739

# The Terminal Amino Groups of α- and β-Caseins[2]

By Edward F. Mellon, Alfred H. Korn and Sam R. Hoover

Dinitrofluorobenzene was used to label the amino groups of the α- and β-fractions of casein. The hydrolyzates of these protein derivatives were assayed for the various dinitrophenylamino acids which were shown to be present by chromatographic methods. The arginine derivative and both mono- and diderivatives of lysine were found in each protein. However, the proportions were different in each protein.

The separation of casein by isoelectric precipitation into two distinct fractions, each completely free from the other, was accomplished by Warner.[3] The same fractions were also obtained by Hipp, *et al.*,[4] by precipitation from aqueous alcohol and aqueous urea. The reproducibility of the separation and of the analytical data of the fractions obtained from various preparations indicated that these materials might be sufficiently homogeneous for use in a study of the end groups of casein.

The dinitrofluorobenzene reagent of Sanger[5] was used. However, satisfactory separations of the dinitrophenylamino acids were not obtained by his chromatographic method and a system of paper chromatograms was used for the separation of the ether-extractable dinitrophenylamino acids. Different solvent pairs than those employed by Sanger were necessary to obtain separations on paper. Although silica gel columns were used for the separation of the non-ether-extractable dinitrophenylamino acids, a variation of the procedure of Bailey[6] which uses formaldehyde-treated silica gel gave a much better separation of the ε-lysine and arginine derivatives than the procedure described by Sanger.

Partial decomposition of some of the dinitrophenylamino acids during the hydrolysis of the protein derivatives makes it necessary to run parallel recovery experiments preferably on a synthetic mixture as closely duplicating the actual situation as possible. The correction factors in some instances are quite large and in a few cases (proline, methionine and cystine) practically complete destruction occurs. Small amounts of dinitroaniline and dinitrophenol were obtained in the experiments reported here but an extraction procedure developed recently by Isherwood and Cruickshank satisfactorily removes these materials from the extract.

## Experimental

**Preparation of Materials.**—The 2,4-dinitrophenyl derivatives of the amino acids for use as control substances

(1) One of the laboratories of the Bureau of Agricultural and Industrial Chemistry, Agricultural Research Administration, U. S. Dept. of Agriculture. Article not copyrighted.

(2) Presented at the 122d Meeting of the American Chemical Society, Atlantic City, N. J., September, 1952.

(3) R. C. Warner, This Journal, **66**, 1725 (1944).

(4) N. J. Hipp, M. L. Groves, J. H. Custer and T. L. McMeekin, *J. Dairy Sci.*, **35**, 272 (1952).

(5) F. Sanger, *Biochem. J.*, **39**, 507 (1945).

(6) K. Bailey, *ibid.*, **49**, 23 (1951).

were prepared from the pure amino acids and 2,4-dinitrochlorobenzene by the procedure of Abderhalden and Blumberg.[7] Recrystallization of the 2,4-dinitrochlorobenzene from ether appears to facilitate the production of crystalline derivatives. All the materials obtained were crystalline and possessed satisfactory analytical and melting point data.

A description of the purification and fractionation of the caseins is given by Hipp, *et al.*,[4] and their analysis has been given by Gordon, *et al.*[8] The 2,4-dinitrophenyl derivatives of these $\alpha$- and $\beta$-caseins were prepared using 2,4-dinitrofluorobenzene. The protein (2.0 g.) was swelled for one-half hour in a small amount of distilled water. The mixture was diluted to 50 ml. and 1 *N* sodium hydroxide was added very slowly with vigorous agitation. The *p*H was kept at 8.0 until the protein dissolved. The 2,4-dinitrofluorobenzene (2 ml.), dissolved in 5 ml. of ether, was added and additional alkali was stirred into the mixture to keep the *p*H at 8.0. The reaction mixture was stirred for one hour after the addition of the reagents had been completed. The mixture was extracted three times with 20-ml. portions of benzene. The emulsion which formed was broken by centrifugation and the benzene layer was separated and discarded. The water layer was adjusted to *p*H 3.8 with dilute hydrochloric acid. The precipitated protein derivative was separated and washed by decantation with water until the washings were free from chloride ion. The derivative was then washed with 50% alcohol, absolute alcohol and ether; then dried overnight at 50°.

**Liberation of Labeled Groups.**—The dinitrophenylcasein derivatives (0.2 g.) were hydrolyzed for 24 hours with 10 ml. of 20% hydrochloric acid. The hydrolyzates were extracted four times with 10-ml. portions of ether and the extracts were evaporated to dryness. This ether extract may contain dinitroaniline, dinitrophenol and any of the dinitrophenyl derivatives of the naturally occurring amino acids except the dinitrophenyl derivatives of histidine, arginine, cystine and both of the mono derivatives of lysine. These excepted derivatives remain in the water layer.

For the identification experiments the residual hydrochloric acid in the ether extract was removed by washing the residue through a water-based silica gel column with a wet ether solvent. The chromatogram spots appear sharper if the spots are also exposed to strong ammonia vapors after they have been applied to the paper.

For the quantitative determination of the $\alpha,\epsilon$-didinitrophenyllysine in the ether extract the interfering dinitroaniline and dinitrophenol were removed by the method of Isherwood and Cruickshank as described in a paper by Schwartz and Lea.[9] The solution obtained by this treatment contained only the $\alpha,\epsilon$-didinitrophenyllysine and, therefore, needed no further separation before the concentration of this material was determined spectrophotometrically at 350 m$\mu$.

**Paper Chromatography.**—The ether extractable dinitrophenylamino acids were identified by separation on ascending chromatograms (Whatman No. 1 paper 11 × 14 inches) run in 6 × 18 inch Pyrex cylinders. A central wick of paper was used to keep the atmosphere saturated with the aqueous layer. The paper to be chromatographed was equilibrated with the water layer overnight before introducing the developing solvent. The solvents used were: A, *n*-butanol saturated with water; B, the organic phase from the mixture—1 volume of *n*-butanol, 2 volumes of *n*-butyl acetate and 3 volumes of 1% concentrated ammonium hydroxide in water; C, the organic phase from the mixture of equal volumes of benzene and 1% glacial acetic acid in water. Solvent B was made fresh each evening before use. Solvents A and C can be stored for some time. With solvent C, the spots have a tendency to elongate but this does not seriously interfere with the identification of the few amino acids for which its use is needed. The $R_f$ values obtained on the control substances are shown in Table I.

**Column Chromatography.**—The aqueous layer from the ether extraction of the hydrolyzates of the dinitrophenyl substituted caseins was evaporated to dryness at 50°. The dry residue was extracted with small amounts of acidified methyl ethyl ketone (0.5 ml. of 1 *N* HCl/liter) until no more color was removed. The extracts were combined and diluted to 2.0 ml. A 0.1-ml. aliquot of this solution was then transferred to the top of the chromatographic column. This column was made by grinding 5 g. of anhydrous silica gel with 2 ml. of a 1:9 formaldehyde (U.S.P.)–water mixture and then suspending this material in a 66% methyl ethyl ketone–ether mixture which had been saturated with water. The suspension was poured into the tube (1 × 40 cm.) and the excess solvent drained off. After the gel had settled to constant volume the test sample was added and the chromatogram developed with a 66% methyl ethyl ketone–ether solvent saturated with water. The concentration of the eluted bands was determined spectrophotometrically at 350 m$\mu$. A very small amount of a very fast moving band was shown to be an artifact and not the histidine derivative by paper chromatograms using solvent A.

**Recovery Experiments.**—Each derivative which was determined quantitatively in the protein hydrolyzates was subjected to recovery experiments. Amounts of these materials approximating the amounts estimated to be present in the original protein were combined with 0.2 g. of the protein. The mixture was hydrolyzed and worked through the previously described separations. The percentage recovered were: $\epsilon$-dinitrophenyllysine, 83%; dinitrophenylarginine, 73%; and $\alpha,\epsilon$-didinitrophenyllysine, 18%. All the data in Table II have been increased by the appropriate factors to account for these losses and the moisture content of the samples. The reproducibility of the measurements with control substances is in the region of 3 to 5% variation. The hydrolysis of the protein derivatives seems to introduce

TABLE I

$R_f$ VALUES OF 2,4-DINITROPHENYLAMINO ACIDS

| Amino acid | Solvent A | Solvent B | Solvent C |
|---|---|---|---|
| (Dinitroaniline) | 90 | 97 | 96 |
| Tyrosine (di-deriv.) | 78 | 90 | 33 |
| Leucine | 74 | 71 | 70 |
| Isoleucine | 73 | 70 | 70 |
| Lysine ($\alpha,\epsilon$-di-deriv.) | 72 | 81 | 11 |
| Phenylalanine | 71 | 70 | 55 |
| Tryptophan | 70 | 68 | 28 |
| Valine | 68 | 47 | 63 |
| Methionine | 65 | 48 | 47 |
| (Dinitrophenol) | 56 | 25 | 99 |
| Alanine | 50 | 18 | 28 |
| Proline | 48 | 17 | 44 |
| Threonine | 43 | 12 | 0 |
| Glycine | 36 | 8 | 7 |
| Serine | 32 | 6 | 0 |
| Glutamic acid | 14 | 0 | 0 |
| Aspartic acid | 12 | 0 | 0 |
| Histidine (di-deriv.)[a] | 35 | 50 | 0 |
| Arginine[a] | 37 | 0 | 0 |
| Lysine ($\epsilon$-deriv.)[a] | 32 | 5 | 0 |
| Lysine ($\alpha$-deriv.)[a] | 33 | 0 | 0 |

[a] Non-ether extractable derivative.

TABLE II

ANALYSIS OF DINITROPHENYLCASEINS[a]

| Group | Alpha | Beta |
|---|---|---|
| Terminal arginine | 9.3 | 4.7 |
| Terminal lysine | 1.3 | 2.1 |
| Chain lysine | 20.7 | 21.9 |
| Free amino groups[b] | 10.7 | 11.4 |
| Dinitrophenyl groups in the ether extract | 2.2 | 3.9 |
| Dinitrophenyl groups in unhydrolyzed protein[c] | 76 | 65 |
| Protein, % | 87 | 89 |

[a] Moles per 100,000 g. of derivative. [b] Van Slyke amino nitrogen on swollen solid sample. [c] Determined as dinitrophenol by spectrophotometry of an alkaline solution of the dinitrophenyl protein.

(7) E. Abderhalden and P. Blumberg, *Z. physiol. Chem.*, **65**, 318 (1910).

(8) W. G. Gordon, W. F. Semmett, R. S. Cable and M. Morris, THIS JOURNAL, **71**, 3293 (1949).

(9) H. M. Schwartz and C. H. Lea *Biochem. J.*, **50**, 713 (1952).

additional variations and a variation as high as 10% has been noticed in some of the values obtained.

## Results and Discussion

The results presented in Table II show that both the α- and β-caseins have the same type of end groups although the proportion present in each is different. Since only lysine and arginine are shown to occupy terminal positions in these purified caseins these results are quite at variance with those obtained by Lea, *et al.*[9,10] They found slight amounts of terminal aspartyl, glutamyl, lysyl, phenylalanyl and valyl residues but did not find any terminal arginyl residues. This discrepancy may be due entirely to the difference in the methods of obtaining the starting proteins. The proteins used in this study were prepared from fresh unpasteurized cows milk and purified and fractionated at low temperatures. The products were dried with alcohol and ether. At no time were they subjected to heating above room temperature. Lea obtained his material by purification of a commercially prepared casein. Caseins prepared commercially usually contain a small amount of carbohydrate and the casein is dried in the presence of this carbohydrate at elevated temperatures. Some interaction of the protein and carbohydrate may be expected under these conditions. The data of Table III show that such interaction may take place preferentially on the terminal arginyl residues. The distribution of the amino nitrogen in the two unfractionated caseins shows that the amount of the epsilon amino groups of lysine is practically identical but the total amount of amino groups is considerably different. The material used in this study shows about nine times the amount of free α-amino groups which are shown by the preparation used by Lea. The total amount of the α-amino nitrogen found by Lea is about the same as the amount of α-aminolysine groups determined in the experiments reported here. The terminal arginyl residues appear to be masked in Lea's casein.

TABLE III

DISTRIBUTION OF AMINO GROUPS[a] IN CASEINS

| Casein | Total | ε-$NH_2$ | α-$NH_2$ |
|---|---|---|---|
| Unfractionated[8] | 60 | 51 | 9 |
| Alpha[8] | 64 | 55 | 9 |
| Beta[8] | 47 | 41 | 6 |
| Lea's[10] | 52–55 | 51 | 1–4 |

[a] As mg. $NH_2$-N/g. protein total N.

The small unmeasurable amounts of the other terminal residues found by Lea are probably due to traces of other proteins or partially split casein molecules.

Although Lea reported almost complete substitution of the ε-amino groups of lysine, the results reported here show, at best, only about two-thirds substitution of these groups. The free amino nitrogen of the dinitrophenylcaseins was determined by the method of Doherty and Ogg[11] and the results for the β-casein derivative are sufficient to account for most of the difference between the determined and expected values for the epsilon groups of lysine. With the α-casein derivative, the number of amino groups fails to approach the amount necessary to give a satisfactory balance for the total number of amino groups accounted for. A sample of the dinitrophenyl α-casein was retreated with dinitrofluorobenzene in the regular manner. This retreated sample contained only a slightly larger amount of dinitrophenylarginine but the amount of ε-dinitrophenyllysine was raised to 47 moles in 100,000 g. of derivative. These results would be sufficient to raise the total amount of amino group accounted for to the expected level. The small amount of this sample produced did not permit a full study of the derivatives obtained from it. It does show, however, that all the amino groups of α-casein are available to the dinitrofluorobenzene reagent, at least during a retreatment.

Since it may be possible for these dinitrophenyl caseins to be substituted on other groups than the amino nitrogens, some measure of the total substitution is necessary, for the protein content of the derivatives must be known before the average chain weight and minimum molecular weight determination of the original protein can be made. The amino acid derivatives which would be expected to be obtained from such substitutions have never been isolated. However, they may be easily decomposed and may be contributing to the amount of dinitrophenol and dinitroaniline which are found in the ether extract of the hydrolyzates. An estimation of the total absorption of the ether extracts of the hydrolyzates was made and the results, calculated as moles of dinitroaniline or dinitrophenol, are reported in Table II. The smallness of the values indicates that there is probably no significant amount of substitution on the hydroxyl or phenol groups in the protein.

An estimation of the total absorption in alkaline solution of the unhydrolyzed derivatives was made and the results, calculated as moles of dinitroaniline or dinitrophenol, are also reported in Table II. The values approximate quite closely the number of free amino groups plus the imidazole groups of the proteins, and it appears that substitution may have occurred only on the nitrogenous groups of the proteins. If each of these groups, which has been substituted, is assumed to add 166 g. to the total weight, an approximate protein content of these materials can be calculated. This value amounts to 87% for the α-casein derivative and 89% for the β-casein derivative.

Table IV gives the estimated results for the number of end groups in the original casein, the minimum molecular weight, and the average chain weight. The values for β-casein seem to be more consistent than those of the α-casein. Both the sum of the terminal amino groups and the sum of all the amino groups approach satisfactorily the values expected from the Van Slyke amino nitrogen and the lysine content of the protein. The average chain weight approaches the value of 14,300 calculated by Hoover, *et al.*[12]

(10) C. H. Lea and R. S. Hannan, *Biochim. Biophys. Acta*, **4**, 518 (1950).

(11) D. G. Doherty and C. L. Ogg, *Ind. Eng. Chem., Anal. Ed.*, **15**, 751 (1943).

(12) S. R. Hoover, E. L. Kokes and R. F. Peterson, *Textile Res. J.*, **18**, 423 (1948).

TABLE IV

THE TERMINAL AMINO GROUPS OF CASEIN FRACTIONS

| | Alpha | Beta |
|---|---|---|
| Terminal arginine[a] | 10.7 | 5.3 |
| Terminal lysine[a] | 1.5 | 2.4 |
| Non-terminal lysine[a] | 23.8 | 24.6 |
| Unassigned amino groups[a,b] (detn.) | 12 | 13 |
| Unassigned amino groups[a,c] (theory) | 33 | 16 |
| Average weight per terminal lysine | 67,000 | 42,000 |
| Av. wt. per terminal amino group | 8,200 | 13,000 |
| Average chain weight[12] | 10,000 | 14,300 |

[a] Moles per 100,000 g. of protein. [b] Van Slyke amino nitrogen on a swollen solid sample of derivative. [c] From Van Slyke amino nitrogen on pure protein sample.

α-Casein gives a slightly higher number of terminal amino groups than is expected from the difference between the Van Slyke amino nitrogen analysis and the lysine content, but the values are still within experimental error. The average chain weight approaches the 10,000 figure calculated by Hoover, *et al.*,[12] and is remarkably close to the 9,000 value which Hoover, *et al.*, find for unfractionated casein. The minimum molecular weight for the α-casein determined by the number of terminal lysines is about twice the value reported for unfractionated casein by Burk and Greenberg.[13] The large correction for decomposition of the didinitrophenyllysine derivative during the analysis and the possible non-homogeneity of the α-casein may account for this discrepancy. The terminal lysine was not determined on the retreated sample of α-casein and an upward revision of the value may be necessary to account for incomplete reactivity of the lysineamino groups. The change required in this value to reach a reasonable molecular weight would be small compared with the total amount of substitution in the α-casein.

In α-casein there now appears to be seven terminal arginyl residues for each terminal lysyl residue and an average weight per terminal group of 8,200. In β-casein there appears to be only two terminal arginyl residues for each terminal lysyl residue and an average weight per terminal group of 13,000.

(13) N. F. Burk and D. M. Greenberg, *J. Biol. Chem.*, **87**, 197 (1930).