Acknowledgments

We thank Patricia S. Mason, Margaret B. Mitchell, and Anthony Fox for their excellent technical assistance, Dr. Robert A. Neal for the use of a fluorometer, Dr. J. Throck Watson and Betty Fox for performing the mass spectral analyses, and Drs. Perry J. Gehring and Bernard A. Schwetz for reviewing the manuscript.

References

Barbin, A., Brêsil, H., Croisy, A., Jacquignon, P., Malaveille, C., Montesano, R., & Bartsch, H. (1975) *Biochem. Biophys. Res. Commun. 67*, 596–603.

Barrio, J. R., Secrist, J. A., & Leonard, N. J. (1972) *Biochem. Biophys. Res. Commun. 46*, 597–604.

Bartsch, H., & Montesano, R. (1975) *Mutat. Res. 32*, 93–114.

Bolt, H. M., & Filser, J. G. (1977) *Environ. Health Perspect. 21*, 107–112.

Bolt, H. M., Kappus, H., Kaufman, R., Appel, K. E., Buchter, A., & Bolt, W. (1975) *INSERM Symp. 52*, 151–164.

Creech, J. L., & Johnson, M. N. (1974) *J. Occup. Med. 16*, 150–151.

Epstein, J., Rosenthal, R. W., & Ess, R. J. (1955) *Anal. Chem. 27*, 1435–1439.

Green, T., & Hathway, D. E. (1975) *Chem.-Biol. Interact. 11*, 545–562.

Green, T., & Hathway, D. E. (1977) *Chem.-Biol. Interact. 17*, 137–150.

Gross, V. H., & Freiberg, J. (1969) *J. Prak. Chem. 311*, 506–510.

Guengerich, F. P. (1977a) *J. Biol. Chem. 252*, 3970–3979.

Guengerich, F. P. (1977b) *Biochem. Pharmacol. 26*, 1909–1915.

Guengerich, F. P. (1978a) *J. Biol. Chem. 253*, 7931–7939.

Guengerich, F. P. (1978b) *Biochemistry 17*, 3633–3639.

Guengerich, F. P., & Strickland, T. W. (1977) *Mol. Pharmacol. 13*, 993–1004.

Guengerich, F. P., & Mason, P. S. (1979) *Mol. Pharmacol. 15*, 154–164.

Guengerich, F. P., & Watanabe, P. G. (1979) *Biochem. Pharmacol. 28*, 589–596.

Guengerich, F. P., Wang, P., Mitchell, M. B., & Mason, P. S. (1979) *J. Biol. Chem.* (in press).

Hefner, R. E., Jr., Watanabe, P. G., & Gehring, P. J. (1975) *Ann. N.Y. Acad. Sci. 246*, 135–148.

Henschler, D., & Bonse, G. (1977) *Arch. Toxicol. 39*, 7–12.

Johnson, M. K. (1967) *Biochem. Pharmacol. 16*, 185–199.

Jones, B. K., & Hathway, D. E. (1978) *Chem.-Biol. Interact. 20*, 27–41.

Kappus, H., Bolt, H. M., Buchter, A., & Bolt, W. (1975) *Nature (London) 257*, 134–135.

Kappus, H., Bolt, H. M., Buchter, A., & Bolt, W. (1976) *Toxicol. Appl. Pharmacol. 37*, 461–471.

Laemmli, U. K. (1970) *Nature (London) 227*, 680–685.

Laib, R. J., & Bolt, H. M. (1977) *Toxicology 8*, 185–195.

Laib, R. J., & Bolt, H. M. (1978) *Arch. Toxicol. 39*, 235–240.

Lee, F. I., & Harry, D. S. (1974) *Lancet 1*, 1316–1319.

Lowry, O. H., Rosebrough, N. J., Farr, A. L., & Randall, R. J. (1951) *J. Biol. Chem. 193*, 265–275.

Malaveille, C., Bartsch, H., Barbin, A., Camus, A. M., Montesano, R., Croisy, A., & Jacquignon, P. (1975) *Biochem. Biophys. Res. Commun. 63*, 363–370.

Maltoni, C., & Lefemine, G. (1974) *Environ. Res. 7*, 387–405.

Mukhtar, H., Elmamlouk, T. H., & Bend, J. R. (1978) *Chem.-Biol. Interact. 22*, 125–137.

Oesch, F. (1973) *Xenobiotica 3*, 305–340.

Schmassmann, H., & Oesch, F. (1978) *Mol. Pharmacol. 14*, 834–847.

Secrist, J. A., Barrio, J. R., Leonard, N. J., & Weber, G. (1972) *Biochemistry 11*, 3499–3506.

Thomas, P. E., Lu, A. Y. H., West, S. B., Ryan, D., Miwa, G. T., & Levin, W. (1977) *Mol. Pharmacol. 13*, 819–831.

Van Duuren, B. L. (1975) *Ann. N.Y. Acad. Sci. 246*, 258–267.

Wagner, E. R., Muelder, W. M., Watanabe, P. G., Hefner, R. E., Jr., Braun, W. H., & Gehring, P. J. (1975) *J. Labelled Compd. 11*, 535–542.

Walling, C., & Fredricks, P. S. (1962) *J. Am. Chem. Soc. 84*, 3326–3331.

Watanabe, P. G., McGowan, G. R., Madrid, E. O., & Gehring, P. J. (1976) *Toxicol. Appl. Pharmacol. 37*, 49–59.

# Primary Structure of Rabbit α-Lactalbumin†

Thomas P. Hopp* and Kenneth R. Woods

ABSTRACT: Rabbit α-lactalbumin was purified from the milk of New Zealand White rabbits. It was found to exist predominantly as a glycoprotein, containing 5 mol of glucosamine per mol of protein, as well as other sugars. The amino acid sequence of the protein was determined by sequenator analysis and carboxypeptidase digestion. There are 122 amino acids in the protein and a single carbohydrate moiety, probably attached to an asparagine residue at position 45. The C terminus of rabbit α-lactalbumin is one residue shorter than that of the other α-lactalbumins. Sequence comparisons indicate that the α-lactalbumin gene has been undergoing more frequent mutation than had previously been thought. A new method of preparative peptide mapping using 2,5-diphenyloxazole (PPO) fluor to detect peptides is described.

α-Lactalbumin is one of the two protein subunits of lactose synthase (EC 2.4.1.22). Sequence studies have shown it to be homologous to the antibacterial glycosidase lysozyme (Brew et al., 1967).

The primary structures of several α-lactalbumins are known, including human (Findlay & Brew, 1972), cow (Brew et al., 1970), goat (MacGillivary et al., 1979), and guinea pig (Brew, 1972). A partial sequence of kangaroo α-lactalbumin has been reported (Brew et al., 1973). Rabbit α-lactalbumin is interesting in that it exists predominantly as a glycoprotein, while

† From the Lindsley F. Kimball Research Institute of the New York Blood Center, New York, New York 10021, and The Department of Biochemistry, Cornell University Medical College, New York, New York 10021. *Received May 4, 1979.* This research was supported in part by National Institutes of Health U.S. Public Health Service Training Grant 5-T01-GM 01803.

0006-2960/79/0418-5182$01.00/0 

the above mentioned α-lactalbumins are either unglycosylated (human, guinea pig) or only partially glycosylated (bovine). Rat α-lactalbumin has recently been reported to be a glycoprotein but its structure is not known (Brown et al., 1977). The primary structure of rabbit α-lactalbumin is also of value in immunochemical studies of the α-lactalbumins (Hopp and Woods, unpublished experiments) since comparative studies using rabbit antisera to other α-lactalbumins may shed light on the role of amino acid substitutions in the antigenic determinants of these homologous proteins.

## Experimental Procedure

*Materials.* TPCK-trypsin, pepsin, and DFP carboxypeptidases A and B were purchased from Worthington Biochemical Corp. *Staphylococcus aureus* protease was from Miles. Citraconic anhydride and 4-vinylpyridine were obtained from Eastman Kodak Co.; 4-vinylpyridine was redistilled prior to use. Dansyl chloride was from Mann Research Laboratories, Inc. Sequence reagents and solvents were from Beckman Instruments, Inc. Polyamide sheets were from Cheng Chin Trading Co., Taipei, Taiwan. (4-Sulfophenyl)isothiocyanate was from Pierce. (Diethylamino)benzaldehyde (Ehrlich reagent) was obtained from Fisher Scientific Co.

*Methods.* Rabbit α-lactalbumin was prepared by the following method. Milk, obtained from New Zealand White rabbits, was defatted by centrifugation and then acidified to pH 4.6 by addition of glacial acetic acid. Precipitated material (curd) was removed by centrifugation, and the supernatant (whey) was adjusted to pH 7.0 with 6 N NaOH. Sufficient solid $(NH_4)_2SO_4$ was added to achieve 30% saturation, and the solution was stirred for 30 min at room temperature. After centrifugation, the supernatant was applied to a Sephadex G-100 column and eluted with 0.05 M $NH_4HCO_3$. In later preparations of rabbit α-lactalbumin, it was found that the $(NH_4)_2SO_4$ step could be omitted, the neutralized whey being directly applied to the Sephadex G-100 column. Rechromatography of the fourth protein peak yielded pure rabbit α-lactalbumin.

Rabbit α-lactalbumin was treated with CNBr by the method of Gross & Witkop (1962) and then reduced and carboxymethylated according to Crestfield et al. (1963). Rabbit α-lactalbumin was reduced and S-pyridylethylated by the method of Friedman et al. (1970). The pyridylethylated (PE)[1] α-lactalbumin was separated from reagents by gel filtration on Sephadex G-75 in 5% acetic acid and then lyophilized. PE rabbit α-lactalbumin was reacted with CNBr by the trifluoroacetic acid method described by Schroeder et al. (1969).

PE rabbit α-lactalbumin was cleaved specifically at arginine residues by the following method. PE rabbit α-lactalbumin (30 mg) was dissolved in 2 mL of 0.05 N NaOH. Citraconic anhydride (40 μL) was added over a 2-h period as the pH was maintained at approximately 8.3 by addition of 6 N NaOH. The mixture was allowed to remain at room temperature for an additional 2 h to allow any residual citraconic anhydride to hydrolyze, and then the pH was adjusted to 8.5 and 0.30 mg of trypsin was added. After incubation at 37 °C for 1 h, an additional 0.30 mg of trypsin was added and the digestion was continued at 37 °C for another hour. At this time, the mixture was acidified to pH 2.5 by addition of 88% formic acid and allowed to stand for 4 h at room temperature to remove the citraconyl groups. The solution was then applied to a column of Sephadex G-75 and eluted with 10% formic acid.

Peptide maps were made on glass plates coated with a 0.5-mm layer of Whatman microgranular cellulose CC41. First-dimension electrophoresis was performed in a pH 2.0 buffer (water–acetic acid–formic acid, 80:15:5) or a pH 5.5 buffer (acetic acid–pyridine–water, 14:40:1946) for 60 min at 300 V; development in the second dimension was carried out by ascending chromatography (1-butanol–pyridine–acetic acid–water, 150:100:30:120). Analytical plates were loaded with 40 nmol of digested material. Peptides were visualized by spraying the plate with a solution consisting of 0.4% ninhydrin in methanol–acetic acid–collidine, 150:50:20. After ninhydrin development, tryptophan-containing peptides were identified with Ehrlich reagent by the following method. Plates were sprayed with a 2% (w/v) solution of (dimethylamino)benzaldehyde in acetone. Following drying at room temperature for 15 min, we sprayed the plates with a solution consisting of 10 mL of concentrated HCl plus 90 mL of acetone. Tryptophan-containing peptides were visible as brownish purple spots while other peptides disappeared as the ninhydrin color was bleached by the HCl.

Preparative peptide maps were made by using 100–200 nmol of digested material. Detection of peptides was by a new method: 2,5-diphenyloxazole (PPO) fluor was added to the chromatography solvent (5 g/L); peptides could be visualized by observing the plates under a UV light (254 nm), where they appeared as darker, colored spots on a bright blue background. Each peptide had a characteristic color, some being dark purple while others were red-brown, ochre, or dark green. It was found that the PPO fluor had no effect on the position of the peptides, their extraction from the plates, their amino acid analysis, or their sequence analysis. Furthermore, the plates could be developed with ninhydrin without interference from the fluor, and, accordingly, ninhydrin was used to double check for missed peptides after removal of spots from preparative plates. Such development showed that the PPO detection method was highly accurate.

Amino acid analyses were performed on a Beckman Model 119 amino acid analyzer. Samples were hydrolyzed in glass-distilled HCl in vacuo at 110 °C for 24, 48, or 72 h. Values for threonine and serine were calculated by extrapolation to zero time; 72-h values were used for isoleucine and valine.

Assays for total carbohydrate were carried out by the phenol–sulfuric acid method (Hirs, 1967). *N*-Acetylglucosamine, mannose, galactose, and fucose were determined by gas–liquid chromatography (Niedermeier, 1971) after deamination of glucosamine (Porter, 1975). Sialic acid was measured by the thiobarbituric acid technique (Spiro, 1966).

N-Terminal amino acids were determined by the method of Woods & Wang (1966). Preparation of dansyl(pyridylethyl)cysteine was made by reacting the free amino acid derivative with dansyl chloride.

All N-terminal sequence analysis was carried out by using a Beckman Model 890B automatic sequencer, without undercut cup or nitrogen blowdown modifications. Protein sequence runs were accomplished with a Beckman double-cleavage quadrol program. Peptides were degraded by using a DMAA program which was adapted from Beckman program No. 11374, making adjustments for the lack of an undercut cup and nitrogen blowdown system.

To enhance their retention in the sequencer cup, we reacted lysine-terminating tryptic peptides with the Braunitzer reagent [(4-sulfophenyl)isothiocyanate] by the following method. The sample, typically 200–400 nmol of peptide, was dissolved in 100 μL of DMAA buffer along with 500 μg of sodium (sulfophenyl)isothiocyanate. The tube was then flushed with

[1] Abbreviations used: CM, carboxymethyl; PE, pyridylethyl(ated); ΔThr, dehydrothreonine; ΔSer, dehydroserine.

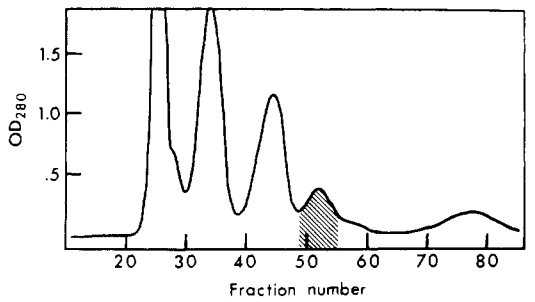


FIGURE 1: Chromatography of rabbit whey on Sephadex G-100. Lyophilized rabbit whey (300 mg) was dissolved in water and applied to the column. The shaded area represents the fractions taken as α-lactalbumin. Final purification was achieved by rechromatography of the α-lactalbumin, yielding a single symmetrical peak. Conditions: a 5 × 30 cm column of Sephadex G-100 equilibrated and eluted with 0.05 M $NH_4HCO_3$, 7.5 mL/fraction.

nitrogen, sealed, and incubated for 1.5–2 h at 50 °C. After this time, the sample was transferred to the sequencer cup and dried, and the program was started.

To minimize sample washout during extractions, 1 mg of succinylated bovine serum albumin was included in the sequencer cup along with some peptide samples. This carrier was tested in a sequence run on carboxymethylated insulin A chain, where it improved yields without causing a significant increase in the levels of background amino acids.

Amino acid anilinothiazolinones obtained from the sequencer were converted to the phenylthiohydantoins (PTH's) by heating at 80 °C for 10 min in 1 N HCl. After extraction with ethyl acetate, the organic- and aqueous-phase derivatives were identified by gas–liquid chromatography and thin-layer chromatography. Data tables for sequence analyses are presented in the supplementary material (see paragraph at end of paper regarding supplementary material). Yields are reported for phenylthiohydantoins identified by gas–liquid chromatography. In most cases, an amino acid was not considered positively identified unless it was clearly detectable by both of the methods used. Where only one method was used to identify the amino acid, it was only assigned if its presence at that position was supported by other evidence, for example, the amino acid composition of small tryptic peptides.

Gas–liquid chromatography was performed on a Hewlett-Packard gas chromatograph equipped with a column of Chromosorb W coated with SP-400 liquid phase, using a programmed temperature gradient. Thin-layer chromatography of organic-phase PTH's was carried out on polyamide sheets according to the method of Summers et al. (1973). The aqueous-phase PTH's of arginine, histidine, and (pyridylethyl)cysteine were identified on polyamide sheets according to a previously described method (Hopp, 1976), using 10% aqueous pyridine. Residues which are not resolved by a particular method are reported in sequence analysis tables as hyphenated pairs.

Carboxypeptidase digestion was by the following method. Pyridylethylated rabbit α-lactalbumin was dissolved at a concentration of 0.5 μmol/mL in 0.1 M triethylamine buffer (adjusted to pH 8.5 with acetic acid) containing 0.5 μmol/mL norleucine. Carboxypeptidase B was added (enzyme/substrate ratio = 1/50) and the digest was incubated at 37 °C. After aliquots were taken at 10, 30, 60, and 120 min, carboxypeptidase A was added, and aliquots were removed at intervals up to 8 h. The aliquots were lyophilized and subjected to amino acid analysis. A parallel control digestion was run, lacking α-lactalbumin. The values obtained were subtracted as background levels from the amino acid values obtained in the α-lactalbumin digest.

Table I: Amino Acid Composition of Rabbit α-Lactalbumin

| AA | by AA analysis | calcd from sequence |
|---|---|---|
| Asp | 18.76 | 19 |
| Thr | 10.05[a] | 10 |
| Ser | 7.95[a] | 8 |
| Glu | 14.21 | 14 |
| Pro | 3.27 | 3 |
| Gly | 5.19 | 5 |
| Ala | 2.24 | 2 |
| 1/2-Cys | 7.11 | 8 |
| Val | 4.23[b] | 4 |
| Met | 1.91 | 2 |
| Ile | 8.07[b] | 8 |
| Leu | 13.26 | 13 |
| Tyr | 2.36 | 2 |
| Phe | 3.03 | 3 |
| His | 2.95 | 3 |
| Lys | 11.89 | 12 |
| Arg | 2.16 | 2 |
| PE-Cys | 7.75[c] | 8 |
| Trp | 4.20[d] | 4 |
| $GlcNH_2$ | 1.87[c] | |

[a] Extrapolated to zero time from 24-, 48-, and 72-h hydrolysates. [b] Value from 72-h hydrolysates. [c] Value obtained from 24-h hydrolysate of PE rabbit α-lactalbumin. [d] Determined spectrophotometrically.

Table II: Carbohydrate Composition of Rabbit α-Lactalbumin[a]

| monosaccharide | residues/mol of protein |
|---|---|
| glucosamine | 4.5 (5)[b] |
| mannose | 3.1 (3) |
| galactose | 2.4 (2) |
| fucose | 0.7 (1) |
| sialic acid | 0.9 (1) |

[a] Determined by gas–liquid chromatography. [b] Numbers in parentheses are nearest integer value, used in calculating the molecular weight of the carbohydrate moiety.

## Results

Figure 1 shows the result of gel filtration of rabbit whey. Rechromatography of the fourth protein peak yielded pure rabbit α-lactalbumin as judged by the presence of a single amino terminus (threonine) by dansylation and an N-terminal sequence free of detectable contaminating residues. When subjected to $NaDodSO_4$–polyacrylamide gel electrophoresis (Figure 2), rabbit α-lactalbumin appeared as two discrete bands with a major component of an apparent molecular weight of approximately 17 800. A fainter band is seen at an apparent molecular weight of 14 700. In contrast, bovine α-lactalbumin yielded a single band with a molecular weight of 14 500. It should be noted that the bovine α-lactalbumin in Figure 2 is reduced but not pyridylethylated; the eight pyridylethyl groups on the rabbit molecule add 840 daltons to its mass. Both bands of rabbit α-lactalbumin were positive when stained by the periodic acid–Schiff method, again showing a major band at 17 800 daltons and a minor band at 14 700 daltons. Since proteins with amino-terminal threonine (or serine) may stain positively with PAS, the 14 700-dalton component may be the nonglycosylated form of rabbit α-lactalbumin; at the least, it probably lacks the major portion of the carbohydrate moiety. The amino acid composition of rabbit α-lactalbumin is presented in Table I, along with the composition derived from the sequence. The carbohydrate composition determined by GLC is given in Table II. The carbohydrate moiety was found to comprise 14.8% of the weight of the protein, by the phenol–sulfuric acid assay.

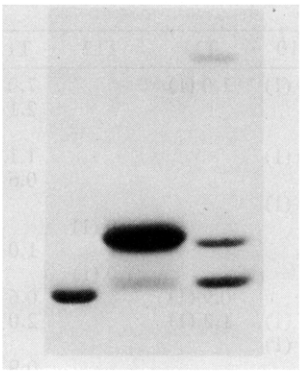

FIGURE 2: Polyacrylamide gel electrophoresis of rabbit α-lactalbumin. PE rabbit α-lactalbumin is in the center lane, with bovine α-lactalbumin in the left-hand lane and standard proteins on the right. Standards are as follows: top band, ovalbumin; middle band, myoglobin; bottom band, chicken lysozyme. A 15% polyacrylamide gel was used, stained with Coomassie blue. All proteins were reduced, but only the rabbit α-lactalbumin sample had been pyridylethylated.

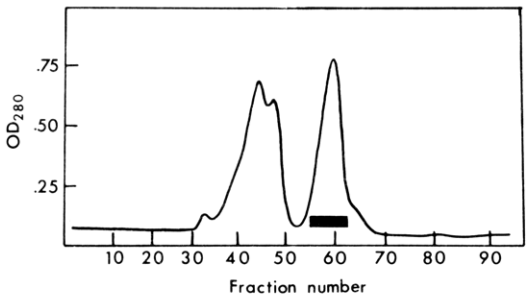


FIGURE 3: Gel filtration of CM-CNBr rabbit α-lactalbumin on Sephadex G-75. Gel filtration was carried out on 50 mg of CM-CNBr rabbit α-lactalbumin using a 1.5 × 90 cm column of Sephadex G-75 in 25% acetic acid. The horizontal bar indicates fractions pooled for ion exchange (Figure 4).

*Sequence Determination. (1) N-Terminal Region, Residues 1–31.* Automated sequence analysis of 670 nmol of native rabbit α-lactalbumin allowed unambiguous assignments for all amino acid residues from the amino terminus through phenylalanine-31, except for the half-cystines at positions 6 and 28 (Table III, supplementary material). The first seven residues were repeated in a run on 100 nmol of pyridylethylated rabbit α-lactalbumin, and (pyridylethyl)cysteine was identified in the aqueous phase of step 6. Several strategically important amino acids were found within the first 31 residues. The location of the two arginines of rabbit α-lactalbumin at positions 5 and 19 led to the use of arginine-specific cleavage to continue the sequence from position 20, while the presence of one of the two methionines at position 23 would enable easy ordering of the three expected cyanogen bromide fragments.

*(2) Cyanogen Bromide Fragments.* Several problems were encountered with the cyanogen bromide reaction. When 50 mg of rabbit α-lactalbumin was reacted by the method of Gross & Witkop (1962), followed by reduction and carboxymethylation, it was found that the expected cleavage at methionine-23 had taken place in extremely low yield. Instead, a larger fragment, containing homoserine but having the amino acid composition expected for the sum of the first and middle cyanogen bromide fragments, was obtained in high yield. Apparently, the cyanogen bromide cleavage at methionine-23 is blocked by serine-24 according to the reaction proposed by Schroeder et al. (1969). The C-terminal fragment CB3, on the other hand, was formed in high yield and could be obtained by gel filtration on Sephadex G-75 in 25% acetic acid, followed by ion exchange on SE-Sephadex C-25 using a pyridine–acetate gradient (Figures 3 and 4).

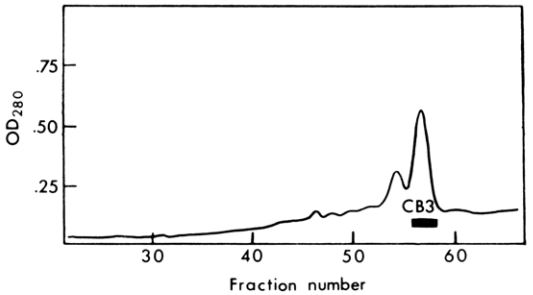


FIGURE 4: Ion exchange of CB3. Sample was applied to a column (0.9 × 90 cm) of Sephadex SE C-25 in 0.1 M pyridine–acetic acid buffer, pH 3.1, and eluted with a gradient made from this buffer (200 mL) plus 200 mL of 2.0 M pyridine–acetic acid, pH 5.0. The horizontal bar indicates fractions pooled for structure analysis.

Table IV: Amino Acid Compositions of Major Fragments[a]

| AA | CB1 | CB3 | CTA | CTB1[b] | CTB2 |
|---|---|---|---|---|---|
| Asp | 2.0 (2)[c] | 3.0 (3) | 18.0 (18) | 8.0 (8) | 10.0 (10) |
| Thr | 2.8 (3) | (0) | 6.4 (7) | 4.9 (5) | 2.1 (2) |
| Ser | 0.9 (1) | 0.8 (1) | 8.0 (8) | 5.2 (7) | 1.3 (1) |
| Glu | 4.0 (4) | 3.9 (4) | 9.6 (10) | 5.4 (5) | 4.7 (5) |
| Pro | (0) | 1.1 (1) | 2.4 (3) | 1.3 (1) | 1.1 (2) |
| Gly | 0.9 (1) | 1.1 (1) | 5.2 (4) | 3.2 (3) | 1.2 (1) |
| Ala | (0) | 1.0 (1) | 2.7 (2) | 1.2 (0) | 1.7 (2) |
| Val | (0) | 0.9 (1) | 4.6 (4) | 2.9 (2) | 1.9 (2) |
| Met | (0) | (0) | 1.3 (2) | 0.5 (1) | 0.9 (1) |
| Ile | 1.0 (1) | 1.9 (2) | 6.8 (8) | 4.0 (5) | 2.8 (3) |
| Leu | 4.2 (4) | 4.0 (4) | 8.8 (9) | 3.4 (3) | 5.7 (6) |
| Tyr | 1.0 (1) | (0) | 1.7 (1) | 1.8 (1) | (0) |
| Phe | (0) | (0) | 2.6 (3) | 1.6 (2) | 1.2 (1) |
| His | (0) | 2.4 (2) | 2.7 (3) | 1.8 (1) | 1.9 (2) |
| Lys | 2.3 (2) | 5.2 (5) | 9.3 (10) | 5.0 (5) | 5.4 (5) |
| Arg | 1.8 (2) | (0) | (0) | (0) | (0) |
| Trp | (0) | 2.3 (3) | | | |
| CM-Cys | | 1.6 (2) | | | |
| PE-Cys | 0.9 (1) | | 4.9 (7) | 1.7 (2) | 4.1 (5) |
| Hse + Hsl | 0.6 (1) | (0) | | | |
| $GlcNH_2$ | | | 2.0 | 1.6 | |
| N-terminal: | Thr | Lys | Asp | Asp | Ile |
| residues: | 1–23 | 94–123 | 20–123 | 20–71 | 72–123 |
| yield (%): | 65 | 70 | 50 | 50 | 50 |

[a] 24-h hydrolysates. [b] CTB1 was slightly contaminated with CTB2. [c] Numbers in parentheses indicate the number of residues expected from the sequence of rabbit α-lactalbumin.

To obtain the N-terminal CNBr fragment CB1, we carried out the cyanogen bromide reaction on 50 mg of PE rabbit α-lactalbumin in 70% trifluoroacetic acid. This solvent was effective in blocking the side reaction with serine-24, and the N-terminal fragment could be obtained by gel filtration on Sephadex G-75 in 5% formic acid (Figure 5, supplementary material). The N-terminal fragment coeluted with the C-terminal fragment but could be obtained pure by further separation on SE-Sephadex C-25. A single peak eluting from the ion-exchange column proved to be the N-terminal fragment, while the C-terminal fragment failed to elute. This behavior of the C-terminal fragment must be a function of the alkylating reagents. The two half-cysteines would acquire two strong negative charges when carboxymethylated; this would increase the solubility of the peptide and serve to repell it from the SE-Sephadex. When pyridylethylated, the fragment would acquire two pyridine rings, which would lower the overall polarity of the peptide and, being weakly basic, might increase the affinity of the peptide for the SE-Sephadex. The amino acid compositions of the N-terminal PE-CNBr fragment (CB1) and the C-terminal CM-CNBr fragment (CB3) are given in Table IV.

Additional support for the validity of the assignment of the first 23 residues of rabbit α-lactalbumin was obtained by

Table V: Tryptic Peptides from Rabbit α-Lactalbumin[a]

| AA | T1 | T2 | T4 | T6A | T7 | T8 | T9 | T10 | T11 | T12 | T13A | T16 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Asp | 4.0 (4) | 0.9 (1) | | 0.9 (1) | | | | 1.0 (1) | 1.0 (1) | | 7.0 (7) | |
| Thr | 2.4 (2) | | | | 1.6 (2) | 1.1 (1) | | | | | 2.1 (2) | |
| Ser | 2.2 (2) | 0.9 (1) | | 1.0 (1) | | | | | | | | 1.0 (1) |
| Glu | 2.2 (2) | 2.0 (2) | | | 1.2 (1) | 1.8 (2) | | 1.0 (1) | | | 1.1 (1) | |
| Pro | | 1.4 (1) | | | | | | | | | 0.6 (1) | |
| Gly | 2.4 (2) | | | | | | | 1.0 (1) | | | | |
| Ala | | | | | | | | | | 1.0 (1) | | |
| Val | 1.1 (1) | | | | | | | | | | 1.0 (1) | 1.2 (1) |
| Met | | | | | | | | | | 1.0 (1) | | |
| Ile | 3.1 (3) | | | 1.0 (1) | | | | | 0.9 (1) | | 0.6 (1) | |
| Leu | | | | | 1.1 (1) | 1.0 (1) | 1.0 (1) | 1.1 (1) | 1.2 (1) | | 2.0 (2) | 0.9 (1) |
| Tyr | 1.1 (1) | | | | | | | 1.0 (1) | | | | |
| Phe | 1.3 (1) | | | | | | | | | | 0.9 (1) | |
| His | | | | | | | | | | | | |
| Lys | 1.0 (1) | 1.1 (1) | 1.0 (1) | | | 0.9 (1) | 1.0 (1) | | 1.2 (1) | 0.8 (1) | 0.9 (1) | 1.1 (1) |
| Arg | | | | | 1.0 (1) | | | 0.8 (1) | | | | |
| Trp | | | | | | | | | | | | +[b] (1) |
| PE-Cys | | | | | | 0.8 (1) | | | | 0.9 (1) | 1.6 (2) | |
| Hse + Hsl | | | | 0.6 (1) | | | | | | | | |
| $GlcNH_2$ | 1.4 | | | | | | | | | | | |
| N-terminal: | Ile | Gln | | Asp | Thr | PE-Cys | Leu | Glu | Ile | PE-Cys | Ile | Leu |
| residues: | 40–58 | 65–70 | 123 | 20–23 | 1–5 | 6–11 | 12–13 | 14–19 | 95–98 | 91–94 | 72–90 | 59–64 |
| yield (%): | 22 | 13 | 54 | 57 | 51 | 60 | 57 | 80 | 71 | 35 | 68 | 66 |

[a] 24-h hydrolysates. [b] Detected by the Ehrlich reaction on peptide maps.

subjecting the N-terminal CNBr fragment to trypsinization and peptide mapping (Figure 6, supplementary material), followed by amino acid analysis of the peptide spots. All of the expected peptides were found, including the C-terminal homoserine-containing peptide. Amino acid compositions are shown in Table V (peptides T6A and T7–T10).

The central CNBr fragment was found to contain the carbohydrate moiety, as determined by the presence of glucosamine in its composition. It was not used for sequence determination. The C-terminal fragment CB3 was subjected to automated sequence analysis using a DMAA peptide program. Good results were obtained through cycle 24, although "histidine preview" appeared at cycle 10 and increased at cycle 14 when the second histidine residue was reached (Table VI, supplementary material). In order to extend the sequence of this CNBr fragment to its C terminus, we subjected it to *S. aureus* protease digestion and peptide mapping (Figure 7, supplementary material). Peptide CB3-SP4 was isolated from preparative peptide maps (Table XV) and was sequenced in its entirety by using the DMAA program with succinylated BSA carrier (Table VII, supplementary material). This information completes the sequence of the C-terminal CNBr fragment, and its extention to the C terminus was verified by carboxypeptidase digests on whole PE-α-lactalbumin (see below). Peptides CB3-SP1, -SP2, and -SP3 were also isolated from the staphylococcal protease maps; they represent the other peptides expected from this digest, and their compositions confirm the sequence of CB3.

(3) *Citraconylation and Trypsinization.* Figure 8 depicts the separation of the products of citraconylated and trypsinized PE rabbit α-lactalbumin on Sephadex G-75 in 10% formic acid. Judging by its elution position, amino acid composition (Table IV), and N-terminal residue (aspartic acid), we found that CTA was the expected large fragment comprising the C-terminal portion of the molecule starting with residue 20. A run of 35 cycles on 300 nmol of this fragment yielded good information from residue 20 through residue 53 (cycle 34), as shown in Table VIII (supplementary material). However, several ambiguities remained. Distinction between asparagine and aspartic acid for residues 43–45 was difficult because there was by that time a significant background of aspartic acid.

Table XV: Amino Acid Compositions of Peptic and Staphylococcal Protease Peptides

| AA | PA1 | CB3-SP1 | CB3-SP2 | CB3-SP3 | CB3-SP4 |
|---|---|---|---|---|---|
| Asp | 4.0 (4) | 1.3 (1) | 1.3 (1) | 1.1 (1) | |
| Thr | 1.6 (2) | | | | |
| Ser | 2.1 (2) | | 0.8 (1) | | |
| Glu | 3.6 (4) | 1.0 (1) | 1.1 (1) | 1.0 (1) | 1.0 (1) |
| Pro | 1.7 (2) | | 1.1 (1) | | |
| Gly | | | 0.9 (1) | 0.3 (0) | |
| Ala | | | 1.2 (1) | | |
| Val | 1.4 (1) | | | | 0.9 (1) |
| Ile | | 1.0 (1) | 1.0 (1) | 0.3 (0) | |
| Leu | | 1.0 (1) | 2.0 (2) | 1.0 (1) | |
| Phe | 1.0 (1) | | | | |
| His | | | 2.0 (2) | | |
| Lys | 2.3 (2) | 1.7 (2) | 0.9 (1) | | 2.0 (2) |
| Trp | +[a] (1) | | + (1) | | + (1) |
| CM-Cys | 3.0 (3) | | 0.7 (1) | | 0.8 (1) |
| residues: | 60–80 | 94–99 | 100–113 | 114–116 | 117–122 |
| yield (%): | 85 | 40 | 32 | 60 | 52 |

[a] Determined by the Ehrlich reaction on peptide maps.

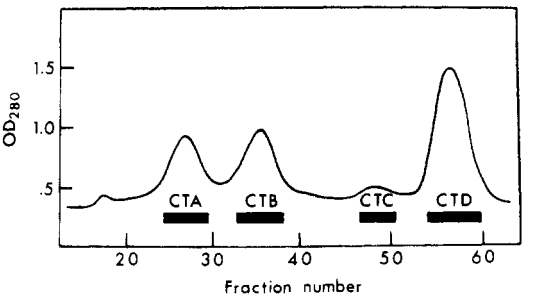


FIGURE 8: Separation of citraconylated trypsinized fragments of PE rabbit α-lactalbumin. Conditions: 30 mg of citraconylated trypsinized PE rabbit α-lactalbumin was applied to a 2.5 × 90 cm column of Sephadex G-75 in 10% formic acid. The horizontal bars indicate fractions pooled.

Serine-47 and threonine-48 were not detected, and differentiation between leucine and isoleucine at position 52 by silylation was impossible, also due to high background levels.

An unanticipated cleavage between positions 71 and 72 occurred in approximately 50% yield during citraconylation

and trypsinization of PE rabbit α-lactalbumin. The two resultant peptides appeared as a single peak (CTB) on gel filtration (Figure 8). Ion exchange on DEAE-Sephadex separated the two peptides. The first, CTB1, contained glucosamine and yielded aspartic acid on N-terminal dansylation analysis and was judged to be the N-terminal half of CTA, while CTB2, lacking glucosamine and having N-terminal isoleucine, proved to be the C-terminal half of CTA. The cause of the cleavage between residues 71 and 72 is unclear. At first, it was suspected that an arginine residue might be responsible; however, the two arginines of rabbit α-lactalbumin had already been assigned to positions 5 and 19. Furthermore, the amino acid composition and sequence of peptic peptide PA1 ultimately ruled out an additional arginine at this position. It may be that the Asn–Ile bond in question is particularly susceptible to hydrolysis during the deblocking incubation at pH 2.5.

A sequenator run of 35 cycles on CTB2 identified all residues from isoleucine-72 to isoleucine-101. In addition, alanine was identified at position 106, and residue 105 was found to be either leucine or isoleucine (Table IX, supplementary material).

This information was confirmed and extended as follows. CTB2 was digested with trypsin and subjected to peptide mapping as seen in Figure 9 (supplementary material). Peptide T13A, comprising the sequence from isoleucine-72 to lysine-90, was isolated from the maps and used to recheck the assignment of aspartic acids and asparagines in positions 82 to 84 and 87 to 88. Approximately 200 nmol was incubated with 40 μg of staphylococcal protease for 18 h at 37 °C in 0.2 M $NH_4HCO_3$ in order to effect cleavage after glutamic acid-78. Following lyophilization, the whole sample was reacted with Braunitzer reagent and applied to the sequencer cup. After the first cycle the N-terminal fragment, now lacking the (sulfophenyl)thiourea group, washed out and did not interfere with the remainder of the analysis. The C-terminal fragment, still derivatized at its terminal lysine, yielded good results through valine-89. Residues 82, 83, 87, and 88 were verified as aspartic acid, and residue 84 was verified as asparagine (Table X, supplementary material).

Peptide T12 was reacted with Braunitzer reagent and sequenced (Table XI, supplementary material), yielding alanine in cycle 2 and methionine in cycle 3. This, together with the N-terminal analysis and the amino acid composition data (Table V), confirms sequence positions 91–94. Similarly (Table XII, supplementary material), peptide T11 gave second and third residues of leucine and aspartic acid, respectively. Consideration of the N-terminal and amino acid analysis, together with this sequence information, confirms residues 95–98.

*(4) Tryptic Peptides of PE Rabbit α-Lactalbumin.* PE rabbit α-lactalbumin (35 mg) was digested with 0.35 mg of trypsin for 18 h at 37 °C and then lyophilized. Figure 10 shows the peptide map obtained by using 0.5 mg of the digested material. Sixteen ninhydrin-positive spots were observed, while only thirteen are expected from the structure determined for rabbit α-lactalbumin, assuming that the lysine–proline bond at position 108 to 109 is not cleaved. Amino acid analysis of peptides isolated from preparative plates resolved this discrepancy. All of the expected peptides were found, while spot T5 yielded only $NH_3$; T3 and T6 were weak ninhydrin spots and accordingly were obtained in low yield. The amino acid compositions of T3 and T6 were nearly identical with that of T1, except that T1 contains several residues of glucosamine, while T3 and T6 have none. Thus, T1, T3, and T6 are different forms of the same peptide, T1 being the

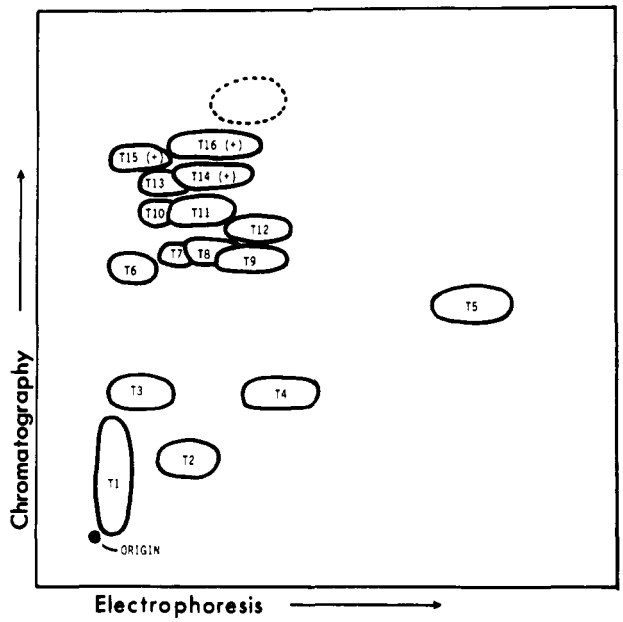


FIGURE 10: Peptide map of tryptic digest of PE rabbit α-lactalbumin. Peptide sample was dissolved in 10 μL of pH 2.0 electrophoresis buffer and spotted on the prewetted electrophoresis plate. Electrophoresis was carried out for 60 min, with the cathode at right. Second-dimension development was by ascending chromatography. The spot indicated by the dashed line was ninhydrin negative but was seen as a dark spot on PPO-containing plates under UV illumination. It yielded no amino acids on hydrolysis, although it was present in all tryptic peptide maps and was useful as a landmark for orienting peptides. A (+) indicates a positive Ehrlich reaction.

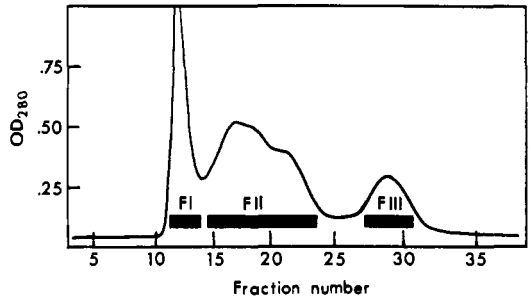


FIGURE 11: Gel filtration of tryptic fragments of PE rabbit α-lactalbumin. Conditions: 35 mg of trypsinized PE rabbit α-lactalbumin was separated on a 1.5 × 90 cm column of Sephadex G-25 in 0.05 M $NH_4HCO_3$. The horizontal bars indicate fractions pooled for peptide mapping.

glycosylated form, while T3 and T6 are unglycosylated, perhaps differing from each other by the presence or absence of the amide on the asparagine residue which normally is linked to the carbohydrate. Amino acid compositions of tryptic peptides relevent to the proof of sequence are listed in Table V.

To simplify the peptide map for isolation of particular peptides, we separated the tryptic digest by gel filtration on Sephadex G-25 (Figure 11) into three fractions. Mapping of the first fraction (FI) revealed only two spots to be present in significant levels; these were T1 and T15. Peptide T1 was isolated from preparative maps, and its amino acid composition and sequence were determined. It was found to contain 19 amino acids, corresponding to positions 40–58 in rabbit α-lactalbumin. It was sequenced in its entirety in the automatic sequencer, using a DMAA program (Table XIII, supplementary material). This information extended the sequence obtained from CTA and clarified several ambiguous residues. The amino acids at positions 43 and 44 were both shown to

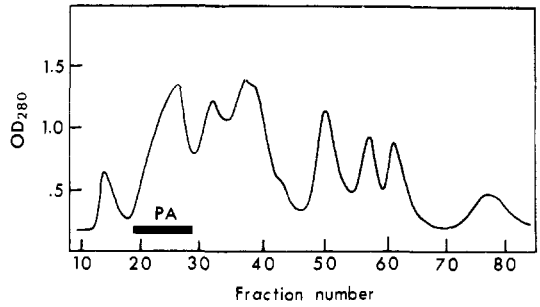


FIGURE 12: Gel filtration of peptic fragments of rabbit α-lactalbumin. A digest of 50 mg of rabbit α-lactalbumin was separated on a 1.5 × 90 cm column of Bio-Gel P-4 in 5% acetic acid. Fraction PA was taken for further purification (see the text).

be asparagines, while residue 47 was identified as serine and residue 48 was identified as threonine. At position 45, a continued elevation in the level of Asp-PTH was seen, while Asn-PTH decreased to near background level. Asp-PTH then decreased dramatically at residue 46. Residue 45 is tentatively assigned as a glycosylated asparagine for the following reasons: peptide T1 clearly contains a carbohydrate moiety as judged by its glucosamine content (Table V); the amount and monosaccharide composition of the carbohydrate portion of rabbit α-lactalbumin are consistent with a single, asparagine-linked glyco moiety of the "complex" type (Spiro, 1973); the requisite sequence for this type of carbohydrate attachment, Asn-X-Thr(or Ser), only occurs at positions 45–47. Finally, the amount of Asp-PTH obtained at cycle 45 is relevant. Our sequencing and conversion conditions result in approximately 50% hydrolysis of Asn-PTH to Asp-PTH; thus asparagine yields two spots of roughly equal intensity on polyamide plates, while aspartic acid yields a spot twice as dark as either spot derived from asparagine. The Asp-PTH spot for position 45 was approximately equal in intensity to the Asp-PTH derived from the previous two asparagine residues, suggesting that it may also be derived from asparagine. However, the Asn-PTH spot is almost undetectable in this cycle. This eliminates the possibility that the Asp-PTH represents tailing from the previous cycles (Asn would tail also) and suggests that the Asp-PTH actually derives from a glycosylated asparagine, since the Asp-PTH spot did not double in darkness, as it should if this residue were free aspartic acid. No unusual spots were seen on the plates for residue 45, suggesting that the glycoasparagine derivative was not extracted from the sequencer cup or, if so, did not survive the conversion conditions.

When FIII was mapped, it was found to contain peptides T11 and T16. T16 was isolated from preparative plates and its structure was determined. The N-terminal residue was found by dansylation to be leucine; the peptide was sequenced after Braunitzer modification (Table XIV, supplementary material), and the single lysine residue, which was not detected in the sequencer fractions due to its derivitization with the Braunitzer reagent, was positioned at the C terminus in accordance with the amino acid composition of the peptide and the specificity of trypsin.

(*5*) *Peptic Peptides.* Peptic peptides were prepared from 100 mg of rabbit α-lactalbumin by digestion for 18 h with 2.5 mg of pepsin, followed by chromatography on Bio-Gel P-4 in 5% acetic acid (Figure 12). The PA peptide obtained from this column was reduced and carboxymethylated and desalted on a Sephadex G-25 column. The peak eluting at the void volume was subjected to ion exchange on DEAE-Sephadex (Figure 13), and the major component, PA1, was used for sequence studies. Its amino acid composition is shown in Table XV.

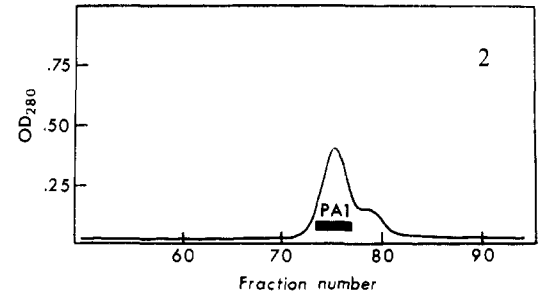


FIGURE 13: Ion exchange of PA peptide. The reduced, carboxymethylated PA peptide was applied to a 0.9 × 60 cm column of DEAE-Sephadex A-50 in 0.01 M Tris-HCl, pH 8.0, and eluted with a gradient of increasing NaCl concentration, from 0 to 0.5 M (400-mL total).

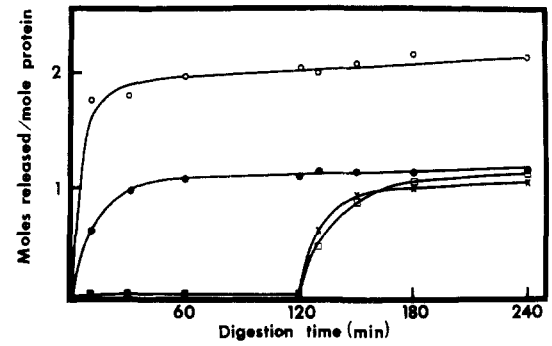


FIGURE 14: Carboxypeptidase digestion of PE rabbit α-lactalbumin. Carboxypeptidase B was added at the zero time point; after the 120-min aliquot was removed, carboxypeptidase A was immediately added. Equivalents of lysine released are indicated by (○), (pyridylethyl)cysteine, by (●), valine, by (×), and tryptophan, by (□). Other amino acids were released at rates slower than that of tryptophan and are not shown.

It was desirable to obtain the sequence of this peptide without use of the Braunitzer reagent so that the two lysine residues could be identified readily. To minimize the effects of sample washout during extractions, we included succinylated bovine serum albumin with the sample in the sequencer cup. Peptide PA1 was sequenced for 15 cycles (Table XVI, supplementary material), and the resulting sequence was sufficient to fill the gap between tryptic peptide T16 and the citraconylated tryptic fragment CTB2. Furthermore, overlap was obtained at both ends of PA1, establishing the order of the three segments as T16, PA1, CTB2.

(*6*) *Carboxypeptidase Digestion.* When PE rabbit α-lactalbumin was incubated with carboxypeptidase A, no amino acids were released over background levels found in control digests lacking substrate. This suggested that lysine or (pyridylethyl)cysteine, a carboxypeptidase B substrate, occupies the C-terminal position. Figure 14 shows the result of sequential digestion by CPB followed by CPA. Digestion with CPB released two lysines and one (pyridylethyl)cysteine from the C terminus of PE rabbit α-lactalbumin. The fact that the 10-min point yields a lysine value more than twice that of (pyridylethyl)cysteine, together with the fact that the lysine value is already greater than 1 equiv/mol of protein, demonstrates that there are two lysines on the C terminus of the molecule, with a (pyridylethyl)cysteine in the third-to-last position. Amino acids released upon addition of CPA are consistent with the structure obtained by sequenator analysis.

Proof of the complete amino acid sequence of rabbit α-lactalbumin is schematized in Figure 15. Overlapping sequence information is available for the entire sequence with the exception of the single peptide bond connecting Lys-58 and Leu-59. This bond is assigned as shown because this ar-

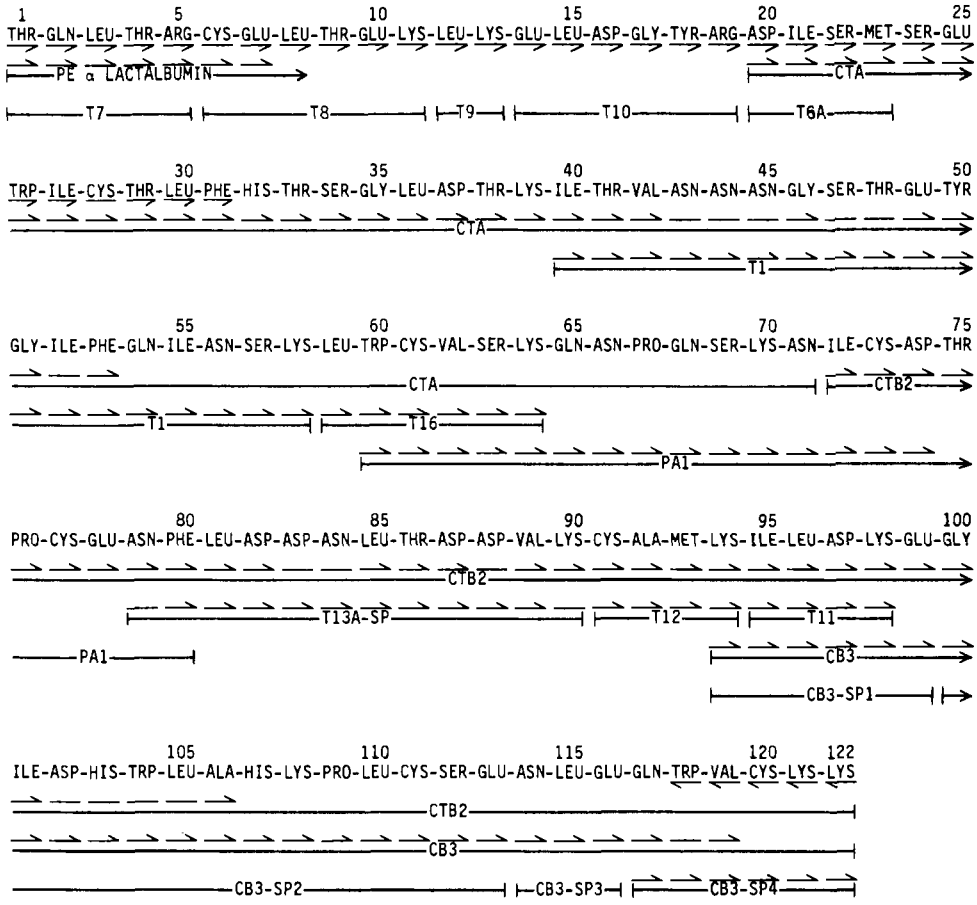


FIGURE 15: Proof of sequence of rabbit α-lactalbumin. Horizontal lines indicate the extent of peptides. A (⇀) below a residue indicates automatic sequencer analysis on whole α-lactalbumin. Residues with a (⇀) were determined by sequencer analysis on peptides. A (⇀) indicates carboxypeptidase digestion. Sequencer cycles where no residue was clearly identified are indicated by (—). Arrows indicate the extent of sequence runs. Peptides, or portions of peptides, lacking arrows were not sequenced but are included where their compositions are valuable in confirming amino acid assignments from sequence runs. Abbreviations used: T, tryptic peptides; CTA and CTB2, citraconylated tryptic peptides; PA1, peptic peptide; SP, staphylococcal protease peptides; CB, cyanogen bromide peptides.

rangement of the two sections of overlapped sequence yields the correct amino-terminal sequence and because insertion of additional amino acids (between Lys-58 and Leu-59) would cause disagreement between the calculated and experimental amino acid compositions. In addition, the sequence as shown is homologous to other α-lactalbumins without additional residues between Lys-58 and Leu-59.

## Discussion

Rabbit α-lactalbumin is a glycoprotein, containing 122 amino acids and a single heteropolysaccharide moiety of approximately 12 sugar residues/mol of protein. The molecular weight of the peptide portion of rabbit α-lactalbumin, calculated from its sequence, is 14 052; that of the carbohydrate moiety is 2259 when calculated from the integral values shown in Table II. This represents a carbohydrate content of 13.8%, which is in accord with the value of 14.8% obtained by the phenol–sulfuric acid assay. The size of the carbohydrate moiety may also be estimated from the polyacrylamide gel electrophoresis results. By subtracting the lower molecular weight electrophoretic band from the higher, we obtained a carbohydrate moiety weight value of 3100. This yields a carbohydrate content of 17.5%, somewhat higher than the others. This value is probably incorrectly high, since glycoproteins are known to appear anomalously heavy on NaDodSO$_4$ electrophoresis (Segrest & Jackson, 1972).

The carbohydrate attachment site occurs within the tryptic fragment encompassing residues 40–58, and our data indicate that asparagine-45 is the specific residue involved. The high degree of glycosylation at this site is especially interesting. Bovine α-lactalbumin can also be glycosylated, although the site of carbohydrate attachment is not known. However, glyco-α-lactalbumin was found to represent at most 7% of the total amount of bovine α-lactalbumin (Barman, 1970) while, in the case of the rabbit, glycosylation appears to approach 100% as judged from the presence of approximately 5 mol of glucosamine per mol of rabbit α-lactalbumin and the positive PAS staining of both of the polyacrylamide gel bands obtained. Brown et al. (1977) demonstrated that rat α-lactalbumin is a glycoprotein and that it contains 2.6 mol of glucosamine per mol of protein, as well as other sugars. They are currently working to localize the site of carbohydrate attachment, and it will be interesting to see if this site is the same in different species.

If asparagine-45 is the site of carbohydrate attachment in bovine α-lactalbumin, the differing degrees of glycosylation might be due to the influence of residue 46, which is glutamine in bovine α-lactalbumin and glycine in rabbit α-lactalbumin. It has been pointed out (Huber et al., 1976) that asparagine glycosylation frequently occurs in β-bend regions of polypeptide chains, and the location in rabbit α-lactalbumin seems to be consistent with this view. Residue 45 of α-lactalbumin corresponds to a β-bend position in the X-ray crystallographic structure of the homologous protein lysozyme, and, furthermore, residues 44–47 are indicated as a β bend by the protein conformation prediction method of Chou & Fasman (1974).

In Figure 16 the amino acid sequences of bovine and human α-lactalbumin and chicken lysozyme are compared to the

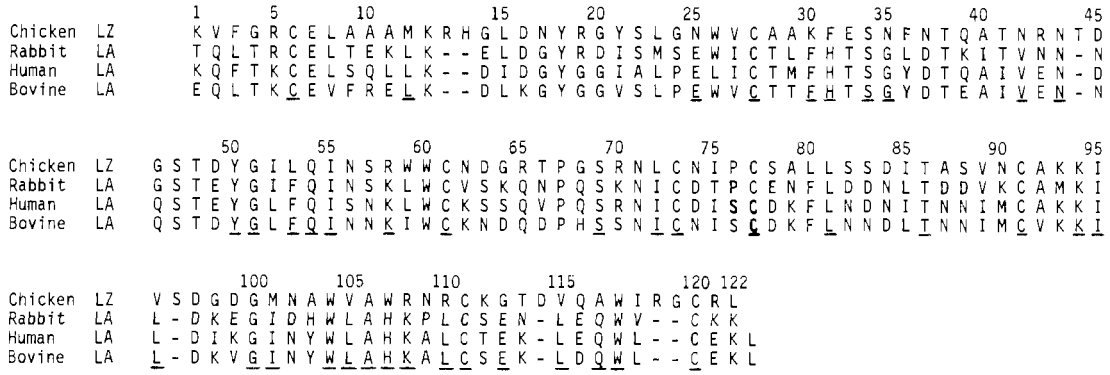


FIGURE 16: Comparison of sequences of $\alpha$-lactalbumins and lysozyme. Numbering is that of rabbit $\alpha$-lactalbumin. Underlined residues are identical in all $\alpha$-lactalbumins, including those of the goat, guinea pig, and kangaroo (not shown). Chicken lysozyme (LZ) sequence was determined by Canfield & Liu (1965). Human $\alpha$-lactalbumin (LA) sequence is that of Findlay & Brew (1972); bovine LA sequence is that of Brew et al. (1970).

sequence of rabbit $\alpha$-lactalbumin. The amino acids underlined below bovine $\alpha$-lactalbumin are conserved in all $\alpha$-lactalbumins sequenced to date, including those of the guinea pig, goat, and kangaroo. Some of these invariant residues probably represent crucial structural and functional groups. For example, the eight half-cystine residues are retained, presumably because of their importance in maintaining the correct three-dimensional structure necessary for the molecule's activity. It has been shown that modification of histidine-32 by diethyl pyrocarbonate (Schindler et al., 1976) inactivated human $\alpha$-lactalbumin in the lactose synthase reaction. If this residue is indeed necessary for catalytic function, then it would be expected to be conserved in all $\alpha$-lactalbumins, and for those analyzed to date, including rabbit $\alpha$-lactalbumin, this is the case. Assuming that $\alpha$-lactalbumin and lysozyme have nearly identical three-dimensional structures (Barel et al., 1972; Browne et al., 1969; Warme et al., 1974), this histidine of $\alpha$-lactalbumin would lie within an area comparable to the substrate cleft of lysozyme. Here, it might be substituted for the catalytic glutamic acid (residue 35) of lysozyme.

With the addition of rabbit $\alpha$-lactalbumin to the number of known $\alpha$-lactalbumin sequences, it appears that the rate of mutation among the $\alpha$-lactalbumins is even greater than the already high level which had previously been perceived (Dayhoff, 1976). Rabbit $\alpha$-lactalbumin varies more from the other $\alpha$-lactalbumins than any of the latter vary from each other. This suggests that the line leading to the rabbit branched off from the common placental mammalian stock significantly earlier than did the lines leading to man, cow, and guinea pig or that the mutation rate has been unusually high in the $\alpha$-lactalbumin gene of the rabbit. The C terminus of rabbit $\alpha$-lactalbumin ends one residue short of the other $\alpha$-lactalbumins. This suggests either a deletion/insertion mutation at the C terminus or a point mutation involving the interchange of an amino acid codon with a terminator codon (several leucine codons are single base mutations from termination codons, and leucine is present at the C terminus of several of the other $\alpha$-lactalbumins). It is not obvious which condition is ancestral, that of the rabbit or that of the other $\alpha$-lactalbumins; however, it may be significant that rabbit $\alpha$-lactalbumin is similar to the lysozymes in having only two amino acids following the cysteine residue at position 120.

It is not clear how the peptide mapping procedure described here works. The darker, colored spots may result from either of two effects. Incoming UV irradiation may be absorbed by the peptides (e.g., at wavelengths like 220 or 280 nm where peptides characteristically absorb) which in turn would result in a decrease in emission by the PPO fluor. Alternatively, the peptides may interact with the PPO fluor on the plate, causing its emission maximum to shift to longer (or possibly shorter) wavelengths, thus generating the observed color changes. Regardless of which effect is operating, the method is probably not as efficient as it might be if conditions are found which optimize the effect. If peptide absorbance is the critical factor, then a wavelength other than 254 nm might be absorbed more completely, resulting in an improvement of the sensitivity of the method. Other fluors might by tested, in order to find a compound whose emission spectrum is shifted farther by peptides or whose absorbance spectrum is more completely blocked by peptide absorbance.

## Acknowledgments

We thank Pearl Shum and Milla Block for technical assistance and Dr. Marilyn S. Horowitz for advice concerning the sequence analysis. Dr. Carey Hanley and Sharon Ryan provided a portion of the rabbit whey used in this study. Dr. R. Mailhammer performed the gas–liquid chromatographic determination of hexoses and hexosamines.

## Supplementary Material Available

Experimental material including sequence analyses (Tables III, VI–XIV, and XVI), a gel filtration plot (Figure 5), and peptide maps (Figures 6, 7, and 9) (11 pages). Ordering information is given on any current masthead page.

## References

Barel, A. O., Prieels, J. P., Maes, E., Looze, Y., & Leonis, J. (1972) *Biochim. Biophys. Acta 257*, 288.

Barman, T. E. (1970) *Biochim. Biophys. Acta 214*, 242.

Brew, K. (1972) *Eur. J. Biochem. 27*, 341.

Brew, K., Vanaman, T. C., & Hill, R. L. (1967) *J. Biol. Chem. 242*, 3747.

Brew, K., Castellino, F. J., Vanaman, T. C., & Hill, R. L. (1970) *J. Biol. Chem. 245*, 4570.

Brew, K., Steinman, H. M., & Hill, R. L. (1973) *J. Biol. Chem. 248*, 4739.

Brown, R. C., Fish, W. W., Hudson, B. G., & Ebner, K. E. (1977) *Biochim. Biophys. Acta 491*, 82.

Browne, W. J., North, A. C. T., Phillips, D. C., Brew, K., Vanaman, T. C., & Hill, R. L. (1969) *J. Mol. Biol. 42*, 65.

Canfield, R., & Liu, A. K. (1965) *J. Biol. Chem. 240*, 1997.

Chou, P. Y., & Fasman, G. D. (1974) *Biochemistry 13*, 222.

Crestfield, A. M., Moore, S., & Stein, W. H. (1963) *J. Biol. Chem. 238*, 622.

Dayhoff, M. O. (1976) *Atlas of Protein Sequence and Structure*, Vol. 5, Suppl. 2, p 6, National Biomedical Research

Foundation, Silver Springs, MD.
Findlay, J. B. C., & Brew, K. (1972) *Eur. J. Biochem. 27*, 65.
Friedman, M., Krull, L. H., & Cavins, J. F. (1970) *J. Biol. Chem. 245*, 3868.
Gross, E., & Witkop, B. (1962) *J. Biol. Chem. 237*, 1856.
Hirs, C. H. W. (1967) *Methods Enzymol. 11*, 411.
Hopp, T. P. (1976) *Anal. Biochem. 74*, 638.
Huber, R., Deisenhofer, J., Colman, P. M., Matsushima, M., & Palm, W. (1976) *Nature (London) 264*, 415.
MacGillivary, R. T. A., Barnes, K., & Brew, K. (1979) *Arch. Biochem. Biophys.* (in press).
Niedermeier, W. (1971) *Anal. Biochem. 40*, 465.
Porter, W. H. (1975) *Anal. Biochem. 63*, 27.
Schindler, M., Sharon, N., & Prieels, J. P. (1976) *Biochem. Biophys. Res. Commun. 69*, 167.
Schroeder, W. A., Shelton, J. B., & Shelton, J. R. (1969) *Arch. Biochem. Biophys. 130*, 551.
Segrest, J. P., & Jackson, R. L. (1972) *Methods Enzymol. 28*, 54.
Spiro, R. G. (1966) *Methods Enzymol. 8*, 3.
Spiro, R. G. (1973) *Adv. Protein Chem. 27*, 349.
Summers, M. R., Smythers, G. W., & Oroszlan, S. (1973) *Anal. Biochem. 53*, 624.
Warme, P. K., Momany, F. A., Rumball, S. V., Tuttle, R. W., & Scheraga, H. A. (1974) *Biochemistry 13*, 768.
Woods, K. R., & Wang, K. T. (1966) *Biochim. Biophys. Acta 133*, 369.

# Increased Thermal Stability of Proteins in the Presence of Sugars and Polyols†

Joan F. Back, David Oakenfull, and Malcolm B. Smith*

ABSTRACT: Sugars and polyols stabilize proteins against heat denaturation. Scanning calorimetry was used to obtain a quantitative estimate of the degree of stabilization. Solutions of ovalbumin, lysozyme, conalbumin, and α-chymotrypsinogen were heated at a constant rate, and the temperature of the maximum rate of denaturation was estimated ($T_m$). Addition of a sugar or polyol raised $T_m$. The magnitude of the stabilizing effect ($\Delta T_m$) depended on both the nature of the protein and the nature of the sugar or polyol, ranging from 18.5 °C for lysozyme at pH 3 in the presence of 50% (w/w) sorbitol to 0 °C for conalbumin at pH 7 in 50% glycerol solution. It is argued that this stabilization is due to the effects of sugars and polyols on hydrophobic interactions. The strength of the hydrophobic interaction was measured in model systems in sucrose and glycerol solutions. Sucrose and glycerol strengthened the pairwise hydrophobic interaction between hydrophobic groups; however, they reduced the tendency for *complete* transfer of hydrophobic groups from an aqueous to a nonpolar environment. The extent of stabilization by different sugars and polyols is explained by their different influences on the structure of water. The difference between the partial molar volume of the sugar or polyol and its van der Waals volume was used as a rough quantitative measure of the structure-making or structure-breaking effect. There was a linear relationship between this quantity and $\Delta T_m$.

It has been known for many years that sugars may protect proteins against loss of solubility during drying and may inhibit heat coagulation (Ball et al., 1943). Simpson & Kauzmann (1953) observed that the extent of denaturation of ovalbumin in urea solutions was reduced in the presence of sucrose. More recently, Gerlsma & Stuur (1972) showed that polyhydric alcohols raised the thermal transition temperatures of lysozyme and ribonuclease and Donovan (1977) observed the stabilizing effect of sucrose on the proteins of egg white.

The technique of scanning calorimetry has been increasingly applied over the last 10 years in studies of thermal transitions in proteins [reviewed by Privalov (1947)] and in investigations of stability changes of proteins caused by specific interaction with metal ions (Donovan & Ross, 1975), by complex formation (Donovan & Ross, 1973; Donovan & Beardslee, 1975), and by change in conformation (Donovan & Mapes, 1976). We have used scanning calorimetry in this investigation to measure the increase in denaturation temperature of several proteins in the presence of various sugars and polyols—an effect apparently related to changes in solvent properties and of opposite sign to the destabilization caused by the addition of monohydric alcohols or urea to protein solutions.

Lakshmi & Nandi (1976) showed that sucrose and glucose decrease the solubility of phenylalanine, tyrosine, and tryptophan in aqueous solution and suggested that this was owing to increased hydrophobic interaction. A similar stabilization of hydrophobic interactions was postulated for proteins in sugar solutions. We have, therefore, also investigated the effects of sucrose and glycerol on hydrophobic interactions in two model systems—ion pair formation by long-chain alkyltrimethylammonium carboxylates (Oakenfull & Fenwick, 1977) and the critical micelle concentrations of alkyltrimethylammonium bromides (Emerson & Holtzer, 1967). These model systems distinguish between two types of hydrophobic interaction which are influenced in different ways by added solutes (Oakenfull & Fenwick, 1979)—pairwise interaction in which hydrophobic groups associate but otherwise remain surrounded by water and complete transfer of a hydrophobic group from an aqueous to a nonpolar environment.

Our results, presented here, indicate a general phenomenon which might be used to throw more light on the relation between solvent structure and the stability of proteins and suggest that reactivity and turnover of proteins in vivo may often be

† From the CSIRO Division of Food Research, Food Research Laboratory, North Ryde, N.S.W. 2113, Australia. *Received April 9, 1979.*

0006-2960/79/0418-5191$01.00/0 