Fiber Formation from Solutions of Collagen. IV. On the Role of the Basic Amino Acid Residues*

HOWARD B. BENSUSAN,† VIRGIL R. MUMAW, AND ANN W. SCANU

From the Benjamin Rose Hospital and Western Reserve University School of Medicine, Cleveland, Ohio

Received September 26, 1961

Tropocollagen was acetylated and guanidinated to varying extents and analyzed to determine the degree of the reaction on lysine and hydroxylysine. The ability of the derivatives to form reconstituted fibers and segmented long-spaced structures was determined. The rate of fiber formation decreased with increased acetylation and increased with increased guanidination. There was a gradual loss of fiber striation with increased guanidination. At 20% guanidination, no striations were seen. There was a rapid loss of striations between 23 and 38% acetylation of lysine residues. Guanidination yielded wider and more uniform segmented long-spaced structures, while acetylation had no apparent effect. Density determinations of segmented long-spaced structures derived from tropocollagen which was 87% acetylated showed the presence of five evenly spaced concentrations of arginyl residues. It is postulated that the overlapping of the two distal positions in end-to-end polymerization would then leave four evenly spaced positions as the sites for the quarter staggering during fiber formation.

Many fibrous proteins polymerize to form distinct structures with characteristic repeating units, visible by electron microscopy. The polymerization of tropocollagen is unique because it can yield different, interconvertible forms (Gross *et al.*, 1954). Illustrations of three of the five structures and the interpretation of the alignment of the monomers for each are shown in Figure 1 (Schmitt, 1959). The form similar to the native fiber may be obtained in several ways. Generally, acidic solutions of tropocollagen are adjusted to *p*H 7–9 and then incubated at 20–37°. The other structure with which we are presently concerned is the segmented long-spaced (SLS) form. This form may be made by the addition of ATP to dilute acetic acid solutions of tropocollagen (Schmitt *et al.*, 1952). Confirmation of the proposed alignment for the natural (640 A spacing) fibers and the segmented long-spaced form has been obtained recently (Hodge and Schmitt, 1960; Kühn *et al.*, 1960). In the segmented long-spaced form, the monomers are arranged in parallel with similar ends in alignment. It follows that the position of certain amino acid residues in the tropocollagen molecule may be demonstrated as bands after these groups have formed complexes with electron-opaque stains. Advantage has been taken of this fact to demonstrate that the basic and acidic amino acid residues occur together in the same discrete positions in the tropocollagen molecule (Hodge and Schmitt, 1960).

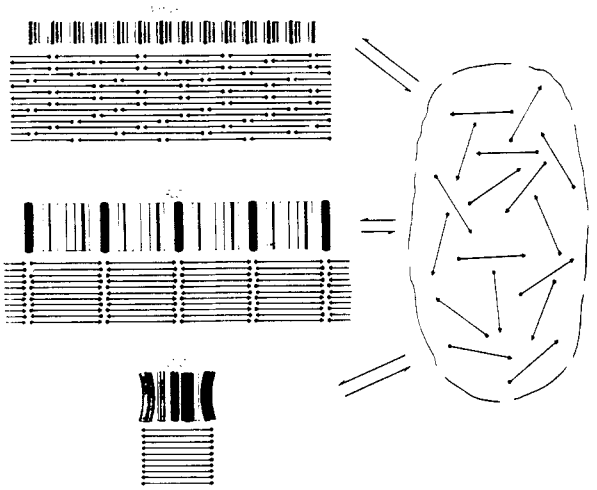

FIG. 1.—Diagrammatic illustration of the various forms obtained from the polymerization of tropocollagen. Taken from an article by Schmitt (1959) with permission of the author.

To form the 640 A–spaced fiber the tropocollagen molecules must be staggered at quarter-length intervals (Fig. 1).

We have been concerned with the mechanism of fibrogenesis and the nature of the specificity of the groups which determine the quarter-staggering. Although the relative contribution of each of the possible intermolecular forces remains obscure, coulombic interactions are known to play a major role (Randall *et al.*, 1955; Gross and Kirk, 1958; Bensusan and Hoyt, 1958; Bensusan, 1960). It has been demonstrated that the ionization of the phenolic group of tyrosyl residues accelerates the rate and decreases the energy of activation of fiber formation (Bensusan and Scanu, 1960). Thus, it appears that the tyrosyl anion is intimately involved. There are four amino acid cations in tropocollagen, namely: lysine, hydroxylysine, histidine, and arginine. There arises the problem of determining the contribution of each. The results

* The research reported in this paper was supported by Grants No. A-1825 and RG-7200 from the United States Public Health Service.

† Address: The Benjamin Rose Hospital, 2073 Abington Road, Cleveland 6, Ohio.

of minimal acetylation and guanidination reported here contribute to the solution of this problem.

Methods

Materials.—Tropocollagen was extracted from calf skin and purified as described (Bensusan and Scanu, 1960). Solutions were prepared by dissolving 1 g of the acetone-dried powder in 100 ml of 0.02 M acetic acid at 4° with gentle stirring. The undissolved residue was removed by centrifugation at 12,000 × *g* for 30 minutes.

Derivatives of Tropocollagen

1. Guanidination (Habeeb, 1960).—A 10-g. sample of 1-guanyl-3,5-dimethylpyrazole nitrate was dissolved in a minimum amount of water, and the solution was adjusted to *p*H 9.5 made up to 80 ml. A similar sample, adjusted to *p*H 7.5, was used for the control. Preliminary experiments showed that treatment at *p*H 7.5 and 9.5 had no effect on the protein in the absence of 1-guanyl-3,5-dimethylpyrazole nitrate. To each sample of reagent was added 115 ml of tropocollagen (approximately 1.4 mg/ml) and the mixture was stirred continuously at 4°. At various intervals of time, 30-ml aliquots were removed from the control and experimental reaction vessels. The experimental samples were adjusted to *p*H 8. All samples were then dialyzed exhaustively against water and finally against 0.01 M acetic acid.

2. Acetylation.—A 15-ml aliquot of tropocollagen (about 1.4 mg/ml) was placed in a small beaker in an ice bath. The beaker was equipped with a magnetic stirring bar and a set of external electrodes from the Beckman Model G *p*H meter. After the addition of 1.5 ml of M K_2HPO_4, the protein solution was adjusted to *p*H 9.3–9.5. Acetic anhydride, diluted with 50 volumes of CCl_4, was added in microliter amounts from a microburet while the *p*H was maintained at 9.40 with N NaOH. After the total amount of acetic anhydride was added the stirring was continued until no further change in *p*H occurred. The samples were then dialyzed extensively against 0.02 M acetic acid. A sham acetylation with an amount of glacial acetic acid equivalent to the maximum amount of acetic anhydride served as a control.

Analytical Methods.—The extent of guanidination was determined by separating the basic amino acids of the acid-hydrolyzed sample on a 6.2-cm column of Amberlite CG-120 as described by Moore *et al.* (1958). Our patterns were similar to those of Shields *et al.* (1959) for guanidinated mercuripapain. The hydroxylysine peak appeared midway between the tyrosine-phenylalanine and lysine peaks. All peaks were completely resolved. We verified the position of the homoarginine peak with a sample which was synthesized by the guanidination of α-acetyl lysine, prepared by the method of Neuberger and Sanger (1943). Ninhydrin determinations were performed by the method of Rosen (1957). Column yields and color yields were determined on pure samples of lysine, hydroxylysine, and arginine. The column yields were quantitative in all cases. The color yield for homoarginine was assumed to be the same as for arginine.

The extent of acetylation was determined in two ways. The per cent acetylation of the total number of ε-amino groups was determined in the usual way by performing the ninhydrin assay on the whole protein. The second method determined the per cent acetylation of lysine and hydroxylysine individually. The protein was heated at 100° for 45 minutes in the presence of the proper proportion of reagents described for the ninhydrin method of Rosen. This treatment destroyed the unacetylated ε-amino groups. The mixture was then dialyzed against running tap water, evaporated *in vacuo*, and dissolved in 6 N HCl (6 ml), and the protein was hydrolyzed at 110° for 18 hours. The samples were then analyzed for the amount of lysine and hydroxylysine by the method of Moore *et al.* (1958). Corrections were made for the fact that only 90% of the lysine and 80% of the hydroxylysine of the unacetylated protein are destroyed by the ninhydrin treatment (unpublished results).

Protein concentrations were determined by digestion followed by treatment with the Nessler reagent. A value of 17.7% nitrogen in collagen was used to calculate the protein concentration (Bensusan and Scanu, 1960).

The samples, containing no more than 1.4 mg of protein/ml, were centrifuged at 60,000 × *g* for 18 hours prior to viscosity determinations. Intrinsic viscosities were determined at 20.0° in Ubbelohde viscosimeters.

Samples for electron microscopic examination were first centrifuged at 60,000 × *g* for 18 hours. To prepare the reconstituted fibers, 0.6 ml of the cold protein solution was diluted with 1.4 ml of borate buffer, *p*H 9.1, to give a final ionic strength of 0.14. The solutions were then warmed to 25–30° until solid clots were formed. The clots were frozen on a freezing microtome and thin snow-like particles were shaved off and picked up on formvar-coated copper grids. The segmented long-spaced form was prepared in 0.2–0.5% ATP (Gross *et al.*, 1954). All samples were stained with phosphotungstic acid and minimally washed to prevent the dissociation of the phosphotungstic acid–ε-amino complex (Kühn, 1958). Electron micrographs were taken in the RCA EMU 2-D electron microscope.

Density tracings along the length of the segmented long-spaced form were obtained by line-selected oscillography (Loeser and Berkley, 1954). Briefly, the electron microscope plates were mounted on a light microscope, which was illuminated from a monochromatic source. The image was viewed by a television camera. One television line, running lengthwise through the segmented long-spaced image, was isolated electronically and the density tracing was projected on an oscilloscope. A sample is shown in Plate IV.

Results

The extent of guanidination and acetylation,

TABLE I

THE ANALYSIS OF THE PRODUCT OF GUANIDINATION OF TROPOCOLLAGEN

Sample	pH	Time (hr.)	Amino Acid Content (μmole/mg protein) Hylys	Lys	Homoarg	% Guanidination[a] Lys	Homoarg	(n)	Clotting Time (min.)
G 1	7.5	1	0.081	0.320	0			20.9	165
G 2	7.5	3	0.077	0.312	0				
G 3	7.5	5	0.071	0.295	0				
G 4	7.5	8	0.067	0.311	0				
G 5	7.5	12	0.068	0.320	0				
G 6	9.5	1	0.066	0.310	0.008	1	3		120
G 7	9.5	3	0.067	0.303	0.008	6	3		75
G 8	9.5	5	0.066	0.277	0.034	13	12	20.0	20
G 9	9.5	8	0.069	0.276	0.051	13	18	20.0	
G 10	9.5	12	0.064	0.260	0.058	19	20	22.6	

[a] Based on the fact that only 90% of the ε-amino groups of lysine are available to acetylation.

TABLE II

THE ANALYSIS OF THE PRODUCT OF ACETYLATION OF TROPOCOLLAGEN

Sample	Ac_2O Added (μmole)	% Acetylation Hylys	Lys	Total[a]	Total[b]	(n)	Clotting Time (min.)
A 29	0	0	0	0	0	15.9	3
A 21	0.5				3.3		6
A 22	1				7.7		10
A 23	2	26	8	12	19		25
A 37	2.3	41	10	16			
A 24	4	45	23	27	30		
A 25	6	72	38	45	42	16.2	>1100
A 26	8	100	50	57	50	17.0	
A 27	10	72	55	59	60	15.2	
A 28	200	79	100	96	95	8.1	

[a] Calculated from the % acetylation of the available hydroxylysine and lysine. [b] From the ninhydrin method on whole protein.

together with the specific viscosity and clotting time, is presented in Tables I and II. It will be noted in Table I that there was no significant guanidination of hydroxylysine under these conditions. As seen in Table II, hydroxylysine acetylates more rapidly than does lysine.

On the basis of the constancy of the intrinsic viscosity measurements, we conclude that the tropocollagen was not denatured by these treatments. The one possible exception was the fully acetylated sample, A 28. The value of 8.1 is too far below the value of approximately 12 cited by Boedtker and Doty (1956) to be considered within experimental error.

It is apparent from the clotting times that the introduction of guanidino groups increases the rate of fiber formation. However, as seen in Plate I, the fibers so formed show a loss of fine structure with increasing amounts of guanidination until at 20% guanidination there is a complete loss of structure. On the other hand, the rate of fiber formation decreases with increasing acetylation. Here too the fibers show a decrease in fiber structure with increased extent of reaction. When the lysyl residues of tropocollagen are about 30–40% acetylated (Plate II), a structureless fiber results. The structureless fibers (part *D* in Plates I and II) appear to have less contrast than the structured fibers despite the fact that the conditions of staining and photography were uniform. No improvement in contrast could be obtained, although micrographs were made from several specimens. Kühn *et al.* (1959) presented photomicrographs of the fibers formed from tropocollagen which had been guanidinated. The extent of guanidination was not reported. They described the fibers as thin and having poor contrast. This is in agreement with our observations on samples which were guanidinated to the extent of 2–16%.

It is apparent in Plate III that both the guanidinated and acetylated samples form segmented long-spaced structures. We have obtained segmented long-spaced structures from the completely acetylated samples. In addition, guanidination yields a wider and more uniform segmented long-spaced structure.

A sample of line-selected oscillography is shown in Plate IV. Figure 2 shows density tracings through three segmented long-spaced structures. Figure 3 shows the tracings through three segmented long-spaced structures obtained from a sample of tropocollagen which was 87% acetylated. Since we used phosphotungstic acid–stained samples, it is very likely that the peaks reflect concentrations of the basic amino acids. With the acetylation of the lysyl and hydroxylysyl residues, the curves in Figure 3 locate the position of the arginyl residues, there being too little histidine to take into account. We were struck by the fact that there were five outstanding peaks located at equidistant positions along the molecule. These peaks are noted by the solid lines.

DISCUSSION

We have obtained the segmented long-spaced form from completely acetylated tropocollagen (sample A 28), demonstrating that the ε-amino groups are not necessary for the formation of the segmented long-spaced structure. The increase in width and uniformity of the segmented-long-spaced structure with increasing guanidination is worthy of note. Kühn *et al.* (1959) postulate that the binding of ATP introduces negative charges into the tropocollagen molecule to replace the carboxyl groups which are protonated at low *p*H. They suggest that, under these conditions, the segmented long-spaced structure is sterically more favorable. It is difficult to understand, on the basis of their interpretation, how the loss of ε-amino groups would have no effect while their conversion to guanidino groups should improve the segmented long-spaced

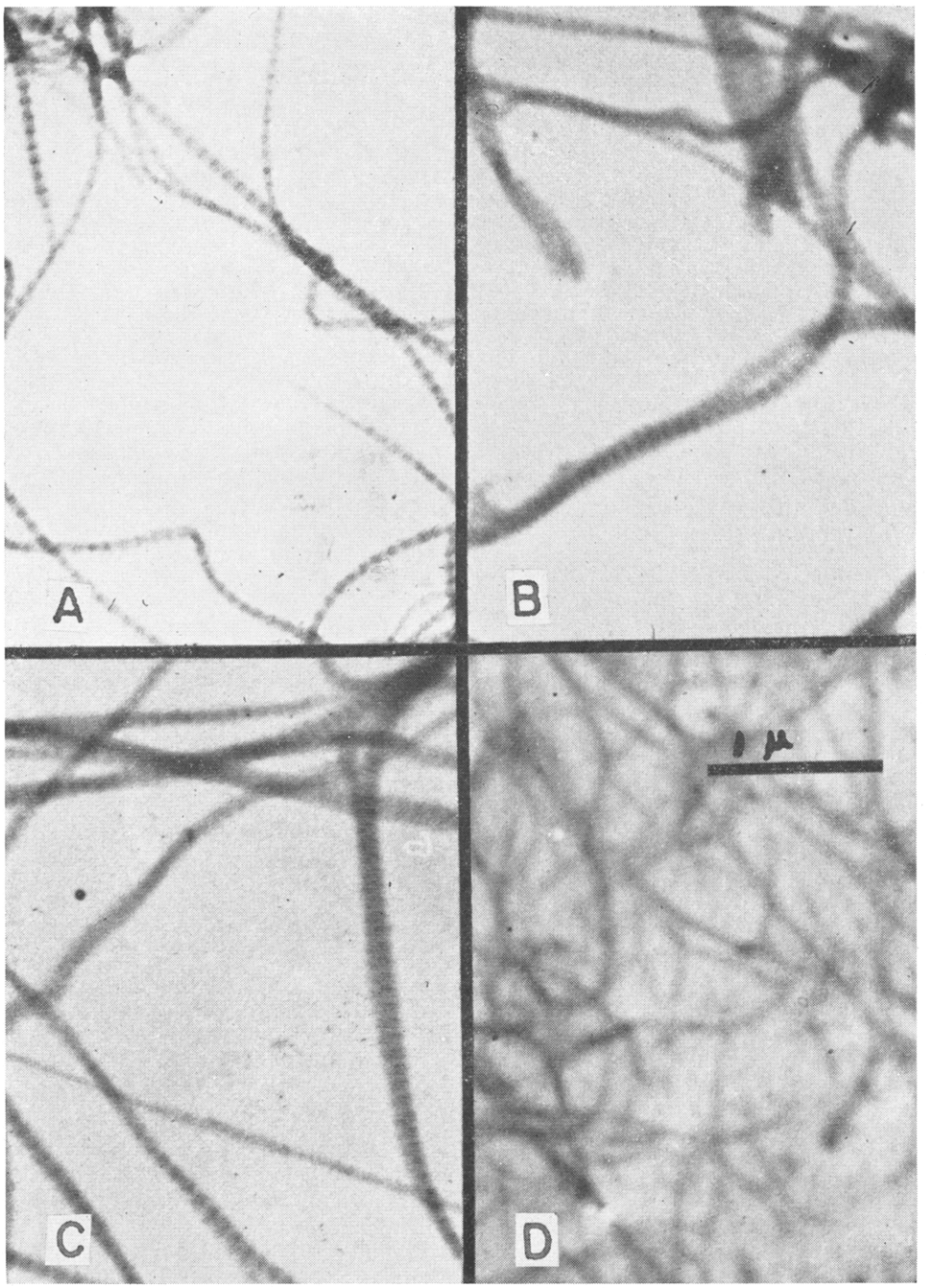

PLATE I.—The fibers formed from guanidinated tropocollagen. The tropocollagen was guanidinated to the extent of 2% in *A*, 13% in *B*, 16% in *C*, and 20% in *D*. ×24,000

structures. Until further information is obtained concerning the binding sites of ATP, no interpretation of our observations can be presented.

It is evident that both guanidination and acetylation yield products which produce nonstriated fibers. However, there is a quantitative difference between the extent of the reactions necessary to produce this effect. The fact that the *p*H of fiber formation (*p*H 9.1) is close to that of the *pK* of the lysyl groups means that many of the unreacted groups are not protonated. This suggests that more than 30–40% of the lysyl groups can be nonionic before a structureless fiber is obtained. We doubt that the mechanism which results in the loss of striations could be the same for both reactions. In one reaction (acetylation) the basic character is lost, whereas with guanidination a basic group is replaced with one that is more

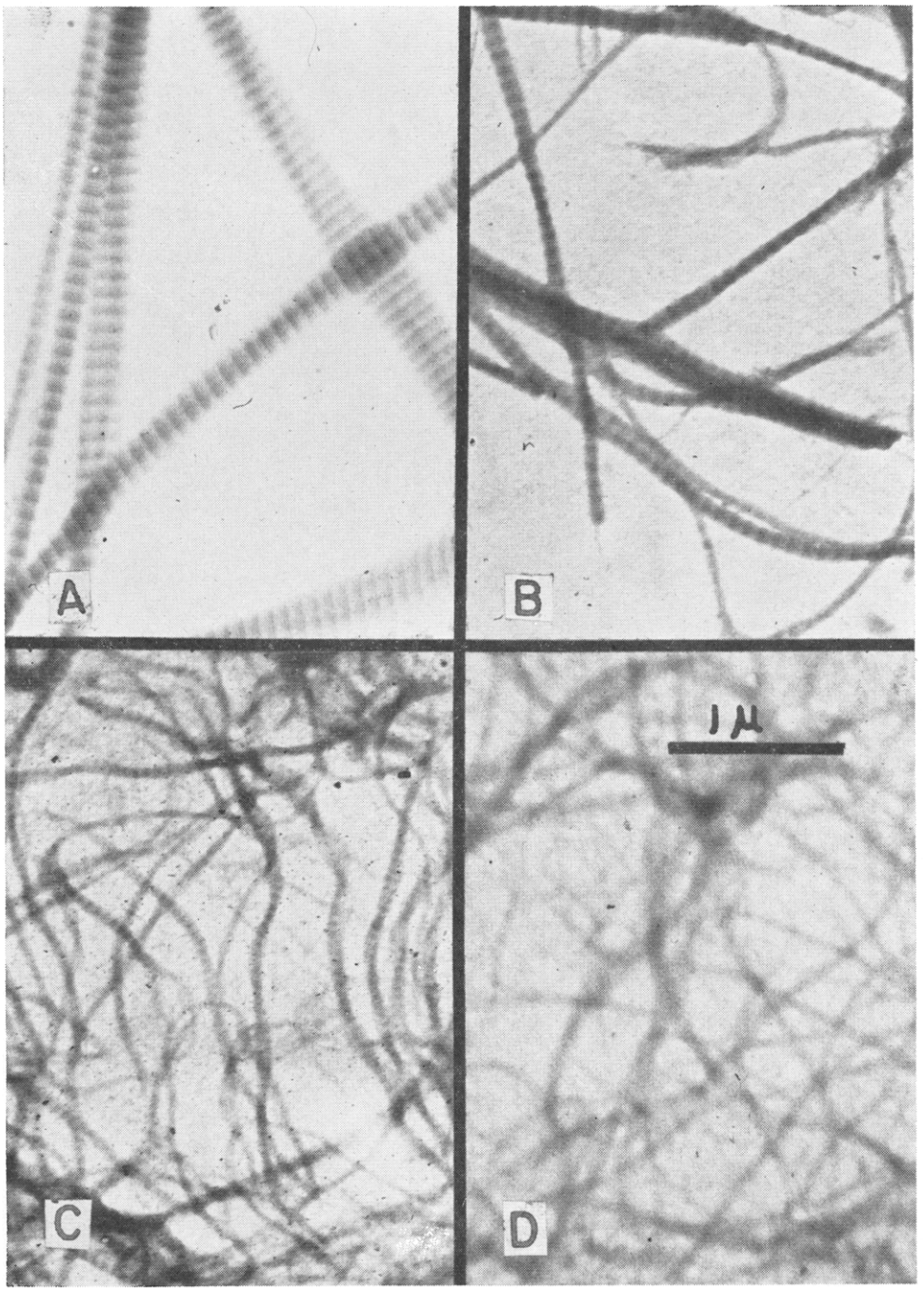

PLATE II.—The fibers formed from acetylated tropocollagen. The lysine in tropocollagen was acetylated to the extent of 0% in *A*, 10% in *B*, 23% in *C*, and 38% in *D*. ×24,000

basic. Our interpretation of the two mechanisms is given below.

If the five main bands of arginyl residues shown in Figure 3 are to be considered significant in determining the quarter-staggered arrangement of the monomers in the formation of the fibers, it must be postulated that the first and last band overlap in the end-to-end polymerization. This overlap would then produce four evenly spaced bands for the sites of attachment (Fig. 4*A*). Hodge and Schmitt (1960) were able to synthesize the characteristic banding of the fiber photographically by successively printing a segmented long-spaced image with a longitudinal displacement between exposures of one fourth the length of the single segment. They have labeled the evenly spaced bands which they used as the loci of attachment $\delta_1 - \delta_4$. The positions of their bands were located

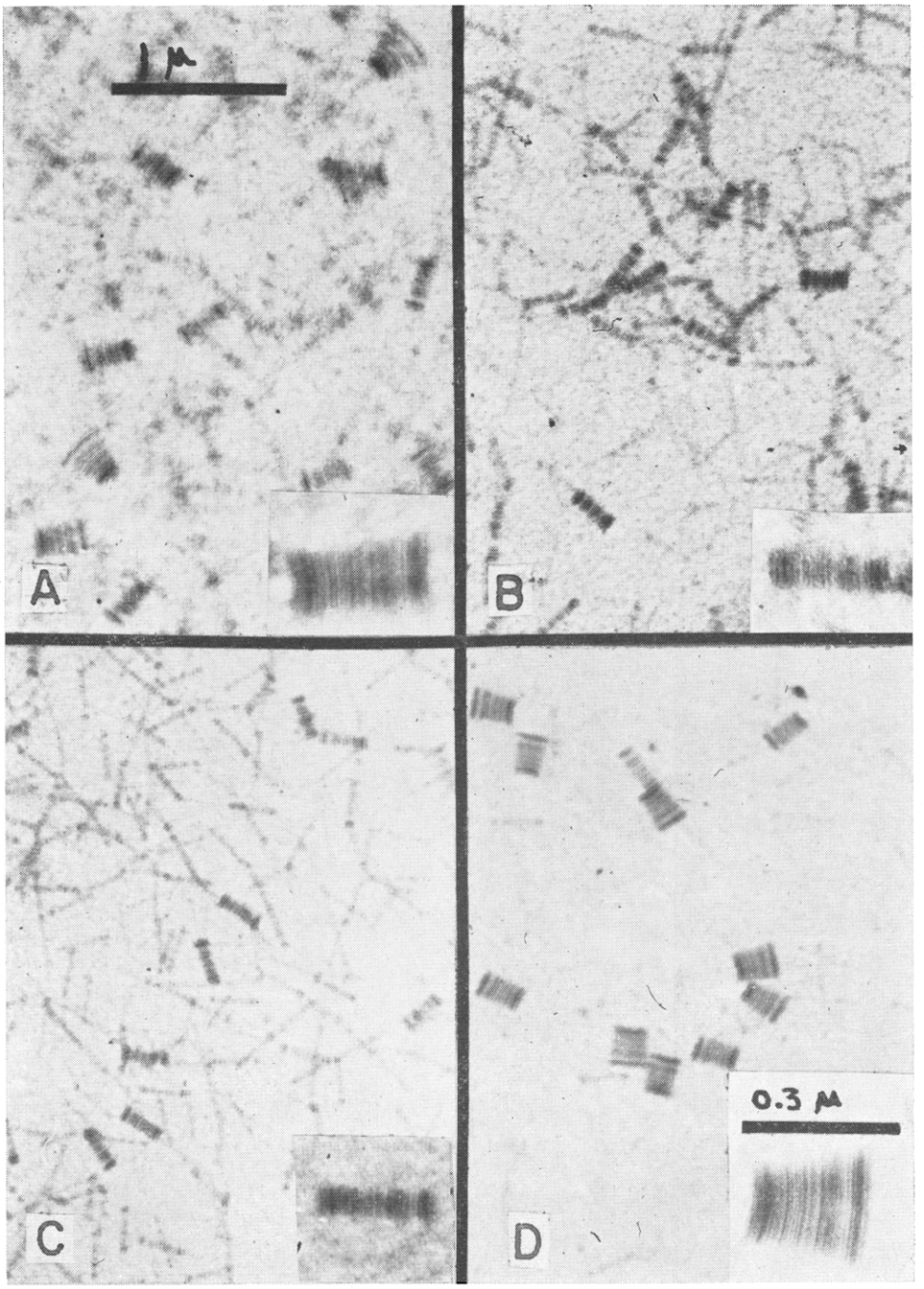

PLATE III.—The segmented long-spaced structures formed from derivatives of tropocollagen using a suboptimal amount of ATP (0.2%). *A* is the untreated control. The protein was acetylated (55% of the available lysyl residues) in *B* and guanidinated to the extent of 13 and 22% in *C* and *D* respectively. ×24,000 (inset ×67,000)

on our monitor photographs and are indicated in Figures 2 and 3 by the broken lines. If we are to conclude that our two outside peaks must overlap in the end-to-end association, the distance between the δ_1 and δ_4 bands must be three fourths of the distance between our two end bands. We have found by measurement that there is agreement within 2%. To the best of our ability to compare our segmented long-spaced structures with those of Hodge and Schmitt (1960) the five positions of high arginine concentration should be labeled β_4^2, β_1^2, β_2^2, β_3^2, and β_4^2 respectively. Hodge and Schmitt (1960) have also shown that acidic and basic residues occur in the same areas within the

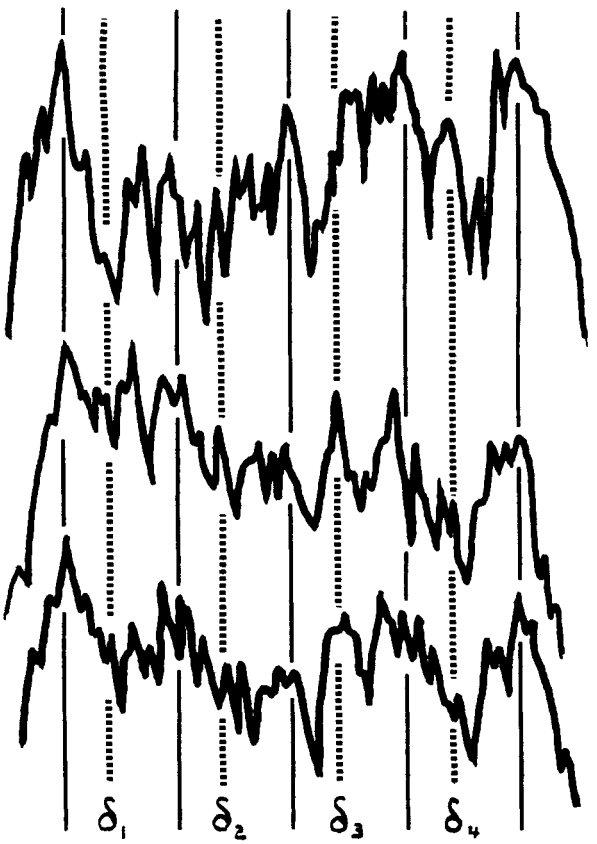

FIG. 2.—The density tracings through three untreated segmented long-spaced structures formed in 0.5% ATP. The solid lines indicate the positions of the five major concentrations of arginyl residues, and the broken lines indicate the positions of the δ bands.

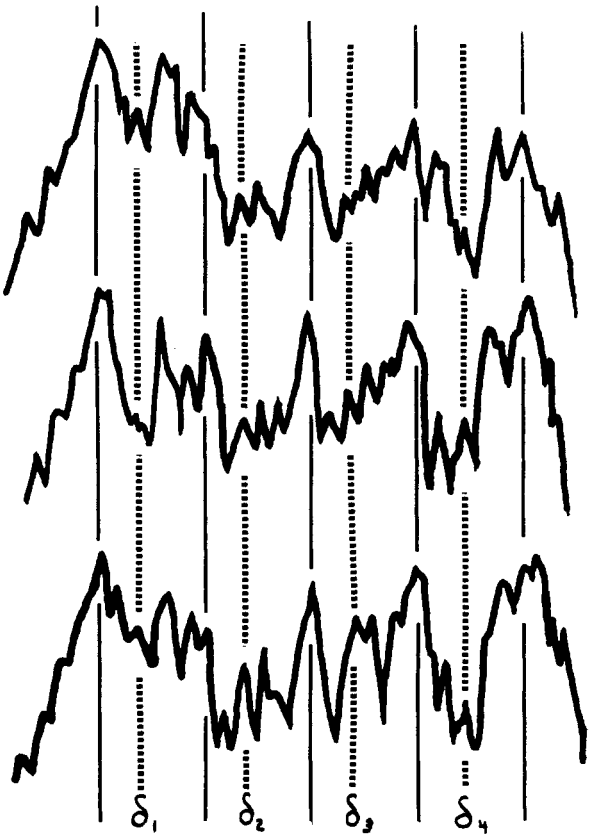

FIG. 3.—The density tracings through three segmented long-spaced structures obtained from tropocollagen which had been 87% acetylated.

molecule. Therefore, we suggest that the formation of ionic bonds between the arginyl residues and the acidic residues in the five equidistant loci could constitute the mechanism for the quarter-staggering during fiber formation.

It now would be evident how guanidination could alter the arrangement of the monomer in the fiber. By the conversion of ε-amino groups to guanidino groups, new loci of interaction would be created which are randomly situated, as diagrammed in Figure 4*B*. Since the homoarginyl residues produced are 1.3 A longer than the arginyl residues, it might well be that these sites become the preferred loci of interaction.

We believe that the mechanism for the formation of nonstriated fibers from acetylated tropocollagen is similar to that for the formation of the structureless fiber at high salt concentration (Gross *et al.*, 1954). We have shown that the decrease in rate of normal fiber formation is the result of the decrease in the over-all activity of charged groups with increased ion-dipole interaction (Bensusan, 1960). With large concentrations of salt it would then be possible to suppress the over-all coulombic interactions to such an ex-

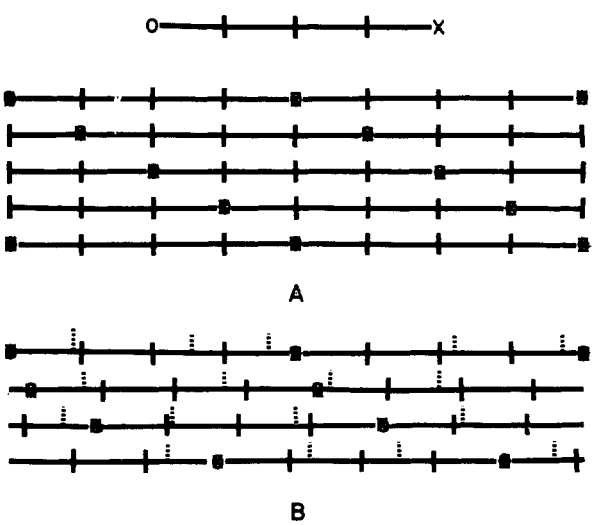

FIG. 4.—The diagrammatic representation of the selective interaction of arginyl residues in the tropocollagen monomer with acidic groups in the same loci to form the quarter-staggered fiber. In *A* the normal arrangement is shown. *B* shows the results of the introduction of additional guanidino groups (broken lines) to form a randomly oriented structure.

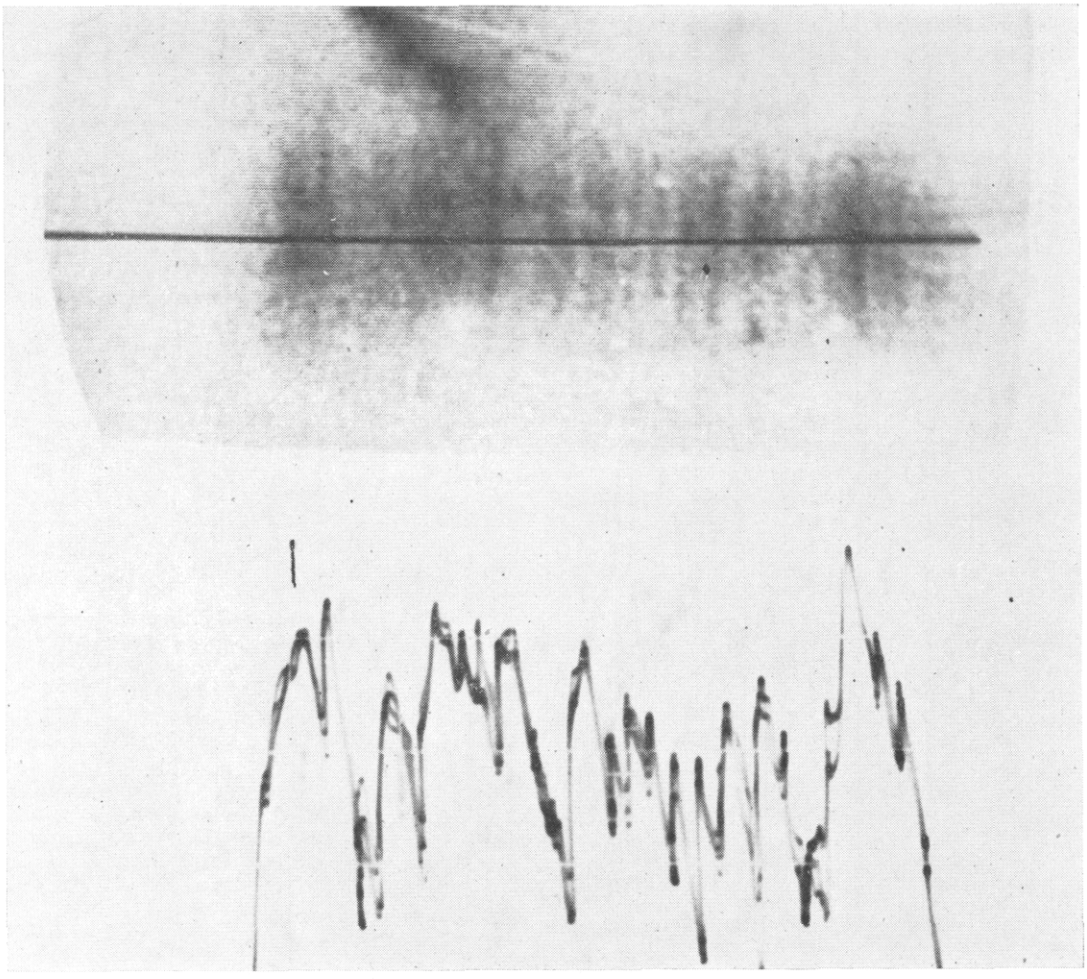

PLATE IV.—The enlarged television picture of the segmented long-spaced structure and the oscillograph tracing of the line indicated in the picture.

tent that a random hydrophobic bonding would predominate. Acetylation would also decrease the over-all coulombic attractions to yield a non-striated fiber in a similar manner. When acetylation proceeds too far, the predominance of anionic groups would prevent fiber formation of any type through electrorepulsion. Thus, it would appear that the ϵ-amino groups are concerned with strengthening coulombic interactions rather than with being the sites for the specific interactions which determine the structure of the fiber. These interpretations are similar to those proposed by Kühn *et al.* (1959) to explain the solubility behavior of methylated and deaminated tropocollagen and the effects of acid mucopolysaccharides and salts on fiber formation.

Martin *et al.* (1961) showed that the residues of histidine had to be nonprotonated before normal fiber formation could take place. They suggested that uncharged histidine residues were responsible for the organization of the normal fiber. Such a conclusion is impossible if one accepts the evidence that normal fiber formation is due to coulombic attractions. We conclude from their evidence that histidine plays no role in the quarter-staggering of the monomer during fiber formation.

ACKNOWLEDGEMENT

We are grateful to Dr. C. N. Loeser for the line-selected oscillography.

REFERENCES

Bensusan, H. B. (1960), *J. Am. Chem. Soc. 82*, 4995.
Bensusan, H. B., and Hoyt, B. L. (1958), *J. Am. Chem. Soc. 80*, 719.
Bensusan, H. B., and Scanu, A. W. (1960), *J. Am. Chem. Soc. 82*, 4990.
Boedtker, H., and Doty, P. (1956), *J. Am. Chem. Soc. 78*, 4267.
Gross, J., Highberger, J. H., and Schmitt, F. O. (1954), *Proc. Nat. Acad. Sci. 40*, 679.
Gross, J., and Kirk, D. (1958), *J. Biol. Chem. 233*, 335.
Habeeb, A. F. S. A. (1960), *Can. J. Biochem. Physiol. 38*, 493.
Hodge, A. J., and Schmitt, F. O. (1960), *Proc. Nat. Acad. Sci. 46*, 186.
Kühn, K. (1958), *Leder 9*, 217.

Kühn, K., Grassmann, W., and Hofmann, U. (1959), *Z. Naturforsch. 14b*, 436.
Kühn, K., Grassmann, W., and Hofmann, U. (1960), *Naturwissenschaften 47*, 258.
Loeser, C. N., and Berkley, C. (1954), *Science 119*, 410.
Martin, G. R., Mergenhagen, S. E., and Scott, D. B. (1961), *Biochim. et Biophys. Acta 49*, 245.
Moore, S., Spackman, D. H., and Stein, W. H. (1958), *Anal. Chem. 30*, 1185.
Neuberger, A., and Sanger, F. (1943), *Biochem. J. 37*, 515.
Randall, J. T., Booth, F., Burge, R. E., Fitton Jackson, S., and Kelly, F. C. (1955), *in* Symposia of the Society for Experimental Biology, No. IX, New York, Academic Press, Inc., pp. 127–147.
Rosen, H. (1957), *Arch. Biochem. Biophys. 67*, 10.
Schmitt, F. O. (1959), *Rev. Modern Phys. 31*, 349.
Schmitt, F. O., Gross, J., and Highberger, J. H. (1952), *Proc. Nat. Acad. Sci. 39*, 459.
Shields, G. S., Hill, R. L., and Smith, E. L. (1959), *J. Biol. Chem. 234*, 1747.

A Stochastic Approach to Statistical Kinetics with Application to Enzyme Kinetics*

ANTHONY F. BARTHOLOMAY

From the Biophysics Research Laboratory, Division of Medical Biology, Department of Medicine, Harvard Medical School, and the Peter Bent Brigham Hospital, Boston

Received January 3, 1962

A new mathematical approach to chemical kinetics which takes into account inherently random aspects of mechanisms of action is applied to the Michaelis-Menten hypothesis. The stochastic difference-differential equation which defines a *stochastic model* for the basic enzyme-substrate reaction is obtained. The probability parameters of the process, which correspond to the ordinary rate constants of the classical equations, are introduced both *axiomatically*, following the general mathematical theory, and by means of *collision theory*. The stochastic Markovian equation gives the rate of change of the probability function of the concentrations of substrate and enzyme-substrate complex; it parallels the system of ordinary differential equations of the classical mathematical model. The classical and stochastic theories are compared in terms of these equations: it is shown that the *deterministic* differential equations may be obtained from the stochastic model by a process of averaging. Thus, the new approach provides a more comprehensive mathematical framework. The stochastic theory predicts randomness, whereas the deterministic theory must add experimental error terms to the basic equations in order to accomodate the random irregularities actually observed in kinetic data. Accordingly, two different types of *irreproducibility* must be distinguished: *experimental* and *inherent*. The latter is characterized by the stochastic model presented in this paper, which brings kinetics into closer relation with the statistical (thermodynamic) treatments of the equilibrium state.

1. INTRODUCTION

During the past 10 years[1] a variety of chemical reaction mechanisms have been examined from the probabilistic point of view of the theory of stochastic processes. The emergent "stochastic models" for various reaction types suggest a general "statistical kinetics" (Bartholomay, 1957). In this paper the new approach will be related to enzyme kinetics.

2. GENERAL AIMS OF THE STOCHASTIC APPROACH TO CHEMICAL KINETICS

The change in approach which the statistical, or stochastic, treatment of kinetics represents relative to the classical deterministic kinetics may be likened to the relationship of statistical mechanics to classical physical mechanics. Whereas classical kinetics has encouraged the conception of the kinetics of a reaction process as the transformation of chemical substances *in bulk*, concentrations forming the units of continuous transformation, the statistical approach grows out of the explicit treatment of these same transformations as compositions of individual molecules and discrete inter- and intramolecular events. The large numbers of such individual molecular events are of course subsumed in the classical treatment and find a place in the "rate constants," but the emergent rate equations have only a *deterministic* context—which excludes the possibility of random fluctuations and cannot, therefore, be directly decomposed into the contributing random events. On the other hand, starting with the more fundamental principles of stochastic models one may pass by averaging processes to the traditional treatment, which therefore finds a place in the new theory in the sense of a central tendency, around which statistical fluctua-

* This work has been supported by the Howard Hughes Medical Institute and a grant-in-aid from the National Institutes of Health of the Department of Health, Education and Welfare, No. H3117(C1).

[1] See bibliography at end of chapter 8 in Bharucha-Reid (1960).