Reprinted from THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 207, No. 1, March, 1954

STUDIES ON THE GROSS STRUCTURE, CROSS-LINKAGES, AND TERMINAL SEQUENCES IN RIBONUCLEASE

BY CHRISTIAN B. ANFINSEN, ROBERT R. REDFIELD, WARREN L. CHOATE, JUANITA PAGE, AND WILLIAM R. CARROLL

(From the Laboratory of Cellular Physiology, National Heart Institute, and the National Institute of Arthritis and Metabolic Diseases, National Institutes of Health, United States Department of Health, Education, and Welfare, Bethesda, Maryland)

(Received for publication, October 2, 1953)

Previous studies (1, 2) have shown an asymmetric labeling of amino acids derived from different points along the chains of radioactive proteins. These results have suggested, as one major possibility, that proteins are synthesized by an assembly mechanism involving the condensation of preformed peptide fragments not in ready equilibrium with the pools of free amino acid. An attractive corollary of this general hypothesis is that such fragments may serve as common building blocks for a number of proteins.

The studies of Sanger and his collaborators (3, 4) have elucidated the amino acid sequences in the insulin molecule and have made available methodology (3, 5) for similar studies on other proteins. The present report deals with the progress to date of our investigations on the fine and gross structures of ribonuclease, particularly in regard to N and C terminal sequences and the nature of the cross-linkages in this single chain protein. It is hoped that, as comparison of the fine structure of these two pancreatic proteins becomes possible, a more rational basis for experiments on protein synthesis in this tissue can be devised.

Materials and Analytical Methods

The homogeneity of each lot of commercial crystalline ribonuclease (Armour and Company or Worthington Biochemical Laboratory) used in these studies was checked ultracentrifugally, and, in most cases, by electrophoresis. These preparations showed a single N terminal amino acid (lysine) by the dinitrofluorobenzene (DNFB) method. Their specific enzyme activity corresponded to the accepted values in the literature (6). Analytical values obtained for phenylalanine (7) (3 residues per mole) and for cysteine estimated as cysteic acid (8 residues per mole) agreed within 5 to 10 per cent with the values obtained on ion exchange columns by Hirs, Moore, and Stein.[1]

The preparations of pepsin, chymotrypsin, and trypsin used for degrada-

[1] Hirs, W., Moore, S., and Stein, W. H., personal communication.

tion were crystalline commercial products. Worthington carboxypeptidase was recrystallized eight times and treated (8) with diisopropyl fluorophosphate (DFP) before use.

In view of the laborious and time-consuming nature of the commonly used methods for ribonuclease assay, a rapid and fairly accurate procedure was developed for use in these studies. Ribonuclease (0 to 14 γ) in 1.50 cc. of 0.1 M acetate buffer, pH 5.0, is added to 1.0 cc. of yeast nucleic acid (Schwarz) dialyzed 48 hours against water, final concentration 0.8 per cent. After incubation for 25 minutes at 25° the reaction is stopped with 0.5 cc. of 0.75 per cent uranium acetate in 25 per cent perchloric acid. Following removal of precipitated protein and substrate by centrifugation, 0.10 cc. of the supernatant fluid is diluted to 3.1 cc. with water and read

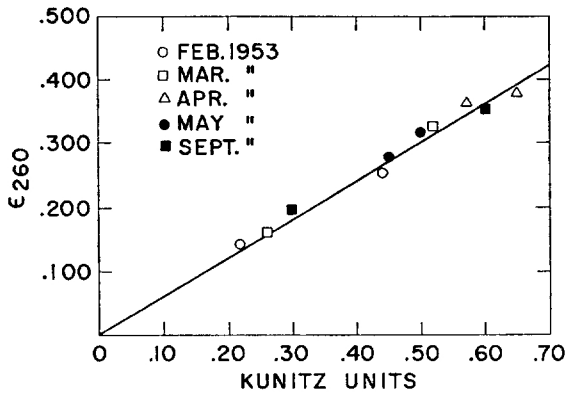

FIG. 1. Calibration curve for the determination of ribonuclease activity

at 260 mμ in the Beckman spectrophotometer. A standard curve with known levels of pure ribonuclease is run with each set of determinations, although this is probably unnecessary in view of the excellent reproducibility of the procedure (see Fig. 1). Correction is made for the reagent blank determined by incubation without enzyme. This blank increases slowly during storage of the substrate in the cold, but appears to have no effect on reproducibility. The method gives linear results up to 0.60 Kunitz unit (45 units per mg. of ribonuclease). The data of Fig. 1 indicate the stability of the reagents employed over a 6 month period.

Results

Gross Structure—The ribonuclease molecule is characterized by a high degree of geometrical symmetry. Its low f/f_0 ratio ((9) and Table I) and unit cell dimensions (30 × 18 × 48 A (10)) indicate the probable presence of an extensive system of cross-linkage. Earlier chemical studies by Brand

and his colleagues (11) and other experiments from this laboratory (12) made possible calculations suggesting the presence of three N terminal groups, and consequently three peptide chains per mole (mol. wt. 13,500). The N terminal and C terminal end-group analyses reported below, however, point almost certainly to a single chain structure, and this conclusion is confirmed by ultracentrifugal and diffusion experiments. Native ribonuclease sediments, under the conditions described in Table I, with an $s_{20,w}$ of 1.9 to 2.2 (see also (9, 12)). Following treatment with performic

TABLE I

Physical and Chemical Data on Native and Oxidized Ribonuclease

Ribonuclease	Chemical data		Physical data				
	—SH groups per mole	$-SO_3^-$ groups per mole	Protein or derivative per cc.	$D_{20,w} \times 10^7$	$s_{20,w} \times 10^{13}$	Calculated mol. wt.	Calculated $\frac{f}{f_0}$
			mg.			*gm. per mole*	
Native	0	0	150	9.9	1.88	15,800	1.31
Oxidized	0	7.8 ± 0.1	90	7.0	1.35	15,800	1.87

Native ribonuclease and oxidized enzyme prepared as described in the text were dialyzed overnight against phosphate buffer, $\Gamma/2 = 0.1$, pH 7.2. $s_{20,w}$ was determined in the Spinco ultracentrifuge and $D_{20,w}$ in the Aminco-Stern electrophoresis apparatus with boundary sharpening by the method of Kahn and Polson (33). The analytical data of Hirs, Moore, and Stein indicate a molecular weight of about 13,500 gm. per mole. Upon performic acid oxidation, this value should increase 500 to 1000 gm. per mole owing to introduction of oxygen into cysteine, methionine, and tyrosine residues. The data above have been chosen from an experiment in which all figures reported were derived from studies on a single batch of enzyme. Calculations of moles are based on finite concentrations of protein. These data, when compared with those of Rothen (9), indicate the desirability of more extensive studies on the physical properties of ribonuclease, particularly in regard to diffusion measurements, and the estimation of frictional ratios; such experiments are being carried out by one of us (W. R. C.).

acid, this constant falls to 1.35. Such a change in sedimentation might be accounted for either by oxidative division of the molecule into two essentially equal fragments or by rupture of cross-linking disulfide bonds, resulting in the production of a derivative so coiled as to impart greater frictional characteristics. Although the first alternative is almost ruled out by the fact that dialysis of oxidized ribonuclease results in no loss of nitrogen from the dialysis sac, this point was more thoroughly established by diffusion measurements. Table I summarizes the data from an experiment in which the molecular weights of native and performic acid-treated ribonuclease are compared. Ribonuclease contains 8 cysteine residues,[2]

[2] The presence of 8 cysteine residues permits the theoretical presence of sixteen

all bound in disulfide linkage as evidenced by the absence of —SH groups when tested by the method of Boyer (14). Solution in 85 per cent formic acid for 30 minutes at room temperature causes minimal, if any, change in the protein, since, upon removal of the solvent *in vacuo*, the enzyme activity is essentially completely recovered. The presence of H_2O_2 during this 30 minute period (1 part of 30 per cent H_2O_2 to 9 parts of HCOOH), however, results in the complete oxidation of cysteine sulfur to the cysteic acid form. Thus, following hydrolysis in a sealed tube with 6 N acid, 8 moles of cysteic acid per mole of protein (Table II) can be separated chromatographically on Dowex 50 columns (H^+ form, pH 7) (13) and determined colorimetrically (15).

The above results strongly suggest that this protein is cross-linked through four disulfide bridges. The physical studies described above and

TABLE II

Number of Cysteine Residues per Mole of Ribonuclease

Experiment No.	Oxidized ribonuclease hydrolyzed	Cysteic acid determined	Cysteic acid per μM ribonuclease
	μM	μM	μM
1	0.0722	0.563	7.8
2	0.504	3.90	7.7
3	0.119	0.940	7.9

In Experiment 1, ribonuclease calculated from the Kjeldahl nitrogen value, assuming 13,400 gm. per mole of ribonuclease and 16.5 per cent nitrogen. In Experiments 2 and 3, ribonuclease calculated from the dry weight of sample. All values orrected for 10 per cent loss of cysteic acid during oxidation of protein (34).

the end-group analysis below lead one to the tentative postulation of a gross structure, such as is depicted in Fig. 2. Some support for this general picture is derived from the x-ray diffraction experiments of Carlisle and Scouloudi (10) which indicated five crystallographic chains.[3] Complete amino acid analyses by Hirs, Moore, and Stein[1] lead to an estimate of 121 amino acid residues per mole of ribonuclease. Thus in this preliminary suggestion, each of the five peptide folds depicted should contain, on the average, 24 amino acid residues with disulfide cross-links as in-

dipeptide sequences of this amino acid in ribonuclease. Using the general method described by Flavin (13), we have, at present, direct, degradative evidence for seven different cysteic acid sequences from a considerably larger family of chromatographically separable di- and tripeptides of this amino acid.

[3] In a more recent paper (16), Carlisle, Scouloudi, and Spier state that further examination of the x-ray data suggests the presence of six crystallographic chains rather than five. We have, nevertheless, schematized the molecule as shown in Fig. 2, with five folds, since the present chemical data are compatible with such a structure.

dicated. It is clear that such a general structure, when further spatially compressed by arrangement of the peptide chain in the α-helix coils suggested by Pauling and Corey (17), would result in a highly compact, symmetrical molecule. The presence of 4 proline residues[1] in a five fold structure is also compatible with the postulated (18) rôle of this amino acid as a center of direction reversal in peptide chains.

N Terminal Residue of Ribonuclease—Dinitrophenyl ribonuclease (DNP ribonuclease) was prepared according to the usual methods for DNP protein (5).

Acid hydrolysis was performed either in concentrated HCl or constant boiling HCl in sealed tubes at 105° for varying intervals of 2 to 18 hours. Identification of the DNP amino acids was made by paper chromatography, by the systems of Blackburn and Lowther (19), Biserte and Os-

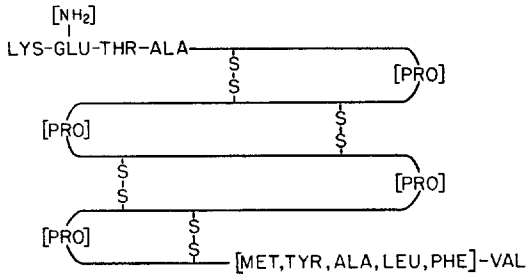

FIG. 2. Generalized gross structure of ribonuclease

teaux (20), Monier and Pénasse (21), and finally the two-dimensional technique of Levy.[4] For quantitative determination, the DNP spots were eluted with 1 per cent sodium bicarbonate and their absorption measured at 350 mμ in the Beckman spectrophotometer (20, 22).

As previously reported (23), bis-DNP lysine was the only DNP amino acid detected in the ether extracts of the hydrolysates. No α-DNP arginine or bis-DNP histidine could be detected in the aqueous phase.

Determination of Moles Bis-DNP Lysine per Mole DNP Ribonuclease—Weighed samples (approximately 0.15 to 0.2 μM of DNP ribonuclease dried over P_2O_5) were submitted to acid hydrolysis at 105° in sealed tubes with 0.5 cc. of constant boiling HCl. After 2, 5, and 8 hours, duplicate samples were diluted with water, extracted with ether, and chromatographed by the two-dimensional technique of Levy.[4] Duplicate 0.20 μM aliquots of a standard bis-DNP lysine solution were subjected to the same procedures, and the unknowns compared to the standards after identical hydrolysis times.

Using Hirs, Moore, and Stein's figures for the amino acid composition

[4] Levy, A. L., personal communication.

of bovine pancreatic ribonuclease,[1] and adding 167 for each ϵ-lysine (ten), N terminal residue (one), *O*-tyrosine (six), and imidazolyl group of histidine (four), as used by Porter (5) and Sanger (22), to the molecular weight of ribonuclease based on Hirs, Moore, and Stein's figures, one arrives at a molecular weight of 13,341 + 21(167) = 16,848 for DNP ribonuclease.

An average of 0.90 mole of bis-DNP lysine per mole DNP ribonuclease was detected. The 2 hour sample was disregarded because solution of the sample was incomplete (Table III).

No spots other than bis-DNP lysine and bis-DNP lysylglutamic acid (see below) were detected. As noted below, since the next residue was found to be threonine, it is a safe assumption that the N terminal group was completely hydrolyzed. In addition to the 5 to 10 per cent error in

TABLE III

Quantitative Determination of N Terminal Amino Acid Residue

Time of hydrolysis	Sample		Amount detected after hydrolysis			
	Weight		Lysine	Lysyl-glutamic	Total	N terminal lysine per mole
hrs.	*mg.*	*μM*	*μM*	*μM*	*μM*	*mole*
5	2.35	0.140	0.100	0.036	0.136	0.97
5	3.35	0.199	0.150	0.037	0.187	0.94
8	3.22	0.191	0.141	0.016	0.157	0.82
8	3.04	0.181	0.141	0.017	0.158	0.87
Average						0.90

the method (22), it is conceivable that the figure of 0.9 rather than 1.0 end-group per mole might be due to a higher rate of destruction when the DNP derivative is present in the protein than that which prevails when the amino acid derivative (24) is free in solution. Alternatively, it may be that the presence of other products of hydrolysis catalyzes the destruction of the derivative.

Determination of N Terminal Sequence—40 mg. of DNP ribonuclease were subjected to partial acid hydrolysis in 11 N HCl for 72 hours at 37°. After dilution, the hydrolysate was extracted with ether, ethyl acetate, and *n*-butanol (25). The extracts were concentrated *in vacuo* and then subjected to electrophoresis on Munktell No. 20 paper with 0.033 M phosphate buffer at pH 7.0 (26) as electrolyte. The main zone present in the ether extract, migrating at the rate of bis-DNP lysine, was extracted with bicarbonate and identified as such by paper chromatography. A fainter zone present in the ether and ethyl acetate extracts proved, upon complete hydrolysis and chromatography, to be bis-DNP lysylglutamic acid.

No zones containing bis-DNP lysine were found on electrophoresis of the *n*-butanol extracts.

A second hydrolysis for 20 hours, rather than 72 hours, showed again only the two components, bis-DNP lysine and bis-DNP lysylglutamic acid, the latter compound being present in greater amount.

Pepsin Hydrolysis of DNP Ribonuclease—40 mg. of DNP-ribonuclease were suspended in 0.01 M HCl, the pH was adjusted to 1.8 with 0.1 M HCl, and the suspension diluted to 4.0 cc. with 0.01 M HCl. Pepsin (Worthington, crystallized four times), 0.28 mg., was dissolved in a drop of water and added to the suspension, which was then incubated with shaking, one-fourth of the initial volume being withdrawn after 4 and after 22 hours. These aliquots were extracted with ethyl acetate until no further color was removed. The extracts were subjected to paper electrophoresis, the zones eluted and hydrolyzed, and the resulting ether-extractable DNP derivatives and the water-soluble residues (27) were chromatographed.

Free bis-DNP lysine, dinitrophenol, and an electrophoretic band containing bis-DNP lysine plus glutamic acid, threonine, and alanine in approximately equimolar concentrations were identified.

In order to determine the sequence of the amino acids in the tetrapeptide component, the remaining material from the above incubation was extracted with ethyl acetate. After removal of the ethyl acetate, the extracted material was hydrolyzed for 18 hours at 37° with 11 N HCl. The hydrolysate was diluted with water and extracted with ethyl acetate. This extract contained bis-DNP lysine (2+) and bis-DNP lysylglutamic acid (4+) determined by the methods used in the previous section. The aqueous portion, after removal of HCl *in vacuo*, was treated with fluorodinitrobenzene in 1 per cent trimethylamine acetate buffer, pH 9.5. The excess reagent was extracted from the alkaline solution, and the solution was dried under high vacuum. A few drops of 6 N HCl were added, and the sample was hydrolyzed in a sealed tube at 105° for 7 hours, diluted with water, and extracted with ethyl acetate. After hydrolysis with 6 N HCl in sealed tubes, the ethyl acetate-soluble and the aqueous portions of the hydrolysate were chromatographed[4] (27). The former contained DNP threonine (4+) and a trace of DNP glutamic acid, and the latter, free alanine (4+).

The over-all findings lead to an N terminal sequence in ribonuclease of Lys-Glu (or Glu-NH_2)-Thr-Ala-. Whether the glutamic acid is in the form of the amide has not been ascertained. The evidence from the migration of the compound during paper electrophoresis is inconclusive, but suggests that the glutamic acid is in the form of its amide, since the peptide moves more slowly than the neutral amino acid derivative, DNP alanine.

C Terminal Residue and Sequence—We have previously reported (23)

the presence of a C terminal valine residue in ribonuclease as determined by carboxypeptidase degradation. These experiments have now been repeated with low levels of DFP-treated carboxypeptidase to permit an estimate of the order of appearance of amino acids during digestion. Aliquots of the incubation mixture were taken initially and after successive intervals and either chromatographed directly after removal of protein (28, 29) or treated with dinitrofluorobenzene (5). After removal of excess reagent, DNP amino acids were extracted from the acidified solution with ether. The extract was freed of dinitrophenol by passage through a silica column with water-saturated chloroform as the moving phase (30), and,

TABLE IV

Amino Acids Released from Native Ribonuclease by DFP-Treated Carboxypeptidase (Estimated As Their DNP Derivatives)

DPN amino acid	Incubation time		
	15 min.	1 hr.	3 hrs.
Valine	0.034	0.064	0.118
Phenylalanine	0.019	0.039	0.098
Leucine	0.024	0.043	0.090
Alanine	0.009	0.021	0.086
Tyrosine	Trace	0.020	0.062
Methionine	Absent	Trace	0.016

The values are expressed as ϵ_{350} readings obtained in 1 cm. cells in the Beckman spectrophotometer on solutions of each DNP derivative extracted from paper chromatograms (Levy, personal communication) with 1 per cent $NaHCO_3$ in a total volume of 2.5 cc. 2 cc. aliquots taken from 6 cc. of phosphate buffer, pH 7.8, 0.1 M, containing 60 mg. of ribonuclease + 0.03 mg. of DFP-treated carboxypeptidase recrystallized eight times.

thereafter, the DNP amino acids were eluted with ether. Two-dimensional chromatography by the method of Levy[4] yielded spots which were cut out and eluted with 1 per cent $NaHCO_3$, the color being read at 350 mμ in the Beckman spectrophotometer. The data in Table IV confirm the presence of valine as the C terminal residue and suggest the subsequent position in the chain of phenylalanine (or leucine and isoleucine), alanine, tyrosine, and methionine, as indicated tentatively in Fig. 2.

The presence of these amino acids, and no others, was confirmed in the paper chromatograms (27) run on aliquots not treated with DNFB. Simultaneous 3 hour control incubations of ribonuclease and carboxypeptidase alone contained no detectable free amino acids, whether examined by visual amino acid chromatography (27) or by the quantitative DNP procedure above.

DISCUSSION

A considerable simplification would be introduced into the study of the general problem of protein synthesis if it could be established that common peptide building blocks were used in the biosynthesis of more than one protein species. The present studies, although obviously rather long term in nature, appear to us to be one approach to the examination of this point, since they will ultimately allow the direct comparison of fine structure in two proteins, insulin and pancreatic ribonuclease, synthesized in the same tissue. Although the ribonuclease molecule appears to be a single cross-linked chain, its general structural features, suggested by the present experiments and tentatively summarized in Fig. 2, make the successful elucidation of its structure a hopeful possibility.

Preliminary studies have been made on a variety of fragments separated from pepsin, chymotrypsin, and trypsin digests of native and DNP ribonuclease, before and after performic acid oxidation. These studies have shown a comforting reproducibility of digestion products and, together with determination of peptide sequence, will be reported in a subsequent communication.

In earlier experiments (12) it was found that digestion of ribonuclease with pepsin yielded a product of which only a relatively small part was dialyzable and which still sedimented in the ultracentrifuge with an $s_{20,w}$ of about 1.6. Treatment with thioglycolate did not change this figure. In the light of similar reducing studies on insulin (31, 32) and in view of our more recent experiments in which this pepsin-produced product has been made completely non-sedimentable by performic acid oxidation, these earlier results can probably now be explained by the incomplete cleavage of disulfide bonds in the pepsin-modified molecule by thioglycolate.

SUMMARY

Physical and chemical studies on ribonuclease indicate that this protein consists of a single chain arranged in a compact, folded structure, cross-linked through four disulfide bridges. The molecule, after performic acid oxidation, still contains only a single chain.

Its N terminal sequence has been shown to be lysylglutamyl (or glutaminyl) threonylalanine. The C terminal position is occupied by a valine residue, followed, back along the peptide chain, by phenylalanine, leucine or isoleucine, alanine, tyrosine, and methionine in an undetermined order.

BIBLIOGRAPHY

1. Steinberg, D., and Anfinsen, C. B., *J. Biol. Chem.*, **199**, 25 (1952).
2. Anfinsen, C. B., and Flavin, M., *Federation Proc.*, **12**, 170 (1953).

3. Sanger, F., and Tuppy, H., *Biochem. J.*, **49,** 463 (1951).
4. Sanger, F., and Thompson, E. O. P., *Biochem. J.*, **53,** 353 (1953).
5. Porter, R. R., in Gerard, R. W., Methods in medical research, Chicago, **3,** 256 (1950).
6. Kunitz, M., *J. Biol. Chem.*, **164,** 563 (1946).
7. Udenfriend, S., and Cooper, J. R., *J. Biol. Chem.*, **203,** 953 (1953).
8. Jansen, E. F., Nutting, M.-D. F., Jang, R., and Balls, A. K., *J. Biol. Chem.*, **179,** 189 (1949).
9. Rothen, A., *J. Gen. Physiol.*, **24,** 203 (1940).
10. Carlisle, C. H., and Scouloudi, H., *Proc. Roy. Soc. London, Series A*, **207,** 496 (1951).
11. Brand, E., *Ann. New York Acad. Sc.*, **47,** 187 (1946).
12. Anfinsen, C. B., *J. Biol. Chem.*, **196,** 201 (1952).
13. Flavin, M., *Nature*, in press.
14. Boyer, P., Abstracts, American Chemical Society, 123rd meeting, Los Angeles, 31C (1953).
15. Moore, S., and Stein, W. H., *J. Biol. Chem.*, **176,** 367 (1948).
16. Carlisle, C. H., Scouloudi, H., and Spier, M., *Proc. Roy. Soc. London, Series B*, **141,** 1 (1953).
17. Pauling, L., and Corey, R. B., *Proc. Roy. Soc. London, Series B*, **141,** 21 (1953).
18. Arndt, U. W., and Riley, D. P., *Nature*, **172,** 245 (1953).
19. Blackburn, S., and Lowther, A. G., *Biochem. J.*, **48,** 126 (1951).
20. Biserte, G., and Osteaux, R., *Bull. Soc. chim. biol.*, **33,** 50 (1951).
21. Monier, R., and Pénasse, L., *Compt. rend. Acad.*, **230,** 1176 (1950).
22. Sanger, F., *Biochem. J.*, **45,** 563 (1949).
23. Anfinsen, C. B., Flavin, M., and Farnsworth, J., *Biochim. et biophys. acta*, **9,** 468 (1952).
24. Desnuelle, P., Rovery, M., and Bonjour, G., *Biochim. et biophys. acta*, **5,** 116 (1950).
25. Woolley, D. W., *J. Biol. Chem.*, **179,** 593 (1949).
26. Biserte, G., *Biochim. et biophys. acta*, **4,** 416 (1950).
27. Redfield, R. R., *Biochim. et biophys. acta*, **10,** 344 (1953).
28. Partridge, S. M., *Nature*, **169,** 496 (1952).
29. Thompson, A. R., *Nature*, **169,** 495 (1952).
30. Sanger, F., *Biochem. J.*, **39,** 507 (1945).
31. Stern, K. G., and White, A., *J. Biol. Chem.*, **117,** 95 (1937).
32. Miller, G. L., and Andersson, K. J. I., *J. Biol. Chem.*, **144,** 465 (1942).
33. Kahn, D. S., and Polson, A., *J. Phys. and Colloid Chem.*, **51,** 816 (1947).
34. Schram, E., Moore, S., and Bigwood, E. J., *Biochem. J.*, in press.