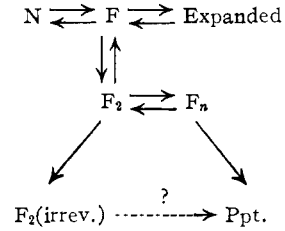
readily salted-out of solution. Cann[39] has also shown that plasma albumin is readily precipitated in acid solution in presence of 0.15 *M* perchlorate. He found a rapid reversible precipitation with a sharp *p*H optimum similar to that in Fig. 3, followed by a slow irreversible precipitation. He considered the rapid precipitation to involve the isomeric form of the protein and the slow precipitation to procede through the expanded form. It appears that his results could equally well be explained through a slow reaction proceeding through the F form.

It is suggested that all of the phenomena discussed above can be explained through a single reaction pattern as illustrated in the reaction diagram, where F_n is a soluble polymer or reversible precipitate, depending on conditions. At room temperature and low ionic strength (for example 0.02 *M* chloride) the expansion equilibrium is favored and



the activity of F form might never be high enough to result in aggregation. At higher ionic strength expansion is strongly repressed[4] and the dimerization equilibrium would be favored, with subsequent slow formation of irreversible dimer. At very high ionic strength (in chloride) or moderately high ionic strength in presence of strongly bound anions such as perchlorate expansion is almost completely repressed and formation of higher aggregates, both reversible and irreversible, favored due to increased screening and reduction of the net positive charge on the protein. At elevated temperatures, as in heat denaturation, all processes are presumably speeded, especially the slow irreversible step.

(38) M. R. Rachinsky and J. F. Foster, *Arch. Biochem. Biophys.*, **70**, 283 (1957).
(39) J. Cann, *J. Phys. Chem.*, **63**, 1545 (1959).

[CONTRIBUTION FROM THE CHEMISTRY DEPARTMENT AND OCEANOGRAPHIC INSTITUTE OF THE FLORIDA STATE UNIVERSITY TALLAHASSEE, FLA.]

The Thermal Copolymerization of Amino Acids Common to Protein[1]

BY SIDNEY W. FOX AND KAORU HARADA

RECEIVED NOVEMBER 19, 1959

By use of an excess of the dicarboxylic amino acids, the dry, eighteen amino acids common to proteins can be simultaneously copolymerized in a few hours at temperatures above 150°. Changes in the constitution of the polymers are described for variations in the conditions of reaction. The properties of the synthetic products are compared with known properties of natural proteins.

Attempts to synthesize protein under hypothetically primitive terrestrial conditions[2] have yielded anhydropolymers which contain all of the eighteen amino acids common to protein and which otherwise resemble protein in many of their properties.[3] (Such materials are referred to as *proteinoids*.) The critical conditions are the maintenance of a hypohydrous state, as by a temperature of above 100°, and the employment of a sufficient excess of dicarboxylic amino acid. The decomposition ordinarily encountered during uncontrolled heating of amino acids above 150° results particularly from the neutral amino acids in the absence of dicarboxylic amino acid.[4] In thermal copolymerizations, the protective effect of excess dicarboxylic amino acid is lost to a significant degree when the temperature exceeds 210°.[5]

One salient feature of the thermal copolymerizations is the finding that the order of amino acid residues in the polymeric products is not completely random, as judged by comparison of total composition and N-terminal composition. Such an effect was first observed most definitively in the pyrocondensation of aspartic acid and glutamic acid.[5] It is of particular interest to compare such effects when the sixteen other amino acids are copolymerized simultaneously ("panpolymerized") with an excess of the dicarboxylic amino acids, inasmuch as such studies may help to explain sequences in natural proteins. Comparisons of this type, similar studies of total amino acid composition, and the effect of other factors on the yield and chemical nature of the products are reported in this paper.

In addition, data presented earlier and data in this paper are reviewed in comparing the synthetic product with natural protein.

Experimental

Proportions of Reactants.—The proportions of all sixteen amino acids exclusive of the dicarboxylic amino acids were the same in all experiments, 0.03 mole each of DL-alanine, L-arginine monohydrochloride, L-cystine, glycine, L-histidine hydrochloride monohydrate, DL-isoleucine, DL-leucine, L-lysine monohydrochloride, DL-methionine, DL-phenylalanine, L-proline, DL-serine, DL-threonine, DL-tryptophan, DL-tyrosine and DL-valine. This mass was intimately ground to a fine powder. Mixture A consisted of 10 g. of L-glutamic acid, 10 g. of DL-aspartic acid and

(1) Presented in part at the 135th Meeting of the American Chemical Society at Boston on April 9, 1959. This research is supported by Grant G-4566 of the National Science Foundation, by the General Foods Corporation, and by Grant C-3971 of the National Institutes of Health, U. S. Public Health Service. Contribution No. 138 of the Oceanographic Institute of The Florida State University.
(2) S. W. Fox and M. Middlebrook, *Federation Proc.*, **13**, 211 (1954).
(3) S. W. Fox and K. Harada, *Science*, **128**, 1214 (1958); S. W. Fox, K. Harada and A. Vegotsky, *Experientia*, **15**, 81 (1959).
(4) S. W. Fox and G. D. Maier, General Petroleum Geochemistry Symposium, Fifth World Petroleum Congress, June 4, 1959, p. 9.
(5) K. Harada and S. W. Fox, *Arch. Biochem. Biophys.*, **86**, 274 (1960).

5 g. of the material described. Mixture B consisted of 5 g. of each. In mixture B, accordingly, there were 16 parts of each acidic amino acid per part of each neutral or basic amino acid.

Typical Panpolymerization.—Ten grams of L-glutamic acid was heated at 175–180° until molten (about 30 min.) after which period it had been largely converted to the lactam. At this time, 10 g of DL-aspartic acid and 5 g. of the mixture of the sixteen basic and neutral (BN) amino acids were added. The solution was then maintained at 170 ± 2° under an atmosphere of nitrogen for varying periods of time. Within a period of a few hours considerable gas had been evolved, and the color of the liquid changed to amber. The vitreous mixture was rubbed vigorously with 75 ml. of water, which converted it to a yellow-brown granular precipitate. After overnight standing, the solid was separated by filtration. This was washed with 50 ml. of ethanol, and as substance S dialytically washed in moving Multidialyzers in water for 4 days, the water being changed thrice daily. The term *dialytic washing* indicates dialytic treatment of a suspension. In some preparations, the solid was dissolved completely in sodium bicarbonate solution and then dialyzed. The dialysis sacs were made of cellulose tubing, 27/32 in., to contain 50 ml. The non-diffusible material was ninhydrin-negative before the fourth day. The non-aqueous contents of the dialysis sac were mainly solid A and a soluble fraction B recovered as solid by concentration in a vacuum desiccator. The mother liquor of S was also dialyzed for 4 days, and then dried to give additional solid C.

Amino Acid Assay of Proteinoid.—Analysis of amino acids as DNP-aspartic acid, DNP-glutamic acid, and as total DNP derivatives of the other sixteen amino acids was executed as in earlier work[6] except that a third band, (DNP-BN) separate from those of each of the dicarboxylic amino acids, was assayed. Colorimetric assays were performed at 430 mμ under acidic conditions to eliminate absorption by dinitrophenol.

One sample of proteinoid was submitted for analysis to the Shankman Laboratories of Los Angeles. Content of water in eight samples varied between 10.7% and 14.6%, as determined by loss of weight in an Abderhalden drying pistol at 80° over phosphorus pentoxide.

Mean Chain Weight of Proteinoid.—These values were estimated as described earlier.[5]

Color Tests.—Color tests were performed in the standard manner.[7]

Elemental Analysis.—C, H, N analyses were performed by the Geller Laboratories. The figures obtained were corrected for the proportions of ash and of water.

Infrared absorption spectra were determined on Nujol mulls by Mrs. Mei-Wei Chen using a Perkin-Elmer model 21 recording spectrophotometer.

Electrophoretic Mobility.—Preliminary studies were kindly performed by Dr. C. S. Vestling of the University of Illinois on a partially purified sample with phosphate sodium chloride buffer at *p*H 7.5 and $\Gamma/2 = 0.1$, according to Miller and Golder.[8]

Proteolyzability.—The effects of two proteolytic enzymes on samples of proteinoid were compared with effects of the same enzymes on casein prepared by Dr. Charles M. Ise from fresh milk.

In the control for pepsin 34 mg. of casein and 3.4 mg. of Mann Research Laboratories pepsin (twice crystallized) were dissolved in 3.4 ml. of 0.1 *M* citrate buffer at *p*H 2.2. In 30 ml. of the same buffer were dissolved 30 mg. of proteinoid and 3.0 mg. of pepsin. Almost the same amounts were used in other tests in which the enzyme was Armour crystalline chymotrypsin in *p*H 8.5 phosphate buffer.

At regular intervals the proteolyzate was sampled, and an aliquot was converted to the DNP derivative by the use of DNFB in aqueous ethanol kept slightly alkaline with sodium bicarbonate. The solvent was removed under reduced pressure and the unchanged DNFB was carefully extracted with ethyl acetate. The extracted aqueous solution was hydrolyzed with pilute hydrochloric acid under reflux for 2–3 hours to eliminate turbidity. The liquid was concentrated until the flask was almost dry. The clear liquid was diluted to 100 ml. with water and the absorption measured at 430 mμ in a Spectronic 20 photometer.

Nutritive Quality.—Tubes containing 5 ml. of peptone broth each were made up for *Lactobacillus arabinosus* 17-5 (ATCC 8014) from a solution of 5 ml. of vitamin mixture, 10 g. of D-glucose, 8 g. of Bacto peptone, 1 g. of Bacto yeast extract, 1 g. of sodium acetate, 5 ml. of salt A solution, 5 ml. of salt B solution and water to 1 l. The vitamin mixture contained 50 mg. of thiamine hydrochloride, 10 mg. of pyridoxine hydrochloride, 10 mg. of calcium pantothenate, 10 mg. of riboflavin, 20 mg. of nicotinic acid, 10 mg. of *p*-aminobenzoic acid, 250 γ of biotin, 50γ of folic acid, 150 ml. of ethanol and water to 200 ml. The salt A solution contained 20 g. each of K_2HPO_4 and KH_2PO_4 in 200 ml. of solution. The salt B solution contained 4.0 g. of $MgSO_4 \cdot 7H_2O$, 0.2 g. of NaCl, 0.2 g. of $FeSO_4 \cdot 7H_2O$ and 0.2 g. of $MnSO_4 \cdot 4H_2O$ in 100 ml. of solution.

In other tubes, peptone was substituted by various preparations of proteinoid.

After sterilization in an autoclave, the tubes were inoculated with washed bacteria and then incubated at 37° for 72 hours. Growth was determined by titration with 0.1 *N* sodium hydroxide and brom thymol blue as indicator.

Antigenicity.—A sample of proteinoid was supplied to Dr. Arthur Cherkin of the Don Baxter, Inc., Laboratories where the material was tested for anaphylactogenicity in whole guinea pigs and in the Schultz-Dale test on isolated uterine strip by R. E. Marshall. Another sample was supplied to Dr. A. W. Turner, of the A. E. Staley Manufacturing Co., who tested it on rabbits.

Results and Discussion

In Table I is found the analysis of the sample of proteinoid submitted to microbial assay of the hydrolyzate.

TABLE I

ASSAY OF AMINO ACIDS IN HYDROLYZATE OF PROTEINOID PREPARED FROM 2:2:1 WEIGHT RATIO OF ASPARTIC ACID: GLUTAMIC ACID:MIXTURE OF BASIC PLUS NEUTRAL AMINO ACIDS (BN)

Isomer detd.	Amino acid	G. amino acid / 100 g. polyamino acid
DL	Alanine	3.0
L	Arginine	1.2
DL	Aspartic acid	67.5
L	Cystine[a]	0.1
DL	Glutamic acid	10.7, 7.8
L	Histidine	1.4
L	Isoleucine	1.0
L	Leucine	1.4
L	Lysine	2.4
DL	Methionine	2.1
L	Phenylalanine	1.7
L	Proline	0.3
L	Serine	.5
L	Threonine	.1
L	Tyrosine	1.8
L	Tryptophan[b]	1.0
L	Valine	0.9
	Glycine	2.0

[a] By hydrolysis in formic acid–hydrochloric acid (1:1). [b] By alkaline hydrolysis.

The total of amino acids recovered is less than 100%, even when allowance is made for moisture and for the fact that many of the amino acids are only partially determined by the relatively stereospecific assay microbes. The finding that aspartic acid is the principal component and glutamic acid the one next most represented in this particular polymer is in essential agreement with the DNP

(6) K. Harada and S. W. Fox, THIS JOURNAL, **80**, 2694 (1958).

(7) P. B. Hawk, B. L. Oser and W. H. Summerson, "Practical Physiological Chemistry," Thirteenth Edition, McGraw-Hill Book Co., Inc., New York, N. Y., 1954, p. 169.

(8) G. L. Miller and R. H. Golder, *Arch. Biochem.*, **29**, 420 (1950).

assays in the later tables. The finding of significant amounts of individual amino acids also verifies the qualitative identification of these units.[3]

In Table II are presented the effects of time of heating on the yields of the various fractions when the reactants are present in weight ratios of 2:2:1 of aspartic acid:glutamic acid:BN amino acids and of 1:1:1. With more extended heating, a larger proportion of product is found to be non-diffusible, which fact suggests an increasing mean chain weight. The proportion of soluble proteinoid is small, being at a maximum of 13% after a 2-hour heating period. All fractions were ninhydrin-negative and biuret-positive with the exception of the 2-hour sample, fraction B, which gave a weak ninhydrin test. Columns S and A in Table II suggest an autocatalytic synthesis. It should be remembered, however, that the weights recorded are those of non-diffusible materials and the picture presented is accordingly not fully representative of the entire product. All B and C fractions were each less than 0.1 g. When the ratios of reactants are 2:2:1 instead of 1:1:1, *i.e.*, the proportions of dicarboxylic amino acid are higher, total yields are mostly relatively larger.

TABLE II

YIELDS OF FRACTIONS IN PANPOLYMERIZATION OF AMINO ACIDS AT 170° AS A FUNCTION OF TIME

Time, hr.	Fraction S, g.	Fraction A, g.
25 g. of 2 DL-aspartic acid:2 L-glutamic acid:1 BN amino acids		
2	2.40	0.56
3⅔	4.01	2.74
6	9.41	6.20
15 g. of 1 DL-aspartic acid:1 L-glutamic acid:1 BN amino acids		
2	1.38	0.31
4	1.42	0.77
6	1.73	1.01

In a simultaneous experiment comparing the two sets of proportions, yield was also higher at the reaction ratio of 2:2:1 (over 70% higher in a 4-hour heating period).

The new ratio included in Table III, 1:5:3.5, was studied because it was found to yield a product in which the proportions of amino acids are similar to those of typical proteins of current organisms.[9] The proportions of reactants are those actually used. In order to compare the yields most meaningfully the amounts recorded in Table III are all calculated to a uniform basis of 25 g. of reactant. The need for a high proportion of the acidic amino acids in providing high yields is again evident. The products in this case are also ninhydrin-negative and biuret-positive, except that those from the 1:5:3.5 ratio were too dark for such tests. As in the absence of H_3PO_4, the hot reaction mixtures were in a single phase. When compared to other results fraction A is found to be four times as large in the H_3PO_4 reaction mixture at four hours as in the absence of the mineral acid. Other comparisons of results show the same trend. Within Table IV it is again seen that the highest specific yield is obtained at a 2:2:1 ratio and the lowest at a 1:5:3.5 ratio. In the third and fourth reactions listed, the only difference is that the aspartic acid in one case is L, in the other DL. Significant differences in yield between these two cases are not observed. Each of the effects described in this paragraph is comparable to that observed with the simple glutamic acid–aspartic acid copolymer.[5]

(9) S. W. Fox and P. G. Homeyer, *Am. Naturalist*, **89**, 163 (1955).

TABLE III

YIELDS[a] OF FRACTIONS IN PANPOLYMERIZATION OF AMINO ACIDS IN THREE PROPORTIONS AT 170°

Proportions of reactants (DL-aspartic acid: L-glutamic acid: BN amino acids) g.	Hr. of heating	Fraction S, g.	Fraction A, g.	Fraction B, g.	Fraction C, g.
10:10:5	10	8.40	7.20	0.07	0.03
5:5:5	10	4.17	3.52	.10	.08
1:5:3.5	7	0.24	0.14	.02	.17

[a] All yields are calculated to a basis of 25 g. of reactant mixture.

TABLE IV

YIELDS[a] OF FRACTIONS IN PANPOLYMERIZATION OF AMINO ACIDS IN THREE PROPORTIONS IN THE PRESENCE OF PHOSPHORIC ACID

Proportions of reactants (DL-aspartic acid: L-glutamic acid: BN amino acids)	Conditions of reaction	Fraction S, g.	Fraction A, g.	Fraction B, g.
1:5:3.5	2.0 ml. 85% H_3PO_4, 170° 3.5 hr.		0.60	0.06
5:5:5	3.0 ml. 85% H_3PO_4, 170° 3.5 hr.	3.12	2.84	0.02
4:4:2	2.5 ml. 85% H_3PO_4, 170° 3 hr.		4.15	0.08
4:4:2[b]	2.5 ml. 85% H_3PO_4, 170° 3 hr.	4.23	4.08	0.05

[a] All yields are corrected to a basis of 15 g. of reactant mixture. [b] L-Aspartic acid instead of DL-aspartic acid.

An additional point of interest is that the phosphate content of the 5:5:5 product was only 0.048%.

The effect of temperatures, in the working range, on the yields of fractions of proteinoid in two sets of proportion of reactants is given in Table V. The effect of increase in temperature is increase of yield up through 190°. Again, there is observed the favorable effect on yield of having twice as great an excess of dicarboxylic amino acid in one case as in the other. Fractions B and C were each less than 0.03 g.

Among the comparisons possible from the data in Table VI is that which concerns the relative proportions of amino acids in the products. A most salient feature is that of the proportions of basic and neutral amino acids when twice as much of these are included in the reaction. The proportion of these is greatly increased in the product, albeit at a sacrifice in total yield (Table III) and apparent purity. This gain in BN amino acids appears to be mainly at the expense of aspartic acid, the major

TABLE V

EFFECT OF TEMPERATURE ON YIELDS OF FRACTIONS IN PANPOLYMERIZATION OF AMINO ACIDS

Proportion of reactants (DL-aspartic acid: L-glutamic acid: BN amino acids)	Temp., °C.	Fraction S, g.	Fraction A, g.
3:3:1.5	160	2.11	1.96
	170	2.29	2.08
	180	2.48	2.35
	190	2.82	2.67
3:3:3	160		0.84
	170		0.91
	180		1.26
	190		1.75

component of the polymer. With increase in time of heating, the proportion of BN amino acids in the 1:1:1 reaction, but not in the 2:2:1 reaction, rises significantly. The proportions in fractions A and B differ particularly in the glutamic acid component, in all cases.

TABLE VI

EFFECT OF TIME OF HEATING ON COMPOSITION OF PROTEINOIDS PREPARED FROM TWO RATIOS OF REACTANTS

Time of heating at 170°, hr.	Proportion in fraction A, % Aspartic acid	% Glutamic acid	% BN amino acids	Proportion in fraction B, % Aspartic acid	% Glutamic acid	% BN amino acids
Reaction ratio of 2:2:1						
2	68	14	18	87	7	6
3⅔	69	14	18	62	21	18
6	71	11	17	59	23	18
Reaction ratio of 1:1:1						
2	60	11	29	67	10	23
4	57	12	31	58	18	24
6	55	11	34	56	17	27

Because of the higher proportion of BN amino acids in the 1:1:1 product, a possibly more favorable ratio, 1:5:3.5, was examined. The product contained 33% aspartic acid, 27% glutamic acid and 40% BN amino acids.

The principal effect of increasing temperature on the composition of the proteinoids is to increase slightly the proportion of BN amino acids concomitantly with decrease in aspartic acid (Table VII). As in the experiment of Table VI, the proportion of BN amino acids is higher in the 1:1:1 reaction product.

TABLE VII

EFFECT OF TEMPERATURE ON PERCENTAGE COMPOSITION OF PROTEINOIDS[a] PREPARED FROM TWO RATIOS OF REACTANTS

Temp., °C.	Aspartic acid	Glutamic acid	BN amino acids
2:2:1 Ratio[b]			
160	66%	13%	22%
170	65	12	23
180	63	13	24
190	56	14	29
1:1:1 Ratio[b]			
160	56	14	30
170	55	13	32
180	55	14	31
190	46	19	35

[a] Fraction A. [b] 6 hr.

The mean chain weight[5] of the proteinoid (Table VIII) is smaller than that of the aspartic acid–glutamic acid copolymer prepared under similar conditions.[5] The effect of increase in temperature on increase in molecular weight is more marked in the case of the proteinoid, perhaps because the molecular weights at the lower temperatures are smaller.

TABLE VIII

EFFECT OF TEMPERATURE ON MEAN CHAIN WEIGHTS OF PROTEINOIDS[a]

Temp., °C.[c]	Mean chain weight[b] 2:2:1 ratio	Mean chain weight[b] 1:1:1 ratio
160	4600	3600
170	4500	3800
180	5500	4100
190	7200	8600

[a] Fraction A after dialytic washing, conversion to sodium salt, and repeated dialysis. [b] By DNP assay. [c] 6 hr.

As Table IX shows, the effect of temperature on the N-terminal composition is marked. The drop in N-glutamic acid with rise in temperature again suggests the possibility that terminal glutamic acid residues are being cyclized[5]; this explanation is unlikely because of the fact that the N-aspartic acid value does not undergo a corresponding increase over this temperature range. Also, with increasing temperature the N-glutamic acid value in the aspartic acid–glutamic acid copolymer increases relatively.[5] The high proportion of basic and neutral amino acids in the N-terminus in all samples is intrinsically noteworthy.

TABLE IX

EFFECT OF TEMPERATURE ON N-TERMINAL AMINO ACID COMPOSITION OF PROTEINOIDS[a] PREPARED FROM TWO RATIOS OF REACTANTS

Temp., °C.	N-Aspartic acid N-amino acids, %	N-Glutamic acid N-amino acids, %	N-BN amino acids N-amino acids, %
2:2:1 Ratio[b]			
160	7	57	36
170	6	46	48
180	11	37	53
190	8	13	80
1:1:1 Ratio[b]			
160	8	42	50
170	7	30	63
180	10	20	71
190	9	10	81

[a] Fraction A. [b] 6 hr. under N_2.

It is of interest that the ratios of amino acids in the products are markedly different from those in the reaction mixtures. The ratios of amino acids in the N-terminal position furthermore are different from the ratios in the total composition. These results suggest that the reactant amino acids themselves contribute to the determination of composition and order of amino acid residues, and that the arrangements cannot be completely random. Such results are of the kind visualized in the working premises at the inception of this investigation.[10]

At least three phenomena observed in copolymerization of amino acids are noteworthy. One of

(10) S. W. Fox, *Am. Scientist*, **44**, 347 (1956); S. W. Fox, M. Winitz and C. W. Pettinga, THIS JOURNAL, **75**, 5539 (1953).

these is the internal effect on composition and arrangement in the product. Another fundamental fact is that amino acids which do not thermally polymerize alone do so together. Finally, the physical properties such as solubility of the copolymers differ from those of the corresponding homopolymers, as also observed in polyamino acids prepared through Leuchs anhydrides.[11]

Behavior of Thermal Polymers in Tests for Protein

Many of the criteria which are usually applied to natural proteins have now been applied to one or more of the synthetic polymers. The results are reviewed in the following paragraph.

Complement of Amino Acids.—The eighteen amino acids which are common to nearly all proteins are found, except for tryptophan, in the acid hydrolyzates of the polymers described. Tryptophan has been found in unhydrolyzed proteinoid by color test and in alkaline digests by microbial assay. The near universality of this roster of amino acids has been noted by a number of authors.[9,12,13] Each of the latter two authors designates twenty amino acids, but they recognize glutamine and glutamic acid separately, and asparagine and aspartic acid separately. These four amino acids reduce to two on total hydrolysis so that their twenty is equivalent to the eighteen studied here.[14]

The finding of seventeen amino acids chromatographically[3] is confirmed by the microbial assay reported.

One remarkable feature of the polycondensation is the ease with which sixteen amino acids enter into peptide linkage with the two dicarboxylic amino acids in excess. This phenomenon is reflected in the ease with which each of the sixteen amino acids individually is copolymerized in substantial proportion with glutamic acid and aspartic acid.[15]

Average Molecular Weight (Mean Chain Weight).—The polymers described here reveal mean molecular weights in the range of 3000–9000 as assessed by peptide chain weight per assayable end-group.[3] Preparations of weight 3,000 have given essentially the same figure in analyses by Vegotsky in the ultracentrifuge,[16] and in dialytic escape time.[17] This molecular weight range is at the lower range of molecular sizes for proteins. Such comparisons are best made on the basis of chain weight defined as mass of peptide per terminal amino acid residue. On this basis, insulin chain A has a weight of 2,700, chain B 3,800.[18]

(11) R. R. Becker and M. A. Stahmann, THIS JOURNAL, **76**, 3707 (1954).

(12) R. L. M. Synge, lecture on "The Occurrence of Amino Acids in Nature," in International Symposium on the Origin of Life on the Earth, Moscow, U.S.S.R., 1957; published by Pergamon Press, London, 1959. p. 224.

(13) M. Yeas, Symposium on Information Theory in Biology, Pergamon Press, New York, N. Y., 1958, p. 70.

(14) Analyses by Dr. Kenneth Woods of an acid hydrolyzate of a polymer prepared by Allen Vegotsky indicates significant proportions of ammonia. This might have arisen from ammonium salt or amide formed during the pyrocondensation. Dr. Woods' analyses also indicate doubtful proportions of serine and theonine.

(15) K. Harada and S. W. Fox, unpublished experiments.

(16) A. Vegotsky, unpublished observations, 1959.

(17) L. C. Craig and T. P. King, THIS JOURNAL, **77**, 6620 (1955).

(18) E. J. Harfenist and L. C. Craig, *ibid.*, **74**, 3087 (1952).

Color Tests.—The standard color tests which are positive for protein include the biuret, xanthoproteic, Hopkins–Cole, and Salkowski tests. The last two are indicative of tryptophan residues. Inasmuch as all four tests were positive for the synthetic material, as for natural proteins, it can be concluded that the thermal synthesis does not result in profound decomposition or in new groups or configurations masking all peptide bonds, benzenoid residues or indole rings. In warm solutions with ninhydrin, the proteinoid is found to yield color (presumably from cleaved amino acids) more slowly than from natural proteins.

Infrared Absorption Spectrum.—The absorption bands found both in the proteinoid and protein are 3300 cm.$^{-1}$ (–NH–), 3080 cm.$^{-1}$ (–NH–), 1650 cm.$^{-1}$ (amide I), and 1550 cm.$^{-1}$ (amide II).[19] The last band is due to carbonyl radicals not observed in diketopiperazines.

In addition, two strong bands at 1720 and 1780 cm.$^{-1}$ indicate imide radicals.[20] Although the presence of a significant proportion of imide bonds may serve to distinguish proteinoid from protein, adequate data on this structural feature in protein are not available, and the imide bond is replaced by the peptide bond by gentle alkaline treatment.[20]

Solubility.—The solubility varies of course with the preparation. A typical proteinoid is nine-tenths insoluble in water and one-tenth water-soluble. Water-insoluble fractions of some preparations can, however, be salted in as by 0.1 *M* phosphate buffer at *p*H 7.9, or in 3.5% sodium chloride solution.

Elemental Analysis.—The C, H, N values found for samples of proteinoid (C = 51.0, H = 4.79, N = 13.2) is in the range of contents of protein, except that the N is slightly low. This can be understood as due to the high content of aspartic acid, polyaspartic acid itself having a theoretical value of 12.4% and analytical values are close to that figure.[20]

Salting-out.—The tendency to be salted in has been mentioned. The synthetic polymer is also salted out in the same way as natural protein by adding several volumes of saturated ammonium sulfate solution to the salted-in solution.

Electrophoretic Mobility.—The synthetic materials indicate their charged state in several ways. One of these is the electrophoretic mobility. One preparation which had not been redialyzed from sodium bicarbonate and might therefore have contained some unchanged diketopiperazine (Fig. 1) showed after twenty minutes a set of essentially neutral components which remained at the position of the initial boundary and a set of anionic components of high mobility. This latter is consistent with the high aspartic acid content in the sample studied.

Proteolyzability.—In Fig. 2 are seen the effects of chymotrypsin and of pepsin on two preparations of polymer. The effects of aliquots of the same solutions of enzyme on casein are also shown. In

(19) L. J. Bellamy, "The Infrared-red Spectra of Complex Molecules," Methuen and Co., Ltd., London, 1958, pp. 223–230.

(20) A. Vegotsky, K. Harada and S. W. Fox, THIS JOURNAL, **80**, 3361 (1958). Data which have accumulated indicate a modification of the inference of likelihood of easy recyclization of opened imide bonds. The imide is now found to be opened by gentle treatment, but reformed only by renewed heating.

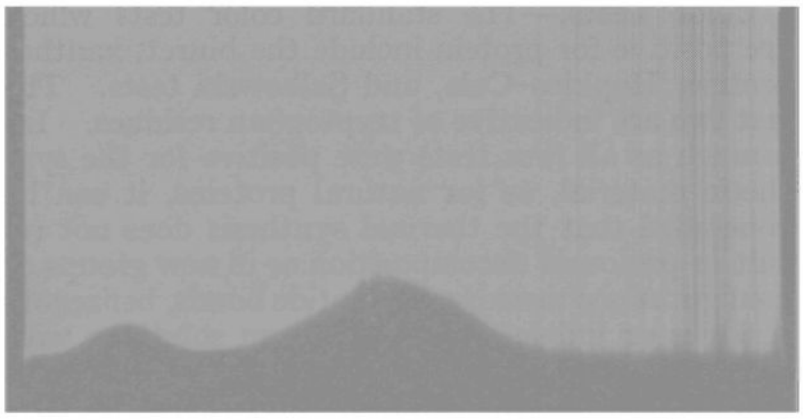


Fig. 1.—Electrophoretogram of proteinoid previously dialyzed once from a partly solid state.

each comparison, casein is more thoroughly digested than is the proteinoid; this result agrees with the fact that a larger proportion of protease-susceptible residues is found in casein. Pepsin is known to be characteristically active in the presence of acidic residues; this behavior may explain why proteinoid is more thoroughly attacked by pepsin than by the other enzyme. It is also true, however, that 1:1:1 proteinoid contains a somewhat higher proportion of the susceptible aromatic groups than does 2:2:1 proteinoid (Table VII). A controlled comparison is accordingly desirable. Rough calculation of the actual proportion of peptide bonds split gives a result of a total of about one bond per average proteinoid molecule in the presence of pepsin after seven days. Control experiments demonstrated no appreciable autodigestion of the proteinoid.

The most salient feature is that the total data indicate that the polymers obtained by pyrosynthesis are attacked by proteolytic enzymes[21] including papain and trypsin.[22] It may be inferred that the thermal synthesis yields at least some peptide bonds and that, furthermore, the heat does not result in additional structures which make all such bonds unavailable to enzyme action.

Nutritive Quality.—The growth of *L. arabinosus* as measured by lactic acid production is presented in Table X. The synthetic polymer fractions are

TABLE X

NUTRITIONAL UTILIZATION OF PROTEINOID BY *L. arabinosus* 17-5

Preparation	0.1 *N* lactic acid produced, ml.
Blank (no peptone, no proteinoid)	2.15 ± 0.15
Blank plus peptone	4.38 ± .03
Blank plus solid 2:2:1 proteinoid	2.38 ± .13
Blank plus solid 1:1:1 proteinoid	3.30 ± .10
Blank plus soluble fraction 1:1:1 proteinoid	3.45

used to a substantial degree by *L. arabinosus*. In other tests, as in this one, the soluble fraction obtained is slightly more active than the solid material from the dialysis. The 1:1:1 material is significantly more utilizable than the 2:2:1 preparations, a result which is in accord with the higher content of amino acids essential for the growth of

(21) S. W. Fox, K. Harada and J. Kendrick, *Science*, **129**, 1221 (1959).

(22) G. Krampitz, *Naturwissenschaften*, **46**, 558 (1959), and private communication.

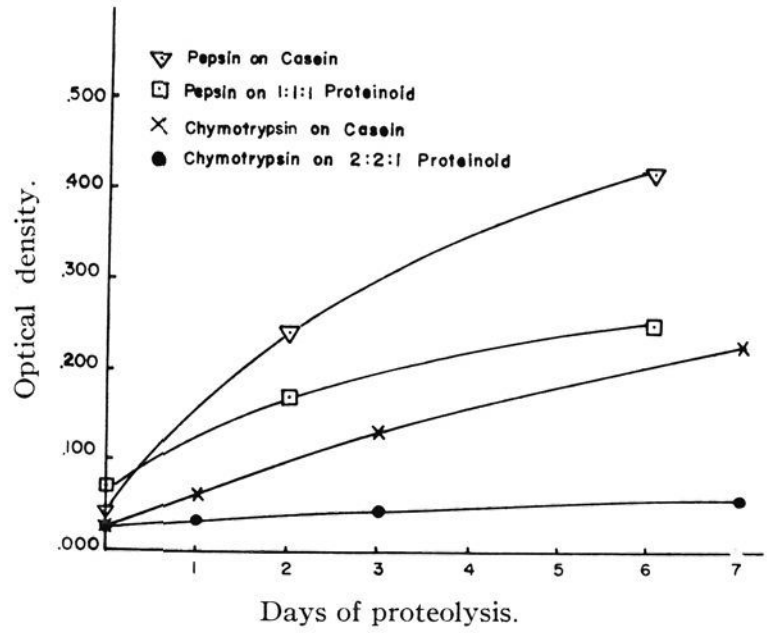


Fig. 2.—Proteolysis of proteinoid by pepsin and by chymotrypsin. Optical density was determined after the hydrolyzate was dinitrophenylated.

L. arabinosus[23] (Table X). The specific activity of the proteinoid is, however, less than that of peptone, and the most active proteinoid yet tested is calculated roughly to be significantly over one-half as nutritionally utilizable as peptone by *L. arabinosus*. This result comports with the fact that on hydrolysis of proteinoid partially racemized glutamic acid is recovered (in one experiment the DNP derivative of recovered glutamic acid had $[\alpha]_D$ of $-40.6°$ as compared to $-80.1°$ for pure DNP-L-glutamic acid).

The bacterial utilizability of the synthetic polymer has significance similar to that of the proteolyzability. Such nutritional quality can hardly be visualized in the light of present knowledge unless the material is attacked by bacterial proteases.

Configuration of Amino Acid Residues.—A most ostensible way in which the synthetic polymer differs from the usual protein is in configuration of residues, inasmuch as those obtained from the former product are partially racemized as judged by glutamic acid. Proteins usually studied, however, are mammalian and some evidence is at hand that bacterial proteins contain substantial proportions of D-amino acids. For example, D-glutamic acid has been found in lactobacilli.[24,25] Substantial proportions of D-amino acids have been shown to be present in *Bacillus brevis* by Konikova and Dobbert.[26] D-Alanine was found in *L. arabinosus*, *Leuconostoc mesenteroides* and *Streptococcus faecalis* by Holden and Snell.[27] Stevens, Halpern and Gigger[28] showed that the D-amino acids of *B. brevis*[26] were largely D-aspartic acid. Work in the same laboratory revealed significant proportions of D-amino acids in *Bacillus subtilis* and *Torula utilis*.[28] Camien learned that D-aspartic acid occurred in

(23) M. S. Dunn, S. Shankman, M. N. Camien and H. Block, *J. Biol. Chem.*, **168**, 1 (1947).

(24) M. N. Camien, A. J. Salle and M. S. Dunn, *Arch. Biochem.*, **8**, 67 (1945).

(25) M. S. Dunn, M. N. Camien, S. Shankman and H. Block, *J. Biol. Chem.*, **168**, 43 (1947).

(26) A. S. Konikova and N. N. Dobbert, *Biokhimiya*, **13**, 115 (1948).

(27) J. T. Holden and E. E. Snell, *J. Biol. Chem.*, **178**, 799 (1949).

(28) C. M. Stevens, P. E. Halpern and R. P. Gigger, *ibid.*, **190**, 705 (1951).

significant proportion in three lactobacilli but not in a fourth.[29] The last author, with Yuwiler and Dunn, showed that the D-glutamic acid of *L. arabinosus* is "tightly bound to insoluble material—probably proteins."[30]

Although the evidence is not complete enough to establish whether the D-amino acids in these cases occur combined in protein, considerable reason for believing not only this but that much of the D-amino acids are present in the bacterial cell walls, as contrasted to cells of higher forms, is at hand. This occurrence is consistent with the large integrated picture suggesting the occurrence of D-amino acids in both primitive proteins[31] and primitive membranes.[21] Although experiments under way offer some promise of yielding thermal polymers of L-amino acids, the possibility that truly primitive protein contained much D-amino acid would nevertheless have to be entertained if these experiments were successful.

Antigenicity.—The guinea pig and rabbit tests do not reveal any antigenicity. The reason for the lack of antigenicity may be low molecular weight relative to other proteins. Although the structural requirements for antigenicity are not definite, peptides which can be shown to have antigenic activity appear to have minimal molecular weights of about 15,000. Other possible explanations are too great a heterogeneity for one molecular species to be present in sufficient proportion, the possibility that a truly primitive protein[3] would not be foreign to any test organism, and the explanation that a thermal polymer lacks a sufficiently subtle structure.

Randomness or Non-randomness of Amino Acid Residues.—A complete answer to the question of whether the amino acid residues are distributed in a random or other arrangement may require a complete assignment of residues in one molecular species of proteinoid plus ancillary data. Some preliminary interpretation is possible, however. In general, the proportions of amino acids found in the product differ markedly from the proportions included in the reaction mixture (Tables VI and VII). Inasmuch as the ratios obtained in the products are consistent in repetitions of any given polymerization, the compositions are influenced by the identity of the reactants and cannot be said to be random on this score. When total compositions are compared with N-terminal compositions (Tables VII and IX) a different kind of marked disparity in comparable proportions is found. If the arrangement were random it would be expected that the proportions in the N-terminal position would be determined by and be the same as the proportions in the total composition. The N-terminal proportions of aspartic acid, glutamic acid or total BN amino acids, however, differ from either the proportions of these in the total composition or of these in the reaction mixtures. The possibility of selective hydrolysis during purification must be considered. Since the process of purification is the same in each case, selective hydrolysis of random arrangements should give the same N-terminal analyses regardless of mode of synthesis, *e.g.*, different temperatures. The data in Table IX, especially when compared with the relatively constant total composition reported in Table VII, show much variation with temperature of synthesis, however. Accordingly, the synthesis cannot be regarded as entirely random.

Quantitative Amino Acid Contents.—The possibility of obtaining panpolymers with proportions of amino acids exactly the same as those of some natural proteins will require many complete analyses before it can be fully evaluated. It is clear from the data in this paper, however, that proportions of amino acids in the polymers can be influenced by proportions in the reaction mixture (Tables VI, VII and IX and text) by temperature (Table VII) and by time of heating (Tables II and VI). Special attention may need to be directed to obtaining more than trace proportions of cystine, serine and threonine.

Branching in the Polymers.—Definitive data on branching in protein molecules and in the thermal polymers might provide an additional basis for comparison. The facts are unclear in each case. Desnuelle has stated, "It is not yet known, in any definite way, whether branching points exist in natural polypeptides or in the peptide chains of proteins."[32] On the other hand, Mechanic and Levy[33] more recently reported an ϵ-lysyl linkage in collagen. Comparison of numbers of N-termini and C-termini in virtually all proteins recorded[34] suggests infrequent or almost no branching. No data are at hand to indicate what proportions of peptide bonds involving aspartic acid and glutamic acid residues may be through the α-carboxyl and the ω-carboxyl groups in the thermal polymers. The aspartic acid residues are involved in imide linkages when first formed. The effect of various conditions in opening the imide linkages to yield free α- or β-carboxyls is being approached experimentally. Each imide linkage that is formed, however, precludes an equivalent amount of branching.

In summary, many kinds of test which are positive on natural proteins of low molecular weight are likewise in the main positive on the thermal panpolymers (proteinoids) of amino acids. Several other types of tests are incompletely evaluated, essentially because data on either the natural or synthetic material, or both, are inadequate. Even in these cases, the information available to data is consistent with the inference that the properties of the materials produced can fall within the range of known properties of proteins.

Acknowledgment.—Thanks are extended to Donna Keith and Jean Kendrick for technical assistance. The authors wish to thank many chemists for volunteered help and solicited advice. Particularly helpful have been Drs. Lyman Craig, C. A. Knight, Carl S. Vestling, Mr. Allen Vegotsky and Dr. Kenneth Woods.

(29) M. N. Camien, *J. Biol. Chem.*, **197**, 687 (1952).

(30) M. N. Camien, A. Yuwiler and M. S. Dunn, *Proc. Soc. Exptl. Biol. Med.*, **94**, 137 (1957).

(31) S. W. Fox, J. E. Johnson and A. Vegotsky, *Science*, **124**, 923 (1956).

(32) P. Desnuelle, in H. Neurath and K. Bailey, "The Proteins," Vol. I, Academic Press, Inc., New York, N. Y., 1953, p. 133.

(33) G. L. Mechanic and M. Levy, THIS JOURNAL, **81**, 1889 (1959).

(34) A. Vegotsky and S. W. Fox, unpublished.