of an α helix—is seen in the structure of the lytic polypeptide melittin (5) and has been postulated for the apoprotein of high-density lipoproteins (6).

The function of crambin is unknown, but it does exhibit extensive sequence homology with a family of membrane-active plant toxins (3, 7, 8)—e.g., viscotoxin and purothionin. Preliminary model building studies (9) suggest that, despite the fact that the toxins are basic and crambin is neutral, the structure of the toxins could contain a similar amphipathic helix. The toxins are also hydrophobic and are isolated from cereal grain in complex with lipid (10). However, it is not clear if this association with lipids is found *in vivo* or if it is an artifact of the petroleum ether isolation procedure. The efficiency of toxin extraction is increased if hexaploid wheat lipids are added to rye (11). A tight association with lipid would be expected for molecules that act by lysing cell membranes. Although crambin does not exhibit similar lytic activity, its sequence and solubility properties suggest that it too might form complexes with lipid molecules. In this study, sedimentation experiments of lipid vesicles formed in the presence of crambin do indeed demonstrate that crambin can complex with lipids.

High-resolution structural details of the protein crambin are known from the crystals grown in ethanol. If the same structure exists in vesicles as in ethanol, it would provide a detailed view of protein structure as found in a membrane bilayer. However, the structures of polypeptides in organic solvents do not necessarily correspond to their structures in vesicles. It has been shown recently that the conformation of the ion-channel gramicidin in crystals formed from ethanol is totally unlike its structure in lipid vesicles (12, 13). Hence, CD spectroscopy has been used to determine (*i*) if the conformation of crambin in 60% ethanol solution resembles its conformation in phospholipid vesicles and (*ii*) how these relate to the structure in the crystals prepared from the ethanol solution.

MATERIALS AND METHODS

Crambin was obtained from C. H. Van Etten as a recrystallized aqueous-acetone extract from defatted seed meal of the plant *Crambe abyssinica* (2). Amino acid analysis indicated it was 99% pure. Phospholipids were obtained from Calbiochem. All reagents and solvents used were of reagent grade or better.

For CD studies, protein concentrations were estimated from the extinction coefficient determined for a 50% ethanol solution, $\varepsilon_{1\,mg}^{277} = 0.79\ M^{-1}\cdot cm^{-1}$ (unpublished results). Vesicle samples were solubilized with 0.4% $NaDodSO_4$ prior to measuring their absorbance. For sedimentation studies, the relative protein concentrations in the gradient fractions were determined by Lowry assay in the presence of 0.1% $NaDodSO_4$, with bovine serum albumin as a standard (14). Lipid concentrations were determined by a modified Fiske–Subbarow phosphate assay method (15).

Small unilamellar vesicles were prepared as described (16) from dried-down solutions that initially contained 2.1 mg of dimyristoyl phosphatidylcholine and 0.084 mg of crambin in 5 ml of $CHCl_3$. The sample was hydrated at 60°C with 0.5 ml of deionized water. The vesicle suspension was sonicated to clarity with a Soniprep 150 ultrasonicator microprobe using 6- to 10-sec bursts at ≈30°C. The resulting suspension was centrifuged at 12,800 × *g* for 6 min. The supernatant contained small unilamellar vesicles. Because protein and lipid are not recovered with the same efficiency, the mole ratio of lipid to protein in the final vesicle preparation was 143 ± 15.

CD spectra were recorded at 21°C on a Cary 60 spectropolarimeter fitted with a model 6001 attachment and a variable position detector. For most experiments, the photomultiplier was placed adjacent to the sample to maximize the light-acceptance angle (90°). To assess the extent of differential light scattering, the specimen-to-detector distance (i.e., the acceptance angle) was varied (17). The wavelength range scanned was 300–190 nm at a scan speed of 5 nm/min and with a time constant of 3 sec. The instrument was calibrated with (+)-10-camphorsulfonic acid at 290 nm. At least three spectra of each sample were collected and averaged. Base lines of either solvent or vesicles without protein were subtracted from each respective sample to yield the contribution

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: NRMSD, normalized standard deviation.

of the protein component. To estimate the correction for absorption flattening effects in the vesicle samples (18, 19), we measured absorption spectra (400–190 nm) of the vesicles before and after solubilization in 0.4% $NaDodSO_4$, using a Cary 15 recording spectrophotometer with the CD cell (0.5-mm path length) placed closest to the detector. Estimates of secondary structure were calculated by a linear least-squares method (18) using a data base for helices, sheets, turns, and random coils derived from water-soluble proteins (20).

Data points at 2-nm intervals between 190 and 240 nm were used for the analyses. A normalized standard deviation (NRMSD) for each curve fitting was expressed as:

$$\text{NRMSD} = \left[\frac{\sum_{i=1}^{N} (\theta_{exp} - \theta_{cal})^2}{\sum_{i=1}^{N} (\theta_{exp})^2}\right]^{1/2},$$

where θ_{exp} and θ_{cal} were the experimental and calculated mean residue ellipticities, respectively, and N was the number of data points used (18). In the calculations, the average number of residues per helical segment was set at eight to correspond to the results found by x-ray diffraction (4). The least-squares fit constrained the fraction of each secondary structure to be positive but did not require that the sum of the fractions be unity. The results were normalized to 100% by multiplying each fraction by a factor equal to the inverse sum of the fractions. It can be shown mathematically and experimentally (18, 21) that this treatment compensates for any errors in concentration but does not change the relative proportions of secondary structural types calculated. The necessity for the concentration correction may arise from the variation in extinction coefficient of the protein with the solvent. In this case, the value used was obtained for ethanol, which may not be applicable for $NaDodSO_4$-solubilized vesicles.

A measure of the difference between the structure determined by CD and that determined by x-rays (ΔF) is expressed as:

$$\Delta F = \sum_{i=1}^{n} |f_c - f_x|,$$

where f_c and f_x are the fractions of secondary structure derived from the calculated CD data and the x-ray structure (based on ϕ, ψ angles) respectively, and where n refers to each secondary structure type (helix, sheet, turn, and coil).

Density gradient centrifugation of 300-μl samples onto 12–20% sucrose was carried out as described (22), with a Beckman L-50 centrifuge with a SW 50.1 rotor operating at 42,000 rpm for 48 hr at 4°C. Parallel gradients with and without crambin were run.

Samples were prepared for electron microscopy on Formvar-coated copper grids at vesicle concentrations of 0.1 mg/ml; subsequently, they were stained with 0.1% uranyl acetate. Grids were examined in a Phillips 300 electron microscope operating at 80 kV and fitted with an anticontamination cold trap.

RESULTS

Characterization of Samples. The protein concentrations in samples used for spectroscopy were established by using an extinction coefficient determined for the protein in ethanol. Because the absorbance of a chromophore depends on the solvent environment and can vary by as much as 1.5- to 2-fold upon changing from water to hydrocarbons, it is possible that this value may not be precise for the protein in a lipid environment. The difference in extinction coefficients would be reflected in the normalization factor calculated for the data (Table 1), which is inversely proportional to the concentration (and hence proportional to the extinction coefficient). The ratio of the two normalization factors is 1.11, indicating that the extinction coefficient in vesicles may be closer to 0.87.

The relative protein concentrations in the gradient fractions were determined by Lowry assay. Although this method does not give an absolute value of the concentration when bovine serum albumin is used as a standard, it was used because its enhanced sensitivity permitted detection of the small amounts of material present in the fractions. Because the purpose of this experiment was only to demonstrate which fractions contained protein, the relative values should be sufficient.

Small unilamellar vesicles were prepared, which appeared by negative-stain electron microscopy to have average diameters of ≈300 Å. Although particulate samples such as liposomes can produce optical artifacts that may significantly distort their CD spectra, it has been shown that such effects for vesicles of this small size are negligible (23). Any differential scattering was corrected for by a sample detection geometry that permits collection of light scattered 45° from the incident direction. Similar CD spectra were obtained for crambin in sonicated vesicles by using acceptance half-angles of either 45° or 1°. This suggests that virtually all the light scattering was near the forward direction and was collected by the detector. The absorption flattening corrections to the CD spectra have the largest magnitudes in the region of high absorption (i.e., from 190 to 195 mm) (23) but can be estimated by comparison of the UV (unpolarized) spectrum of $NaDodSO_4$-solubilized and unsolubilized specimens (18, 19); they are small for these vesicles (≈2% of the measured ellipticities) and can be neglected.

Crambin Is Incorporated into Phospholipid Vesicles. To demonstrate the incorporation of crambin, vesicles were prepared in the presence of the protein and examined by equilibrium density centrifugation on sucrose gradients. The resulting gradients had densities ranging from 1.03 to 1.11 g/ml, which permitted the separation of crambin-containing vesicles from those that did not contain crambin. The crambin-containing vesicles migrate to a density of 1.074, while the vesicles alone are found at a density of 1.064; free protein would be located near the bottom of the tube. All of the protein was observed to bind to vesicles (Fig. 1, peak centered at fraction 7), although not all of the vesicles had protein bound (peak centered at fraction 11).

Thus, crambin formed a reasonably tight complex with the lipid molecules, which was not dissociated by the low ionic strength conditions of the gradient. Therefore, the molecule,

Table 1. Calculated secondary structures and parameters for crambin in vesicles, in solution, and in crystals

Sample	Method	Helix	Sheet	Turn	Random	NRMSD	Normalization factor
Vesicles	CD*	0.36	0.22	0.17	0.25	0.077	0.81
60% ethanol	CD*	0.36	0.23	0.18	0.23	0.075	0.91
60% ethanol (crystals)	X-ray†	0.37	0.17	0.09	0.37		

*λ = +3 nm.
†Using values from ϕ, ψ angles.

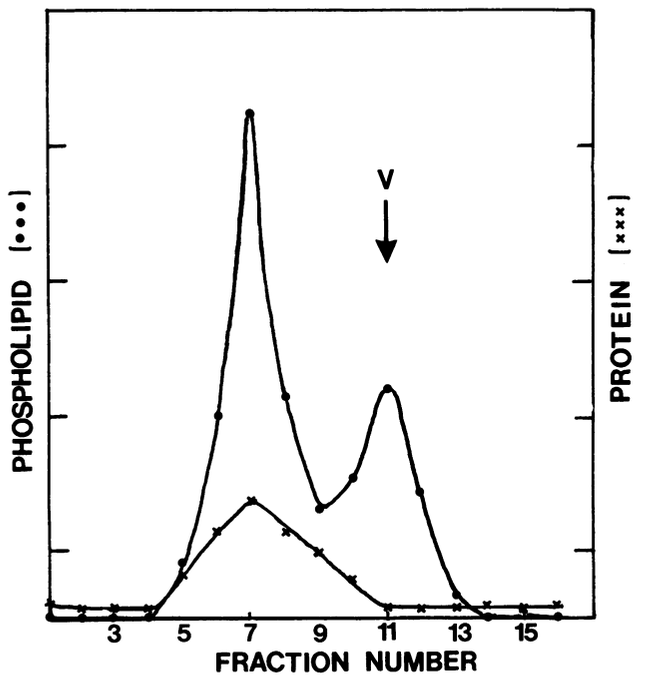

FIG. 1. Sedimentation of phospholipid vesicle samples containing crambin into 12–20% continuous sucrose density gradients. Fractions were analyzed for phospholipid (phosphate) and protein (Lowry method). The position of migration of vesicles without crambin is indicated (V).

which is insoluble in water, is capable of existing in a stable aqueous suspension of phospholipid vesicles.

The Conformations of Crambin in Vesicles and in 60% Ethanol Solutions Are Nearly Identical. Because the x-ray structure of crambin has been determined for crystals prepared from 60% ethanol, it should be of interest to compare the structure of the protein in 60% ethanol solution with that in vesicles, anticipating that, if there were a close correspondence between the two, one could utilize the crystal information. Furthermore, such a comparison could give insight into whether the protein had a stable secondary structure or whether its conformation was significantly influenced by its solvent environment. Therefore, CD spectra were obtained for crambin in vesicles and in ethanol solution.

Because none of the protein in the sample was found to be unassociated with vesicles (see above), the only peptide signal that would be detected in unfractionated specimens would be from the vesicle-bound form. Ethanol solution and vesicle samples could be directly compared (Fig. 2) because light scattering and absorption flattening effects from the vesicles were negligible (see above). The spectra obtained were identical within the precision of the measurements, differing at most by 5%. This is an indication that the net average backbone conformation of the molecules in the two environments is the same. Of course, because the spectra obtained are merely a sum of the secondary structures, these results could also have been obtained for two different conformations that happened to have the same net amount of each secondary structure; this is an unlikely possibility.

Finally, it should be recognized that the spectra could arise from a mixture of the conformers; if this were so, it would seem that the same mixture of conformers would be present in the two solvent systems.

The λ_{max} of the Protein Is Shifted in a Solvent of Low Polarizability. The position of the absorption bands of proteins can be affected by the environment of the peptide chromophores. $\pi \rightarrow \pi^*$ transitions are generally blue-shifted going from aqueous to nonaqueous solvents because the higher polarizability of the aqueous solvent results in more destabilization of the excited state, which usually has a larger dipole moment.

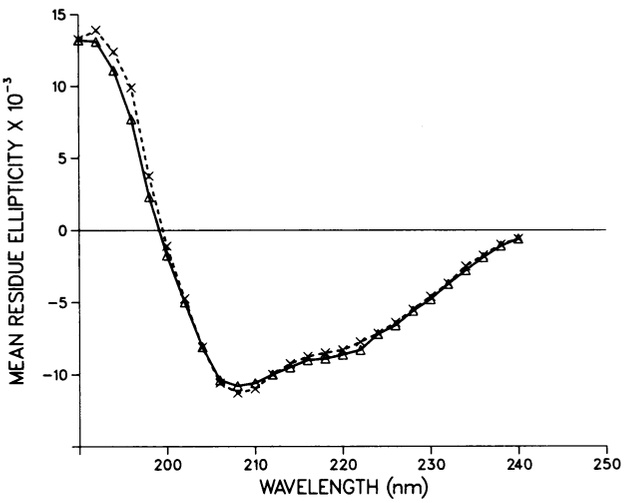

FIG. 2. Normalized CD spectra of crambin in 60% ethanol (×) and of crambin in dimyristoyl phosphatidylcholine vesicles (△). The protein-to-lipid mole ratio in the vesicle sample was 1:143 ± 15.

Thus, the crambin data in ethanol might be expected to be blue-shifted when compared to the reference data set, which was derived from aqueous-soluble proteins. To test if this were the case, the protein CD data collected was shifted by 1-nm increments in the positive and negative directions and compared to the calculated best fit; a better fit of the calculated data to actual data (lower NRMSD) would be expected as the wavelength approached the value equal to the solvent-induced shift. Upon doing so, we found that a +3-nm shift gave the best fit to the data (Fig. 3), which is consistent with a solvent polarizability effect.

Crambin Conformation in 60% Ethanol Solution Is Very Similar to That in Crystals. In order to utilize the structural information from the crystals, one must have some idea if the crystals represent the structure in solution. This can be examined by comparing the secondary structure calculated from the CD data of the ethanol solution with that determined by x-ray diffraction of the crystals (Table 1). Using the best fit of the CD data to the calculated secondary structure (lowest NRMSD), one can compare these values with the x-ray results by the statistical parameter ΔF. It appears that the same wavelength shift that gives the lowest NRMSD is also the same value that gives the best fit to the crystal data (Fig. 3).

The correspondence between the crystal and solution conformations (Table 1) is very good, and compares favorably with the correspondences for aqueous-soluble globular proteins. The differences detected can be ascribed largely to uncertainties in the methods. First of all, the appropriateness of the CD reference data set must be considered. The fractional composition of secondary structure types was determined by using least-squares fitting of the experimental CD spectrum to a reference set of CD spectra obtained from aqueous-soluble proteins of known secondary structure. The analysis assumes that the optical activity in the wavelength range of 190–240 nm is predominantly due to peptide transitions, with negligible contribution from aromatic amino acids. This condition is met for crambin, which has a low aromatic amino acid content (ref. 3; one phenylalanine and one tyrosine). The accuracy of the analysis depends on the applicability of the reference data set (24), which could potentially be a problem as it compares a hydrophobic protein with aqueous-soluble proteins. However, it does appear to give accurate

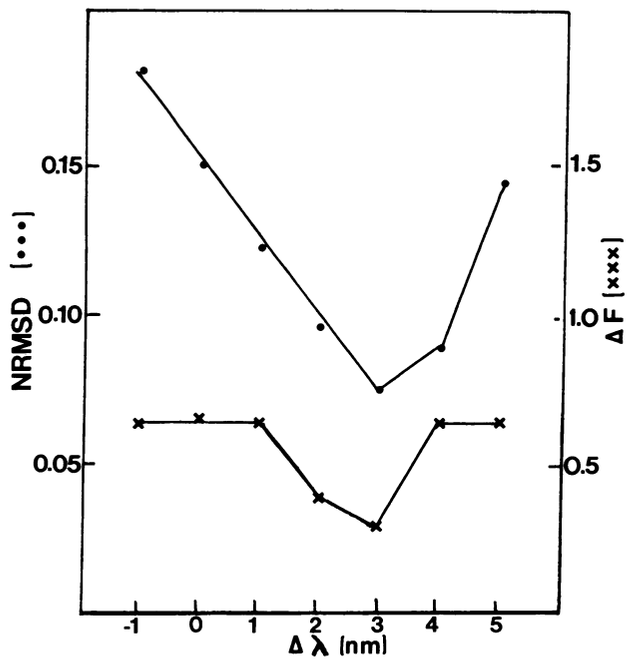

FIG. 3. Plot indicating the extent of agreement between calculated secondary structures and measured CD and x-ray data for crambin in 60% ethanol as a function of wavelength shift applied to the reference data set. NRMSD, a statistical parameter correlating CD data and calculated results, and ΔF, a measure of the difference between the calculated CD results and the x-ray structure, indicate the effect of solvent dipole moment on the λ_{max} of the optically active absorption bands.

estimates for bacteriorhodopsin (18), the only membrane protein whose structure is known with any precision.

Second, the definition of secondary structure in the crystal is not completely straightforward. The most appropriate criteria appears to be use of the Ramachandran ϕ, ψ angles. Conservative decisions were made about whether or not to include NH_2- and COOH-terminal residues in the helix. In addition, the number of residues to include in a turn is somewhat subjective. If the criteria of hydrogen-bond patterns had been used to define the secondary structure, a significantly different set of values for these fractions would have been obtained.

Finally, CD measures the average of all conformations that the molecule adopts in solution. X-rays sample a conformation that has fewer degrees of freedom. If crambin is a flexible molecule, one might expect more conformations to be present in solution and, thus, a poorer correspondence.

Despite all these constraints, the agreement between CD and x-ray fractions of secondary structure is remarkably close.

DISCUSSION

The sedimentation equilibrium experiments have demonstrated that crambin can be incorporated into phospholipid vesicles and that, when bound to these vesicles, crambin can exist in aqueous solution—an environment in which it is otherwise insoluble. This tight binding to vesicles would certainly be a basic requirement for any molecule that acts as the homologous plant toxins do—i.e., hemolysis of erythrocytes or lysis of other cell membranes (25). While these studies indicate that crambin is capable of being solubilized by and associating with lipids, they do not demonstrate that crambin is associated with the plant cell membrane *in vivo*, although this would be consistent with the solubility properties of the molecule. By analogy with the amphipathic helices of melittin and apolipoproteins, one might expect an orientation of the molecule such that the helix would align along the aqueous solution/membrane interface. However, the present experiments do not bear on the orientation question but provide a basis for model building of this and other plant toxins. The strong sequence homology of crambin, viscotoxin, and purothionin has already been shown (3). These studies indicate crambin has yet another structural feature in common with plant toxins—its solubility and complexation with lipids—even though it does not appear to have the same biological activity. Parallel CD studies on the plant toxins to examine their secondary structures, may show the manner in which they differ from crambin and may give insight into the nature of their toxicity.

The spectroscopic studies in this work indicate that the structures of the protein in vesicles and in 60% ethanol solution are nearly identical. Such was not the case for another membrane-associated protein, the ion-channel gramicidin (12, 13), which adopts very different conformations in organic solvents and vesicles. The larger size and more compact, disulfide-crosslinked globular structure of the crambin molecule may account for its enhanced stability as it is transferred between these environments. CD spectroscopy also suggests that the crambin x-ray structure, determined for crystals prepared from 60% ethanol, is closely related to the average structure in the ethanol solution. This suggests two points: (*i*) crystal packing forces do not grossly distort the secondary structures at the interfaces between these small proteins, and (*ii*) there exists one major conformation in ethanol solution, rather than a mixture of several significantly different conformers. The differences observed between the CD- and x-ray-detected structures are within the limits of measurement and the interpretations of the methods.

Because the protein conformation in the crystal is essentially the same as in the 60% ethanol solution and the ethanol solution and vesicle spectra are nearly identical, this work indicates that the crambin crystal structure could be related to the average conformation the molecule adopts in membranes. Furthermore, it suggests that crambin may be useful as a model protein for folding studies and spectroscopic standards in hydrophobic environments because its structure is known to such high resolution.

We thank D. Mao for helpful discussions and for programs for fitting the CD data. This work was supported by National Institutes of Health Grant GM27292 (to B.A.W.) and by National Science Foundation Grant PCM-8003929 (to M.M.T.). B.A.W. is the recipient of a Hirschl Career Scientist Award.

1. Van Etten, C. H., Nielsen, H. C. & Peters, J. E. (1965) *Phytochemistry* **4**, 467–473.
2. Teeter, M. M. & Hendrickson, W. A. (1979) *J. Mol. Biol.* **127**, 219–224.
3. Teeter, M. M., Mazer, J. A. & L'Italien, J. J. (1981) *Biochemistry* **20**, 5437–5443.
4. Hendrickson, W. A. & Teeter, M. M. (1981) *Nature (London)* **290**, 107–113.
5. Terwilliger, T. C., Weissman, L. & Eisenberg, D. (1982) *Biophys. J.* **37**, 353–361.
6. Segrest, J. P., Jackson, R. L., Morrisett, J. D. & Gotto, A. M. (1974) *FEBS Lett.* **38**, 247–253.
7. Mak, A. S. & Jones, B. L. (1976) *Can. J. Biochem.* **54**, 835–842.
8. Samuelsson, G., Seger, L. & Olson, T. (1968) *Acta Chem. Scand.* **22**, 2624–2642.
9. Teeter, M. M. & Whitlow, M. (1982) *Biophys. J.* **37**, 94a (abstr.).
10. Redman, D. G. & Fisher, N. J. (1982) *J. Sci. Fd. Agric.* **19**, 651–655.
11. Hernández-Lucas, C., Carbonero, P. & García-Olmedo, F. (1978) *J. Agric. Food Chem.* **26**, 794–796.
12. Wallace, B. A., Veatch, W. R. & Blout, E. R. (1981) *Biochemistry* **20**, 5754–5760.

13. Wallace, B. A. (1983) *Biopolymers* **22,** 397–402.
14. Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. (1951) *J. Biol. Chem.* **193,** 265–275.
15. Fiske, C. H. & Subbarow, Y. (1925) *J. Biol. Chem.* **66,** 375–400.
16. Wallace, B. A. & Blout, E. R. (1979) *Proc. Natl. Acad. Sci. USA* **76,** 1775–1779.
17. Schneider, A. S. & Harmatz, D. (1976) *Biochemistry* **15,** 4158–4162.
18. Mao, D., Wachter, E. & Wallace, B. A. (1982) *Biochemistry* **21,** 4960–4968.
19. Gresalfi, T. J. & Wallace, B. A. (1984) *J. Biol. Chem.* **259,** in press.
20. Chang, C. T., Wu, C.-S. & Yang, J. T. (1978) *Anal. Biochem.* **91,** 13–31.
21. Mao, D. (1984) Dissertation (Columbia Univ., New York).
22. Rossi, J. D. & Wallace, B. A. (1983) *J. Biol. Chem.* **258,** 3327–3331.
23. Mao, D. & Wallace, B. A. (1984) *Biochemistry* **23,** in press.
24. Baker, C. C. & Eisenberg, I. (1976) *Biochemistry* **15,** 629–634.
25. Carrasco, L., Vasquez, D., Hernandez-Lucas, C., Carbonero, P. & Garcia-Olmedo, F. (1981) *Eur. J. Biochem.* **116,** 185–189.