Kowalski, D., Leary, T. R., McKee, R. E., Sealock, R. W., Wang, D., and Laskowski, M., Jr. (1974), in Bayer-Symposium V, Berlin, Springer-Verlag, p 311.
Lawson, W. B., Gross, E., Foltz, C. M., and Witkop, B. (1962), *J. Am. Chem. Soc. 84,* 1715-1718.
Lin, L. J., and Foster, J. F. (1975), *Anal. Biochem. 63,* 485-490.
Link, T. P., and Stark, G. R. (1968), *J. Biol. Chem. 243,* 1082-1088.
Naider, F., and Bohak, Z. (1972), *Biochemistry 11,* 3208-3211.
Neumann, N. P., Moore, S., and Stein, W. H. (1962), *Biochemistry 1,* 68-75.
Odani, S., and Ikanaka, T. (1973), *J. Biochem. (Tokyo) 74,* 857-860.
Rubinstein, M., Shechter, Y., and Patchornik, A. (1976), *Biochem. Biophys. Res. Commun. 70,* 1257-1263.
Schultz, J. (1967), *Methods Enzymol. 11,* 255-263.
Schwert, G. W., and Takenaka, Y. (1955), *Biochim. Biophys. Acta 16,* 570-575.
Shechter, Y., Burstein, Y., and Patchornik, A. (1972), *Biochemistry 11,* 653-660.
Shechter, Y., Burstein, Y., and Patchornik, A. (1975), *Biochemistry 14,* 4497-4503.
Shechter, Y., Patchornik, A., and Burstein, Y. (1973), *Biochemistry 12,* 3407-3413.
Sluyterman, L. A. E., and Wijdenes, J. (1974), *Methods Enzymol. 34,* 544-547.
Spande, T. F., Witkop, B., Degani, Y., and Patchornik, A. (1970), *Adv. Protein Chem. 24,* 97-260.
Stark, G. R., and Stein, W. H. (1964), *J. Biol. Chem. 239,* 3755-3761.
Tang, J., and Hartley, B. S. (1967), *Biochem. J. 102,* 593-599.
Tsai, H. J., and Williams, G. R. (1965), *Can. J. Biochem. 43,* 1409-1415.

# Primary Structure of Rat Lysozyme†

Thomas J. White,* George A. Mross, Elliott F. Osserman, and Allan C. Wilson

ABSTRACT: For evolutionary reasons, we determined the primary structure of rat lysozyme. The chymotryptic peptides from the reduced and carboxymethylated protein were sequenced and aligned by homology with the sequence of human lysozyme. Overlaps were confirmed by partial structures of tryptic peptides and an automatic sequencer run on the whole protein. By comparing this lysozyme sequence with those of human and baboon and taking into account paleontological estimates of the times of divergence of these species from one another, an approximate estimate of the average rate of lysozyme evolution was made. This rate is not significantly different from the average rate of lactalbumin evolution in mammals—a finding which is at variance with Dickerson's [Dickerson, R. E. (1971), *J. Mol. Evol. 1,* 26] and Dayhoff's [Dayhoff, M. O., Ed. (1972), Atlas of Protein Structure and Sequence, Vol. 5, Silver Spring, Md., The National Biomedical Research Foundation] conclusion that lactalbumin evolution has been faster than lysozyme evolution. Our finding raises the possibility that the gene duplication event responsible for the origin of lactalbumin from lysozyme was more ancient than is generally supposed. Furthermore, from comparison of the rates of lysozyme evolution in rodents and primates, it is suggested that generation time is not a key factor in lysozyme evolution.

To understand the mechanism of evolution, it is essential to study the rates at which evolutionary change has taken place in the sequences of macromolecules. It is important from this point of view to examine cases in which a protein has undergone a radical evolutionary change in biological function. Does the amino acid sequence evolve much faster than normal under these conditions and, if so, by how much? By focusing on this question, one may gain a better understanding of the driving force for sequence change in proteins. Lysozyme, a bacteriolytic enzyme, and lactalbumin, a regulator of lactose synthesis, provide an opportunity to study this problem. Sequence evidence indicates that the lactalbumin gene arose by duplication of the gene for lysozyme (Hill and Brew, 1975). The claim that the lactalbumin gene arose at the outset of mammalian evolution and subsequently experienced accelerated sequence evolution has been widely accepted (Dickerson, 1971; Hood et al., 1975; Florkin, 1975). However, an alternative model, consistent with similar rates of change in the two proteins since their divergence, suggests that the duplication event may be much more ancient.

Support for the recent divergence model and a speed-up in lactalbumin evolution originally came from the observation that the rate of sequence change in mammalian lactalbumins was two or three times that of bird lysozymes (Dayhoff, 1972). However, later immunological work indicated the rate of lysozyme evolution in mammals was more rapid than in birds (Hanke et al., 1973). To check on this point we decided to sequence rat lysozyme. By comparing the rat sequence with that of other mammalian lysozymes and considering the probable times of divergence involved, we estimated that the average rate of lysozyme evolution among mammals is comparable to

† From the Department of Biochemistry, University of California, Berkeley, California 94720 (T.J.W., G.A.M., and A.C.W.), and the Institute of Cancer Research, Columbia University, New York, New York 10032 (E.F.O.). *Received July 16, 1976.* This work was supported in part by Grant GM 21509 from the National Institutes of Health and Grant CA02332 from the National Cancer Institute. A preliminary report of this work has been presented at the 66th Annual Meeting of the American Society of Biological Chemists, Atlantic City, April 1975, and at the Pacific Slope Biochemical Conference, Honolulu, June 1975.

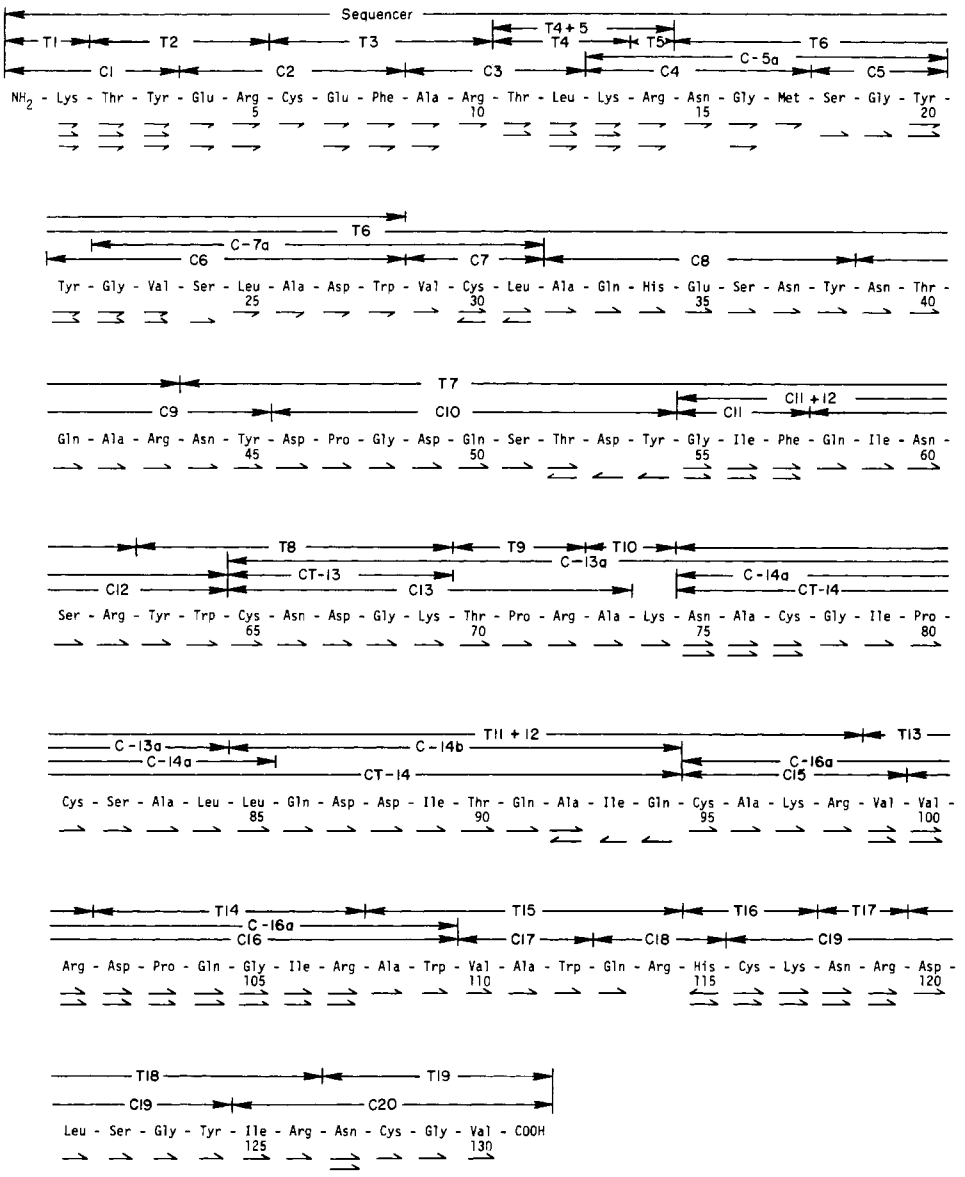


FIGURE 1: The amino acid sequence of rat lysozyme. The various peptides are indicated by double-headed arrows. Residues identified by manual Edman degradation (—), sequencer analysis and solid-phase techniques (—), and carboxypeptidase A hydrolysis (—) are so indicated. Abbreviations used are: T, tryptic peptide; C, chymotryptic peptide; CT, peptide isolated from a tryptic digest of a chymotryptic peptide. For each peptide the residues sequenced are given in parentheses as follows: C1 (1-3); C2 (4); C3 (9); C4 (13); C5 (18-20); C5a (13); C6 (21); C7 (29-31); C7a (22-24, 30-31); C8 (32-38); C9 (39-45); C10 (46-54); C11 (55-57); C12 (58-64); C13 (65-69); C13a (65); CT14 (75-94); C14a (75-77); C14b (85); C15 (95-99); C16 (100-109); C16a (95); C17 (110-112); C18 (113-115); C19 (116-124); C20 (125-130); T1 (1); T2 (2); T3 (6); T4 (11-13); T4 + 5 (11); T5 (14); T6 (15); T7 (44); T8 (64); T9 (70-72); T10 (73-74); T11 + 12 (75); T13 (99-101); T14 (102-107); T16 (115-117); T17 (118-119); T18 (120); T19 (127); Sequencer (1-28); solid-phase technique (1-17). For additional information, see paragraph concerning supplementary material at the end of this paper.

that of lactalbumin. As there is no evidence for a speed-up in lactalbumin evolution, the gene duplication event may be more ancient than is commonly supposed.

The rat lysozyme sequence also enabled us to examine the claim that generation time affects the rate of molecular evolution (Laird et al., 1969; Kohne, 1970; Benveniste and Todaro, 1976). Since sequence information is available for the lysozymes of some primates, the sequence of rat lysozyme enabled us to compare the rates of lysozyme evolution in primates and rodents. The results indicate that generation time is not a major factor in lysozyme evolution.

Rat lysozyme is of interest also because of its possible role as a mediator of the antitumor functions of macrophages (Osserman et al., 1973). Knowledge of its structure might contribute to our understanding of the role of lysozyme in leukemia and cell surface phenomena.

## Experimental Procedure

*Rat Lysozyme Preparation.* Lysozyme was purified from the urine of Wistar/Furth rats bearing the transplantable chloroleukemic tumor described by Klockars et al. (1974). The procedure used by Canfield et al. (1971) and Canfield and McMurray (1967) to isolate human leukemic lysozyme was followed with two additional steps. The preparation obtained by the Canfield procedure was applied to a 2.5 × 75 cm column containing Sephadex G-50 Fine and eluted with 0.1 M ammonium acetate buffer pH 9.1 at 23 °C. The fractions containing lysozyme activity were then chromatographed on a 2.5 × 51 cm column containing Bio-Rex 70 at 23 °C using 0.2 M sodium phosphate buffer, pH 7.1. The lysozyme eluted from

TABLE I: Properties of the Chymotryptic Peptides[a] of Rat Lysozyme: Amino Acid Composition, Number of Residues, Yield, N Terminus,

| Composition[b] Amino Acid | C-1 | C-2 | C-3 | C-4 | C-5 | C-6 | C-7 | C-8 | C-9 | C-10 | C-11 | C-12 | C-13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Lys | 113.6 (1) | | | 6.8[f] (1) | | | | | | | | | 10.6 (1) |
| His | | | | | | | | 17.6 (1) | | | | | |
| Arg | | 25.6 (1) | 21.1 (1) | 10.1 (1) | | | | | 68.6 (1) | | | 9.9 (1) | 8.7 (1) |
| Asp | | | | 12.0 (1) | | 6.4 (1) | | 20.6 (1) | 159.8 (2) | 31.7[f] (3) | | 7.8 (1) | 16.9 (2) |
| Thr | 108.2 (1) | | 19.0 (1) | | | | | | 76.5 (1) | 11.9 (1) | | | 6.0 (1) |
| Ser | | | | | 14.0 (1) | 6.2 (1) | | 17.6 (1) | | 13.6 (1) | | 6.4 (1) | |
| Glu | | 47.0 (2) | | | | | | 36.9 (2) | 77.1 (1) | 14.0 (1) | | 6.9 (1) | |
| Pro | | | | | | | | | | 9.1 (1) | | | 6.2 (1) |
| Gly | | | | 10.2 (1) | 17.0 (1) | 7.5 (1) | | | | 14.7 (1) | 22.0 (1) | | 8.5 (1) |
| Ala | | | 20.0 (1) | | | 7.5 (1) | | 18.3 (1) | 82.1 (1) | | | | 8.5 (1) |
| CM-Cys | | 25.6 (1) | | | | | 30.9[e] (1) | | | | | | 6.4[e] (1) |
| Val | | | | | | 6.4 (1) | 38.1 (1) | | | | | | |
| Met | | | | 6.7 (1) | | | | | | | | | |
| Ile | | | | | | | | | | | 26.0 (1) | 6.0 (1) | |
| Leu | | | 21.1 (1) | | | 7.2 (1) | 46.1 (1) | | | | | | |
| Tyr | 112.8 (1) | | | | 15.8 (1) | 7.2 (1) | | 19.6 (1) | 74.5 (1) | 14.9 (1) | | 6.7 (1) | |
| Phe | | 23.3 (1) | | | | | | | | | 21.1 (1) | | |
| Trp | | | | | | 5–10 (1) | | | | | | ~7 (1) | |
| No. of residues | 3 | 5 | 4 | 5 | 3 | 8 | 3 | 7 | 7 | 9 | 3 | 7 | 9 |
| Yield (μmol) | 0.22 | 2.50 | 2.00 | 0.04 | 1.50 | 0.30 | 0.15 | 1.80 | 2.10 | 3.00 | 2.00 | 0.70 | 0.08 |
| N terminus | Lys | Glu | Ala | Lys | Ser | Tyr | Val | Ala | Asp | Asp | Gly | Glu | CM-Cys |
| Net charge[c] at pH 6.4 | +1 | −2 | +1 | +2 | 0 | −1 | −1 | 0 | +1 | −3 | 0 | +1 | 0 |
| Method of purification[d] | S25, Dowex, $E_2$ | S25, Dowex | S25, Dowex | S25 Dowex, $E_2$ | S25, Dowex | S25, Dowex | S25, Dowex, $E_2$ | S25, Dowex | S25, Dowex, $E_2$ | S25, Dowex, $E_2$, TLC | S25, Dowex | S25, Dowex | S25, S15, Dowex |

[a] Peptides are numbered from the N terminus of lysozyme. C, chymotryptic peptide; CT, chymotryptic–tryptic peptide. [b] The amino acid composition of each peptide is given in nanomoles. Integers in parentheses are the number of residues per amino acid determined by the minimum integer method. [c] Calculated from the mobility relative to aspartic acid at pH 6.4 (Offord, 1966). A minus sign indicates movement toward the anode at pH 6.4; a plus sign, movement toward the cathode; a zero, no movement; ND, not done. [d] The symbols refer to: S15.

the latter column as a single peak of constant specific activity. Dansyl N-terminal analysis of the preparation gave only bislysine. Rat lysozyme migrated as a single band toward the cathode during electrophoresis at pH 8.6 on cellulose acetate membranes. Electrophoresis was carried out at 0 °C for 40 min at 250 V with 0.08 M sodium diethyl barbiturate buffer using a Beckman Microzone electrophoresis apparatus. The yield of purified lysozyme was 120 mg from 130 g (wet weight) of total solids from a 60% saturated ammonium sulfate fractionation of rat urine.

*Edman sequential degradation* of peptides was carried out manually on 0.05 to 0.2 μmol. The thiazolinone and phenylthiohydantoin derivatives were hydrolyzed in 6 N HCl for 20 h at 140 °C and subjected to amino acid analysis. The phenylthiohydantoin derivatives of asparagine, aspartic acid, glutamine, glutamic acid, serine, threonine, tryptophan, and carboxymethylcysteine were identified by thin-layer chromatography on Cheng Chin polyamide sheets as previously described (Jeppsson and Sjöquist, 1967) or by the method of Summers et al. (1973). The Edman degradation was also used in combination with the dansyl amino terminal identification technique (Hartley, 1970), and the subtractive Edman procedure. Tryptophan residues located at the C terminus of chymotryptic peptides were determined by amino acid analysis of the unhydrolyzed peptide following cleavage of the penultimate residue. The amino acid cleaved at each step was usually identified by two different methods.

*Automatic Sequencer.* One hundred nanomoles of *S*-carboxymethyllysozyme was sequenced through 28 cycles on a Beckman Model 890 sequencer using the Fast Protein-Quadrol program in the laboratory of R. D. Cole.

*Solid-Phase Sequencing.* The solid-phase thioacetylthioglycolic acid sequencing method of Mross and Doolittle (1971) was used for two peptides. The procedure was automated by G. Mross in the laboratory of A. C. Wilson. This solid-phase sequencer was very similar to that of Laursen (1971).

## Results

*Amino Acid Composition.* The amino acid composition of rat lysozyme agreed with published values (Mulvey et al., 1974).

*Chymotryptic Peptides.* Sixty milligrams of rat *S*-carboxymethyllysozyme was digested with chymotrypsin for 16 h. The amino acid composition, yield, N terminus, net charge at pH 6.4, and method of purification of each peptide are shown in Table I. The amino acid sequence of peptides C-5 to C-20 was determined. The complete amino acid sequence of

Net Charge at pH 6.4, and Method of Purification.

| CT-14 | C-15 | C-16 | C-17 | C-18 | C-19 | C-20 | C-5a | C-7a | CT-13 | C-13a | C-14a | C-14b | C-16a |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 23.6 (1) | | | 129.0 (1) | 19.6 (1) | | 50.5 (1) | | 20.2 (1) | 9.3 (2) | | | 21.9 (1) |
| | 24.7 (1) | 30.6 (2) | | 165.1 (1) | 17.0 (1) | 50.8 (1) | 58.4 (1) | | | 4.8 (1) | | | 46.8 (3) |
| 34.7 (3) | | 19.3 (1) | | | 43.5 (2) | 78.7 (1) | 61.4 (1) | 9.7 (1) | 42.9 (2) | 15.5 (3) | 25.9 (1) | 10.8 (2) | 22.9 (1) |
| 15.9 (1) | | | | | | | | | | 5.7 (1) | | 4.8 (1) | |
| 13.4 (1) | | | | | 19.6 (1) | | 55.6 (1) | 8.9 (1) | | 6.4 (1) | 17.1 (1) | | |
| 42.8 (3) | | 17.3 (1) | | 145.0 (1) | | | | | | | | 14.3 (3) | 19.2 (1) |
| 13.9 (1) | | 18.0 (1) | | | | | | | | 10.3 (2) | 18.1 (1) | | 18.1 (1) |
| 16.4 (1) | | 19.0 (1) | | | 22.1 (1) | 78.9 (1) | 124.1 (2) | 11.1 (1) | 21.7 (1) | 11.6 (2) | 27.1 (1) | | 20.0 (1) |
| 35.0 (3) | 25.8 (1) | 16.9 (1) | 91.7 (1) | | | | | 12.5 (1) | | 15.2 (3) | 39.8 (2) | 5.2 (1) | 35.3 (2) |
| 20.5 (2) | 21.9 (1) | | | | 17.6 (1) | ~25[e] (1) | | 10.1 (1) | 11.2 (1)[e] | ~5 (3)[e,f] | 28.9 (2)[e] | | 8.0 (1)[e] |
| | 21.9 (1) | 9.5[f] (1) | 85.8[f] (1) | | | 83.1 (1) | | 19.1 (2) | | | | | 29.2 (2) |
| | | | | | | | 51.2 (1) | | | | | | |
| 30.8 (3) | | 14.1 (1) | | | | 52.7[f] (1) | | | | 3.4 (1) | 21.5 (1) | 8.1 (2) | 15.5 (1) |
| 24.8 (2) | | | | | 22.1 (1) | | | 23.4 (2) | | 5.3 (1) | 35.5 (2) | 4.0 (1) | |
| | | | | | 20.0 (1) | | 52.8 (1) | 3.6 | | | | | |
| | | 16.8 (1) | ~100 (1) | | | | | ~10 (1) | | | | | (1)[g] |
| 20 | 5 | 10 | 3 | 3 | 9 | 6 | 8 | 10 | 5 | 20 | 11 | 10 | 15 |
| 0.30 | 0.40 | 0.32 | 0.09 | 1.30 | 0.40 | 1.00 | 0.50 | 0.30 | 0.15 | 0.05 | 0.10 | 0.10 | 0.05 |
| Asp | CM-Cys | Val | Val | <Glu[h] | CM-Cys | Ile | Lys | Gly | CM-Cys | CM-Cys | Asn | Leu | CM-Cys |
| −3 | +1 | +1 | 0 | +1 | 0 | 0 | +2 | −2 | −1 | ND | ND | ND | ND |
| S25, S15 | S25, Dowex | S25, Dowex | S25, Dowex, map | S25, Dowex, $E_2$ | S25, Dowex | S25, Dowex, $E_2$ | S25, Dowex, $E_2$ | S25, Dowex, TLC | S25, S15, map | S25, S15, map | S25, Dowex, map | S25, S15, map | S25, Dowex, map |

S25, Sephadex G-15 and G-25 chromatography, respectively; Dowex, Dowex-50 chromatography; $E_2$, high voltage electrophoresis at pH 6.4; TLC, thin-layer chromatography; map, peptide map. [e] Low recovery of CM-cysteine from paper. [f] Low recovery of N-terminal amino acid after staining with ninhydrin. [g] Identified with the Ehrlich stain. [h] 5-Pyrrolidone-2-carboxylic acid.

rat lysozyme and its chymotryptic peptides are shown in Figure 1.

*Alignment.* Chymotryptic peptides were aligned by homology with human lysozyme or the partial sequence of mouse lysozyme (Canfield et al., 1971; Riblet, 1974). This alignment was confirmed by amino acid compositions, N-terminal analysis, and partial structure determinations of the tryptic peptides of rat lysozyme. The positions of the tryptic peptides in the sequence, T-1 to T-19, are shown in Figure 1.

*Sulfhydryl Analysis.* No sulfhydryl groups were found in rat lysozyme just as none have been found in any other lysozymes *c* whose sequences are known. Hence the residues indicated as cysteine in Figure 1 must exist in disulfide linkages in the native molecule.

*Sequence Difficulties.* Although chymotryptic peptides sufficient to account for all the amino acid residues of rat lysozyme were isolated and sequenced, two weaknesses in the proposed sequence exist. First, tryptic peptide T-15 (see Figure 1) was not isolated. Occupying positions 108–114, T-15 overlaps chymotryptic peptides C-16, C-17, and C-18 and would have been useful in confirming glutamine-113. However, the absence of this peptide does not seriously detract from the definitive nature of the proposed sequence. Residues 108–112 are invariant in all lysozymes of known sequence. Furthermore, the assignment of arginine to position 114 is consistent with the sequence of peptide T-16.

The second weakness in the sequence data is due to the lack of a sequenced overlap peptide at positions 73–74. The presence of two residues of lysine in the chymotryptic overlap peptide C-13a and the likelihood of a C-terminal alanine residue in peptide C-13 (see peptide T-9, Figure 1) strongly suggest the assignment of alanyllysine at positions 73–74. This dipeptide also occurs at positions 96–97 in the sequence. Only the presence of this dipeptide at positions 73–74 is consistent with release of the free dipeptide (T-10) following tryptic digestion of rat lysozyme.

*Homologies with Other Lysozymes.* Although rat lysozyme is clearly homologous in amino acid sequence to lysozymes of the *c* type, it differs by substitutions at many positions from those whose primary structure is known. For the number of sequence differences from various lysozymes *c*, see Table II. As is evident from the table, rat lysozyme differs by 33–37 substitutions from primate lysozymes and by 52–61 substitu-

[1] Abbreviations used are: CM, carboxymethyl; S15, S25, Sephadex G-15 and G-25 chromatography, respectively; $E_1$ and $E_2$, high voltage electrophoresis at pH 1.9 and 6.4, respectively; TLC, thin-layer chromatography.

TABLE II: Amino Acid Differences and Minimal Mutation Distances among Animal Lysozymes and Lactalbumins.[a]

| Species Compared | Minimal Mutation Distances: Lysozymes C | BWQ | JQ | Tur | GF | Duck II | Duck III | Cha | Bab | Hum | Rat | α-Lactalbumins Hum | Cow | GP |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Lysozymes | | | | | | | | | | | | | | |
| Chicken | | 4 | 7 | 7 | 10 | 27 | 27 | 30 | 69 | 70 | 82 | 103 | 108 | 120 |
| Bobwhite quail | 4 | | 11 | 11 | 8 | 28 | 28 | 34 | 70 | 71 | 83 | 104 | 109 | 121 |
| Japanese quail | 6 | 10 | | 10 | 16 | 29 | 29 | 32 | 69 | 72 | 79 | 106 | 113 | 123 |
| Turkey | 7 | 11 | 10 | | 16 | 27 | 27 | 32 | 72 | 72 | 84 | 109 | 114 | 126 |
| Guinea fowl | 10 | 8 | 15 | 16 | | 33 | 33 | 32 | 72 | 75 | 84 | 107 | 112 | 124 |
| Duck II | 22 | 23 | 25 | 22 | 29 | | 7 | 36 | 68 | 69 | 77 | 107 | 113 | 121 |
| Duck III | 23 | 24 | 26 | 23 | 30 | 6 | | 36 | 66 | 67 | 71 | 106 | 113 | 121 |
| Chachalaca | 27 | 31 | 30 | 28 | 28 | 30 | 31 | | 70 | 75 | 75 | 105 | 111 | 123 |
| Baboon | 50 | 51 | 52 | 53 | 54 | 50 | 49 | 50 | | 14 | 45 | 108 | 117 | 117 |
| Human | 52 | 53 | 54 | 54 | 57 | 53 | 52 | 55 | 14 | | 51 | 106 | 112 | 115 |
| Rat | 58 | 59 | 57 | 60 | 61 | 56 | 52 | 57 | 33 | 37 | | 108 | 113 | 117 |
| α-Lactalbumins | | | | | | | | | | | | | | |
| Human | 81 | 82 | 83 | 86 | 84 | 86 | 87 | 84 | 80 | 83 | 85 | | 35 | 44 |
| Cow | 82 | 83 | 84 | 84 | 85 | 86 | 87 | 84 | 84 | 87 | 83 | 32 | | 56 |
| Guinea pig | 89 | 89 | 91 | 92 | 91 | 91 | 93 | 93 | 87 | 90 | 92 | 37 | 45 | |

[a] The number of amino acid differences between any two lysozymes or lactalbumins is given in the lower left-hand section of the matrix, while the minimal mutation distances appear in the upper right-hand section. The source of the lactalbumin sequences was Hill et al. (1974). The lysozyme sequences are from Jollès et al. (1976) and references therein. Each deletion of an amino acid in the lactalbumin or bird lysozymes relative to mammalian lysozymes has been counted as an amino acid difference but has been ignored in the computation of minimal mutation distances. C, chicken; BWQ, bobwhite quail; JQ, Japanese quail; Tur, turkey; GF, guinea fowl; Cha, chachalaca; Bab, baboon; Hum, human; GP, guinea pig.

TABLE III: Relative Rates of Sequence Evolution in Lysozyme and α-Lactalbumin in Mammals.

| Comparison | Corr % Change, $m$[a] | Divergence Time (millions of years) | Unit Evolutionary Period (millions of years) |
|---|---|---|---|
| α-Lactalbumin | | | |
| Human–guinea pig | 36 | 75 | 2.1 |
| Human–cow | 26 | 75 | 2.9 |
| Cow–guinea pig | 41 | 75 | 1.8 |
| Mean | | | 2.3 |
| Lysozyme | | | |
| Human–rat | 33.5 | 75 | 2.2 |
| Baboon–rat | 29.3 | 75 | 2.6 |
| Human–baboon | 11.4 | 30 | 2.6 |
| Mean | | | 2.5 |
| Birds (mean for 21 pairs of species)[b] | | | 6 |

[a] To correct for multiple evolutionary substitutions at the same site, the corrected percent change, $m$, was calculated from the following formula (Dickerson, 1971): $m/100 = -\ln [1 - (n/100)]$ where $n$ is the observed percent change in amino acid sequence. [b] From Prager et al. (1972). As pointed out by Wilson et al. (1977), the bird rate may not be as slow as that estimated by Prager et al. (1972).

tions from the lysozymes of birds.

A moderately close relationship appears to exist between rat and mouse urinary lysozyme. We have used partial sequence data and peptide compositions produced by R. Riblet and A. C. Wang to align the amino acid residues of mouse lysozyme with the rat sequence. There is a minimum of 14 substitutions at the 106 positions compared. This strong homology is consistent with the results of immunological tests (Glynn and Parkman, 1967; E. M. Prager and A. C. Wilson, unpublished results[2]).

## Discussion

*Rates of Lysozyme and Lactalbumin Evolution.* By considering the number of sequence differences among various lysozymes (Table II) in relation to fossil evidence regarding divergence times, one can estimate approximately the average rates at which amino acid substitutions have accumulated. As shown in Table III, lysozyme has evolved more rapidly in the mammals studied than in birds by a factor of about 2. This is consistent with immunological evidence for 19 primate lysozymes (Hanke et al., 1973), and the partial sequence information for mouse lysozyme referred to above. Indeed, mammalian lysozymes appear to have evolved at a rate that is not significantly different from the lactalbumin rate. The mean unit evolutionary period, i.e., the time required for a 1% sequence difference to accumulate between two proteins, is about 2.3 million years for mammalian lactalbumins and about 2.5 million years for mammalian lysozymes. This finding stimulates us to question the widespread supposition that lactalbumin underwent accelerated evolution following its origin from lysozyme (Dickerson, 1971; Hood et al., 1975).

*Time of Gene Duplication Event.* Two phylogenetic models have been proposed for the evolution of lactalbumin from lysozyme (Dickerson and and Geis, 1969; Dayhoff, 1972). They are illustrated in Figure 2. Model I postulates that the duplication event took place at the outset of mammalian evolution when the mammary gland originated, approximately 175 million years ago. Consistent with this model, neither lactose

[2] We plan to discuss elsewhere the bearing of the rat lysozyme sequence on our understanding of antigenic cross-reactivity among lysozymes.

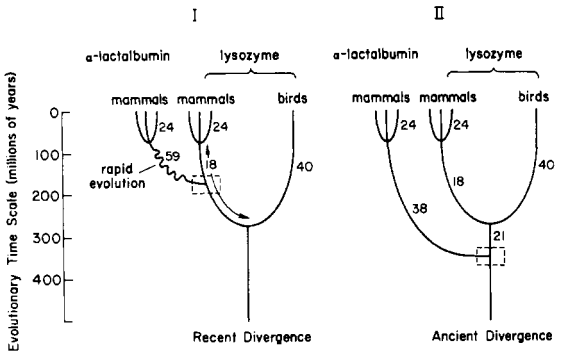


FIGURE 2: Alternative models for the divergence of the genes coding for lysozyme and α-lactalbumin. The relative number of minimal mutation distances per lineage are shown for both models. The dashed rectangle indicates the hypothetical gene duplication event.

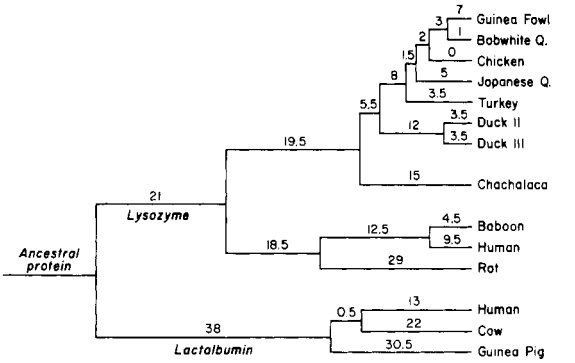


FIGURE 3: Phylogenetic tree for lysozymes and α-lactalbumins, as determined by the Farris method (1972). The lysozyme tree, excluding the rat sequence, is from Jollès et al. (1976). The data are from Table II. The Farris method was especially designed for evolutionary rate analysis. The numbers on the limbs represent the minimum number of nucleotide substitutions calculated to have occurred along each limb since a given divergence point. The abbreviation Q stands for quail.

nor lactalbumin have been found in nonmammalian vertebrates. Whereas the search for lactose encompassed a wide variety of species (Elsner, 1941), we are not aware of a comparable search for lactalbumin-like proteins in nonmammals. It is therefore justifiable to consider model II, which postulates an ancient duplication predating the last common ancestor of birds and mammals (Figure 2). According to this model, lactalbumin-like proteins may exist in nonmammals.

Both models can explain the observation (see Table II) that bird and mammal lysozymes are more alike in sequence than either is to lactalbumin. Model I does so by coupling a recent duplication with accelerated evolution in the lactalbumin lineage. Model II achieves a similar result by assuming an ancient duplication with no acceleration of sequence change in the lactalbumin lineage.

It is difficult, however, for model I to explain the observation (see Table II) that lactalbumin is not significantly closer in sequence to mammalian lysozymes than to bird lysozymes. Yet, model II is easily compatible with the latter observation, as the gene duplication occurred before the bird–mammal split.

A detailed phylogeny, consistent with model II and embracing all the lysozymes *c* and lactalbumins of known sequence, is shown in Figure 3. This phylogeny was constructed from the matrix of minimal mutation distances shown in Table II by use of the Farris (1972) method. This method was chosen because it makes no assumptions about the homogeneity of evolutionary rates along different lineages of a phylogenetic tree.

*Years or Generations.* The phylogeny in Figure 3 allowed us to examine the relative importance of years and generations in the evolution of lysozyme and lactalbumin. As pointed out by Carlson et al. (1977) and Wilson et al. (1977), although the primate and rodent lineages are equally old, 75 million years, the rodent lineage has experienced about eight times as many generations as that leading to higher primates such as man. It is clear from Figure 3, however, that the rodent lineage has not accumulated eight times as many amino acid substitutions as has the human lineage. Instead of an 8-fold effect, we estimate an effect that is 1.3-fold for lysozyme and 2-fold for lactalbumin.

If one considers the primate lineages in the lysozyme phylogeny shown in Figure 3, more sequence change has occurred along the human lineage than along the baboon lineage. Yet the lineage leading to Old World monkeys has experienced more generations than has the lineage leading to man (Sarich and Wilson, 1973). Thus the relative amounts of sequence change along these two primate lineages are not consistent with a generation-time effect.

We conclude that years may be a more important factor than generations for sequence evolution in lysozyme and lactalbumin. A similar conclusion emerges from analogous studies of serum albumin (Sarich, 1972; Sarich and Wilson, 1973) and cytochrome *c* (Carlson et al., 1977).

## Acknowledgments

We thank M. Hsiang for operating the automatic sequencer and R. D. Cole and C. Olsen for use of equipment. We are also grateful to S. S. Carlson, E. M. Prager, and I. Fung for helpful discussions and to G. Stark for typing assistance.

## Supplementary Material Available

Detailed information regarding experimental procedures, amino acid composition of rat lysozyme, properties of the tryptic peptides, sequence results for chymotryptic peptides C-5 through C-20, and the results of applying the automatic sequencer and solid-phase sequencing methods to the whole protein (12 pages). Ordering information is given on any current masthead page.

## References

Banyard, S. H., Blake, C. C. F., and Swan, I. D. A. (1974), in Lysozyme, Osserman, E. F., et al., Ed., New York, N.Y., Academic Press, p 71.

Bennett, J. C. (1967), *Methods Enzymol. 11,* 330.

Benveniste, R. E., and Todaro, G. J. (1976), *Nature (London) 261,* 101.

Canfield, R. E., Kammerman, S., Sobel, J. H., and Morgan, F. J. (1971), *Nature (London), New Biol. 232,* 16.

Canfield, R. E., and McMurray, S. (1967), *Biochem. Biophys. Res. Commun. 26,* 38.

Carlson, S. S., Mross, G. A., Wilson, A. C., Mead, R. T., Wolin, L. D., Bowers, S. F., Foley, N. T., Muijsers, A. O., and Margoliash, E. (1977), *Biochemistry 16* (following paper in this issue).

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

Dayhoff, M. O., Ed. (1972), Atlas of Protein Structure and Sequence, Vol. 5, Silver Spring, Md., The National Biomedical Research Foundation.

Dickerson, R. E. (1971), *J. Mol. Evol. 1,* 26.

Dickerson, R. E., and Geis, I. (1969), The Structure and Ac-

tion of Proteins, Menlo Park, Calif., Benjamin Inc.
DiSabato, G., Pesce, A., and Kaplan, N. O. (1963), *Biochim. Biophys. Acta 77,* 135.
Edelhoch, H. (1967), *Biochemistry 6,* 1948.
Edman, P. (1970), in Protein Sequence Determination, Needleman, S. B., Ed., New York, N.Y., Springer-Verlag, p 253.
Elsner, H. (1941), Grundriss der Kohlenhydrat Chemie, Berlin, Verlag von Paul Parey, p 130.
Farris, J. S. (1972), *Am. Nat. 106,* 645.
Florkin, M. (1975), *Compr. Biochem. 29B,* 130.
Glynn, A. A., and Parkman, R. (1967), *Immunology 7,* 724.
Hanke, N., Prager, E. M., and Wilson, A. C. (1973), *J. Biol. Chem. 248,* 2824.
Hartley, B. S. (1970), *Biochem. J. 119,* 805.
Hill, R. L., and Brew, K. (1975), *Adv. Enzymol. 43,* 411.
Hill, R. L., Steinman, H. M., and Brew, K. (1974), in Lysozyme, Osserman, E. F., et al., Ed., New York, N.Y., Academic Press, p 55.
Hirs, C. H. W. (1967), *Methods Enzymol. 11,* 325.
Hood, L. E., Wilson, J. H., and Wood, W. B. (1975), Molecular Biology of Eucaryotic Cells, Menlo Park, Calif., Benjamin Inc.
Jensen, K. A., and Pederson, C. (1961), *Acta Chem. Scand. 15,* 1087.
Jeppsson, J. O., and Sjöquist, J. (1967), *Anal. Biochem. 18,* 264.
Jollès, J., Schoentgen, F., Jollès, P., Prager, E. M., and Wilson, A. C. (1976), *J. Mol. Evol. 8,* 59.
Klockars, M., Azar, H. A., Hermida, R., Isobe, T., Hsu, C. C. S., Ansari, H., and Osserman, E. F. (1974), *Cancer Res. 34,* 47.
Kohne, D. E. (1970), *Q. Rev. Biophys. 33,* 327.
Laird, C. D., McConaughy, B. J., and McCarthy, B. J. (1969), *Nature (London) 224,* 149.
Laursen, R. A. (1971), *Eur. J. Biochem. 20,* 89.
Liu, T.-Y., and Chang, Y. H. (1971), *J. Biol. Chem. 246,* 2842.
Marvel, C. S., de Radzitzky, P., and Brader, J. J. (1955), *J. Am. Chem. Soc. 77,* 5997.
Moore, S., and Stein, W. H. (1954), *J. Biol. Chem. 211,* 907.
Mross, G. A., and Doolittle, R. F. (1971), *Fed. Proc., Fed. Am. Soc. Exp. Biol. 30,* 1241.
Mulvey, R. S., Gualtieri, R. J., and Beychok, S. (1974), *Biochemistry 13,* 782.
Offord, R. E. (1966), *Nature (London) 221,* 591.
Osserman, E. F., Klockars, M., Halper, J., and Fischel, R. E. (1973), *Nature (London) 243,* 331.
Parry, R. M., Jr., Chandan, R. C., and Shahani, K. M. (1965), *Proc. Soc. Exp. Biol. Med. 119,* 384.
Prager, E. M., Arnheim, N., Mross, G. A., and Wilson, A. C. (1972), *J. Biol. Chem. 247,* 2905.
Riblet, R. (1974), in Lysozyme, Osserman, E. F., et al., Ed., New York, N.Y., Academic Press, p 89.
Sarich, V. M. (1972), *Biochem. Genet. 7,* 205.
Sarich, V. M., and Wilson, A. C. (1973), *Science 179,* 1144.
Schroeder, W. A. (1967), *Methods Enzymol. 11,* 351.
Summers, M. R., Smythers, G. W., and Oroszlon, S. (1973), *Anal. Biochem. 53,* 624.
Wilson, A. C., Carlson, S. S., and White, T. J. (1977), *Annu. Rev. Biochem. 46* (in press).