

2501

Analysis of Proteolytic Digests of Genetic Variants of α_{s1} -Casein

EDWIN B. KALAN, RAE GREENBERG, AND MARVIN P. THOMPSON

Eastern Regional Research Laboratory¹, Philadelphia, Pennsylvania 19118

The genetic variants, A, B, and C, of α_{s1} -casein from cow's milk were digested under controlled conditions with trypsin, chymotrypsin, and pepsin. The resulting peptides were examined by a mapping procedure and the patterns were compared within each set of digests. The B and C variants are known to differ in composition by a single amino acid substitution—a glutamic acid residue in B being replaced by a glycine in C. Probable difference peptides containing this substitution were observed in each set of proteolytic digestions. The results of specific staining for arginine, histidine, methionine, tyrosine, and tryptophan residues suggest that the presumed difference peptides are related and indicate the presence of several of these amino acids in the vicinity of the amino acid substitution.

The α_{s1} -casein A variant, which differs from the other two proteins quite substantially in amino acid composition, revealed peptide patterns almost identical to those of B and C, with the main difference being the absence of one or two major spots in A. The hypothesis is proposed that α_{s1} -casein A, which represents an unusual type of genetic variant, is devoid of a portion of the molecule, with the remaining sequence similar to the B protein.

The term " α_{s1} -casein" (23) refers to those components of the α -casein complex of cow's milk that are precipitated by calcium and stabilized by κ -casein against precipitation by calcium. This protein has been found to occur in three genetic forms designated A, B, and C (22). The polymorphism has been extensively studied and reported by Kiddy *et al.* (15); Thompson and Kiddy (20); and Thompson *et al.* (23). The α_{s1} -A variant is evidently very rare and so far has been found only in a single line of Holstein cattle (15). The α_{s1} -B and -C proteins have been found in all breeds of cattle examined. The mode of inheritance of the α_{s1} -casein polymorphism is controlled by three allelic autosomal genes with no dominance (2); that is, the phenotype of a protein synthesized by a heterozygous animal is represented by the occurrence of two α_{s1} -casein variants in the milk.

¹Eastern Utilization Research and Development Division, Agricultural Research Service, U.S. Department of Agriculture.

End-group analyses revealed identical amino-terminal residues, namely, arginine, and identical carboxyl-terminal sequences, -leucyl-tryptophan, for all three variants (14). Molecular weights of 26,000 for A and 31,000 for B and C calculated from the amount of C-terminal tryptophan released by carboxypeptidase A (14), are in the same range as the value of 27,300 from light scattering reported by Dreizen *et al.* (4) for $\alpha_{s1,2}$ -caseins (presumed to be a mixture of two α_{s1} -variants), and 28,000 for A and 28,600 for B and C recently reported by Gordon *et al.* (6) from amino acid composition studies. The latter investigation, carried out with highly purified, twice-chromatographed α_{s1} -caseins, revealed that B and C are virtually identical in composition and differ only by what appears to be a glutamic acid/glycine substitution, but A is distinguished by numerous and quite substantial amino acid differences in comparison to the other variants. The glutamic acid/glycine difference between B and C has also

been reported by de Koning and van Rooijen (3).

Enzymic digestions were performed under controlled conditions with several proteolytic enzymes. The peptides produced thereby were examined by a peptide mapping technique, and the patterns obtained from the individual variants were compared. This paper reports the results of this further characterization of the α_{s1} -caseins A, B, and C and the probable location of difference peptides containing the aforementioned amino acid substitution between α_{s1} -caseins B and C.

MATERIALS AND METHODS

α_{s1} -Caseins. The genetic α_{s1} -casein variants A, B, and C were prepared from milks of individual cows typed homozygous for each protein. The isolation and purification procedures have been described in detail previously (20). The purified materials were examined by polyacrylamide-gel electrophoresis at pH 9.1–9.3 in Tris- Na_2EDTA -borate buffer in 4.5 M urea; a single band was obtained for each protein.

Proteolytic Digestions

Tryptic and chymotryptic. All enzymic digestions were carried out under nitrogen in a Radiometer² recording pH-stat equipped with a jacketed reaction vessel maintained at 37°. For each digestion, 100 mg of protein was dissolved in approximately 10 ml of water and the solution was adjusted to pH 8.2. After transfer to the reaction vessel, 0.25 ml of a solution of enzyme (trypsin or chymotrypsin, Worthington Biochemical Corporation;² crystallized and lyophilized; dissolved in water at 4 mg per milliliter) was added. The pH of 8.2 was maintained by the automatic addition of 0.1 N NaOH. At the end of 3 hours for tryptic digestion and 4 hours for chymotryptic digestion, another 0.25 ml of enzyme solution was added and the reaction was allowed to continue one additional hour. Immediately thereafter the digests were frozen and lyophilized.

Peptic. Samples of each α_{s1} -casein (100 mg) were suspended in about 10 ml of water and adjusted to pH 2.2. After several hours of mechanical stirring, clear solutions were obtained and transferred to the reaction vessel of the pH-stat. Digestion was begun with the addition of 0.25 ml of a solution of pepsin (Armour,² crystalline; 4 mg per

milliliter) and the desired pH was maintained by the automatic addition of 0.5 N HCl. Again, after 3 hours another 0.25 ml aliquot of enzyme solution was added. The reaction was permitted to continue for an additional hour and was terminated by freezing; the sample was lyophilized.

Peptide patterns. Peptide mapping was achieved by a two-dimensional combination of high-voltage paper electrophoresis and chromatography with 2 mg samples of each α_{s1} -casein digest. Electrophoresis was carried out on Whatman No 3 MM paper for 2 hours at 40 V per centimeter in a horizontal water-cooled plate apparatus; pyridine-acetic acid-water buffer (pH 6.4) was used; the procedure was that of Ingram (11). After air drying, the papers were subjected to ascending chromatography in the second dimension with the solvent system *n*-butanol-acetic acid-water (3:1:1). Peptides were detected by dipping the dried papers into a solution of 0.2% ninhydrin in acetone, drying again, and heating 2–3 minutes in an oven at 110°.

Specific staining. Replicate sheets were treated with specific color reagents to locate peptides containing arginine, histidine, methionine, tryptophan, and tyrosine. In some cases the peptides were first detected with ninhydrin, their locations marked, and the ninhydrin color subsequently decolorized with 1.0 N HCl-acetone (1:3). Several specific tests could then be successfully applied in sequence to these patterns. Arginine peptides were located with the modified Sakaguchi reagent described by Fraenkel-Conrat *et al.* (5); histidine peptides with the Pauly diazotized sulfanilic acid reagent; methionine with the chloroplatinic reducing sulfur stain of Toennies and Kolb (24); tryptophan with the Ehrlich reagent; and tyrosine with acidic α -nitroso- β -naphthol.

RESULTS

Preteolytic digestions of α_{s1} -caseins. The uptake curves for tryptic and chymotryptic digestion of the α_{s1} -casein variants were quite normal. The initial rates and total base uptake (0.5 ml) were essentially the same for the three proteins upon tryptic digestion. This is in qualitative agreement with the identical combined arginine and lysine content of 24 residues for each variant, as shown in Table I, which includes a partial listing of the amino acid composition presented by Gordon *et al.* (6). The digestion proceeded rapidly and appeared to be complete in less than 30 minutes with no increase in base uptake after the addition of the second aliquot of enzyme.

² It is not implied that the U.S. Department of Agriculture recommends the above company to the exclusion of others in the same business.

TABLE I
AMINO ACID COMPOSITION OF α_{s1} -CASEINS^a
Residues amino acid per mole protein.

Amino Acid	A	B	C
Aspartic acid	16.8	18.1	18.2
Threonine	6.7	6.0	6.1
Serine	17.8	17.3	17.6
Glutamic acid	46.6	46.4	45.5
Proline	20.6	20.3	20.4
Glycine	10.7	10.7	11.8
Alanine	9.9	10.8	10.8
Valine	11.9	13.4	13.6
Methionine	5.9	5.7	5.7
Isoleucine	13.6	13.1	13.3
Leucine	17.3	20.3	20.5
Tyrosine	12.1	11.6	11.7
Phenylalanine	7.5	9.6	9.7
Lysine	18.1	17.0	17.0
Histidine	6.2	6.1	6.1
Arginine	6.2	7.2	7.2
Tryptophan	2.8	2.7	2.8

^a Table taken from Gordon *et al.* (6).

For chymotryptic digestion, the observed decreased initial base uptake of A as compared with the other two variants correlates with the difference in phenylalanine and leucine content of the proteins. In this case, upon addition of more enzyme, a slight increase in the rate of digestion was noted. The total base uptake was 0.65 ml for B, 0.62 ml for C, and 0.52 ml for the A protein.

Peptic digestion produced virtually identical acid uptake curves (total uptake, 0.16 ml) for all three α_{s1} -caseins.

Peptide patterns and specific staining. In addition to the peptide maps (Figs. 1-3), the top panel of each figure shows the electrophoresis patterns before ascending chromatography. All the peptide maps could be consistently reproduced, as attested to by the replicate runs used for specific staining. On overall examination of the figures it is evident that within each set of digests the maps for the three α_{s1} -variants are strikingly similar. This is especially true for the B and C variants where it is to be expected, but occurs even with α_{s1} -A the composition differences of which from the others might suggest greater peptide variation.

A comparison of the B and C tryptic digest maps (Fig. 1) shows only a single pair of

difference peptides. Peptide 19, present in B, is absent in C and presumably replaced there by 23. If this pair is assumed to contain the glutamic acid/glycine substitution, the decreased anodic electrophoretic migration of 23 may be explained by the replacement of a neutral residue for an acidic group. The occurrence of tyrosine in both these peptides would also support this conjecture. Peptide 19 is also present in the map of α_{s1} -A; this protein possesses the same number of glycine and glutamic acid residues as the B variant. It is possible that the sequences in A and B are identical in the area of the amino acid substitution between B and C. The α_{s1} -A pattern is similar to the others with the addition of peptide 24 and the absence of 22. As seen from Table II, further differences may be recognized by the specific staining techniques. Peptide composition differences, especially with respect to uncharged residues, are possible despite the fact that the spots occupy seemingly identical locations. Examples of this are peptide 6, which gives tests for arginine and methionine in B and C and only arginine in A, and peptide 12-A, which contains histidine and methionine while the 12 region in the other variants gives no positive test. In some cases where the ninhydrin spots were very light, specific staining gave results too weak to indicate the definite presence or absence of the amino acid in that area. These cases were also recorded as test negative, e.g., tryptic 21-A and peptic 5-C. It would seem likely that peptide 22, absent in the A variant but present in B and C, represents the additional arginine residue found by amino acid analysis (Table I). Ideally, 25 peptides are expected from tryptic digestion of the α_{s1} -caseins while 22 were actually observed.

The chymotryptic patterns (Fig. 2) present a similar picture. All comments made concerning the tryptic difference pair are valid for the chymotryptic peptides 24 and 28 which occur in the α_{s1} -casein B and C digests, respectively. In this case, both peptides also contain histidine, methionine and tryptophan. Peptide 29, which appears only in C and gives positive tests for methionine and tryptophan (Table III) has no counterpart in B and is not compensated for by

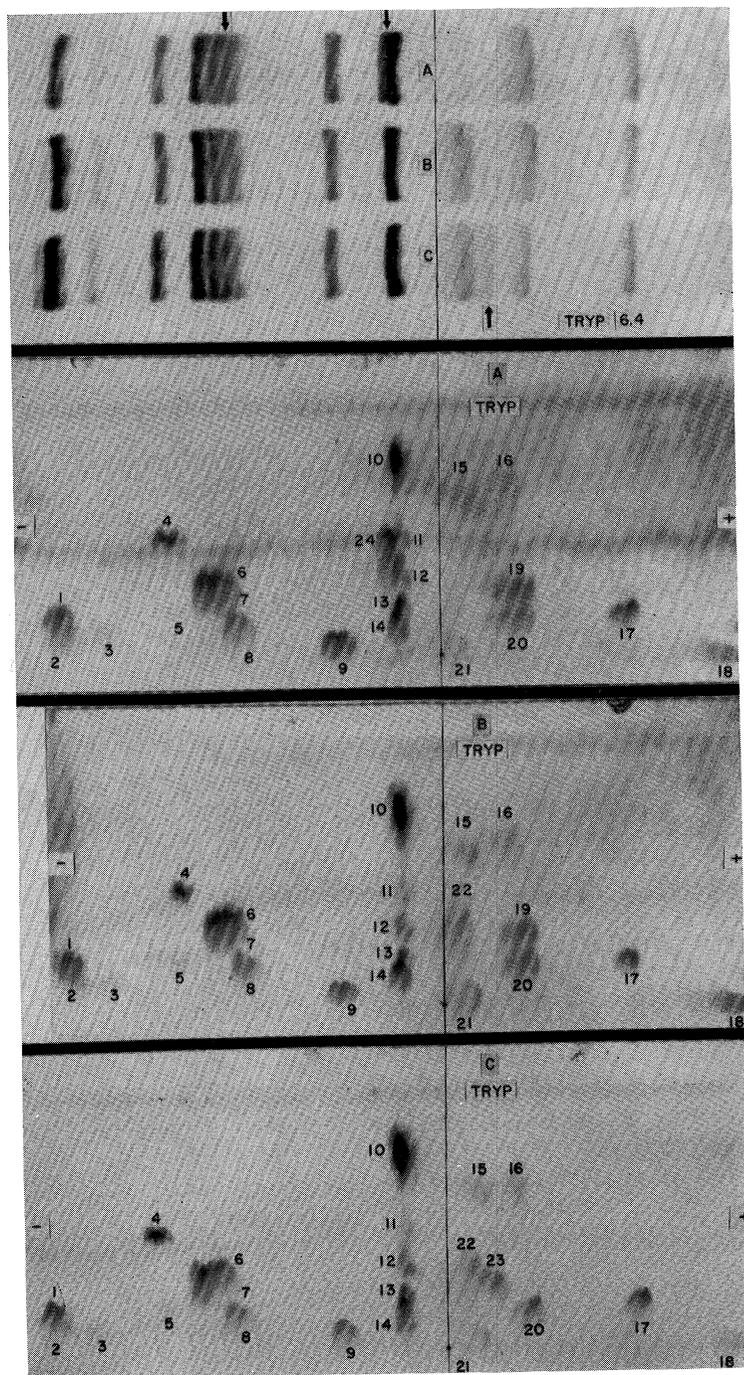


FIG. 1. Peptide patterns from tryptic digest of α_{s1} -caseins A, B, and C. The top panel shows one-dimensional high-voltage paper electrophoresis at pH 6.4, 40 V/cm, 2 hours, with arrows marking differences. The lower three panels are maps produced by two-dimensional electrophoresis and chromatography. See text for experimental conditions. Peptides located in equivalent positions are given the same number.

TABLE II
SPECIFIC STAINING OF PEPTIDES FROM A TRYPTIC DIGEST OF α_{s1} -CASEINS A, B, AND C

Peptide No. ^a	A	B	C
1	arg	arg	arg
6	arg	arg, met	arg, met
7	his	his	his
8	his	his	his
9	his	his	his
10	arg, his, met, try, tyr	arg, his, met, try, tyr	arg, his, met, try, tyr
11	try	try	try
12	his, met	0 ^b	0
15	try, tyr	try, tyr	try, tyr
16	arg, met, tyr	arg, met, tyr	arg, met, tyr
17	arg	arg	arg
18	met	met	met
19	tyr	tyr	—
20	arg, met, tyr	arg, met, try, tyr	arg, met, try, tyr
21	0	arg, his	arg, his
22	—	arg, his, met	arg, his, met
23	—	—	tyr

^a From Fig. 1.

^b 0 indicates test negative.

missing residues in any other spot. Therefore, it is considered the product of an incomplete split. The map from A shows the absence of peptides 4, 8, and 26, with the appearance of a new spot, 27. The specific color tests (Table III) show little variation among equivalent peptides. From this Table it should also be noted that peptide 8, present in B and C and absent in A, contains an arginine residue. The total numbers of peptides obtained by chymotryptic digestion are 25 for A, 26 for B, and 27 for C. The greater number of peptides found in the maps of the chymotryptic digests, in contrast to the tryptic digests, can be ascribed to the lesser specificity of chymotrypsin.

Pepsin is the least specific of the proteolytic enzymes employed but the digestions are equally consistent (Fig. 3). Peptides 30 in B and 32 in C appear to be the difference peptides. Their relationship to each other in regard to electrophoretic mobility is the same as that of the difference pairs existing in the other digests. In addition, both peptides contain methionine. Aside from the additional peptide 33, present in the B variant only, the remainder of the B and C digest patterns are identical. However, peptides 5 and 15 in B give positive specific tests

for arginine and tryptophan, respectively, while the corresponding regions in C are negative. The α_{s1} -A map is lacking spots 4 and 24 but is otherwise in agreement with the contiguous patterns. The dark spot above the “+” sign is an artifact. Peptides detected total 29, 32, and 31, respectively, for α_{s1} -caseins A, B, and C. As deduced from Table IV, where nine or ten peptides appear to contain arginine, as opposed to a theoretical content of six or seven residues, peptic digestion produces many overlapping peptides. However, it is evident that the locations of bonds split are the same for the three variants and that the digestion is reproducible.

DISCUSSION

Genetic polymorphism has been observed in all the major caseins, α_{s1} -, β -, and κ -, of cow's milk. In the case of κ -casein (17, 18, 26), two variants have been found, designated A and B, which differ in that the A variant contains an excess of one residue of aspartic acid and threonine and is lacking one residue of alanine and isoleucine as compared to the B protein (27). The α_{s1} - and β -caseins are each known to exist as at least three genetic variants, the occurrence

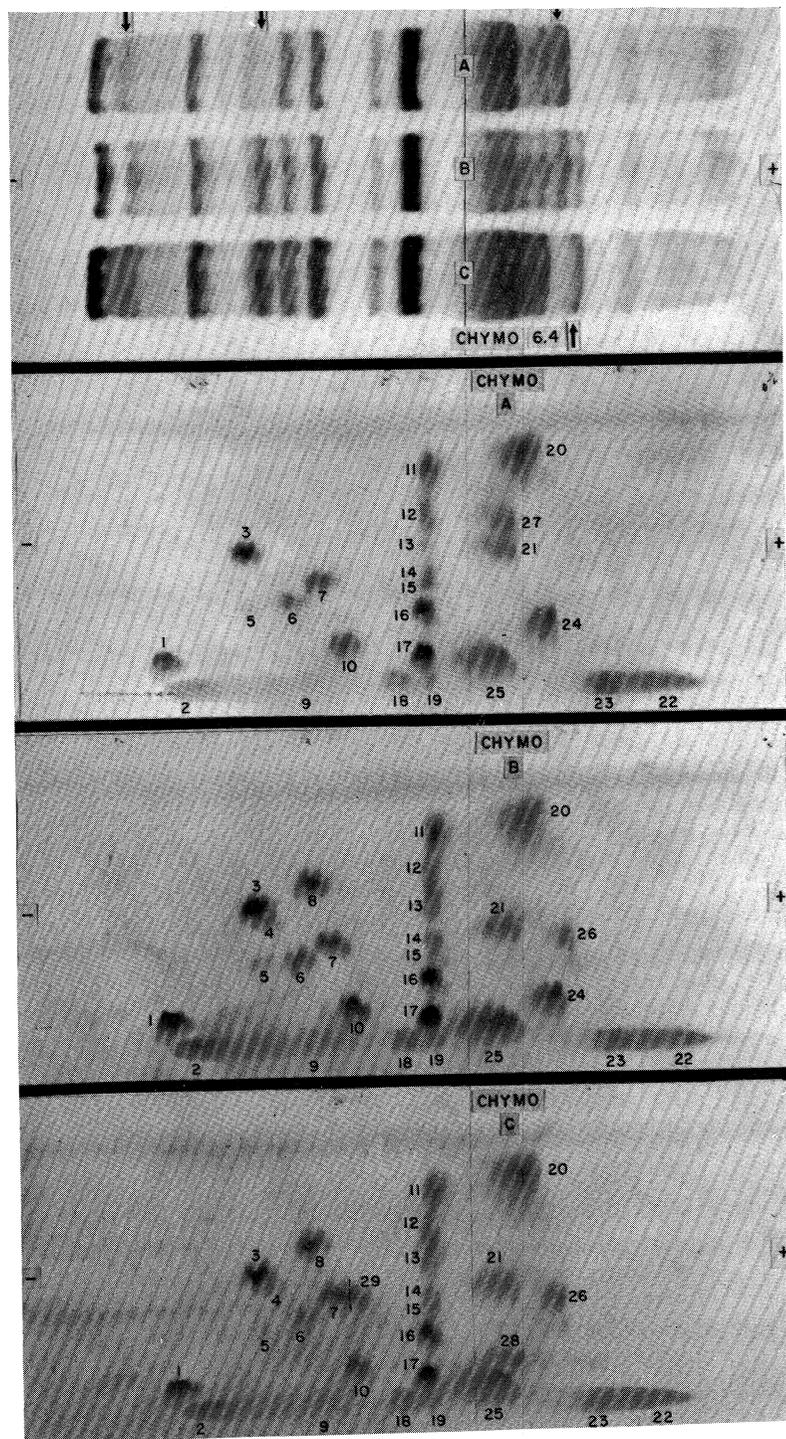


FIG. 2. Peptide patterns from chymotryptic digests of α_{s1} -caseins A, B, and C. Conditions of electrophoresis, chromatography, and peptide numbering are the same as in Fig. 1.

TABLE III
SPECIFIC STAINING OF PEPTIDES IN A CHYMOTRYPTIC DIGEST OF α_{s1} -CASEINS A, B, AND C

Peptide No. ^a	A	B	C
1	his, tyr	his, tyr	his, tyr
2	arg, his	arg, his	arg, his
3	arg	arg	arg
6	arg	arg	arg
7	arg, his, tyr	arg, tyr	arg, tyr
8	—	arg	arg
10	his, met	his, met	his, met
11	his, try, tyr	his, try, tyr	his, try, tyr
12	arg, tyr	tyr	tyr
13	try, tyr	try, tyr	try, tyr
16	tyr	tyr	tyr
17	met	met	met
18	met	met	met
20	his, tyr	his, met, tyr	his, met, tyr
21	his, try, tyr	his, try, tyr	his, try, tyr
22	his, met, tyr	his, met, tyr	his, met, tyr
23	met	met	met
24	his, met, try	his, met, try	—
25	arg, his, tyr	arg, his, tyr	arg, his, tyr
28	—	—	his, met, try
29	—	—	met, try

^a From Fig. 2.

of which is breed specific (1, 15, 21). In addition, it has recently been shown that these two caseins do not occur independently (7, 16). Amino acid compositional data show that the β -casein as well as the α_{s1} -casein variants differ significantly (19). In the α_{s1} -caseins, as discussed earlier, the A variant possesses the greatest number of differences, while B and C differ by only a single amino acid pair.

This amino acid substitution has now been probably located in a pair of difference peptides in each of three sets of proteolytic digests of the B and C variants. In the case of tryptic, chymotryptic, and peptic digests, the B protein gives rise to a peptide which is acidic at pH 6.4 and absent in patterns obtained from the C protein. Instead, a less acidic peptide is present in C patterns but absent in the B maps. This difference in mobility can be explained on the basis of an excess of glutamic acid in B, and the presence of a glycine residue in its place in C. The aspartic acid/glycine peptide difference pair in the β -lactoglobulins exhibits similar electrophoretic behavior at pH 6.4 (13). The identical results in the specific staining

reactions given by the α_{s1} -casein B and C difference peptides (Tables II-IV) suggest that they are indeed related and indicate the presence of tyrosine, methionine, histidine, and tryptophan in the vicinity of the amino acid substitution. Lysine rather than arginine is presumably also located in this vicinity. Final proof must await isolation and compositional analysis of the peptides in question. The glutamic acid/glycine substitution has been found to occur in hemoglobin (10), tobacco mosaic virus coat protein (25), and the reverse substitution in tryptophan synthetase (8). It is represented by the triplets GAA/GGA of the genetic code (12, 28).

Gordon *et al.* (6), in their discussion of the relationship of α_{s1} -casein A to the other polymorphs, have suggested from consideration of the numerous amino acid differences that the A variant is not a mutant of B or C or both in the sense that the term is usually applied. A system analogous to the β -lactoglobulin case, for which no genetic mechanism has been proposed, where the A and B variants differ by two amino acids, and not just one, probably also exists in the

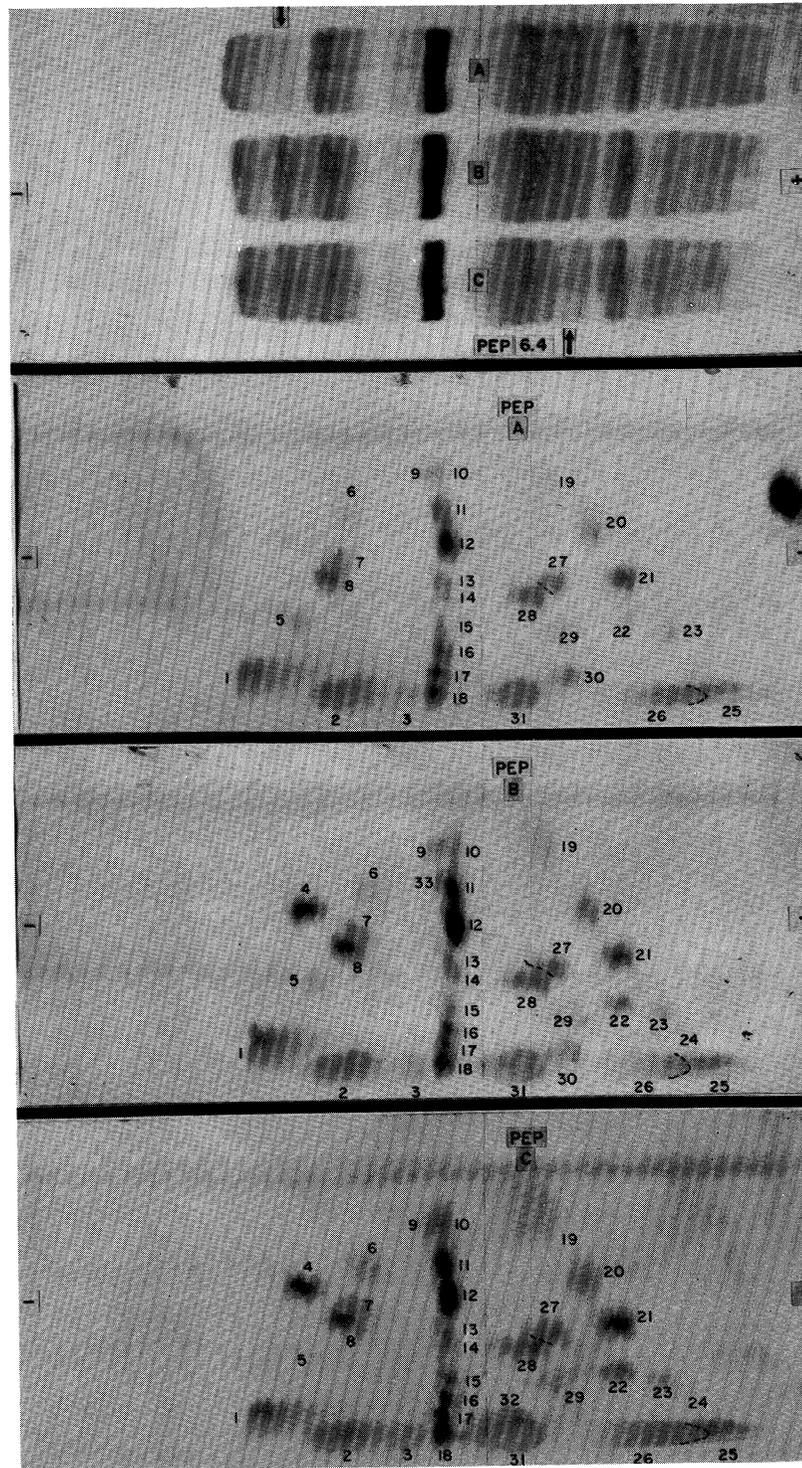


FIG. 3. Peptide patterns from peptic digests of α_{s1} -caseins A, B, and C. Conditions of electrophoresis, chromatography, and peptide numbering are the same as in Fig. 1.

TABLE IV
SPECIFIC STAINING OF PEPTIDES IN A PEPTIC DIGEST OF α_{s1} -CASEINS A, B, AND C

Peptide No. ^a	A	B	C
1	arg, tyr	arg, tyr	arg, tyr
2	arg, his	arg, his	arg, his
4	—	arg	arg
5	arg	arg	0 ^b
6	arg, tyr	arg, tyr	arg, tyr
7	arg, his, tyr	arg, his, tyr	arg, his, tyr
8	arg	arg	arg
9	try	try	try
11	tyr	tyr	tyr
12	tyr	tyr	tyr
13	try	try	try
15	try	try	0
18	arg, his, met	arg, his, met	arg, his, met
20	tyr	tyr	tyr
25	his	his	his
26	arg, met	arg, met	arg, met
27	tyr	tyr	tyr
28	try, tyr	try, tyr	try, tyr
29	tyr	tyr	tyr
30	met	met	—
31	arg, his, met	arg, his, met	arg, his, met
32	—	—	met

^a From Fig. 3.

^b 0 indicates test negative.

case of the α_{s1} -variants. Yet, the fact must be considered that all three α_{s1} -caseins have many similar chemical and physical properties (9, 20).

If the many amino acid deficiencies occurring in A were distributed throughout the molecule, the peptide maps produced might be expected to differ vastly from those of the other variants. As reported here, however, this is not the case. All the α_{s1} -casein peptide maps within each set of digests are quite similar, with the outstanding characteristic of the A maps being the absence of one or two major spots. These results suggest the hypothesis that the A variant is devoid of one portion or perhaps several portions of the molecule, with the remaining sequence similar to the B protein. Identity of end-groups among the three α_{s1} -caseins suggest that the segment(s) is missing from the internal sequence of the protein. However, this does not exclude the possibility that the missing portion is equivalent to the N-terminal peptide of B or C while still another arginine residue occupies the amino-terminal position in the A protein. This must be con-

sidered, since there is one less residue of arginine in α_{s1} -A compared with B or C.

The use of preparative scale column chromatography and related techniques should provide sufficient material for identification and study of the key peptides involved in both the α_{s1} -A hypothesis and the amino acid substitution between B and C. In conclusion, if the interpretation of the localization of the amino acid differences in α_{s1} -A proves correct, this protein provides an example of a type of mutation involving not a one or two amino acid substitution, but the deletion from the molecule of an intact fragment without causing gross changes in the physical and chemical properties.

REFERENCES

1. ASCHAFFENBURG, R., *J. Dairy Res.* **30**, 251 (1963).
2. ASCHAFFENBURG, R., *J. Dairy Sci.* **48**, 128 (1965).
3. DE KONING, P. J., AND VAN ROOIJEN, P. J., *Biochem. Biophys. Res. Commun.* **20**, 241 (1965).
4. DREIZEN, P., NOBLE, R. W., AND WAUGH, D. F., *J. Am. Chem. Soc.* **84**, 4938 (1962).

5. FRAENKEL-CONRAT, H., HARRIS, J. I., AND LEVY, A. L., *Methods Biochem. Anal.* **2**, 359 (1955).
6. GORDON, W. G., BASCH, J. J., AND THOMPSON, M. P., *J. Dairy Sci.* **48**, 1010 (1965).
7. GROSCLAUDE, F., GARNIER, J., RIBADEAU-DUMAS, B., AND JEUNET, R., *Compt. Rend. Acad. Sci.* **259**, 1569 (1964).
8. HENNING, U., AND YANOFSKY, C., *Proc. Natl. Acad. Sci. U.S.* **48**, 1497 (1962).
9. HERSKOVITS, T. T., *Abstracts of 146th Meeting Am. Chem. Soc.* p. 22A (1964).
10. HILL, R. L., AND SCHWARTZ, H. C., *Nature* **184**, 641 (1959).
11. INGRAM, V. M., *Biochim Biophys. Acta* **28**, 539 (1958).
12. JUKES, T. H., *Biochem. Biophys. Res. Commun.* **19**, 391 (1965).
13. KALAN, E. B., GORDON, W. G., BASCH, J. J., AND TOWNEND, R., *Arch. Biochem. Biophys.* **96**, 376 (1962).
14. KALAN, E. B., THOMPSON, M. P., AND GREENBERG, R., *Arch. Biochem. Biophys.* **107**, 521 (1964).
15. KIDDY, C. A., JOHNSTON, J. O., AND THOMPSON, M. P., *J. Dairy Sci.* **47**, 147 (1964).
16. KING, J. W. B., ASCHAFFENBURG, R., KIDDY, C. A., AND THOMPSON, M. P., *Nature* **206**, 324 (1965).
17. NEELIN, J. M., *J. Dairy Sci.* **47**, 506 (1964).
18. SCHMIDT, D. G., *Biochim. Biophys. Acta* **90**, 411 (1964).
19. THOMPSON, M. P., AND GORDON, W. G., *J. Dairy Sci.* **47**, 671 (1964).
20. THOMPSON, M. P., AND KIDDY, C. A., *J. Dairy Sci.* **47**, 626 (1964).
21. THOMPSON, M. P., KIDDY, C. A., JOHNSTON, J. O., AND WEINBERG, R. M., *J. Dairy Sci.* **47**, 378 (1964).
22. THOMPSON, M. P., KIDDY, C. A., PEPPER, L., AND ZITTLE, C. A., *Nature* **195**, 1001 (1962).
23. THOMPSON, M. P., TARASSUK, N. P., JENNESS, R., LILLEVIK, H. A., ASHWORTH, U. S., AND ROSE, D., *J. Dairy Sci.* **48**, 159 (1965).
24. TOENNIES, G., AND KOLB, J. J., *Anal. Chem.* **23**, 823 (1951).
25. WITTMANN, H. G., AND WITTMANN-LIEBOLD, B., *Cold Spring Harbor Symp. Quant. Biol.* **28**, 589 (1963).
26. WOYCHIK, J. H., *Biochem. Biophys. Res. Commun.* **16**, 267 (1964).
27. WOYCHIK, J. H., KALAN, E. B., AND NOELKEN, M. E., *Biochemistry* (in press.)
28. YANOFSKY, C., *Cold Spring Harbor Symp. Quant. Biol.* **28**, 581 (1963).