

Stomach Lysozymes of Ruminants

II. AMINO ACID SEQUENCE OF COW LYSOZYME 2 AND IMMUNOLOGICAL COMPARISONS WITH OTHER LYSOZYMES*

(Received for publication, December 15, 1983)

Pierre Jollès‡, Françoise Schoentgen, and Jacqueline Jollès

From the Laboratory of Proteins, University of Paris V, 45 rue des Saints-Pères, F-75270 Paris Cedex 06, France

Deborah E. Dobson§, Ellen M. Prager, and Allan C. Wilson¶

From the Department of Biochemistry, University of California, Berkeley, California 94720

The complete sequence of 129 amino acids has been determined for one of three closely related lysozymes *c* purified from cow stomach mucosa. The sequence differs from those known for 17 other lysozymes *c* at 39–60 positions, at one of which there has been a deletion of 1 amino acid. The glutamate replacement at position 101 and the deletion of proline at position 102 eliminate the aspartyl-prolyl bond that is present between these positions in all other mammalian lysozymes *c* tested. This bond appears to be the most acid-sensitive one in such lysozymes at physiological temperature. Of the 40 positions previously found to be invariant among lysozymes *c*, only one has undergone substitution in the cow lineage. This modest number of changes at novel positions is consistent with the inference, based on tree analysis and antigenic comparisons, that the tempo of evolutionary change in the cow lysozyme lineage has not been radically different from that in other lysozyme *c* lineages. The mutations responsible for the distinctive catalytic properties and stability of cow lysozyme *c* could be a minor fraction of the total that have been fixed in the cow lineage.

The lysozyme *c* of ruminants provides an opportunity to examine both the driving force for protein evolution and the structural basis of altered protein function. As the result of a major regulatory change, this lysozyme appears to have lost its old function and acquired a new one. According to the hypothesis offered by Dobson *et al.* (1979, 1984), ruminant lysozyme *c* no longer works at or near neutral pH as a shield against bacterial infection in many tissues and secretions (e.g. white blood cells, tears, egg white, and milk), as do the lysozymes *c* of many mammals and birds (Feeney and Allison, 1969; Osserman *et al.*, 1974). Instead, in ruminants, this enzyme appears to be produced only by the stomach mucosa and to function exclusively as a major digestive enzyme (Dobson *et al.*, 1979, 1984) in the presence of pepsin and at a lower

pH than is usually the case for the lysozymes of typical mammals and birds.

An important step toward understanding the structural basis for functional change is to determine the amino acid sequence of a ruminant stomach lysozyme and to compare it with the sequences known for conventional lysozymes *c* from other animals. Seventeen complete sequences are known for non-ruminant lysozymes *c* (Jollès *et al.*, 1979a; Jung *et al.*, 1980; Kondo *et al.*, 1982; and references therein) and partial sequences are available for seven more (White, 1976; Jollès *et al.*, 1979a, 1979b; Gavilanes *et al.*, 1982; and references therein). We have, therefore, characterized and sequenced a lysozyme *c* from the stomach mucosa of a typical ruminant, the domestic cow. Furthermore, by conducting a tree analysis of the sequence data and making quantitative immunological comparisons of cow lysozyme *c* to other lysozymes, it has been possible to find out whether the rate of sequence evolution has been affected by the functional shift.

EXPERIMENTAL PROCEDURES¹

Lysozymes and Tissues—The purified lysozymes,² mammalian stomach extracts, and primate milks used in this study are described by Dobson *et al.* (1984).

Enzymes and Reagents—Trypsin (EC 3.4.21.4) and carboxypeptidases A and B (EC 3.4.17.1, EC 3.4.17.2) were purchased from Worthington and *Staphylococcus aureus* V8 protease (EC 3.4.21.19) from Miles. Sephadex G-10, G-25 (fine), and G-50 (fine) were obtained from Pharmacia, and Bio-Gel P-60 from Bio-Rad. Cyanogen bromide was obtained from Merck. All other reagents (analytical grade) were purchased from Merck or Prolabo; those employed for the Sequencer were from Merck (Sequanal grade). For work other than sequencing, reagents were of standard analytical grade.

Reduction, Alkylation, Citraconylation, and Enzymatic and Chemical Cleavage—Bovine stomach lysozyme 2 was reduced with 2-mercaptoethanol and alkylated with iodoacetamide according to Jollès *et al.* (1972). Reduced and alkylated lysozyme (20 mg) dissolved in 4 ml of 10 mM NaOH to which 4 ml of 50 mM *N*-ethylmorpholine were then added was citraconylated at pH 8.5 for 2 h at 20 °C (Maley *et al.*

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ Supported by grants from the Centre National de la Recherche Scientifique (Equipe de Recherche No. 102) and the Institut National de la Santé et de la Recherche Médicale (Unité U-116). This is the 119th article on lysozyme from this author's laboratory, the 118th being Berthou *et al.* (1983).

§ Present address, Dana-Farber Cancer Institute and Department of Pharmacology, Harvard Medical School, Boston, MA 02115.

¶ Supported by National Institutes of Health Grant GM-21509.

¹ Portions of this paper (including part of "Experimental Procedures," part of "Results," part of "Discussion," additional Figs. 1–6, Tables 1–7, and additional references) are presented in miniprint at the end of this paper. Miniprint is easily read with the aid of a standard magnifying glass. Full size photocopies are available from the Journal of Biological Chemistry, 9650 Rockville Pike, Bethesda, MD 20814. Request Document No. 83M-3553, cite the authors, and include a check or money order for \$6.80 per set of photocopies. Full size photocopies are also included in the microfilm edition of the Journal that is available from Waverly Press.

² Since the lysozymes which are the focus of this report are of the *c* type, the term "lysozyme" shall denote lysozyme *c*. The terms "cow" and "bovine" are used interchangeably to refer to domestic cattle of the species *Bos taurus*.

al., 1979); the tryptic digestion was subsequently performed during 24 h at 37 °C with an enzyme/substrate ratio of 1:50. Trypsin was pretreated for 16 h with 0.0625 M HCl at 37 °C. Decitraconylation was carried out in 30% acetic acid for 2 h at 20 °C. The reduced and alkylated bovine lysozyme (15 mg) was also subjected to tryptic digestion (5 h, 37 °C) without citraconylation in 0.1 M ammonium bicarbonate with an enzyme/substrate ratio of 1:40. Digestion with *S. aureus* V8 protease of the reduced and alkylated lysozyme (15 mg) was carried out after solubilization of the latter in 0.1 M ammonium bicarbonate containing 1.5% sodium dodecyl sulfate (7 ml); after dialysis against 0.1 M ammonium bicarbonate containing 0.01% sodium dodecyl sulfate for 24 h at 4 °C, the protease (0.4 mg) was added and the digestion carried out for 24 h at 37 °C; after a new addition of the protease (0.2 mg), the digestion was continued for 18 h at 37 °C. Chemical cleavage was carried out by solubilizing reduced and alkylated bovine stomach lysozyme (15 mg) in 70% formic acid (5 ml), adding cyanogen bromide (140 mg), and allowing the reaction to proceed for 24 h at 20 °C; 30% acetic acid (20 ml) was then added and the solution concentrated by lyophilization.

Peptide Purification and Analysis—Filtrations on Sephadex G-10, G-25, G-50, and Bio-Gel P-60 (200–270 × 1.2–2-cm columns) were performed with 30% acetic acid as the eluant (12 ml/h). The peptides were detected at 280 nm or by the fluorescamine procedure (Nakai *et al.*, 1974) with a Jobin and Yvon JY 3D spectrofluorimeter. Preparative paper chromatography (Whatman No. 1) in *n*-butyl alcohol/pyridine/acetic acid/water (15:10:3:12, v/v/v/v) or paper electrophoresis (Whatman No. 1; 50 V/cm; 45 min) at pH 6.5 in pyridine/acetic acid/water (100:3.5:900, v/v/v) was the final purification step for several peptides. The amino acid composition of the peptides after total hydrolysis (5.6 M HCl containing 1:2000 2-mercaptoethanol; 18, 48, and 72 h; under vacuum) was determined with a Biotronik Autoanalyzer.

Sequence Determination—Automated Edman degradation was carried out in a Beckman Sequencer 890 C for longer peptides by the 1 M Quadrol double-cleavage method (Edman and Begg, 1967) and for shorter peptides by the 0.2 M Quadrol single-cleavage method in the presence of Polybrene (Tarr *et al.*, 1978). The phenylthiohydantoins were identified by thin-layer chromatography with chloroform/methanol (90:10, v/v) and neat chloroform as solvents and by high performance liquid chromatography with a Waters chromatograph (model ALC/GPC-204; μ Bondapak C₁₈ column) employing the acetate system: 40 mM sodium acetate at pH 4.4 mixed with methanol in a ratio of 9:1 (v/v) for buffer A and in a 1:9 ratio for buffer B.

The COOH-terminal amino acids were determined by digestion with carboxypeptidases A and B at 37 °C for different time intervals in 0.1 M ammonium bicarbonate and in the presence of 2 mM diisopropyl fluorophosphate; the digests were then analyzed on an amino acid autoanalyzer to determine released COOH-terminal amino acids.

RESULTS

Cow Stomach Lysozymes Are of the *c* Type—The three, non-allelic lysozymes that Dobson *et al.* (1984) purified from the cow stomach mucosa were shown in the present study to be of the *c* type as regards molecular size, extinction coefficient, amino-terminal sequence, and reactivity with chitotetraose; in addition, they are extremely closely related to one another in amino acid composition (see Miniprint and Dobson, 1981). In these respects, they differ sharply from lysozymes of the *g* type, which appear to be present at very low levels in non-stomach tissues of the cow (Dobson *et al.*, 1984). The most abundant of the three lysozymes *c*, namely cow lysozyme 2, was subjected to detailed sequence analysis as follows.

NH₂-terminal Sequence and COOH-terminal Amino Acid—The NH₂-terminal sequence was established by subjecting 140 nmol of the reduced and alkylated protein to 46 cycles on a sequenator (Fig. 1). Lysine was the NH₂-terminal amino acid. The different tryptic peptides contained in this 46-residue sequence were characterized again when the tryptic digests were studied.

Leucine (0.2 residue) was the unique COOH-terminal amino acid obtained from digestion of reduced and alkylated

bovine lysozyme for 24 h at 37 °C with carboxypeptidases A and B.

Tryptic Peptides from Citraconylated Bovine Lysozyme and Their Alignment—As this lysozyme contains only 3 arginine residues per molecule, it was decided to submit the reduced and alkylated enzyme (20 mg) to tryptic digestion after citraconylation. The digest was decitraconylated and filtered on Sephadex G-25. Four peaks (CT1–CT4) were resolved and the four peptides contained therein were characterized. These four peptides add up to a total of 129 amino acids in the molecule and are in accordance with the arginine content. Peptides CT3, CT4, and CT2 were sequenced with an automated sequenator (Fig. 1): the last fragment corresponds to the COOH-terminal sequence of the bovine lysozyme, as it is devoid of a basic amino acid. Since CT3 and CT4 were contained in the NH₂-terminal sequence of the bovine lysozyme, the alignment of the four tryptic peptides of the citraconylated enzyme is CT3 → CT4 → CT1 → CT2 (Fig. 1).

Cyanogen Bromide, Tryptic, and Protease Peptides: Completion of the Bovine Stomach Lysozyme Sequence—For peptide CT1 (Fig. 1) only a 17-amino acid NH₂-terminal sequence was determined by automated analysis; an extensive study of diverse peptides encompassing the 75 COOH-terminal residues of CT1 was thus necessary in order to establish the complete sequence of cow stomach lysozyme. It was completed by examination of some of the peptides obtained from the reduced and alkylated protein by cyanogen bromide cleavage (CN peptides), by digestion with trypsin without citraconylation (T peptides), and by digestion with *S. aureus* V8 protease (SP peptides).

Three peaks (CN1–CN3) were isolated by gel filtration of the digest obtained after cyanogen bromide treatment of bovine lysozyme, which contains only 1 methionine residue/molecule (position 84, Fig. 1). Besides CN1 and CN2, the NH₂- and COOH-terminal moieties of the protein, respectively, a third peak was obtained, CN3, containing two peptides which were formed by cleavage after tryptophan residues. Lability at tryptophan residues in general during CNBr treatment complicated use of this fragmentation method (see also Blumenthal *et al.*, 1975; Braunitzer and Aschauer, 1975; Ozols and Gerard, 1977). Peptide CN3a permitted determination of a portion in the center of fragment CT1, while peptide CN3b encompassed the COOH-terminus of CT1 and all of CT2 (Fig. 1).

From the tryptic digest of bovine lysozyme, five peaks (T1–T5) were characterized after gel filtration. Five peptides situated in fragment CT1 were automatically sequenced.

At this point, the entire sequence had been established (Fig. 1) but there was no overlap encompassing the junction of peptides T1b and T5. We therefore characterized six fractions (SP1–SP6) from the *S. aureus* V8 protease digest after gel chromatography, with peptide SP2 being isolated from peak SP2. Its sequence (Fig. 1) allowed the alignment of tryptic peptides T1b, T5, and T1a. The primary structure of fragment CT1 and therefore of bovine stomach lysozyme was thus established (Fig. 1).

Immunological Cross-reactions—Since polyclonal rabbit antisera are known to provide an approximate measure of sequence relatedness among lysozymes *c*, when tested by the micro-complement fixation method (Benjamin *et al.*, 1984), we used this method to compare the stomach lysozymes of the cow and other mammals. Micro-complement fixation tests confirmed that the three lysozymes (1, 2, and 3) purified from cow stomach are very closely related to one another antigenically. This is evident from tests made with antisera to cow lysozymes 1 and 2 (Table I). With both antisera, the immu-

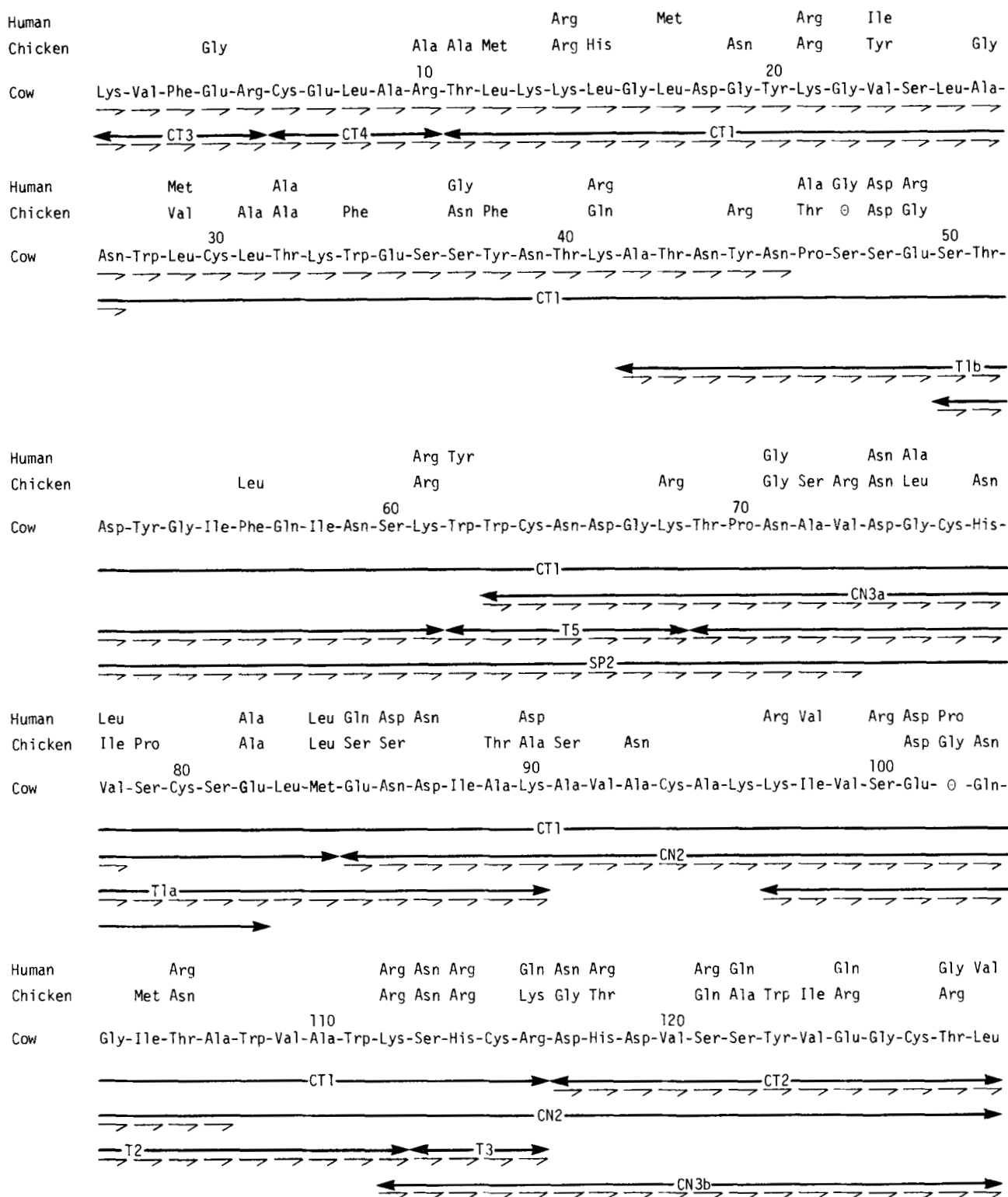


FIG. 1. Amino acid sequence of cow lysozyme 2. The first 46 residues were determined directly with a sequenator. The four tryptic peptides obtained after citraconylation of the lysozyme are indicated on the figure, as well as the various cyanogen bromide, non-citraconylated tryptic, and V8 protease peptides necessary for the establishment of the sequence. —, residue determined by automated Edman degradation; $\ominus$, deletion. The sequences of chicken and human lysozymes are given at those positions where the enzymes differ. To facilitate discussion in relation to previously reported sequences, the residues are numbered according to the chicken sequence; due to the mammalian addition between positions 47 and 48 and the bovine deletion at residue 102, these numbers are out of register with those applied to the actual cow and human sequences.

TABLE I

Immunological differences among ungulate lysozymes measured with antisera to cow lysozymes 1 and 2

Lysozyme	Immunological distance ^a	
	Cow 1	Cow 2
Pure lysozymes		
Cow 1	0	8
Cow 2	9	0
Cow 3	16	8
Stomach extracts ^b		
Cow ^c	7	2
Goat	20	19
Sheep	24	16
Bighorn sheep	28	21
Black-tailed deer	39	29
Axis deer	40	31
Roe deer	54	41
Camel	67	56
Pig	75	70
Horse ^d	≈70	≈65

^a The immunological distances were determined with the quantitative micro-complement fixation method (Champion *et al.*, 1974; *cf.* Miniprint). Immunodiffusion tests gave results consistent with the immunological distances.

^b For all stomach extracts the peak of the micro-complement fixation curve occurred at a dilution consistent with the levels of lysozyme reported from activity measurements (see Dobson *et al.*, 1984).

^c The immunological distances for the cow stomach extract are commensurate with expectations, given the values above for the pure enzymes and the observation that cow 2 predominates in such extracts, with lesser amounts of cow 1 and little cow 3 (Dobson *et al.*, 1984).

^d Values estimated from immunodiffusion reactions (*cf.* Prager and Wilson, 1971; Prager *et al.*, 1976) comparing the pig and horse with the axis deer. Technical problems prevented direct determination by micro-complement fixation.

nological distances among these three lysozymes are smaller than those between cow lysozymes and those of sheep and goats, which are the closest relatives of the cow tested. More distantly related species of hoofed animals have stomach lysozymes showing progressively greater differences from cow lysozymes in their antigenic behavior. This is evident from comparing the immunological distances among lysozymes to the order of branching of the lineages in the evolutionary tree for those species and to the times of divergence of those lineages from the cow lineage (see Miniprint; Fig. 6). Surprisingly, some non-ruminants (*e.g.* pig and horse) have stomach lysozymes *c* that are about as closely related immunologically to cow lysozyme *c* as is the camel enzyme.

Immunological distances between cow lysozymes *c* and those from primates and rodents are more than twice as large as the maximum distances in Table I. An extremely weak reaction was seen with langur stomach and squirrel monkey milk lysozymes and no reactions were detected with pure human, baboon, and rat lysozymes in immunodiffusion tests with antisera to cow lysozymes. In a reciprocal comparison, an antiserum from one of four rabbits immunized with baboon lysozyme reacted very weakly with cow lysozyme.

DISCUSSION

Adaptive Shifts and Sequence Evolution—It is important to obtain a molecular understanding of one of the generalizations emerging from comparisons of proteins, *viz* that the rate of amino acid substitution in a given type of protein (*e.g.* lysozyme *c*) is about the same along the various lineages leading to present day species of vertebrates (Wilson *et al.*, 1977). Our study of cow lysozyme may offer an unusual opportunity to find out whether adaptively significant changes in this protein

cause deviations from its average rate of sequence change. Hemoglobin is the only other protein which has offered a similar opportunity. We refer especially to its loss of old functions (*i.e.* the binding of organic phosphate, chloride, and carbamino CO₂) and gain of a new function (*i.e.* bicarbonate binding) on the lineage leading from an ancestral terrestrial reptile to the aquatic crocodilians. This functional shift represents an adaptively significant evolutionary response to the problem of blood acidity, which develops during prolonged periods under water, and can be explained by five amino acid substitutions (Perutz *et al.*, 1981). The question raised here is whether the lineage leading to crocodilian hemoglobins has experienced more sequence change than have lineages retaining the standard set of functional properties. The answer, although apparently positive (Perutz *et al.*, 1981), is ambiguous because the time elapsed since the divergence of the crocodilian lineage from other hemoglobin lineages is great enough to allow not only multiple substitutions at many sites but also intergenic exchange events to occur at this multigene locus and thus obscure the record of point mutational divergence (*cf.* Martin *et al.*, 1983). In the absence of a clearcut answer from this hemoglobin example, the case of ruminant lysozymes *c* is of particular interest.

Extent of Evolutionary Change—As a first approach to finding out how divergent the cow lysozyme sequence is, we calculated the number of amino acid differences between this lysozyme and others of known sequence (see Table II); there is a minimum of 38 amino acid substitutions and the one deletion (see below) from the other mammalian sequences and at least 56 differences from the bird sequences. In terms of both this measure and the minimum number of base substitutions needed to account for the amino acid differences, the cow lysozyme is not unusually divergent. Indeed, the cow generally differs by a little more than primates do and a little less than the rat does from birds as regards minimal mutation distance.

Phylogenetic analysis shows that the cow lysozyme is closely related to other mammalian lysozymes *c* (Fig. 2), in the same way that the species themselves are thought to be related (Fitch and Langley, 1976); this implies that the sequences have an orthologous relationship (Wilson *et al.*, 1977). The phylogenetic tree in Fig. 2 suggests that the cow lineage has accumulated a modest number (30.8) of base changes leading to amino acid changes, as compared to the rat (48.9), human (37), and baboon (29.1) lineages. Thus, despite having lost an old function and gained a new one, which affected its catalytic properties and possibly its resistance to pepsin (Dob-

TABLE II

Amino acid sequence differences and minimal mutation distances among animal lysozymes

The number of amino acid differences between any two lysozymes is given in the upper right-hand section of the matrix, while the minimal mutation distance appears in the lower left-hand section. Each addition and deletion has been counted as 1 amino acid difference and as one mutation in the computation of minimal mutation distances.

Lysozymes compared	Cow 2	Baboon	Human	Rat	Chicken	Duck	Chachalaca
Cow 2		39	41	55	58	57	56
Baboon	47		14	33	50	49	50
Human	48	14		37	52	52	55
Rat	70	45	51		58	55	57
Chicken	76	70	71	83		21	27
Duck 1 ^a	75	68	69	77	26		29
Chachalaca	73	71	76	76	30	35	

^a From Kondo *et al.* (1982), one of five different duck lysozymes sequenced.

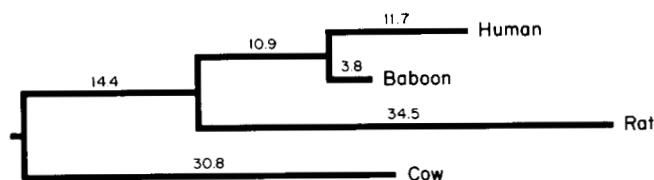


FIG. 2. Tree showing genealogical relationships among mammalian lysozymes of known amino acid sequence and the number of mutations needed to account for the divergence of these sequences from a common ancestor. This tree is based on the parsimony principle, which means that it accounts for the sequence diversity among the four mammalian lysozymes with fewer mutations than does any other branching order (cf. Ferris *et al.*, 1981). A tree with the same branching order as shown here and very similar branch lengths resulted when the minimal mutation distances in Table II were subjected to phylogenetic analysis by the Farris (1972) method.

son *et al.*, 1984) and stability in acid (see below), the cow lysozyme lineage appears not to have undergone accelerated evolution; cow lysozyme is not unusually far from the ancestral mammalian state in amino acid sequence (Fig. 2).

The immunological results (Table I) reinforce this view: the degree of antigenic difference between cow lysozyme and the lysozymes of pig and horse are hardly greater than the distance from the camel to the cow. Yet, the camel is a ruminant with a high stomach lysozyme level and a low pH optimum, while the pig and horse are non-ruminants with low levels of stomach lysozymes that do not function optimally at low pH (Dobson *et al.*, 1984). One possible implication of these findings is that only a small number of mutations was required to produce a ruminant stomach lysozyme from a conventional lysozyme.

Location and Nature of Amino Acid Replacements—The positions at which the cow sequence has diverged are for the most part unremarkable. There is an apparent amino acid substitution at only one of the 40 positions previously found to be invariant in lysozymes *c*: aspartate for asparagine at position 74. The residue at this position in chicken lysozyme probably has no contact with the hexasaccharide substrate according to Smith-Gill *et al.* (1984). Of the 25 positions where the cow enzyme has a unique amino acid or gap, 24 of them are likely to be either fully external or on the surface of the molecule (Browne *et al.*, 1969; Lee and Richards, 1971), if it is assumed that the three-dimensional structure is similar to that of chicken and human lysozyme (Artymiuk and Blake, 1981). Of these, five positions (75, 101, 102, 112, and 113) have been implicated among 30 contact residues. In contrast, there are only 12 positions (11 external or surface, two among contact residues) where either or both primate lysozymes (human and baboon) have an amino acid not seen in other vertebrate lysozymes; for the two rodent lysozymes (rat and mouse taken collectively) the corresponding value is 17 (15 external or surface, five among contact residues). This analysis suggests that cow lysozyme may have a slightly more specialized overall sequence to go with its presumed new function as a digestive enzyme.

Possible Adaptive Significance of the Changes at Positions 101 and 102—The only specific changes in cow lysozyme that we have succeeded in tentatively relating to its novel function are the deletion of proline at position 102, along with the adjacent substitution of glutamate for aspartate at position 101. These two mutations eliminate the aspartyl-prolyl bond, which is more sensitive to acid hydrolysis at physiological temperature than any other bond in lysozyme (Jauregui-Adell and Marti, 1975; Landon, 1977; Inglis, 1983). The lysozymes *c* of all five monogastric mammals tested have an aspartyl-

prolyl bond at positions 101–102. For the human, baboon, rat, and mouse, this is known from sequence analysis; for the horse (a close relative of the ruminants), this is evident from end group studies on the two peptides into which horse lysozyme is cleaved by acid at 40 °C (Jauregui-Adell and Marti, 1975). From this, we infer that the ancestral mammalian lysozyme *c* probably had this acid-sensitive bond and that it was lost on the lineage leading to ruminants. A possible implication is that one or the other of these two mutations may have been fixed as the result of selection pressure for the ability of lysozyme to survive and function after exposure to low pH (in the presence of substrates, fatty acids, pepsin, surfactant mucins, and other chemicals) not only in the stomach but, perhaps, in the duodenum as well (see Dobson *et al.*, 1984).

An additional feature of the cow lysozyme sequence that may be relevant to stability at low pH is its relatively low content of aspartyl residues and amide groups. The amide and aspartyl peptide bonds are acid-labile, although not to the extent of the aspartyl-prolyl bond. In contrast to conventional lysozymes *c*, which have an average of 23 residues (range, 20–28) of aspartate, asparagine, and glutamine, cow lysozyme 2 has only 17 such residues.

Net Charge and Charge Distribution—On the basis of electrophoretic evidence, Dobson *et al.* (1984) suggested that the ruminant lysozymes are relatively non-basic and that this might be responsible for their low catalytic activity at neutral and high pH values under physiological salt conditions. In agreement with this suggestion, from the sequence of cow lysozyme 2 we calculate a net charge of zero at pH 7.65 (uncertainty, 7.40–7.83),³ which contrasts with the high positive charge borne by most lysozymes at this pH. The lysozyme of known sequence that is most similar to cow 2 in net charge is that of the baboon, which also has a low pH optimum under physiological salt conditions (Dobson *et al.*, 1984). The difference in net charge between cow lysozyme and the conventional lysozymes *c* of rat, human, and birds at pH 8 is from 6 to 10 units. This difference is due more to an increase in the number of negative charges than to a decline in the number of positive charges. A minimum of six mutations would thus be required to produce a lysozyme having a fully ruminant-like net charge from a conventional lysozyme.

Assuming that cow lysozyme has the same general three-dimensional structure as that of chicken and human lysozyme (Artymiuk and Blake, 1981), we calculate that, at pH 8, every face⁴ of cow lysozyme bears more negative charges than does the corresponding face in chicken or human lysozyme: averages of -8.7 , -5.3 , and -6.2 , respectively. The number of positive charges per face, in contrast, is less for the cow (9.5) than for the chicken or human (each 11.9). From all aspects, then, cow lysozyme has less net positive charge and so will be less attracted than conventional lysozymes to the negatively charged bacterial substrate. This result fits with the explanation offered by Dobson *et al.* (1984) for the low catalytic

³ For this calculation we used the pK values of the chargeable amino acid residues of chicken lysozyme given by Imoto *et al.* (1972). Our calculated pI value of 7.65 ± 0.22 agrees with the value of 7.5 ± 0.1 observed by Pahud and Widmer (1982) for the major component (D) of calf stomach lysozyme by isoelectric focussing.

⁴ Six faces of the lysozyme molecule were considered: front, back, left, right, top, and bottom, as defined by Smith-Gill *et al.* (1982). At pH 8, the side chain of glutamate 35 was assigned a charge of -0.95 , histidyl residues $+0.06$, amino termini $+0.33$, all other carboxyl groups -1 , and all lysyl and arginyl residues $+1$ (Imoto *et al.*, 1972). The approximation was made that any charge-bearing amino acid residue contributed its charge to all faces on whose surface it appeared according to Smith-Gill *et al.* (1982).

activity of cow lysozyme c at high pH.

Acknowledgments—The excellent technical assistance of Ly Quan Le is gratefully acknowledged. We thank S. J. Smith-Gill, C. R. Mainhart, and R. J. Feldmann for providing crystallographic outlines of chicken lysozyme and S. M. Beverley, R. D. Cole, W. S. Davidson, H. Fujio, M. F. Hammer, P. V. Hornbeck, J. Jauregui-Adell, M. McClelland, J. A. Rupley, V. M. Sarich, E. L. Smith, S. J. Smith-Gill, C.-B. Stewart, and P. Van Soest for helpful discussion.

REFERENCES

- Artymiuk, P. J., and Blake, C. C. F. (1981) *J. Mol. Biol.* **152**, 737–762
- Benjamin, D. C., Berzofsky, J. A., East, I. A., Gurd, F. R. N., Hannum, C., Leach, S. J., Margoliash, E., Michael, J. G., Miller, A., Prager, E. M., Reichlin, M., Sercarz, E. E., Smith-Gill, S. J., Todd, P. E., and Wilson, A. C. (1984) *Annu. Rev. Immunol.* **2**, 67–101
- Berthou, J., Lifchitz, A., Artymiuk, P., and Jollès, P. (1983) *Proc. R. Soc. Lond. B Biol. Sci.* **217**, 471–489
- Blumenthal, K. M., Moon, K., and Smith, E. L. (1975) *J. Biol. Chem.* **250**, 3644–3654
- Braunitzer, G., and Aschauer, H. J. (1975) *Z. Phys. Chem.* **356**, 473–474
- Browne, W. J., North, A. C. T., Phillips, D. C., Brew, K., Vanaman, T. C., and Hill, R. L. (1969) *J. Mol. Biol.* **42**, 65–86.
- Champion, A. B., Prager, E. M., Wachter, D., and Wilson, A. C. (1974) in *Biochemical and Immunological Taxonomy of Animals* (Wright, C. A., ed) pp. 397–416, Academic Press Inc. Ltd., London
- Dobson, D. E. (1981) Ph.D. dissertation, University of California, Berkeley
- Dobson, D. E., Dayan, E., and Wilson, A. C. (1979) *Fed. Proc.* **38**, 674
- Dobson, D. E., Prager, E. M., and Wilson, A. C. (1984) *J. Biol. Chem.* **259**, 11607–11616
- Edman, P., and Begg, G. (1967) *Eur. J. Biochem.* **1**, 80–91
- Farris, J. S. (1972) *Am. Nat.* **106**, 645–668
- Feeney, R. E., and Allison, R. G. (1969) *Evolutionary Biochemistry of Proteins*, John Wiley and Sons, New York
- Ferris, S. D., Wilson, A. C., and Brown, W. M. (1981) *Proc. Natl. Acad. Sci. U. S. A.* **78**, 2432–2436
- Fitch, W. M., and Langley, C. H. (1976) *Fed. Proc.* **35**, 2092–2097
- Gavilanes, J. G., González de Buitrago, G., Martínez del Pozo, A., Pérez-Castells, R., and Rodríguez, R. (1982) *Int. J. Pept. Protein Res.* **20**, 238–245
- Imoto, T., Johnson, L. N., North, A. C. T., Phillips, D. C., and Rupley, J. A. (1972) *Enzymes*, 3rd Ed., (Boyer, P. D., ed) Vol 7, pp. 665–868, Academic Press, New York
- Inglis, A. S. (1983) *Methods Enzymol.* **91**, 324–332
- Jauregui-Adell, J., and Marti, J. (1975) *Anal. Biochem.* **69**, 468–473
- Jollès, J., van Leemputten, E., Mouton, A., and Jollès, P. (1972) *Biochim. Biophys. Acta* **257**, 497–510
- Jollès, J., Ibrahim, I. M., Prager, E. M., Schoentgen, F., Jollès, P., and Wilson, A. C. (1979a) *Biochemistry* **18**, 2744–2752
- Jollès, J., Schoentgen, F., Crozier, G., Crozier, L., and Jollès, P. (1979b) *J. Mol. Evol.* **14**, 267–271
- Jung, A., Sippel, A. E., Grez, M., and Schütz, G. (1980) *Proc. Natl. Acad. Sci. U. S. A.* **77**, 5759–5763
- Kondo, K., Fujio, H., and Amano, T. (1982) *J. Biochem. (Tokyo)* **91**, 571–587
- Landon, M. (1977) *Methods Enzymol.* **47**, 145–149
- Lee, B., and Richards, F. M. (1971) *J. Mol. Biol.* **55**, 379–400
- Maley, G. F., Bellisario, R. L., Guarino, D. U., and Maley, F. (1979) *J. Biol. Chem.* **254**, 1288–1295
- Martin, S. L., Vincent, K. A., and Wilson, A. C. (1983) *J. Mol. Biol.* **164**, 513–528
- Nakai, N., Lai, C. Y., and Horecker, B. L. (1974) *Anal. Biochem.* **58**, 563–570
- Osserman, E. F., Canfield, R. E., and Beychok, S. (eds) (1974) *Lysozyme*, Academic Press, New York
- Ozols, J., and Gerard, C. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74**, 3725–3729
- Pahud, J.-J., and Widmer, F. (1982) *Biochem. J.* **201**, 661–664
- Perutz, M. F., Bauer, C., Gros, G., Leclercq, F., Vandecasserie, C., Schnek, A. G., Braunitzer, G., Friday, A. E., and Joysey, K. A. (1981) *Nature (Lond.)* **291**, 682–684
- Prager, E. M., and Wilson, A. C. (1971) *J. Biol. Chem.* **246**, 7010–7017
- Prager, E. M., Fowler, D. P., and Wilson, A. C. (1976) *Evolution* **30**, 637–649
- Smith-Gill, S. J., Rupley, J. A., Pincus, M. R., Carty, R. P., and Scheraga, H. A. (1984) *Biochemistry* **23**, 993–997
- Smith-Gill, S. J., Wilson, A. C., Potter, M., Prager, E. M., Feldmann, R. J., and Mainhart, C. R. (1982) *J. Immunol.* **128**, 314–322
- Tarr, G. E., Beecher, J. F., Bell, M., and McKean, D. J. (1978) *Anal. Biochem.* **84**, 622–627
- White, T. J. (1976) Ph.D. dissertation, University of California, Berkeley
- Wilson, A. C., Carlson, S. S., and White, T. J. (1977) *Annu. Rev. Biochem.* **46**, 573–639

Additional references are found on p. 11625.

Supplementary Material to "Stomach Lysozymes of Ruminants
II. Amino Acid Sequence of Cow Lysozyme 2 and
Immunological Comparisons with Other Lysozymes"

Pierre Jollès, Françoise Schoentgen, Jacqueline Jollès, Deborah E. Dobson, Ellen M. Prager, and Allan C. Wilson

EXPERIMENTAL PROCEDURES

All experimental procedures described in the Supplementary Material are those employed by the Berkeley laboratory. For determination of amino acid compositions, amino-terminal sequencing, and gel filtration, procedures which were conducted in Paris as well as in Berkeley, details of the methods used by the Jollès laboratory appear in the main text.

Electrophoresis

Denaturing gel electrophoresis through 15% polyacrylamide with sodium dodecyl sulfate as the denaturant was carried out as described by Dobson et al. (1984). Following the procedure of Laemmli (1970), a standard curve relating size to electrophoretic mobility for proteins of known molecular weight was constructed and used to estimate the molecular weight of cow stomach lysozymes.

Gel Filtration

A Sephadex G-75 column (57 x 2 cm, 180 ml) equilibrated in 0.2% acetic acid was calibrated with proteins of known molecular weight. Samples of purified cow lysozyme 2 were chromatographed to determine elution position, and molecular weight was calculated from the standard curve. The data were analyzed as described by Fischer (1969): for each protein, the retardation coefficient K_{av} was calculated as $K_{av} = (V_e - V_0)/(V_e - V_0)$, where V_e = measured elution volume, V_0 = void volume, and V_e = total column volume.

Extinction Coefficient

Purified, lyophilized lysozymes were weighed, dissolved in 0.05 M sodium phosphate, pH 7, and the absorbance at 280 nm measured with a Cary spectrophotometer. Chicken lysozyme α served as a control, and all measured values were normalized to an $E_{280}^{1\%}$ of 26.7 for chicken lysozyme α (Amheim et al., 1969, and references therein).

Amino Acid Analysis

Lysozymes were subjected to two successive rounds of reduction and carboxymethylation based on the procedure described by Amheim et al. (1969). The denaturant was 5 M guanidine hydrochloride and the solutions were dialyzed against 1 M ammonium hydroxide rather than against water and then lyophilized. The proteins were hydrolyzed under vacuum in constant-boiling HCl at 105–110°C for 24, 48, or 72 h, and the amino acid compositions were determined with a Beckman 121 automatic amino acid analyzer. Serine and threonine were determined by extrapolation to zero time of hydrolysis, the valine and isoleucine values were taken from the 72-h hydrolysates, and 1/2-cysteine was determined as S-carboxymethylcysteine. Tryptophan was measured by the procedure of Edelhoch (1967) with 5 M guanidine hydrochloride as the denaturant. Chicken lysozyme served as a control.

Amino-Terminal Sequence Analysis

The method of Edman (1970) as described by Ibrahim (1977) was used. Degradations were carried out on 100–200 nanomoles of protein. The phenylthiohydantoin derivatives were hydrolyzed to the free amino acids in constant-boiling HCl at 140°C for 15 h under vacuum. The amino acids were identified with a Beckman automatic amino acid analyzer as described above. Chicken lysozyme was included as a control.

Carbohydrate Analysis

Sufficient material was used to detect one residue or less of sugar per molecule of protein. The procedures outlined by Amheim et al. (1969) were followed to detect neutral sugars and sialic acid; amino sugars were detected in acid-hydrolyzed samples using a Beckman amino acid analyzer, with glucosamine as the standard.

Antisera, Antigens, and Immunological Methods

The rabbit antisera to baboon and human lysozymes (Hanke et al., 1973) and rat lysozyme (Prager et al., 1978) have been described before. Antisera to cow stomach lysozymes were produced essentially by the method of Prager and Wilson (1971). Two groups of four Dutch-Belted rabbits were immunized, one with pure cow lysozyme 1 and the other with pure cow lysozyme 2; on day 1, each rabbit received 0.2 mg of lysozyme in supplemented Freund's complete adjuvant. Series of intravenous booster injections were begun on days 47, 131, 192, and 235; each rabbit received 0.2 mg of lysozyme per injection. The antisera used in this work were from post-boost bleedings taken after 7 months of immunization. The four individual antisera were pooled as described by Champion et al. (1974). For use as antigens in both micro-complement fixation and immunodiffusion tests, extracts of stomach and other tissues containing low concentrations of lysozyme were sometimes lyophilized and redissolved in isotonic buffer (Champion et al., 1974; Hornbeck and Wilson, 1984).

Immunodiffusion was done as described by Prager et al. (1976). To measure antigen concentrations, wells (2 mm in diameter), placed 2 mm apart, received undiluted antiserum or various dilutions of antigen. To detect weak reactions between antigen and antiserum, larger wells (5.5 mm in diameter) were sometimes used.

Quantitative micro-complement fixation (MCF) was done as described by Champion et al. (1974). The degree of antigenic difference in the MCF assay is given in terms of immunological distance, which is equal to 100 times the log of the factor by which the antiserum concentration must be raised for a heterologous antigen to produce a complement fixation curve whose peak height is equal to that produced by the homologous antigen (the immunogen). For series of the globular proteins lysozyme, ribonuclease, azurin, myoglobin, and serum albumin (Benjamin et al., 1984), a linear relationship has been observed between percent amino acid sequence difference and immunological distance, with a correlation coefficient of about 0.9 between these parameters. Besides estimating degree of antigenic difference by comparing peak heights, we estimated antigen concentrations by comparing the antigen dilutions required to get peak fixation.

RESULTS

Physical and Chemical Characterization of Cow Stomach Lysozymes

Molecular Weight

In agreement with the results of Pahud and Widmer (1982), the molecular weights of cow stomach lysozymes were found by two methods to be typical of lysozymes α . In denaturing polyacrylamide gels (Fig. 1), the values were: 14,600 (1), 14,500 (2), and 14,300 (3), the standard deviation for reference proteins in these experiments being 4%. The molecular weight of cow lysozyme 2 was found by gel filtration (Fig. 2) to be 15,000 $\pm$ 1400. For comparison, the molecular weight of cow lysozyme 2 calculated from the amino acid sequence is 14,398. Thus, these stomach lysozymes are smaller than cow milk lysozyme, whose molecular weight is 18,000 (Eitenmiller et al., 1976).

Extinction Coefficients

Like previously described lysozymes α , the stomach lysozymes of the cow have high extinction coefficients ($E_{280}^{1\%}$), namely 29.6 (1), 28.2 (2), and 23.0 (3). By contrast, the $E_{280}^{1\%}$ reported for cow milk lysozyme is much lower (7.1; Chandan et al., 1965), as is the case for lysozymes β of known amino acid sequence (Schoentgen et al., 1982).

Absence of Carbohydrate

No neutral sugars, amino sugars, or sialic acid were detected in cow stomach lysozymes, although control proteins gave the expected values. To date, no lysozyme has been found to contain any carbohydrate.

Amino-Terminal Sequences

The amino-terminal sequences of both cow stomach lysozymes 1 and 2 were found to be the same as that of chicken and human lysozymes α : Lys-Val-Phe.

Amino Acid Composition

In amino acid content, the cow stomach lysozymes resemble lysozymes α of known sequence (Jollès et al., 1979b) in having high 1/2-cysteine residues and a high proportion of tryptophan, which is consistent with their high extinction coefficients. The most notable difference from conventional lysozymes α is a low content of arginine (Table 1). The compositions support the view that the three cow stomach enzymes are closely related to one another and distantly, or not at all, related to cow milk lysozyme.

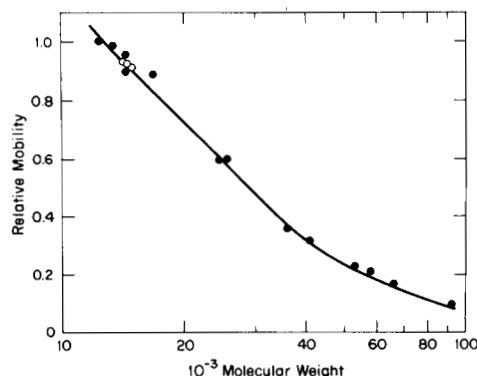


Fig. 1. Molecular weight determination of cow stomach lysozymes based on denaturing gel electrophoresis. A standard curve was constructed in which the mobility of each protein (relative to cytochrome c in 15% denaturing polyacrylamide gels) is plotted against the molecular weight. The relative mobilities of cow stomach lysozymes are shown as open circles. Protein standards in order of increasing size were cytochrome c , human lysozyme α , chicken lysozyme α , ribonuclease, myoglobin, chymotrypsinogen, gamma globulin, glyceraldehyde-3-phosphate dehydrogenase, alcohol dehydrogenase, glutamic dehydrogenase, catalase, bovine serum albumin, and phosphorylase a .

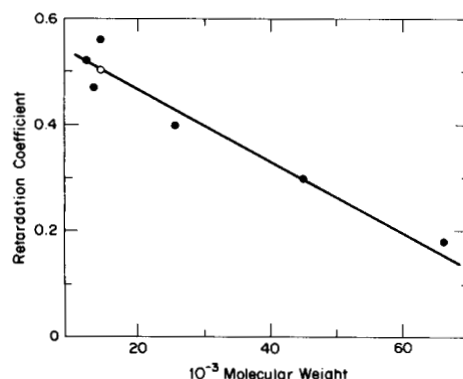


Fig. 2. Molecular weight of cow stomach lysozyme 2 based on gel chromatography. The indicated line was fitted by eye to the solid circles, which represent the values observed for proteins of known molecular weight: cytochrome c , chicken lysozyme α , ribonuclease, chymotrypsinogen, ovalbumin, and human serum albumin. The value of the retardation coefficient for cow lysozyme 2 is shown as an open circle.

TABLE 1
Amino acid composition of cow and chicken lysozymes

Amino acid	Residues per molecule ^a			Cow milk ^b	Chicken ^c
	1	2	3		
Cysteine (Cn)	7.8 (8)	7.8 (8)	8.0 (8)	6	7.8 (8)
Aspartic acid	16.5 (15)	16.8 (15)	16.0 (14)	22–23	20.6 (21)
Threonine	7.9 (8)	8.5 (8)	8.4 (8)	8	6.6 (7)
Serine	11.4 (11)	12.9 (13)	12.8 (13)	8	9.1 (10)
Glutamic acid	11.2 (11)	10.3 (10)	10.7 (10)	14	5.4 (5)
Proline	2.3 (2)	2.2 (2)	3.1 (3)	9	2.0 (2)
Glycine	9.1 (9)	8.0 (8)	8.1 (8)	11	10.8 (12)
Alanine	10.3 (10)	10.1 (10)	9.6 (10)	5	11.6 (12)
Valine	9.8 (9)	10.0 (9)	10.1 (9)	6	6.2 (6)
Methionine	1.1 (1)	1.0 (1)	1.6 (2)	2	1.9 (2)
Isoleucine	4.8 (5)	4.7 (5)	4.4 (5)	10	5.6 (6)
Leucine	9.1 (9)	9.1 (9)	8.9 (9)	5–6	7.7 (8)
Tyrosine	4.6 (5)	4.5 (5)	4.2 (5)	7	2.7 (3)
Phenylalanine	2.1 (2)	2.1 (2)	1.8 (2)	7	2.8 (3)
Histidine	2.3 (3)	2.3 (3)	3.0 (4)	7	0.7 (1)
Lysine	9.8 (12)	9.8 (12)	10.3 (12)	10	5.0 (6)
Arginine	3.2 (3)	3.2 (3)	3.8 (4)	15	10.2 (11)
Tryptophan	6.3 (6)	6.1 (6)	4.7 (5)	1	6.0 (6)
Total	129	129	131 ^d	154	129

^a Integral values are given in parentheses. Those for cow lysozyme 2 were taken from the sequence and the integral values for cow 1 and 3 proportionally adjusted for those amino acids where the values for cow 2 determined here disagree with those determined from the complete sequence analysis.

^b From Eitenmiller et al. (1976).

^c The experimental values were determined as part of this study and the integral values taken from the sequence (Canfield, 1963).

^d The slightly higher number of residues (131 compared to 129) is not considered indicative of a length difference in cow 3; it results from the steps involved in rounding off to integral values. The experimental values of residues per molecule total 129.6, 129.4, and 129.5 for cow 1, 2, and 3, respectively. See also Molecular Weight, above, and Fig. 2 of the Supplementary Material of Dobson et al. (1984).

Determination of the Sequence of Cow Stomach Lysozyme 2

Peptide Purification

The chromatographic separations of the peptides generated from reduced and alkylated cow lysozyme subjected to citraconylation and tryptic digestion (Fig. 3) or to cyanogen bromide cleavage (Fig. 4) are shown.

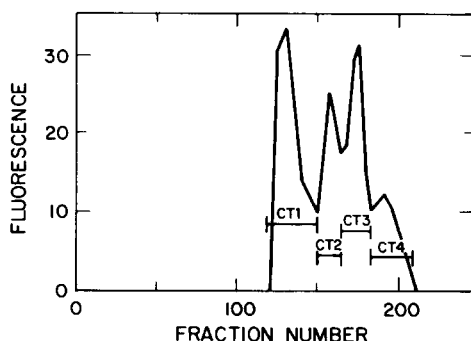


Fig. 3. Filtration on Sephadex G-25 (200 x 2 cm) of the tryptic digest of 20 mg of reduced, alkylated, and citraconylated bovine stomach lysozyme. The eluent was 30% acetic acid and the fraction size 2 ml. Fluorescence (arbitrary units) was measured by the emission at 480 nm following excitation at 390 nm.

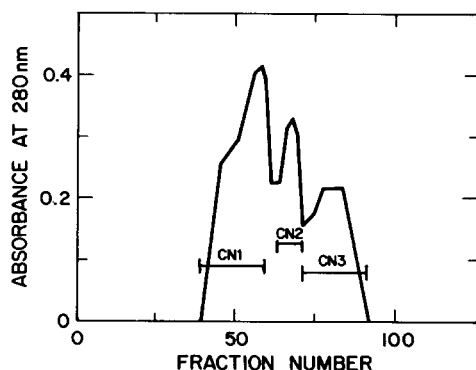


Fig. 4. Filtration on Sephadex G-50 (240 x 1.2 cm) of the cyanogen bromide fragments of 15 mg of reduced and alkylated bovine stomach lysozyme. The eluent was 30% acetic acid and the fraction size 1.5 ml. Peak CN3 contained two peptides, CN3a and CN3b, which were purified by paper chromatography prior to sequencing.

The tryptic digest of noncitraconylated, reduced, and alkylated bovine lysozyme was chromatographed on Sephadex G-25 (Fig. 5). Five peaks (T1 to T5) were studied and five peptides derived from within T1 sequenced (cf. Tables 4 and 6, below). Three of these peptides were first purified by preparative paper chromatography (T1a and T1b) or electrophoresis (T2). Peptide T5 was greatly retarded on the Sephadex column because two of its seven residues are tryptophans. The *Staphylococcus aureus* V8 protease digest of the reduced and alkylated lysozyme was likewise chromatographed on a 270 x 1.4 cm column of Sephadex G-25; peptide SP2 was isolated, characterized, and sequenced as reported in the main text and in Tables 4 and 6 below.

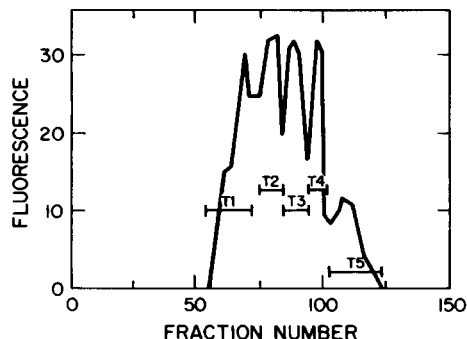


Fig. 5. Filtration on Sephadex G-25 (270 x 1.4 cm) of the tryptic digest of 15 mg of reduced and alkylated bovine stomach lysozyme. The eluent was 30% acetic acid and the fraction size 2 ml. Fluorescence (arbitrary units) was measured by the emission at 480 nm following excitation at 390 nm.

Peptide Properties

Tables 2-4 give the amino acid compositions, paper chromatographic and electrophoretic mobilities, and yields of the peptides isolated and sequenced in the course of determining the complete cow lysozyme sequence. Residue number localizations in these tables are given in accordance with Fig. 1 of the main text, which employs the numbering system of the chicken lysozyme sequence. For this reason the total number of amino acid residues per peptide (Tables 2-4) does not agree with the number one would calculate from the localization for peptides CN1, CN2 (Table 3) and T1b, T2 (Table 4).

TABLE 2

Properties of the tryptic peptides of reduced, alkylated, and citraconylated bovine stomach lysozyme: amino acid composition, R_f , m at pH 6.5, and yield. Values are based on hydrolysis for 18 h. The number of residues per peptide is given in parentheses to the nearest integer. Values obtained by linear extrapolation to zero time (Ser and Thr) or after 48 and 72 h of hydrolysis were taken into consideration in arriving at the integral values. 1/2-Cystine was determined as S-carboxymethylcysteine. Tryptophan was identified by the Ehrlich reaction and during the automated degradation. $m = 0$ for Gly, +1 for Arg, and -1 for cysteic acid. ND, not determined.

Amino acid	Peptide				Total
	CT1	CT2	CT3	CT4	
Aspartic acid	11.5 (13)	2.0 (2)			15
Threonine	6.2 (7)	1.0 (1)			8
Serine	9.0 (11)	2.0 (2)			13
Glutamic acid	6.2 (7)	0.9 (1)	1.0 (1)	1.0 (1)	10
Proline	2.0 (2)				2
Glycine	6.0 (7)	1.0 (1)			8
Alanine	7.5 (9)		0.9 (1)	0.9 (1)	10
Valine	5.0 (6)	1.6 (2)			9
1/2-Cystine	4.8 (6)	0.9 (1)		1.0 (1)	8
Methionine	0.8 (1)				1
Isoleucine	3.8 (5)				5
Leucine	7.0 (7)	1.0 (1)		1.0 (1)	9
Tyrosine	4.0 (4)	0.8 (1)			5
Phenylalanine	0.7 (1)		0.6 (1)		2
Lysine	10.1 (11)		0.8 (1)		12
Histidine	2.0 (2)	1.0 (1)			3
Arginine	1.0 (1)		0.9 (1)	1.0 (1)	3
Tryptophan	+ (6)				6
Total	106	13	5	5	129
R_f	ND	0.35	0.36	0.30	
Mobility, m	ND	-0.58	+0.40	0	
Yield (%)	30	42	61	50	
Localization, residue numbers	11-116	117-129	1-5	6-10	

TABLE 3

Amino acid compositions, R_f and m values, and yields of the cyanogen bromide fragments of reduced and alkylated bovine stomach lysozyme. Values were measured as described in Table 2. Methionine was identified as homoserine. ND, not determined.

Amino acid	Peptide		Total	Peptide	
	CN1	CN2		CN3a	CN3b
Aspartic acid	9.5 (11)	4.2 (4)	15	3.4 (4)	2.2 (2)
Threonine	5.5 (6)	2.0 (2)	8	1.0 (1)	1.0 (1)
Serine	8.2 (9)	3.9 (4)	13	2.0 (2)	2.7 (3)
Glutamic acid	6.0 (6)	3.8 (4)	10	1.0 (1)	0.9 (1)
Proline	1.5 (2)		2	0.6 (1)	
Glycine	5.2 (6)	2.1 (2)	8	1.8 (2)	1.1 (1)
Alanine	4.0 (4)	4.6 (6)	10	1.0 (1)	
Valine	3.0 (4)	3.5 (5)	9	1.9 (2)	1.5 (2)
1/2-Cystine	4.5 (5)	2.9 (3)	8	2.2 (3)	1.7 (2)
Methionine	+ (1)		1	+ (1)	
Isoleucine	1.4 (2)	1.8 (3)	5		
Leucine	7.2 (8)	1.4 (1)	9	0.8 (1)	0.8 (1)
Tyrosine	4.0 (4)	1.0 (1)	5		0.7 (1)
Phenylalanine	1.4 (2)		2		
Lysine	7.1 (8)	3.2 (4)	12	0.9 (1)	0.8 (1)
Histidine	0.8 (1)	1.8 (2)	3	1.0 (1)	1.6 (2)
Arginine	1.8 (2)	1.0 (1)	3		0.7 (1)
Tryptophan	+ (4)	+ (2)	6	+ (1)	
Total	85	44	129	22	18
R_f	ND	ND		0.10	0.15
Mobility, m	ND	ND		-0.04	0
Yield (%)	56	62		20	20
Localization, residue numbers	1-84 ^a	85-129 ^a		63-84	112-129

^aSee the text portion of Peptide Properties for an explanation of why the total number of residues differs from that calculated from the localization.

TABLE 4

Amino acid compositions, R_f and m values, and yields of the tryptic (noncitraconylated) and V8 protease peptides of reduced and alkylated bovine stomach lysozyme used for determination of the complete sequence. Values were measured as described in Table 2. ND, not determined.

Amino acid	Peptide					
	T1a	T1b	T2	T3	T5	SP2
Aspartic acid	4.2 (4)	4.1 (4)			2.1 (2)	6.0 (6)
Threonine	0.7 (1)	1.3 (2)	1.0 (1)			1.8 (2)
Serine	1.9 (2)	3.2 (4)	0.8 (1)	1.0 (1)		3.5 (4)
Glutamic acid	2.0 (2)	1.8 (2)	2.0 (2)			1.9 (2)
Proline	0.8 (1)	0.8 (1)				1.0 (1)
Glycine	1.1 (1)	1.0 (1)	0.9 (1)		1.2 (1)	2.7 (3)
Alanine	1.7 (2)	1.0 (1)	1.8 (2)			0.8 (1)
Valine	1.6 (2)		1.8 (2)			1.6 (2)
1/2-Cystine	1.8 (2)			0.7 (1)	1.0 (1)	2.6 (3)
Methionine	0.9 (1)					
Isoleucine	0.7 (1)	1.5 (2)	1.6 (2)			1.4 (2)
Leucine	1.0 (1)					
Tyrosine		1.6 (2)				0.8 (1)
Phenylalanine		0.6 (1)				0.6 (1)
Lysine	1.0 (1)	1.0 (1)	1.6 (2)		1.0 (1)	1.6 (2)
Histidine	1.0 (1)			1.0 (1)		1.0 (1)
Arginine				1.0 (1)		
Tryptophan			+ (2)		+ (2)	+ (2)
Total	22	21	15	4	7	33
R_f	0.20	0.25	0.25	0.12	0.14	ND
Mobility, m	-0.38	-0.27	0	+0.68	0	ND
Yield (%)	30	33	16	22	43	30
Localization, residue numbers	69-90	42-61 ^a	97-112 ^a	113-116	62-68	50-82

^aSee Footnote a to Table 3.

Automated Edman Degradation

Table 5 details the course of determination of the first 46 residues of cow lysozyme by automated sequential degradation.

TABLE 5
Automated sequential degradation of reduced and alkylated bovine stomach lysozyme

Step	Amino acid	Identification methods ^a	Yield ^b (nmol)	Step	Amino acid	Identification methods	Yield (nmol)
1	Lys	T, H	94	24	Ser	T, H	NC
2	Val	T, H	98	25	Leu	T, H	13
3	Phe	T, H	87	26	Ala	T, H	15
4	Glu	T, H	62	27	Asn	T, H	4
5	Arg	T, H	18	28	Trp	T, H	3
6	Cys	T, H	NC ^c	29	Leu	T, H	16
7	Glu	T, H	48	30	Cys	T, H	NC
8	Leu	T, H	77	31	Leu	T, H	18
9	Ala	T, H	50	32	Thr	T, H	NC
10	Arg	T, H	10	33	Lys	T, H	3
11	Thr	T, H	14	34	Trp	T, H	2
12	Leu	T, H	54	35	Glu	T, H	6
13	Lys	T, H	22	36	Ser	T, H	NC
14	Lys	T, H	27	37	Ser	T, H	NC
15	Leu	T, H	21	38	Tyr	T, H	1
16	Gly	T, H	14	39	Asn	T, H	1
17	Leu	T, H	19	40	Thr	T, H	NC
18	Asp	T, H	15	41	Lys	T, H	1
19	Gly	T, H	12	42	Ala	T, H	3
20	Tyr	T, H	14	43	Thr	T, H	NC
21	Lys	T, H	12	44	Asn	T, H	0.5
22	Gly	T, H	12	45	Tyr	T, H	1
23	Val	T, H	16	46	Asn	T, H	0.3

^aThe phenylthiohydantoin amino acids were identified by thin-layer chromatography (T) and by high performance liquid chromatography (H).

^bRepetitive yield, 94%.

^cCysteine was recovered as S-carboxamidomethylcysteine.

^dNC, not calculated.

Thirteen peptides were sequenced in the course of determination of the identity and order of all 129 amino acids of cow lysozyme 2. Table 6, in a fashion analogous to Table 5, details the course of Edman degradation of these peptides. The repetitive yields obtained during automatic sequencing of the longer peptides (cf. Fig. 1 of the main text and Tables 2-4 here) were as follows: CT1, 76%; CH2, 88%; T1a, 89%; T1b, 75%.

TABLE 6

Automated sequential degradation of peptides of bovine stomach lysozyme. The peptides are identified as to location in the molecule in Fig. 1 of the main text and in Tables 2-4 here. The number of nanomoles subjected to degradation is indicated immediately after the peptide name. The phenylthiohydantoin amino acids were identified and quantitated by high performance liquid chromatography. Yields are given in nanomoles. Cysteine was recovered as S-carboxamidomethylcysteine. AA, amino acid.

Step	AA	Yield	Step	AA	Yield	Step	AA	Yield	Step	AA	Yield
<u>CT1-110 nmol</u>			<u>CH2-90 nmol</u>			<u>CH3b (continued)</u>			<u>T2-40 nmol</u>		
1	Thr	72	1	Glu	49	13	Val	23	1	Lys	20
2	Leu	89	2	Asn	24	14	Glu	15	2	Ile	25
3	Lys	40	3	Asp	38	15	Gly	9	3	Val	25
4	Lys	45	4	Ile	35	16	Cys	6	4	Ser	4
5	Leu	50	5	Ala	36	17	Thr	1	5	Glu	6
6	Gly	35	6	Lys	28	18	Leu	3	6	Gln	5
7	Leu	40	7	Ala	17	19	Gly	1	7	Gly	2
8	Asp	30	8	Val	16	20	Val	12	8	Ile	2
9	Gly	12	9	Ala	16	21	Thr	5	9	Thr	0.3
10	Tyr	21	10	Cys	5	1	Thr	5	10	Ala	3
11	Lys	7	11	Ala	10	2	Pro	14	11	Trp	1
12	Gly	5	12	Lys	7	3	Asn	15	12	Val	2
13	Val	6	13	Lys	7	4	Ala	24	13	Ala	2
14	Ser	1	14	Ile	6	5	Val	23	14	Trp	0.5
15	Leu	4	15	Val	5	6	Asp	16	15	Lys	0.2
16	Ala	3	16	Ser	1	7	Gly	13			
17	Asn	0.5	17	Glu	3	8	Cys	9			
<u>CT2-80 nmol</u>			18	Gln	2	9	His	12	<u>T3-40 nmol</u>		
1	Asp	52	19	Gly	1	10	Val	13	1	Ser	12
2	His	24	20	Ile	2	11	Ser	4	2	His	6
3	Asp	47	21	Thr	1	12	Cys	4	3	Cys	6
4	Val	45	22	Ala	2	13	Ser	4	4	Arg	14
5	Ser	9	<u>CH3a-30 nmol</u>			14	Glu	12	<u>T5-70 nmol</u>		
6	Ser	5	1	Trp	12	15	Leu	11	1	Trp	20
7	Tyr	36	2	Cys	12	16	Met	3	2	Trp	15
8	Val	32	3	Asn	10	17	Glu	4	3	Cys	20
9	Glu	21	4	Ala	15	18	Asn	2	4	Asn	17
10	Gly	14	5	Asp	10	19	Ala	2	5	Asp	16
11	Cys	10	6	Lys	12	20	Ile	4	6	Gly	5
12	Thr	2	7	Lys	12	21	Ala	3	7	Lys	2
13	Leu	4	8	Thr	5	22	Lys	0.5	<u>SP2-70 nmol</u>		
<u>CT3-50 nmol</u>			9	Pro	7	1	Ala	48	1	Ser	10
1	Lys	28	10	Ala	12	2	Thr	13	2	Thr	11
2	Val	35	11	Val	8	3	Asn	28	3	Asp	42
3	Phe	20	12	Asp	4	4	Tyr	35	4	Tyr	40
4	Glu	16	13	Gly	2	5	Asn	25	5	Gly	30
5	Arg	8	14	Cys	1	6	Pro	15	6	Ile	36
<u>CT4-40 nmol</u>			15	His	1	7	Ser	4	7	Phe	34
1	Cys	14	16	Val	1	8	Ser	2	8	Gln	27
2	Glu	28	8	Pro	7	9	Leu	14	9	Ile	24
3	Lys	25	9	Asn	7	10	Ser	2	10	Asn	19
4	Ala	20	10	Ala	12	11	Thr	1	11	Ser	4
5	Arg	6	11	Val	8	12	Asp	9	12	Lys	13
<u>CH3b-90 nmol</u>			12	Asp	4	13	Tyr	8	13	Trp	8
1	Lys	60	13	Gly	2	14	Gly	4	14	Trp	7
2	Glu	22	14	Cys	30	15	Ile	4	15	Cys	7
3	His	45	15	Asp	25	16	Phe	4	16	Asn	5
4	Cys	30	16	Phe	4	17	Gln	3	17	Asp	6
5	Arg	25	17	His	30	18	Ile	2	18	Gly	3
6	Asp	36	18	Ile	2	19	Asn	1	19	Lys	4
7	His	30	19	Val	35	20	Ser	0.2	20	Thr	8
8	Asp	37	20	Ser	4	21	Lys	2	21	Pro	1
9	Val	35	21	Lys	2	22	Asn	0.5	22	Asn	0.5
10	Ser	4	22	Ser	30	23	Ala	1	23	Ala	1
11	Ser	2	<u>CH3c-90 nmol</u>			24	Val	1	24	Val	1
12	Tyr	30	1	Lys	60	25	Val	1	25	Val	1

Evolutionary Relationships

Fig. 6 shows the evolutionary relationships and divergence times among the ungulates whose stomach lysozymes were tested with antisera to cow stomach lysozymes 1 and 2 (cf. Table 1 of the main text).

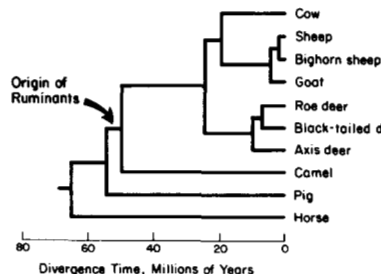


Fig. 6. Evolutionary relationships of ruminants and related ungulates. Divergence times are based on organismal (Romer, 1966; Novacek, 1982; Savage and Russell, 1983) and biochemical (Carlson et al., 1978; V. M. Sarich and A. Bennett, personal communication) evidence.

DISCUSSION

Phylogenetic Tree for Mammalian Lysozymes

To build the tree shown in the main text, relating the four mammalian lysozymes of known sequence (Fig. 2), we used three highly divergent bird lysozymes as an outgroup. Although 79 amino acid positions are variable among these seven lysozymes, only the 15 positions listed in Table 7 proved to be relevant as regards determining the order of branching of the mammalian lineages by the parsimony method. In each of these 15 cases, the number of mutations required to account for the amino acid differences depended on the order of branching in the tree (Table 7). The most parsimonious tree (A in Table 7) accounted for the amino acid differences at these 15 positions with 48 mutations at the DNA level. The alternative trees (B, C, and D) differ in branching order and required from 49 to 56 mutations (Table 7). At each of the remaining 64 of the 79 variable positions, the number of mutations required to account for the amino acid differences was independent of the tree topology (A, B, C, or D).

TABLE 7

Phylogenetically informative amino acid differences among mammal and bird lysozymes

Position	Amino acid							Minimum number of events per tree ^a			
	Cow	Baboon	Human	Rat	Chicken	Duck	Cha ^b	A	B	C	D
3	Phe	Phe	Phe	Tyr	Phe	Tyr	Tyr	3	2	3	3
15	Leu	Leu	Leu	Asn	His	Leu	Tyr	4	4	5	4
21	Lys	Arg	Arg	Tyr	Arg	Arg	Arg	5	4	5	4
34	Trp	Trp	Trp	His	Phe	Tyr	Tyr	6	4	6	6
49	Glu	Gln	Arg	Gln	Gly	Gly	Gly	3	4	4	4
62	Trp	Tyr	Tyr	Tyr	Trp	Trp	Trp	2	4	4	4
73	Val	Val	Val	Lys	Arg	Lys	Lys	5	3	5	5
85	Glu	Gln	Gln	Ser	Ser	Arg	Gly	4	1	2	5
98	Ile	Val	Val	Val	Ile	Ile	Ile	1	2	2	2
112	Lys	Arg	Arg	Gln	Arg	Arg	Arg	3	3	2	3
113	Ser	Asn	Asn	Arg	Asn	Asn	Lys	4	4	3	4
117	Asp	Asn	Asn	Asn	Gly	Gly	Gly	2	3	3	3
124	Val	Val	Val	Ile	Ile	Ile	Ile	2	1	2	2
128	Thr	Gly	Gly	Gly	Arg	Arg	Lys	3	4	4	4
129	Leu	Val	Val	Val	Leu	Leu	Leu	1	2	2	2
Total								48	49	53	56

^aTree A is shown in Fig. 2 of the main text. Tree B differs in branching order by allyling the primate and cow lineages most closely. Tree C associates the rat and cow lineages most closely. Tree D assumes a three-way split among the primate, rat, and cow lineages. All four trees preserve the branching order shown by Jollès et al. (1976, 1979a) for the bird lysozymes, in which the duck and chicken lineages are more closely related than either is to the chachalaca lineage.

^bCha, chachalaca.

The next step was to assign the mutations at each of the 79 positions to particular lineages on tree A. At some positions, this was a simple task, as pointed out, for example, in the main text for the deletion of proline at position 102 on the cow lineage. At other positions, there is more than one way of apportioning the mutations. For example, at position 3 (Table 7), there are three alternative solutions, each requiring three mutations. First, there could be three T → A substitutions, on the rat, duck, and chachalaca lineages, in each case generating a phenylalanine to tyrosine replacement. Second, there could be three A → T substitutions, on the lineages leading to the chicken, the cow, and the common ancestor of primates. Third, there could be one A → T change on the chicken lineage, one T → A change on the rat lineage, and a T → A change on the lineage lying between the common ancestor of mammals and the common ancestor of birds. The three alternative solutions are then averaged and the resulting estimates of numbers of mutations per lineage at position 3 are as follows: cow (0.33), baboon (0), human (0), rat (0.57), chicken (0.67), duck (0.33), chachalaca (0.33), common ancestor of primates (0.33), common ancestor of mammals to common ancestor of birds (0.33). By proceeding in this manner, we apportioned the 48 mutations at the 15 informative positions and the 151 mutations at the 64 uninformative positions on the tree for seven species. The results for the mammalian part of the tree appear in Fig. 2 of the main text.

REFERENCES

- Armstrong, N., Prager, E. M., and Wilson, A. C. (1969) *J. Biol. Chem.* 244, 2085-2094.
 Canfield, R. E. (1963) *J. Biol. Chem.* 238, 2698-2707.
 Carlson, S. S., Wilson, A. C., and Maxson, R. D. (1978) *Science* 200, 1183-1185.
 Chander, R. C., Parry, R. M., Jr., and Shahani, K. M. (1965) *Biochim. Biophys. Acta* 110, 389-398.
 Edelhoch, H. (1967) *Biochemistry* 6, 1948-1954.
 Edman, P. (1970) In *Protein Sequence Determination: A Sourcebook of Methods and Techniques* (Needleman, S. B., ed) pp. 221-255, Springer-Verlag, New York.
 Eitenmiller, R. R., Friend, B. A., and Shahani, K. M. (1975) *J. Dairy Sci.* 59, 834-839.
 Fischer, L. (1969) In *Laboratory Techniques in Biochemistry and Molecular Biology*, Vol. 1 (Work, T. S., and Work, E., eds) pp. 151-396, American Elsevier, New York.
 Hanke, N., Prager, E. M., and Wilson, A. C. (1973) *J. Biol. Chem.* 248, 2824-2828.
 Hornbeck, P. V., and Wilson, A. C. (1984) *Biochemistry* 23, 998-1002.
 Ibrahim, I. M. (1977) Ph.D. Thesis, University of California, Berkeley.
 Jollès, J., Schoentgen, F., Jollès, P., Prager, E. M., and Wilson, A. C. (1976) *J. Mol. Evol.* 8, 59-78.
 Laemmli, U. K. (1970) *Nature* 227, 680-685.
 Novacek, M. J. (1982) In *Macromolecular Sequences in Systematic and Evolutionary Biology* (Goodman, M., ed) pp. 3-41, Plenum Press, New York.
 Prager, E. M., and Wilson, A. C. (1971) *J. Biol. Chem.* 246, 5978-5989.
 Prager, E. M., Wilson, A. C., Péron, J.-P., and Jollès, P. (1978) *Immunochimistry* 15, 577-583.
 Romer, A. S. (1966) *Vertebrate Paleontology*, University of Chicago, Chicago.
 Savage, D. E., and Russell, D. E. (1983) *Mammalian Paleofaunas of the World*, Addison-Wesley, London.
 Schoentgen, F., Jollès, J., and Jollès, P. (1982) *Eur. J. Biochem.* 123, 489-497.