pared from 6-trimethylsilyl-2-hexanone by treatment with methylmagnesium bromide in the usual manner. The authentic alcohol had b.p. 93° (8 mm.), $n^{20}D$ 1.4376, d^{20} 0.8290.

Infrared spectra of the authentic alcohol and the alcohol derived from the adduct were identical.

UNIVERSITY PARK, PENNA.

[CONTRIBUTION NO. 2148 FROM THE GATES AND CRELLIN LABORATORIES OF CHEMISTRY, CALIFORNIA INSTITUTE OF TECHNOLOGY]

The Chromatographic Separation and Identification of Some Peptides in Partial Hydrolysates of White Turkey Feather Calamus

BY W. A. SCHROEDER, LOIS M. KAY, NANCY MUNGER, NANCY MARTIN AND JOAN BALOG

RECEIVED DECEMBER 11, 1956

Chromatographic investigations of partial acidic hydrolysates of unoxidized and oxidized white turkey feather calamus have resulted in the isolation and identification of 59 peptides. The sequences which have been observed suggest that unusual sequences do not occur in feather keratin. The possible relation of some of the more prominent sequences to the spatial structure of the proteins is discussed. The isolation and identification of cysteic acid peptides by means of ion-exchange chromatography are described.

Many published investigations treat the feather as a more or less homogeneous unit despite its visually apparent and morphologically distinct parts and despite the fact that the parts of a feather are known to develop from specific portions of the papilla.[1] Recent analyses[2,3] have shown that the amino acid composition is not identical in all parts of a feather, so that a distinct and properly characterized part only should be utilized for chemical investigations. The relatively good orientation which is observed in certain feather parts by X-ray diffraction might lead one to conclude that they are composed of a reasonably homogeneous protein, but cell walls and cell debris from the original growing feather certainly are not absent. Purification of the protein without alteration unfortunately is impossible because of its insolubility, so one is forced to use the material as it is obtained. Woodin[4] has obtained soluble keratin from whole feathers in an apparently homogeneous state with a molecular weight of about 10,000. This material is not unaltered protein inasmuch as it has been isolated after either oxidation or reduction. In the present paper, we describe the isolation and identification of peptides from partial hydrolysates of unaltered protein and of oxidized protein from white turkey feather calamus. Although the rachis has usually been used when well-oriented feather protein is desired, we have chosen the calamus for four reasons: (1) rachis and calamus develop from the same portion of the papilla, (2) the amino acid compositions are essentially identical, (3) the X-ray diffraction patterns for most of the calamus and for the rachis are identical,[5] and (4) necessary amounts of calamus are easy to obtain whereas the separation of the rachis from the medulla is laborious.

Experimental

Source of Calamus and Preparation of Oxidized Calamus.—The sample of white turkey feather calamus was prepared in the way previously described.[2]

Oxidized calamus for the investigation of cysteic acid peptides was prepared in the following manner which combines features of the methods of Schram, Moore and Bigwood[6] and of Hirs.[7] Preformed performic acid was prepared by mixing 9 volumes of 90% formic acid and one volume of 30% hydrogen peroxide. After an hour at room temperature, the mixture was cooled to 0° in an ice-bath and 25 ml. was added to 500 mg. of calamus in the form of pieces about 1 mm. square and 0.1 mm. thick. The reaction mixture was maintained at 0° for 16 hr. during which time the pieces of calamus became swollen but did not dissolve to any obvious extent. The reaction mixture was transferred with 50 ml. of ice-cold water to 350 ml. of ice-cold water in a lyophilizing bulb. The solution was then frozen and the solvents were removed by lyophilization. The process was repeated after the addition of 80 ml. of water. The residue contained pieces in the form of original calamus as well as the usual voluminous residue which soluble substances form on lyophilizing and it was evident that considerable solution had occurred during some part of the procedure.

Partial Hydrolysis of Calamus and Oxidized Calamus.—A 300-mg. sample of calamus and 3 ml. of concd. hydrochloric acid (Baker's Analyzed Reagent Grade) were placed in a 5-ml. volumetric flask. The stem of the flask was attached parallel to a shaft which was inclined at 45° and rotated at 25 r.p.m. In this way, the reaction mixture was constantly agitated for 48 hr. at 37°. Within a few minutes after hydrolysis was begun, the pieces of calamus began to swell and continued to do so until the liquid was almost filled. Gradual dissolving took place and was complete in 3 to 3.5 hr. to give a light brown solution which darkened somewhat during the remainder of the hydrolysis. After about 24 hr., a very small amount of dark insoluble material had become attached to the flask at the surface of the liquid; it was somewhat greasy and had a density less than that of water. It may be composed of sterols or fatty acids.

A 240-mg. sample of oxidized calamus in 2.5 ml. of concd. hydrochloric acid was hydrolyzed at 37° for 48 hr. with constant agitation as above. Most of the material dissolved almost immediately and the remainder within an hour to give a light brown solution. The same type of insoluble residue appeared after about 24 hr. At the end of the partial hydrolysis, the reaction mixture was transferred with 100 ml. of water to a lyophilizing bulb, the solution was frozen, and

(1) F. R. Lillie, *Biol. Rev.*, **17**, 247 (1942).

(2) W. A. Schroeder and L. M. Kay, THIS JOURNAL, **77**, 3901 (1955). A definition of terms relating to feathers is presented in this reference.

(3) W. H. Ward, C. H. Binkley and N. S. Snell, *Textile Res. J.*, **25**, 314 (1955).

(4) A. M. Woodin, *Biochem. J.*, **57**, 99 (1954).

(5) S. Krimm, private communication. Krimm has observed with turkey feathers that the proximal tip of the calamus is unoriented, that good orientation is present but changes in direction within the first 6 mm. from the tip and then is unchanged in direction and character throughout the remainder of the calamus and rachis.

(6) E. Schram, S. Moore and E. J. Bigwood, *Biochem. J.*, **57**, 33 (1954).

(7) C. H. W. Hirs, *J. Biol. Chem.*, **219**, 611 (1956).

the solvent was removed. The residue was then taken up in 10 ml. of water which was again removed by lyophilizing.

Ion-exchange Chromatography of the Partial Hydrolysate of Calamus.—The partial hydrolysate of calamus was chromatographed on a column of Dowex 50-X4 that was 2.2 × 100 cm. in dimension. The resin (200–400 mesh, high porosity, Lot No. 3198-42) was purified in the manner that Moore and Stein[8] describe for Dowex 50-X8 and the column was packed as they suggest. The proportion of mesh sizes in the resin for this column was 4 parts of through-200-mesh resin and one part of 120–200 mesh resin. The detergent BRIJ35 and thiodiglycol were omitted from the buffers which otherwise were prepared as prescribed.[8]

The partial hydrolysate of calamus prepared as above was made up for chromatography by diluting with 42 ml. of *p*H 6.7 buffer after which the *p*H was brought to 2.1 with 3.8 ml. of 6 *N* sodium hydroxide. The 48 ml. of solution that was placed on the column contained the equivalent of 268 mg. of moisture- and ash-free calamus.

Development was begun with *p*H 3.42 buffer at a temperature of 55°. This temperature was maintained through the *p*H 4.25 buffer, but was reduced to 37° during the *p*H 6.7 and *p*H 8.3 buffers, and finally to room temperature when 0.2 *N* sodium hydroxide was used (see Fig. 1). Deaerated buffers were required at 55°. The flow rate of developer was 50 ml. per hr. and the fraction size was 5 ml. The ninhydrin procedure[9] was applied to 0.5-ml. portions of alternate fractions in order to determine the progress of the chromatogram.

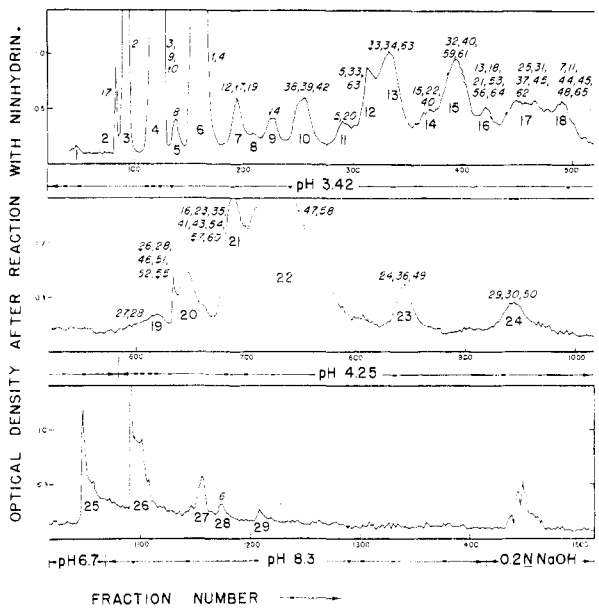

Fig. 1.—Chromatogram of a partial hydrolysate of unoxidized white turkey feather calamus on Dowex 50-X4. Zones are numbered consecutively. Italicized numerals identify compounds in each zone (see Table I).

Fractions containing a zone were pooled, neutralized to *p*H 6.5–7.5, evaporated at 35° to one-tenth their volume on a rotary evaporator,[10] finally evaporated to dryness in a vacuum desiccator over concd. sulfuric acid, and stored at −10° until investigated

Ion-exchange Chromatography of the Partial Hydrolysate of Oxidized Calamus.—The partial hydrolysate of oxidized calamus was chromatographed on a column of Dowex 1-X8 that was 1.6 × 100 cm. in dimension. The resin (200–400 mesh; Lot No. 3589-27; chloride form; medium porosity) was purified in the following way. A volume of about 300 ml. of resin was suspended in 750 ml. of water, the resin was allowed to settle for an hour, and the cloudy supernatant liquid was decanted; the process was repeated 4 times. The resin was then filtered, suspended in 200 ml. of 2 *N* hydrochloric acid, filtered and washed on the filter with 800 ml. of 2 *N* hydrochloric acid; this treatment was repeated once and the filter cake was finally washed with 500 ml. of water. The Dowex 1 was now passed wet through a 120-mesh sieve.[11] Fine particles were further removed at this point by seven suspensions and settlings in 500 ml. of water. The resin was converted to the chloroacetate form by suspending it in 600 ml. of *N* sodium chloroacetate, filtering, repeating this procedure twice and then washing on the filter with 1200 ml. of *N* sodium chloroacetate. This process was repeated with *N* chloroacetic acid and with 0.1 *N* chloroacetic acid and the resin was finally suspended in 600 ml. of the latter. The column was packed in the usual way and conditioned with 675 ml. of 0.1 *N* chloroacetic acid.

For chromatography, the residue that remained after the lyophilization of the partial hydrolysate was dissolved in 7 ml. of 0.01 *N* chloroacetic acid and placed on the column. Inasmuch as the moisture and ash content of the oxidized calamus was not determined, the amount of material that was placed on the column can only be estimated to be equivalent to 200–220 mg. of moisture- and ash-free protein.

The chromatogram was developed throughout with 0.1 *N* chloroacetic acid at room temperature (approximately 22°). The fraction size was 3 ml. and the flow rate of developer was 12 ml. per hr. to fraction no. 150, 18 ml. per hr. from there to fraction no. 400, and 24 ml. per hr. thereafter. The progress of the chromatogram was assessed by applying the ninhydrin procedure[9] to 0.5-ml. portions of alternate fractions.

Pooled fractions that contained a given zone were frozen and stored at −10°. For further investigation, the frozen solution was thawed and the chloroacetic acid was extracted with 2 × 25 ml. and 7 × 10 ml. of ether. Finally, aliquot portions of the extracted solution were removed and individually taken to dryness at 40° on a rotary evaporator.[10]

Further Separation and Identification of the Peptides.—The simpler mixtures of peptides which were obtained from the complex partial hydrolysate by means of the ion-exchange chromatography on Dowex 50-X4 were separated and identified by methods which did not differ appreciably from those previously utilized[12]; that is, the peptides in the pooled fractions were dinitrophenylated, separated by chromatography on silicic acid–Celite, estimated spectrophotometrically, and identified after complete hydrolysis. For peptides longer than dipeptides, the complete sequence was usually not determined; where the complete sequence is known it usually was obvious from the composition (thus, thr-val-val) or the peptide was difficult to hydrolyse and smaller peptides were isolated and identified.

Dinitrophenylation according to Levy and Chung[13] in aqueous solution is more satisfactory than in the aqueous alcoholic medium which was used heretofore.[12] Specifically, the procedure now used is this. The residue of buffer salts and peptides from the evaporation of the pooled fractions was dissolved in 10 to 15 ml. of water and the *p*H was adjusted to 8.5 to 9.5. A 20-mg. portion of sodium carbonate and 0.05 ml. of 2,4-dinitrofluorobenzene (DNFB) were then added and the reaction mixture was vigorously stirred at 40° for 3 hr. Extraction then followed the scheme previously used.[14] This method has the advantage that the buffer salts remain in solution but, what is more important, extraction of soln. 7[14] may not be necessary. Because of the absence of alcohol, the miscibility of phases during the extraction of excess DNFB from the basic reaction mixture (soln. 1[14]) is decreased and DNP-compounds are less likely to be carried over into soln. 7. If a yellow color is present in soln. 7 but disappears on acidification, it is caused by dinitrophenol and the solution may be discarded. Comparison of the two methods with test substances showed that the dinitrophenylation was equally quantitative in both.

Attempts to Separate DNP-Cysteic Acid Peptides Chromatographically.—Although definite separation of cysteic acid peptides was achieved by the use of Dowex 1, it is

(8) S. Moore and W. H. Stein, *J. Biol. Chem.*, **192**, 663 (1951).

(9) S. Moore and W. H. Stein, *ibid.*, **176**, 367 (1948).

(10) L. C. Craig, J. D. Gregory and W. Hausmann, *Anal. Chem.*, **22**, 1462 (1950).

(11) Only about one-fourth of this lot of Dowex 1 would pass a 200-mesh sieve.

(12) W. A. Schroeder, L. M. Kay, J. LeGette, L. Honnen and F. C. Green, THIS JOURNAL, **76**, 3556 (1954).

(13) A. L. Levy and D. Chung, *ibid.*, **77**, 2898 (1955).

(14) See Fig. 1 of ref. 12.

hardly to be expected that the separation would be complete. Because DNP-peptides of other kinds can be so readily separated by chromatography on silicic acid–Celite, experiments with DNP-cysteic acid as test substance were begun in an attempt to find chromatographic conditions which might apply to DNP-cysteic acid peptides.

DNP-Cysteic acid itself can be prepared by dinitrophenylation in aqueous alcohol or in purely aqueous solution as above and can be extracted from the acidified reaction mixture with *n*-butanol. The product, however, is oily and less homogeneous than when the following procedure is used. To a solution of 10 mg. of cysteic acid (Nutritional Biochemicals Corpn.) in 5 ml. of water was added 0.25 ml. of 25% aqueous reagent trimethylamine (Matheson, Coleman and Bell). The *p*H was about 10. Then 0.05 ml. of DNFB (Eastman) was added and the reaction mixture was stirred vigorously at 40° for 3 hr. The reaction mixture was extracted with 4 × 25 ml. of ether to remove excess DNFB, then evaporated to dryness at 40° on a rotary evaporator, and finally evaporated twice to dryness after the addition of 5 ml. of water.

Unsatisfactory chromatograms resulted when DNP-cysteic acid was chromatographed on silicic acid–Celite by methods similar to those of Green and Kay.[15] DNP-Cysteic acid is very strongly adsorbed so that strong developers such as 7AA-35A-B and 8AA-40A-B are necessary. However, two zones are present on the chromatogram; the more strongly adsorbed when rechromatographed again produces two zones whereas the less strongly adsorbed is unchanged in its behavior on rechromatography. Apparently, the adsorbent forces a conversion to the less strongly adsorbed form.

Inasmuch as DNP-cysteic acid is highly polar and easily soluble in water, chromatography in systems in which the developer contained water were studied. These efforts for the most part have been fruitless. The separation of DNP-cysteic acid from DNP-cysteic acid peptides could be achieved (a separation was actually already accomplished with the undinitrophenylated compounds on Dowex 1) but the DNP-peptides did not separate from each other. The results of the experiments will be described briefly. Chromatography on Dowex 2-X10 in a manner similar to that for cysteic acid itself[6] apparently resulted in much destruction of the very strongly adsorbed DNP-cysteic acid. On Dowex 50,[8] the DNP-cysteic acid moved with the front of *p*H 3.42 buffer and although the DNP-peptides were retarded there was little evidence of separation. DNP-Cysteic acid is rather strongly adsorbed on starch when the developer is *n*-butanol–benzyl alcohol–water.[16] DNP-Cysteic acid peptides are less strongly adsorbed and some separation occurred but the procedure was not pursued further because of the apparent deleterious effect of butanol which has been mentioned above.

Many experiments were made to try to chromatograph DNP-cysteic acid on cellulose columns with developers patterned after those which are used in the paper chromatography of DNP-compounds. For this purpose, the following general procedure was used. Ten grams of Solka-Floc BW-200 (Brown Company, New York) was suspended in 50 ml. of developer, stirred vigorously for 30 min., filtered, and the process repeated. Finally, the Solka-Floc was suspended in 50 ml. of developer, the column was poured in the manner of starch columns[16] and then 25 ml. of developer was passed through. In a tube of 10 mm. inner diameter, most columns were about 30 cm. in length. The flow rate under identical conditions was dependent on the developer and was adjusted by positive or negative pressure to 5 to 10 ml. per hr. The following are typical of the developers which were studied: (1) 1:2:1 acetic acid–*n*-butanol–water,[17] (2) 1.5 *M* phosphate buffer at *p*H 6.0,[18] (3) *M* citrate buffer at *p*H 6.2,[19] (4) *t*-amyl alcohol saturated with 3% aqueous ammonia,[20] and (5) *t*-amyl alcohol saturated with 0.1 *M* phthalate buffer at *p*H 6.0.[21] With these developers in the order given, DNP-cysteic acid moves through the column with increasing slowness; thus, about 30 ml. of citrate buffer moves DNP-cysteic acid through a column of the above size whereas about 300 ml. are required with *t*-amyl alcohol–phthalate buffer. These experiments showed that several of these systems would separate DNP-cysteic acid from its DNP-peptides. However, there was no evidence that mixtures of DNP-peptides would separate. Development with citrate buffer is generally suitable for the identification of DNP-cysteic acid. Our experience has been that DNP-cysteic acid and its DNP-peptides tend to decompose rather readily.

Other Methods for Identifying the Cysteic Acid Peptides.—The degree of heterogeneity of the residue from particular pooled fractions of the Dowex 1 chromatogram was assessed by hydrolyzing a portion of the residue completely, dinitrophenylating the products of hydrolysis, extracting other DNP-amino acids from DNP-cysteic acid with ethyl acetate, and identifying the DNP-amino acids. When this procedure was applied to zones 0–4, 0–5 and 0–7, it was apparentt hat zones 0–4 and 0–5 contained mixtures of several peptides whereas zone 0–7 contained mostly (ser, cya[22]). For further investigation, another portion of zone 0–7 was dinitrophenylated prior to hydrolysis. From this hydrolysate, DNP-serine equivalent to about 20% of the quantity of peptide was isolated.[23] Although this result could be interpreted to mean that a mixture of ser-cya and cya-ser was present, this possibility was excluded by dinitrophenylating the amino acids in the extracted hydrolysate of the peptide: no DNP-serine was detected and it was concluded that ser-cya is the main constituent. Inasmuch as similar difficulty might be anticipated in the hydrolysis of other zones after dinitrophenylation, the final investigation of the cysteic acid peptides was made by hydrolyzing the undinitrophenylated residues from the pooled fractions, and identifying the constituent amino acids and estimating the amounts by chromatography on 100-cm. columns of Dowex 50.[8] In this way, if the several peptides in the mixture were present in sufficiently disparate amounts, the composition could be deduced by comparing the molar amounts of the constituents.

Results

The separations which were achieved by chromatographing the partial hydrolysate of unoxidized calamus on Dowex 50-X4 are depicted in Fig. 1. Although more satisfactory separations were obtained in preliminary small-scale experiments especially in the region of zones 11 to 18, there is relatively minor overlapping of the peptides from zone to zone.

All zones have been investigated and some compound has been identified in each with the exception of zones 1, 8, 25, 26, 27 and 29. The qualitative and quantitative results are presented in Table I. Minor components, identification of which was not attempted, were also present in most of the zones.

The anomalous behavior of asp-gly was again observed in this chromatogram: it was present in the non-adjacent zones 2 and 7. A similar unexplained behavior was observed in partial hydrolysates of Tussah silk fibroin.[24]

(15) F. C. Green and L. M. Kay, *Anal. Chem.*, **24**, 726 (1952).

(16) W. H. Stein and S. Moore, *J. Biol. Chem.*, **176**, 337 (1948).

(17) D. H. Peterson and L. M. Reineke, *ibid.*, **181**, 95 (1949).

(18) A. L. Levy, *Nature*, **174**, 126 (1954).

(19) M. Rovery and C. Fabre, *Bull. soc. chim. biol.*, **35**, 541 (1953).

(20) H. Fraenkel-Conrat, J. I. Harris and A. L. Levy in "Methods of Biochemical Analysis," Vol. 2, edited by D. Glick, p. 375.

(21) S. Blackburn and A. L. Lowther, *Biochem. J.*, **48**, 126 (1951).

(22) The commonly accepted abbreviations for the amino acids use the first three letters of the name but because of obvious duplication with cystine and cysteine, cysteic acid is usually abbreviated cySO₃H. Further difficulty arises with glutamic acid and glutamine and with aspartic acid and asparagine. In the interest of simplicity, we would like to suggest the following usages: cysteic acid, cya; cysteine, cye; glutamine, glm; and asparagine, asg.

(23) Our experience has been that greater destruction of DNP-serine compared to other DNP-amino acids is common during the hydrolysis of DNP-peptides although the destruction usually is not as great as observed here. A. L. Levy and C. H. Li (*J. Biol. Chem.*, **213**, 487 (1955)) recovered only 10 to 20% of the expected N-terminal DNP-serine from ACTH.

(24) L. M. Kay, W. A. Schroeder, N. Munger and N. Burt, THIS JOURNAL, **78**, 2430 (1956).

TABLE I

PEPTIDES AND FREE AMINO ACIDS ISOLATED FROM A PARTIAL ACIDIC HYDROLYSATE OF UNOXIDIZED WHITE TURKEY FEATHER CALAMUS

Designation in Fig. 1	Amino acid or peptide	Isolated from zone no.	Amt. isold. μmoles/g.[a]	Designation in Fig. 1	Amino acid or peptide	Isolated from zone no.	Amt. isold. μmoles/g.[a]
1	alanine	6	77	*35*	ser-(i)leu	21	26.0
2	aspartic acid	3	118	*36*	ser-phe	23	4.3
3	glutamic acid	4	25	*37*	ser-pro	17	2.4
4	glycine	6	346	*38*	ser-ser	10	36.2
5	(iso)leucine[c]	11, *12*[b]	49	*39*	ser-thr[d]	10	1.2
6	lysine	28	1.2	*40*	thr-ala	14, *15*	48.9
7	phenylalanine	18	34	*41*	thr-(i)leu	21	10.6
8	proline	5	20	*42*	thr-ser[d]	10	2.0
9	serine	4	350	*43*	tyr-gly[d]	21	3.5
10	threonine	4	18	*44*	val-gly	18	49.3
11	tyrosine	18	11				
12	valine	7	8	*45*	ala-(gly,val)	*17*, 18	38.2
				46	gly-(gly,(i)leu)	20	7.5
13	ala-ala	16	5.1	*47*	gly-((i)leu,pro)	22	7.8
14	ala-asp	9	14.2	*48*	(i)leu-(ala,ser)[d]	18	3.1
15	ala-gly	14	7.5	*49*	phe-(glu,pro)	23	2.0
16	ala-(i)leu[e]	21	7.1	*50*	pro-(i)leu-(i)leu	24	18.9
17	asp-gly	2, *7*	7.1	*51*	ser-(ala,(i)leu)	20	11.0
18	asp-(i)leu	16	1.2	*52*	ser-(gly,(i)leu)	20	5.5
19	glu-asp	7	9.9	*53*	thr-(ala,val)	16	9.5
20	glu-gly	11	3.9	*54*	thr-pro-(i)leu[e]	21	14.2
21	glu-pro	16	3.9	*55*	thr-((i)leu,pro)[e]	20	6.3
22	gly-gly	14	17.3	*56*	thr-val-val	16	8.7
23	gly-(i)leu	21	30.3	*57*	val-ala-(i)leu	21	13.8
24	gly-phe	23	35.4	*58*	val-((i)leu,val)	22	5.5
25	gly-pro	17	41.0				
26	(i)leu-ala	20	4.7	*59*	ser-(ala,gly,val)[d]	15	5.1
27	(i)leu-glu	19	3.1	*60*	ser-(glu,phe,pro)	21	11.0
28	(i)leu-gly	*19*, 20	29.1	*61*	ser-(gly,(i)leu,pro)[d]	15	4.3
29	(i)leu-(i)leu	24	2.4	*62*	thr-(ala,(i)leu,pro)	17	13.8
30	phe-gly	24	34.6	*63*	thr-(ala,gly,val)	12, *13*	22.4
31	pro-glu	17	3.2	*64*	thr-(gly,(i)leu,pro)	16	10.2
32	ser-ala	15	50.8				
33	ser-glu	*12*, 13	15.4	*65*	ser-(asp,glu,phe,pro)	18	3.1
34	ser-gly	13	55.0				

[a] Recalculated to one g. of moisture- and ash-free calamus. The quantities of the DNP-peptides were determined spectrophotometrically. [b] The number of the zone which contained the major portion is italicized. [c] Inasmuch as the procedures did not distinguish between leucine and isoleucine, the designations "(iso)leucine" and "(i)leu" indicate that either or both may be present. [d] The presence of these peptides though probable is not definitely established. [e] Sequence may be identical; they were isolated from adjacent zones and in only one was the actual sequence identified.

Although zone 22 is the largest zone on the chromatogram, the quantity of peptides is insignificant and the main component is ammonia. Ammonia was not identified as such in this zone because much was inadvertently lost during the isolative procedures before it was realized that a volatile component was present. However, under these same conditions of chromatography, ammonia in a control chromatogram emerged in the same region. The exact point of emergence of this zone is somewhat dependent upon uncontrolled chromatographic conditions: in some chromatograms it has been well separated from a zone equivalent to zone 21 and has partially overlapped zone 23. It is rather surprising that ammonia emerges from the column so early in the chromatogram because Moore and Stein[16] report that it emerges near arginine with *p*H 11.0 buffer after several intermediate increases in the *p*H of the buffer. In the present experiments, ammonia emerged soon after the change to *p*H 4.25 buffer and the conditions differed from those of Moore and Stein essentially only in the use of about twice the volume of *p*H 3.42 buffer and a temperature of 55° instead of 37°.

Figure 2 shows the chromatogram which resulted when the partial hydrolysate of oxidized calamus was chromatographed on Dowex 1-X8. Most of the material in the hydrolysate emerges almost immediately from the column as would be expected; three poorly separated peaks can be detected in this region of the chromatogram. This chromatogram was continued only until shortly after the appearance of cysteic acid (zone 0–12). Preliminary chromatograms on 0.9 × 100-cm. columns failed to reveal any substances which were more strongly adsorbed than cysteic acid. Thus, in these chromatograms, development with 0.1 *N* chloroacetic acid was continued to about fraction no. 800 and was then changed to a linear gradient between 0.1 *N* and *N* chloroacetic acid. The apparatus described by Bock and Ling[25] in their Fig. 6 was used. By the use of 300 ml. in each reser-

(25) R. M. Bock and N. S. Ling, *Anal. Chem.*, **26**, 1543 (1954).

voir of equal cross-sectional area, the influent concentration would be N chloroacetic acid when all had been used. The equilibrium is established somewhat less rapidly on Dowex 1 than on Dowex 50. Actually, the gradient was not apparent in the effluent until 200 fractions more had been collected. The concentration of chloroacetic acid in the effluent was about 0.5 N and in the mixer was about 0.75 N when the chromatogram was stopped 450 fractions after the change to gradient elution.

The peptides of cysteic acid which were tentatively identified are listed in Table II. As an example of the interpretation required to come to the conclusions given there, consider zone 0–5. Ion-exchange chromatography showed the presence of 6.5 μmoles of cysteic acid, 3.6 μmoles of serine, 4.0 μmoles of proline, 1.5 μmoles each of aspartic acid, glutamic acid and glycine, 1.2 μmoles of alanine, 0.8 μmole of (iso)leucine, and a trace of valine. Serine and proline are present in essentially equivalent amount (there is some destruction of serine during hydrolysis) and in more than twice the amount of any other amino acid except cysteic acid. If we assume that zone 0–5 must contain cysteic acid peptides, then it is reasonable to conclude that (cya, pro, ser) is present. If this is accepted, then about 2.5 μmoles of cysteic acid peptide are not accounted for and we have about 1.5 μmoles each of aspartic acid, glutamic acid, glycine and alanine unaccounted for. Hence, two other cysteic acid tripeptides may be present but it is futile to suggest their composition because so many possibilities arise. In this way, the peptides listed in Table II have been identified with reasonable certainty.

Table III presents an accounting of the percentages of the amino acids in free and peptide form in these hydrolysates. If we calculate the amount of material in each of the zones on the basis of the ninhydrin color and compare this with the amount actually isolated and identified, a wide variety of values results. In some zones, nothing was identified whereas in others the amount approaches or slightly exceeds 100%. In most of the major zones in Fig. 1, the amount accounted for on this basis was usually between 50 and 75%. These data must be treated with considerable reserve because the color yields of the peptides which constitute the zones are unknown. Approximately 10% of the protein has been accounted for in the form of free amino acids and about 20% in the form of peptides.

TABLE II

TENTATIVE IDENTIFICATION OF CYSTEIC ACID PEPTIDES FROM A PARTIAL ACIDIC HYDROLYSATE OF OXIDIZED WHITE TURKEY FEATHER CALAMUS

Zone no. in Fig. 2	Compound	Approx. amt., μmoles/g.
0–1, 0–2, 0–3	Contain cysteic acid peptides in mixture with other peptides of hydrolysate	
0–4	Much overlap from zone 0–3. Possibly contains (phe, cya)	9
0–5	(cya, pro, ser)	18
0–6	(cya, glu, pro)	19
0–7	(cya, ser) probably ser, cya	50
0–8	(cya, glu, pro)	12
0–9	(cya, glu)	10
0–10	(cya, glu)	7
0–11	(cya, ser, glu?)[a]	..
0–12	cya	80

[a] Cysteic acid and serine are definitely present but the identification of glutamic acid is uncertain. A zone was present which emerged in the region where glutamic acid is expected but more rapidly than observed in other chromatograms.

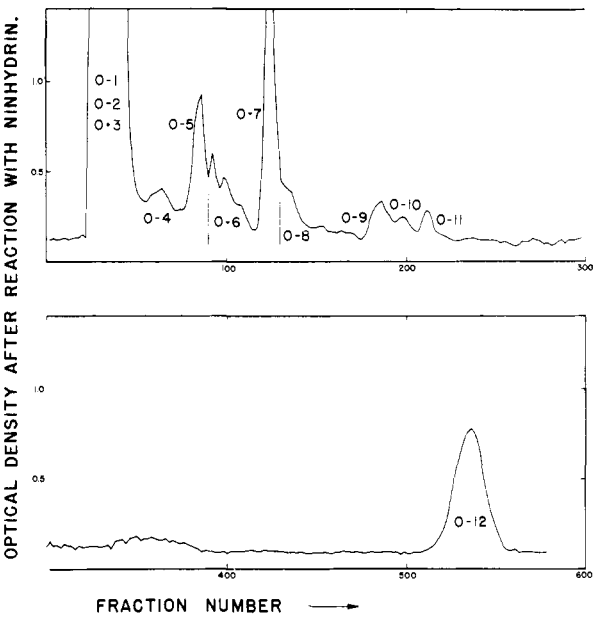

Fig. 2.—Chromatogram of a partial hydrolysate of oxidized white turkey feather calamus on Dowex 1-X8.

TABLE III

PERCENTAGE OF AMINO ACIDS IN WHITE TURKEY FEATHER CALAMUS ACCOUNTED FOR BY PEPTIDES AND AMINO ACIDS OF TABLES I AND II[a]

Amino acid	% accounted for by Free amino acids	Peptides	Total % accounted for
Ala	10	33	43
Asp	22	7	29
Glu	4	9	13
Gly	27	34	61
(I)leu	5	26	31
Lys	2	0	2
Phe	10	26	36
Pro	2	15	17
Ser	24	19	43
Thr	5	37	42
Tyr	5	2	7
Val	1	23	24
Cya	12	44[b]	56
Arg			0
Cys			0
Hist			0
Met			0

[a] Of the data from Table II, only cysteic acid is included here. If other amino acids from the cysteic acid peptides were added, possible double accounting might result because two hydrolysates are involved. [b] This amount was detected in peptide form but the peptides themselves were not all identified.

Discussion

In proteins such as *Bombyx* silk fibroin which are largely composed of a few amino acids, one expects to find and does find a rather restricted variety of peptides in a partial hydrolysate. On the other hand, the feather proteins contain none of the amino acids in unusual amount and peptides of all kinds are to be anticipated. Tables I and II show that many kinds may be isolated. Even if many more peptides had been isolated any attempt to define exactly the amino acid sequence in feather proteins would be unlikely to succeed. In the fibrous proteins, such as these feather proteins, one does not have a well-defined material of known molecular weight as in the case of insulin and corticotropin. One cannot know, therefore, the number of residues of any given amino acid in a molecule and one cannot decide when all sequences containing this amino acid have been identified. With fibrous proteins of unique amino acid composition such as the silk fibroins much information has been gained but, with the keratins, general rather than specific information about the sequences and structure is about all that one can expect to obtain.

In a preliminary experiment with oxidized calamus, a 48-hr. hydrolysate was chromatographed on a smaller scale on Dowex 50 under the conditions of Fig. 1 and the results were compared with those obtained from a 48-hr. hydrolysate of unoxidized calamus. As would be anticipated, cysteic acid and its peptides were essentially unretarded under these conditions and emerged almost with the front of the developer. Throughout the remainder of the chromatograms, the shapes and sizes of the peaks were almost identical and one may conclude that the presence of cysteic acid rather than cystine in the protein has had little influence on the course of the acidic hydrolysis. This type of comparison, admittedly, is not very exact because the components of the individual zones may not be identical. On the other hand, when hydrolysates of separate individual feather parts were compared, there were slight differences in the chromatograms which could be correlated with differences in the amino acid composition. Thus, 48-hr. hydrolysates of turkey feather barbs, calamus, medulla and rachis were made as described in the experimental part of this paper and were chromatographed under the conditions of Fig. 1. The barbs, in contrast to the other parts, required about 12 hr. instead of about 3 hr. for complete solution. The calamus and rachis are essentially identical in amino acid composition[2] and X-ray pattern[5] and the chromatograms of their partial hydrolysates were not significantly different. On the other hand, the chromatogram of the barb hydrolysate showed much smaller amounts of the zones equivalent to zones 9, 15 and 21 of Fig. 1 and two zones were present in the position of zone 23; the medulla hydrolysate showed these same differences to a lesser extent. It may be that the slower rate of solution of the barbs is responsible for these differences. However, zones 9 and 15 contain alanyl peptides and the alanine content of barbs is much lower than that of calamus while that of medulla is intermediate. Likewise, zone 21 contains peptides of leucine and/or isoleucine and the feather parts differ in their content of these amino acids. The possible relation is less obvious in this case because leucine and isoleucine were not distinguished from each other in the identification of the peptides.

The 32 dipeptides which have been isolated from the hydrolysate of unoxidized calamus are an appreciable fraction of the total number of dipeptides which are possible from combinations of the 11 amino acids in peptide form. The quantities are very dissimilar and this fact is probably partly a reflection of the ease or difficulty of hydrolysis of the particular bonds which must be hydrolyzed to give the peptide. It is unprofitable to speculate upon the sequences in the tri-, tetra- and pentapeptides as the following examples will show. The sequence in the peptide, thr-pro-(i)leu, was definitely established but the isolation of thr-(i)leu and not thr-pro would have led to the suggestion that the more probable sequence was thr-(i)leu-pro. Likewise, the sequence in gly-(gly,(i)leu) might equally well be either gly-(i)leu-gly or gly-gly-(i)leu on the basis of the amounts of gly-(i)leu, (i)leu-gly and gly-gly which were isolated.

The chromatographic separation of the cysteic acid peptides is not entirely satisfactory because zones 0–1 to 0–3 which contain the main bulk of the other peptides also contain cysteic acid peptides in about half the amount to be found in zones 0–4 to 0–11. Probably the chromatogram should be started with a developer such as 0.5 *M* acetic acid[26] in order to separate the neutral and weakly acidic compounds and then changed to 0.1 *N* chloroacetic acid in order to remove the more strongly acidic ones. On Dowex 1, the cysteic acid peptides are all appreciably less strongly adsorbed than cysteic acid itself. In contrast, on Dowex 50 peptides generally tend to emerge after any of the constituent amino acids.

Six cysteic acid peptides have been identified with some certainty. In zones 0–6 and 0–8, the main peptide has the composition (cya, glu, pro) and in zones 0–9 and 0–10, the composition is (cya, glu). Inasmuch as the separation of peptides of identical composition but different sequence is frequently observed on Dowex 50 chromatograms, it is safe to assume that in these instances also the peptides differ in sequence. It is interesting to observe that in these peptides cysteic acid and hence cystine has glutamic acid, serine and proline as close neighbors. Likewise, in the A chain of insulin,[27] glutamic acid and serine are adjacent to cystine. In feather keratin, proline also seems to be closely associated with cystine; possibly the close proximity of a cross-link between chains and of an amino acid of restricted structure such as a proline plays an important role in the structure.

When a protein contains as much proline as does feather keratin in which about one residue in ten is proline, it seems probable that this amino acid, by virtue of its definite spatial state, must play an important role in determining the spatial arrange-

(26) C. H. W. Hirs, S. Moore and W. H. Stein, THIS JOURNAL, **76**, 6063 (1954).

(27) F. Sanger and E. O. P. Thompson, *Biochem. J.*, **53**, 366 (1953).

ment of the amino acids in the polypeptide chains. Because the ring system imposes such restrictions on proline, this amino acid will have a profound effect on the direction of the polypeptide chain whenever a prolyl residue is present in the sequence. Indeed, it is unlikely that the polypeptide chains in feather keratin would assume any of the pleated sheet or simple helical configurations[28,29] which have been suggested thus far. Is there any evidence from the present data or from other data that the amino acids adjacent to proline are restricted in any way? In the present investigation, 15 prolyl peptides have been identified. Of those peptides of determined sequence, gly-pro, and thr-pro-(i)leu are the most predominant. One may conclude tentatively that, in a sequence such as -X-pro-, X tends to be an amino acid with a relatively short side chain, for example, glycine, serine and threonine and that amino acids with longer side chains such as glutamic acid are less likely to be found in this sequence. This suggestion receives some support from the literature. Thus, in gelatin[12] and collagen,[30] -gly-pro- is an important sequence and -ala-pro- and -ser-pro- have been found. Likewise, in a melanophore-stimulating peptide, the sequences -gly-pro- and -ser-pro-pro- are present.[31,32] By comparison, a much different situation obtains in the corticotropins[33,34] which contain the sequences -lys-pro-val-, -arg-pro-val-, -tyr-pro-asp-, and -phe-pro-leu-. If, indeed, the proline residues are important in determining the spatial configuration of the polypeptide chains, it may be suggested that the adjacent residues in conjunction with the proline residues assist in maintaining certain specific arrangements which, in turn, are responsible for some of the characteristic properties of the proteins themselves.

Acknowledgments.—This investigation was supported in part by a contract between the Quartermaster Corps., U. S. Army, and the California Institute of Technology.

(28) L. Pauling, R. B. Corey and H. R. Branson, *Proc. Nat. Acad Sci.*, **37**, 205 (1951).

(29) L. Pauling and R. B. Corey, *ibid.*, **37**, 251 (1951).

(30) T. D. Kroner, W. Tabroff and J. J. McGarr, THIS JOURNAL, **77**, 3356 (1955).

(31) J. I. Harris and P. Roos, *Nature*, **178**, 90 (1956).

(32) I. I. Geschwind, C. H. Li and L. Barnafi, THIS JOURNAL, **78**, 4494 (1956).

(33) P. Bell, *ibid.*, **76**, 5565 (1954).

(34) W. F, White and W. A. Landmann, *ibid.*, **77**, 1711 (1955).

PASADENA, CALIFORNIA

[CONTRIBUTION FROM THE LABORATORIES OF THE SLOAN-KETTERING DIVISION OF CORNELL UNIVERSITY MEDICAL COLLEGE]

Pyrimidine Nucleosides. II. The Synthesis of 1-β-D-Arabinofuranosylthymine ("Spongothymidine")[1]

BY JACK J. FOX, NAISHUN YUNG AND AARON BENDICH

RECEIVED DECEMBER 19, 1956

1-D-Ribofuranosylthymine (I) was converted by a series of reactions to 1-β-D-arabinofuranosylthymine. Since this epimerization can occur only with nucleosides containing the β-D-glycosyl structure, the configuration at the glycosyl center of I and of other thymine nucleosides was firmly established as beta. 1-β-D-Arabinofuranosylthymine was shown to be identical with the naturally occurring nucleoside, "spongothymidine."

In Part I of this series[2] a method was described for the facile synthesis of thymine nucleosides by the condensation of poly-O-acylglycosyl halides with dithyminylmercury followed by removal of the protecting acetyl or benzoyl groups. In this manner 1-D-ribofuranosylthymine (I) and 1-D-xylofuranosylthymine (II) were prepared. The former was identical with a product prepared enzymically by Lampen[3] and also was shown to be similar to a compound isolated recently by Fink and co-workers[4] from rat liver slices incubated with radiothymine.

Spectral and metaperiodate-oxidation studies[2] with I and II showed that the sugar–pyrimidine linkage was at position 1 of thymine and, further, that the lactol ring structures were of the furanose type. It was demonstrated that I, II and 1-D-glucopyranosylthymine gave the same dialdehyde upon oxidation with metaperiodate. On the basis of the reasonable assumption that 1-D-glucopyranosylthymine (III) is a β-nucleoside,[5] it was tentatively concluded that I and II are nucleosides of the β-configuration.

An unequivocal determination of the structure at the glycosyl center of I would serve not only for the assignment of configuration to the biochemically produced materials of Lampen[3] and Fink,[4] but it would also help to provide a basis for future generalizations about the structure of nucleosides obtained by this new synthetic procedure with other pyrimidines. In this paper, a proof is given for the structure of 1-D-ribofuranosylthymine. As a corollary to this proof, a first synthesis and unequivocal proof of structure of "spongothymidine" are presented. A preliminary report on this study has appeared.[6]

A rigorous method for the determination of the

(1) This investigation was supported in part by funds from the National Cancer Institute, National Institutes of Health, Public Health Service (Grant No. C-2329 and C-471), and from the Ann Dickler League.

(2) J. J. Fox, N. Yung, J. Davoll and G. B. Brown, THIS JOURNAL, **78**, 2117 (1956).

(3) J. O. Lampen, in W. D. McElroy and B. Glass, "Phosphorus Metabolism," Vol. II, The Johns Hopkins Press, Baltimore, Md., 1952, p. 368.

(4) K. Fink, R. E. Cline, R. B. Henderson and R. M. Fink, *J. Biol. Chem.*, **221**, 425 (1956).

(5) See discussion in footnote 12 of Part 1 of this series.

(6) J. J. Fox and N. Yung, *Federation Proc.*, **15**, 254 (1956).