Critical Reviews in Biochemistry and Molecular Biology, 32(4):255–306 (1997)

Molecular Divergence of Lysozymes and α-Lactalbumin*

*Pradman K. Qasba[1,**] and Soma Kumar[2]*

[1]Structural Glycobiology Section, Laboratory of Experimental and Computational Biology, National Cancer Institute, N. I. H., P.O. Box B, Bldg. 469, Room 221, Frederick, MD 21702–1201; [2]Department of Chemistry, Georgetown University, 37th and O Streets, N. W., Washington, D. C. 20057–1227

* The writing of this review was initiated during Dr. Kumar's sabbatical in the Structural Glycobiology Section of the Laboratory of Experimental and Computational Biology, NCI. Just before this manuscript was to be sent for publication another review by McKenzie (1996) came to our attention. The thrust of that review is quite different from ours.

** To whom correspondence should be addressed. email address: qasba@helix.nih.gov; Tel. no. 301–846–1933 or –1934; Fax no. 301–846–7149.

Referee: Dr. K. Brew, University of Miami

ABSTRACT: The vast number of proteins that sustain the currently living organisms have been generated from a relatively small number of ancestral genes that has involved a variety of processes. Lysozyme is an ancient protein whose origin goes back an estimated 400 to 600 million years. This protein was originally a bacteriolytic defensive agent and has been adapted to serve a digestive function on at least two occasions, separated by nearly 40 million years. The origins of the related goose type and T4 phage lysozyme that are distinct from the more common C type are obscure. They share no discernable amino acid sequence identity and yet they possess common secondary and tertiary structures. Lysozyme C gene also gave rise, after gene duplication 300 to 400 million years ago, to a gene that currently codes for α-lactalbumin, a protein expressed only in the lactating mammary gland of all but a few species of mammals. It is required for the synthesis of lactose, the sugar secreted in milk. α-Lactalbumin shares only 40% identity in amino acid sequence with lysozyme C, but it has a closer spatial structure and gene organization. Although structurally similar, functionally they are quite distinct. Specific amino acid substitutions in α-lactalbumin account for the loss of the enzyme activity of lysozyme and the acquisition of the features necessary for its role in lactose synthesis. Evolutionary implications are as yet unclear but are being unraveled in many laboratories.

KEY WORDS: α-lactalbumin, lysozymes, galactosyltransferase, evolution, molecular divergence, lactose synthetase, protein-protein interactions, gene organizations, sequence similarity.

Abbreviations: HEW, hen egg white; **GEW,** goose egg white; **LZ,** lysozyme, **α-LA,** α-lactalbumin; **UDP-Gal,** uridine diphosphate galactose; **NAG** or **GlcNAc,** *N*-acetyl-glucosamine; **GalNAc,** *N*-acetyl-galactosamine; **GT** or **galactosyltransferase,** UDP-galactose: *N*-acetyl-glucosamine β–1,4-galactosyl tranferase (E.C. 2.4.90).

1040-9238/97/$.50

I. INTRODUCTION

Evolution of the vast number of currently living, diverse, and complex organisms was made possible because the evolving species possessed the ability to acquire the means to perform an increasing number of daunting tasks essential for their survival in a continuously changing environment. This was achieved largely through the generation of a corresponding array of new proteins capable of performing the requisite functions. It is clear from the available amino acid/DNA sequences and the three-dimensional structures of proteins that a majority of the late-appearing proteins were not created *de novo* but were the products of new genes created from a limited number of ancestral ones (Doolittle, 1985; Hartl, 1991). Two features are of prime importance in the generation of a new functional protein: (1) the stability of the three-dimensional structure that the linear polymer of amino acids assumes on folding, and (2) the presence of specific and appropriately placed amino acid residues in the native conformation of its molecule (Richardson and Richardson, 1989). It is already apparent that during the evolution of currently existing proteins their tertiary structures were more highly conserved than were the amino acid sequences (Bajaj and Blundell, 1984; Overington et al., 1990).

The ancestral genes encoding the amino acid sequences of those proteins that perform a role common to all living organisms, the so-called housekeeping proteins, have been subjected to continuous and generally conservative mutations. This has resulted in genes that produce families of species-specific proteins (Chothia and Lesk, 1987; Doolittle, 1990; Chothia, 1992; Creighton, 1993). Each family has retained sequence homology as well as structural and functional identities throughout the evolutionary process. Examples of such proteins are the glycolytic enzymes, the well-studied cytochrome *C* and globins. Although these proteins perform the same function in all species many of them have been adapted to meet particular needs of individual species. Minor changes in amino acid sequences result in subtle alterations in the tertiary structure of the proteins, which in turn affect their interactions with specific ligands. Hemoglobin provides a good example of a protein whose molecule has undergone adaptation in many ways to suit the special oxygen transport requirements imposed by their respective habitats on fishes, amphibia, reptiles, birds, and mammals (Perutz, 1984). Examples of other proteins that are known to have adapted to environmental changes have been provided by Gillespie (1991).

The list of functions that proteins perform in the present day eukaryotes, already impressive, continues to grow in length at an astonishing rate. A variety of natural processes have contributed to the evolution of this immense repertoire of extant proteins. Briefly, they are

A. Multifunctionality of Existing Proteins

Fortuitously, it seems, quite a few of the proteins that fulfilled specific roles in an earlier period of biological evolution possessed the capacity to perform additional functions, unnecessary then, but which were put to good use later. For instance, lens crystallins, a group of proteins needed in the eye for the refraction of light, have been established to constitute proteins that perform entirely different functions in other tissues. α-Crystallin, one of the major lens proteins of the vertebrate eyes also serves as a molecular chaperon (Horwitz, 1992; Sax and Piatigorsky, 1994). Many of the other

lens proteins have been reported to be identical to a variety of enzymes of more primitive origin (Piatigorsky and Wistow, 1991; Wistow, 1993). Evidently, the same gene is expressed in different tissues to produce proteins that serve different roles. A widely distributed enzyme, protein disulfide isomerase, discovered at first to be involved in the folding of many proteins has been shown to possess not only thioredoxin-like activity but other activities as well (Freedman et al., 1994; Ramasarma, 1994). It also contributes two of the four subunits of prolyl-4-hydroxylase (Koivu et al., 1987). Glyceraldehyde-3-phosphate dehydrogenase has been reported to act as an ADP-ribosylating enzyme (Pancholi and Fischetti, 1993) as well as a t-RNA binding protein (Singh and Green, 1993). A substantial increase in the functional capability of proteins is seen to have been achieved by simply taking advantage of their versatility (Ramasarma, 1994).

B. Gene Duplications to Generate Families of Proteins

Certain ancestral genes got duplicated one or more times. Each of these, subjected to independent mutations, resulted in families of proteins with altered properties even while retaining sequence homologies (McLachlan, 1987; Jörnvall et al., 1987; Fothergill-Gilmore, 1986). Alterations in amino acid sequences have also resulted in peptide chains that associate with each other to form polymeric proteins capable of regulatory functions. Hemoglobins are tetramers formed by the association of a pair each of two different peptide chains that undoubtedly evolved from a common ancestral myoglobin gene (Goodman et al., 1987). According to evidence currently available, all 20 aminoacyl-tRNA synthetases have evolved from only two unrelated genes (Nagel and Doolittle, 1991).

C. Generation of Proteins by the Extension of Amino- and Carboxy-Terminals

New genes were created by the simple addition of specific DNA sequences to the ancestral genes at the two ends. Human glutaminyl-tRNA synthetase contains 1440 amino acid residues, which is twice as many as in yeast and three times as many as the enzyme in bacteria. Its core stretch of 360 amino acid residues, positions 125 to 485, is concerned with its primary function and is homologous with the bacterial enzyme. The amino acid sequence at the carboxyl terminal consists of three repeats of a stretch of 57 amino acids followed by a stretch of polar amino acids. This motif is similar to that of the highly conserved region of the α subunit of the translation elongation factor, EF1 (Fett and Knippers, 1991). It is believed that the core amino acid sequence of the enzyme has been extended using sequences derived from EF1. The roles of the extensions of both ends of the core enzyme have yet to be established.

D. Generation of Proteins by Gene Duplication and Tandem Fusion

Ancestral DNA sequences, presumably giving rise to basic structural building blocks of a functional peptide, duplicated and got fused in tandem (Ohno, 1987; McLachlan, 1987). The new sequence now encoded a product that was better adapted for its function. *Peptococcus aerogenes* ferrodoxin

consists of 54 amino acids, which are two repeats of 27 amino acids each (Creighton, 1993). Similarly, calmodulin consists of essentially four similar sequences the result of two duplications of the same original sequence, followed by their fusion (Means, Putkey, and Epstein, 1988). Whey acidic protein, found in rodent milk, is a protein with two repeat domains. The arrangement of the half cystines and the amino acids of these are homologous with that of one of the two domains of a red sea turtle protease inhibitor (Dandekar et al., 1982).

E. New Proteins Generated by the Combination of Duplicated DNA Sequences in Different Ways

This process, widely referred to as exon shuffling, has created a variety of new genes by the simple fusion of copies of preexisting DNA sequences arranged differently. Proteins produced in this manner have sequence homology of segments found in others. Functionally, they are quite different. Examples of the evolution of proteins by this mechanism are many (Doolittle, 1985). An excellent illustrative example of a protein generated by this process is the low-density lipo-protein receptor. Its gene consists of 18 exons of which 13 have been established to be similar to those found in the genes of other known proteins (Südhoff et al., 1985).

F. Fusion of Functional Genes to Form Multifunctional Protein Complexes

Fatty acid synthase, Type I, found in animal tissues is a complex consisting of two identical and long peptide chains each of which is made up of at least seven distinct domains (Wakil, 1989). These represent separable enzymes and an acyl carrier protein in Type II synthase found in prokaryotes and plants (Stumpf, 1984). With the help of the acyl carrier protein, these enzymes catalyze the various partial reactions required for the elongation of the two acetyl carbons to form long chain saturated fatty acids. Type I complex, the fusion protein, has all the partial activities distributed in the different domains of each of the two peptide chains. Type I enzyme in yeast has an even more complex organization comprising six copies each of two nonidentical peptide chains that has the partial activities distributed in one or the other of them (Kuziora et al., 1983). Fatty acid synthase provides a singular example of the fusion of a large number of discrete genes to create the more efficient multienzyme complex (Kuziora et al., 1983; McCarthy and Hardie, 1984). Pyruvate dehydrogenase is another multienzyme complex produced by a similar fusion of the genes of component proteins. A further association of the proteins derived from the fused genes produce even larger complexes (Patel and Roche, 1990). Prolyl- and glutamyl-tRNA synthetases are distinct enzymes in *Escherichia coli*. However, they become the amino and carboxyl terminal domains of a bifunctional protein in *Drosophila*. The central domain of this protein consists of six repeats of 46 amino acid residues of an as yet unknown function (Cerini et al., 1991). This fusion protein is itself a component of a much larger complex comprising many other aminoacyl-tRNA synthetases (Yang et al., 1985; Cerini et al., 1991).

G. *De Novo* Synthesis of Proteins

A novel method of generating new proteins was by "overprinting", a process in

which they were created during translation by starting from a hitherto unutilized or noncoding open reading frame (Keese and Gibbs, 1992). Overlapping cellular genes are known to express different proteins in bacteria, mitochondria, nuclei, and viruses.

H. Gene Duplication Followed by Independent Mutations Resulted in Proteins with Different Functions

This process is a variation of the phenomenon described in Section B. However, the duplicated genes, instead of producing a family of proteins whose members fulfill a similar role, followed a different course such that one line continued to encode the protein that performed the original function, while the other followed an independent course of mutations to generate a protein that acquired the ability to perform an entirely different function. This review is concerned with such a process.

Hen egg white lysozyme and α-lactalbumin provided the first example of a pair of proteins that fulfilled entirely different biochemical roles but which were shown to possess extensive homology (Brew, Vanaman, and Hill, 1967; Brew, Vanaman, and Hill, 1968: Brew, 1970). The former lysed Gram-positive bacteria by hydrolyzing their cell wall peptidoglycans. The latter with no catalytic activity of its own turned out to be an essential component of lactose synthase, an enzyme complex active only in lactating mammary glands (Brodbeck et al., 1967). An amino acid sequence identity amounting to 40% observed between these two proteins was astonishing at that time because it was then implicitly held that each protein had a unique, native three-dimensional structure that was dictated by its amino acid sequence. Because the biological activity depended on its tertiary structure, each functional protein was expected to have a unique primary structure. The surprise was compounded even further when it was found that the similarity between the two proteins extended to their three-dimensional structures as well (Acharya et al., 1989). The belief that the specific function of a particular protein requires a unique structure did not appear to hold at least for these two proteins. Since then other instances of pairs of proteins have come to light that carry out different activities even though they have similar tertiary structures. Mandelate racemase and muconate lactonizing enzyme, despite low sequence similarities, have not only superimposable tertiary structures but identical quaternary structure as well (Neidhart et al., 1990). Yet, they carry out mechanistically different reactions on structurally distinct substrates. A comparison of the crystal structure of enolase with pyruvate kinase indicates a striking structural similarity in the location of a substantial proportion of their C_α-atoms (Lebioda and Stec, 1988). A related phenomenon is that of pairs of proteins that share identical structure in only certain domains or parts of the molecule. *Penicillium citrinum* P1 nuclease, for instance, possesses 56% spatial structural identity with phospholipase C of *Bacillus cereus*. Of the 249 C_α-atoms 145 were superimposable even though sequence identity of the superimposable segments was only 18% (Volbeda et al., 1991). Similarly, despite low sequence similarity, the structural identity between the monomers of fructose-1,6-bisphosphatase and inositol monophosphatase was pronounced (Zhang, Liang, and Lipscomb, 1993). The behavior described above is distinct from those of enzymes that catalyze mechanistically similar reactions but that use different substrates and that therefore possess common structural motifs. Examples of these are the nicotinamide-dependent dehydrogenases (Jörnvall et al., 1987) and serine-proteases

(Higaki et al., 1987), which clearly share the same ancestries.

II. LYSOZYME FAMILY

The antiinfective activity of egg white and unboiled milk was recognized and had been used in what is now called folk medicine for over two millennia in various cultures throughout the world. The reason for their beneficial effects became clear only after the investigations initiated in 1922 by Alexander Fleming demonstrated the presence of a bacteriolytic activity in the nasal secretion, tear, milk, and extracts of various tissues (Fleming, 1974; McKenzie and White, 1991). Fleming, it appears, considered lysozyme for therapeutic use that was mocked in a cartoon that appeared in *Punch*, the British humor magazine (Figure 1). This was years before the discovery of penicillin. In the following decades the enzyme shown to be responsible for this lytic action, which Fleming named lysozyme, was purified from a variety of sources, and its physical and chemical characteristics were established (McKenzie and White, 1991). A number of recent studies suggest that this enzyme has roles other than its bacteriolytic action (Mason, Jones, and Goodnow, 1992; Pepys et al., 1993; Taylor, Cripps, and Clancy, 1995). The three-dimensional structure of the enzyme obtained from hen's egg white was elucidated from the electron density maps created from the X-ray diffraction patterns of its crystal (Blake et al., 1965). The molecule was seen to consist of two domains separated by a deep cleft that stretched across its entire width. One of the domains contains most of the helices and the other most of the β-sheets. A helical segment joins the two domains. Subsequent work refined the original structure considerably. The contacts between the atoms of side chain residues in the cleft and those of the substrate (or inhibitor) were also established (Phillips, 1974).

Investigation of the nature of lysozyme from different sources has revealed that this ancient enzyme has undergone mutational

FIGURE 1. A cartoon by J. H. Dowd showing the collection of "Tear Antiseptic" from young "volunteers". It appeared in an issue of *Punch*, a British humor magazine, in 1922. Actually, a drop of lemon juice was applied to the eye, a much milder procedure, for the secretion of the tear. (This magazine ceased publication a few years ago. Attempts to contact Punch Publications, Ltd., London, for permission to reproduce this figure were not successful.)

alterations beyond the normal changes during speciation to generate at least four distinct types of proteins. Two of these were designated chicken and goose, or C and G, types, respectively, following the characterization of the enzyme isolated from its richest sources, egg whites of hen and of goose (Canfield and McMurry, 1967). The other two forms are those found in the digestive systems of ruminants and of leaf-eating monkeys (Dobson et al., 1984; Jollès et al., 1984; Stewart et al., 1987). Recently, a similar type of lysozyme has been reported to be present in the foregut of hoatzin, a leaf-eating bird inhabiting certain regions of South America (Kornegay et al., 1994). The bacteriolytic action of lysozymes C and G provides the protection to the eggs that Fleming had originally attributed to tears and other secretions. The same lytic activity of the enzyme serves a digestive purpose instead in the fore-gut chambers of ruminants and leaf-eating monkeys. It lyses the profuse microfloral cells that grow in the rumen and fore-gut by the fermentation of cellulose and releases valuable nutrients that the microflora contain. At present, the C type of lysozyme appears to be the most widely distributed form of the enzyme. It is found in the tissues of most species, including the non-stomach tissues of ruminants. Besides these four types of enzymes, an enzyme possessing a similar activity is encoded in the genome of *E. coli* phage, T4, which has no sequence homology with the lysozymes found in animal tissues. This enzyme is used by the newly produced phage particles for their own release from the host cells by the lysis of the cell wall from within (Tsugita, 1970).

The ancient animal lysozyme gene appears to have been recruited, even before its duplication and adaptation to produce digestive enzymes, to fulfill yet another and entirely different role. This undoubtedly played a significant role in the evolution of mammals. The new gene generated, after what now appears to have been a long period of apparent dormancy, codes for α-lactalbumin, a constituent protein of milk. Its primary structure, when determined by Brew, Vanaman, and Hill (1967), showed a surprising similarity with hen egg white lysozyme. α-Lactalbumin was not known then to have any specific biochemical role other than a presumptive nutritional one. No two proteins that were functionally so distinct had been observed to have such a sequence similarity. The interest in these proteins was stimulated further when it was discovered later that α-lactalbumin was one of the two components of lactose synthase (Brodbeck and Ebner, 1966; Brodbeck et al., 1967). Lactose, the product of this enzyme system, is one of the valuable nutrients in milk that a nursing female provides to its neonate.

Lactose is found in nature only in milk. Its content is between 4 and 6% in the milks of a large number of species (Oftedal and Iverson, 1995). The possible advantages of providing lactose as a nutrient are believed to be the following (Newburg and Neubauer, 1995). (1) As a disaccharide it has half the osmolarity of a monosaccharide and this may be helpful in maintaining the appropriate osmolarity of the contents of intestinal lumen. (2) The $\beta(1 \rightarrow 4)$ linkage between galactose and glucose of lactose promotes the growth of only those microorganisms in the neonatal intestine that can hydrolyze this particular glycosidic link. Lactose is, therefore, likely to prevent the proliferation of those organisms that hydrolyze the more abundant polymers of glucose linked by $\alpha(1 \rightarrow 4)$ bonds. (3) Finally, D-galactose formed by the hydrolysis of lactose during its digestion can be utilized directly for the synthesis of galactolipids needed for optimal myelination and brain development. The importance of lactose (and consequently of α-lactalbumin), in at least the murine spe-

cies, is clear from the results of recent investigations involving mice deficient in, or lacking, α-lactalbumin gene (Stinnekre et al., 1994; Stacey et al., 1995). The litters of the mutant mice as well as of the wild type failed to survive on the milk of α-lac⁻/α-lac⁻ homozygote nursing females. The introduction of human α-lactalbumin gene at the same site restored the adequacy of the mutant milk (Stacey et al., 1995). Whether this reflects a nutritional need is uncertain as even the elimination of β-casein, a major component of milk proteins, by the deletion of its gene had no deleterious effects on the growth and development of the nursing young mice (Kumar et al., 1994). The essentiality of lactose as a nutrient appears also not to be universal as several species of mammals, including polar bears, whales, dolphins, and seals, contain little or no carbohydrate in their milk (Oftedal and Iverson, 1995). Moreover, the identities of the carbohydrates in the milk of many species are yet to be established. Marsupial milk, for instance, has been shown to contain not lactose but oligosaccharides larger than the disaccharide (Oftedal and Iverson, 1995).

Lactose synthesis requires the formation of a glycosidic bond between the two sugar residues, D-galactose and D-glucose, a reaction that is mechanistically the reverse of the hydrolytic reaction that lysozyme carries out. The synthesis is achieved by utilizing the glycosidic bond-forming activity of a specific galactosyltransferase that is normally involved in the synthesis of the carbohydrate component of glycoproteins in a wide variety of cells. The activity of this transferase is modified for the synthesis of lactose in the following novel manner. Galactosyltransferase by itself catalyzes the transfer of the galactose moiety of uridine diphosphate galactose to *N*-acetyl glucosamine (NAG)[a] residues to form the lactosamine residues of glyco-polymers. Galactose can also be transferred to free NAG. In the presence of α-lactalbumin, however, the galactosyl acceptor specificity of galactosyltransferase is altered from NAG to glucose and the product is lactose (Brew, Vanaman, and Hill, 1968). α-Lactalbumin, not surprisingly, is present in the secretory cells of mammary glands only after the induction of lactation. Like its relative lysozyme, α-lactalbumin has also been implicated in more than this one function (Hall and Campbell, 1986; Hakansson et al., 1995; Do, K. Y., Do, S.-I. and Cummings, R. D., 1995).

This review attempts to describe the differences between the members of this family of proteins at the molecular level, relate them to the mode of their actions, and, to the extent possible, trace their evolutionary relationships. Repetitions of the materials covered in other excellent reviews (McKenzie and White, Jr., 1991; Hall and Campbell, 1986) are kept to a minimum. Lysozyme and α-lactalbumin have served as models for the investigation of protein folding. Literature on this is extensive (Dobson, Evans, and Radford, 1994; Kronman, 1989; Sugai and Ikeguchi, 1994). This topic, important as it is, is also not covered in this review.

III. MECHANISMS OF ACTION

A. Lysozyme

Lysozyme hydrolyzes the glycosidic bond between specific sugar residues of the polymers out of which bacterial cell walls are constructed. These polysaccharides contain alternating units of *N*-acetyl muramic acid and 2-deoxy-*N*-acetyl glucosamine linked to each other by means of (1 → 4) bonds. It can also hydrolyze chitins, homopolymers of NAG. Lysozyme binds its

natural substrate in the cleft it contains, which can accommodate a hexamer section (rings A to F) of its natural polymeric substrate (Plate 1).* The binding is such that ring D, normally a NAM residue, is at the catalytic site of the enzyme. The cleavage occurs at the bond between rings D and E, the latter being a NAG residue. The interactions between the various atoms of the substrate and of the enzyme, deduced from the crystal structures of the substrate-, and its analog-, bound enzymes are shown in Plate 1 and Table 5. The hydrolytic action of the enzyme, it is proposed, follows general acid catalysis involving the protonated γ-carboxyl group of Glu-35 (HEW numbering). Asp-52 β-carboxylate ion is believed to aid the reaction by stabilizing the carbo-cation intermediate. The mechanism of the reaction, now widely accepted, is shown in Figure 2. As originally proposed, this mechanism lacked direct experimental evidence. It was based on models built using crystallographic data that indicated that there was insufficient space in the D-ring binding site of the protein to accommodate the C–5 hydroxymethyl group of the sugar residue in the energetically favorable equatorial orientation of the chair conformation. A number of considerations, including the binding characteristics of appropriate substrate analogs, led to the suggestion that the D-ring of the substrate assumes a near half-chair conformation after binding at its specific site that alters the orientation of the hydroxymethyl group from equatorial to a more axial one. The needed space then becomes available (Phillips, 1974). The catalytic amino acid residues were also found to be in the appropriate environments. Lack of conclusive evidence, despite extensive work, prompted suggestions of alternate mechanisms not involving a distortion of the D-ring (Post and Karplus, 1986; Sinnot, 1990). Direct evidence for these were also lacking.

The central feature of the mechanism (Figure 2) is the essentiality of Glu-35, the proton donor, the distortion of the NAM ring when bound at site D of the enzyme and the presence of Asp-52 in a polar environment that stabilizes the oxonium ion of the reaction intermediate. The life-time of this reaction intermediate needs to be long enough for rings E and F to diffuse away. The indispensability of Glu-35 was established when it was found that the substitution of Glu-35 of HEW lysozyme with glutamine by means of site-directed mutagenesis abolished the enzyme activity completely (Malcolm et al., 1989). However, the mutant with asparagine substituting for Asp-52 retained an initial activity that was 5% of that of the wild-type enzyme. The distortion of ring D, questioned by many workers, was also found to occur in high-resolution crystallographic studies of the trimer, NAM-NAG-NAM, bound to HEW lysozyme (Strynadka and James, 1991). The enzyme altered the conformation of the NAM residue, when bound at site D from a normal chair to one in which the ring atoms became more planar and the primary alcohol side chain, $C(6)H_2OH$, adopted the predicted near-axial orientation (Strynadka and James, 1991). The distortion of the sugar residue concerned was also observed in the substrate or its analogs when bound to a mutant of T4 lysozyme (Kuroki, Weaver, and Matthews, 1993). Regarding the catalytic amino acids, whereas there is general agreement on the role glutamic acid plays as a proton donor, that of aspartate, corresponding to residue number 52 of HEW and of residue number 20 of T4 lysozymes, is far from clear. The absence of this acid in a corresponding position in goose lysozyme adds to the questionable role of this residue (Weaver, Grütter, and Matthews, 1995).

* Plate 1 appears after page 266.

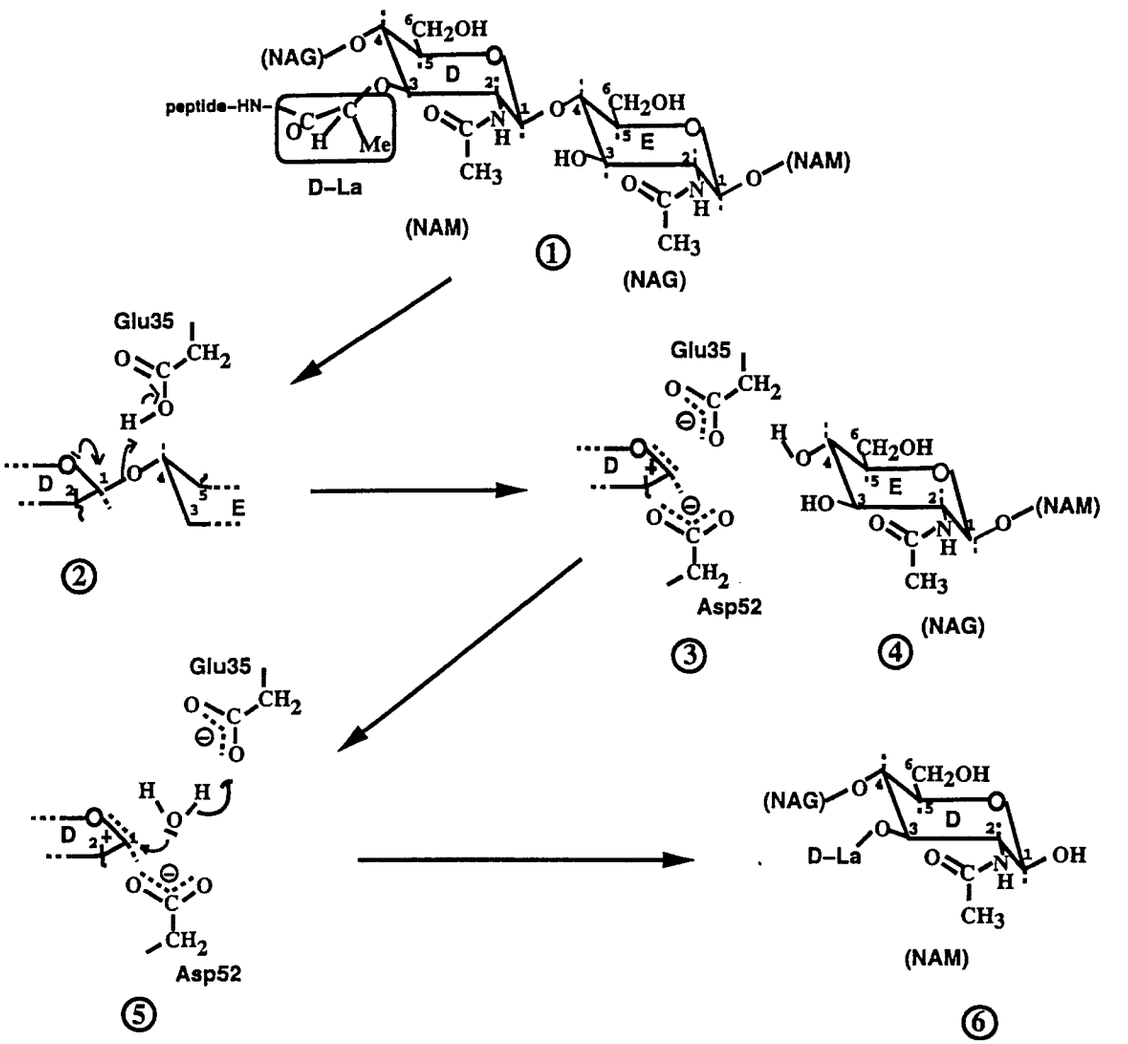

FIGURE 2. Mechanism proposed for the action of lysozyme binding of the substrate to the enzyme (1) results in the distortion of ring D of NAM to a less stable half chair conformation (2) due to steric factors. This in turn weakens the scissile C_1-O glycosidic bond. The protonated carboxyl group of Glu 35, 3 Å away surrounded by non-polar groups, provides the proton to the glycosidic O-atom to form the oxonium ion (3), which makes ring E (4) a good leaving group. The carbocation formed at C_1 and ring O-atom (3) is stabilized by Asp-52 carboxylate ion, also located at a distance of 3 Å in a polar environment, until it acquires a hydroxyl group (5) from a water molecule to form (6). Glu-35 carboxylate then acquires a proton form the medium. (Adapted from W. N. Lipscomb, 1972.) Abbreviations: *N*-acetylmuramic acid, NAM; *N*-Acetylglucosamine, NAG; D-lactyl, D-La.

B. α-Lactalbumin

Synthesis of lactose requires two protein components (Figure 3). One of these is the ubiquitous enzyme galactosyltransferase and the other is α-lactalbumin. Galactosyltransferase in tissues other than lactating mammary glands but including the mammary glands of pregnant females transfers the galactosyl moiety of UDP-galactose to terminal *N*-acetylglucosamine residues of the carbohydrate component to form larger glycoproteins. This occurs in the Golgi apparatus. The enzyme *in vitro* can add galac-

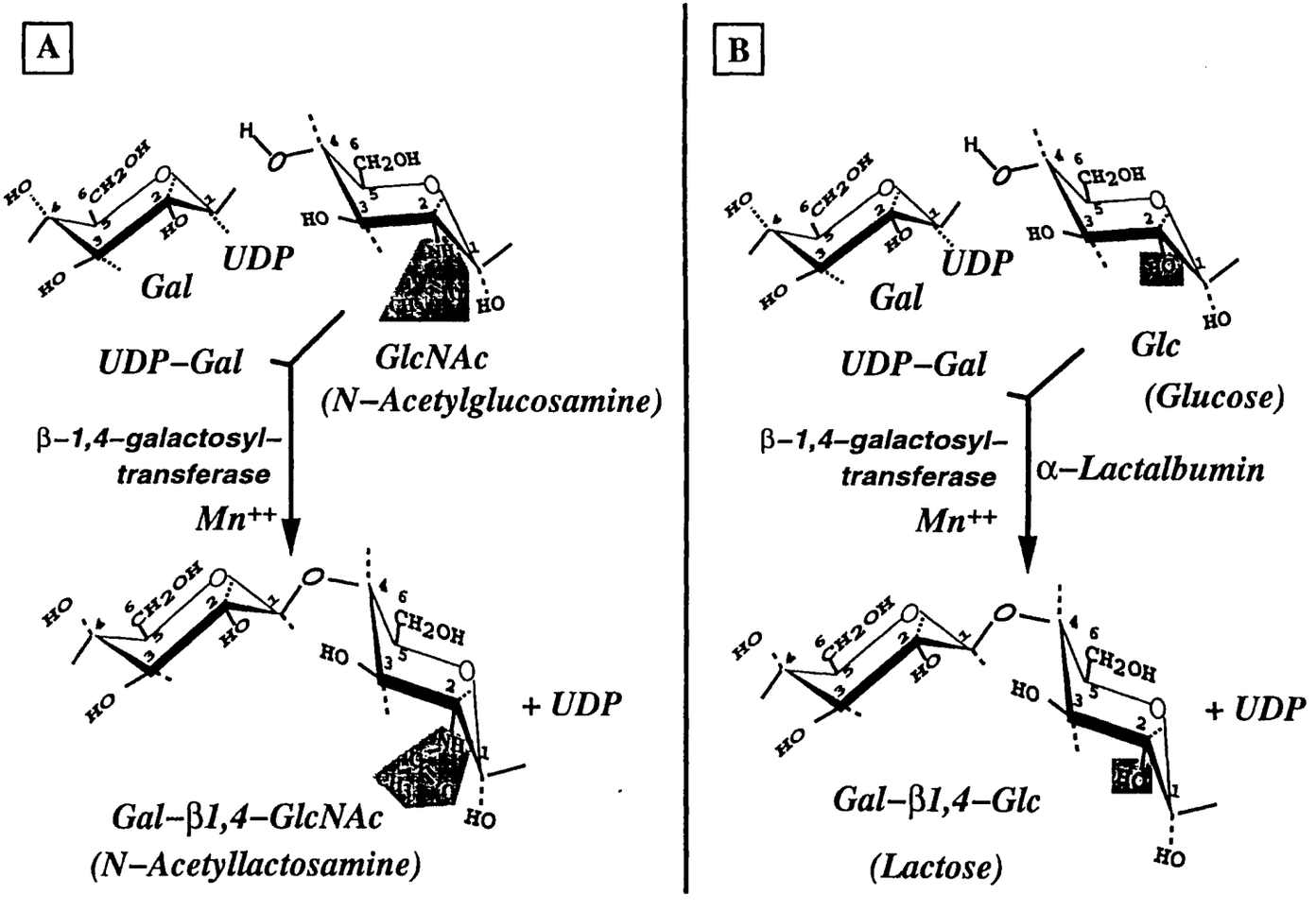

FIGURE 3. Reactions catalyzed by galactosyltransferase (GT) and α-lactalbumin. (A) In the absence of α-lactalbumin galactosyl-transferase transfers the D-galactosyl moiety of UDP-galactose to C–4 of free *N*-acetylglucosamine, GlcNAC, or an appropriate GlcNAC containing polysaccharide. (B) In the presence of α-lactalbumin the galactosyl acceptor specificity of the transferase for GlcNAC is decreased or abolished, depending on the concentration of α-lactalbumin, with a corresponding increase in the effectiveness of glucose as the galactosyl acceptor.

tose from its natural donor to free *N*-acetylglucosamine, however, at a much reduced rate, to form lactosamine. The reaction is shown in Figure 3A. For the synthesis of lactose, the same galactosyltransferase is employed. However, in the presence of α-lactalbumin, known to be present only in the secretory cells of the lactating mammary gland, the acceptor specificity of transferase is altered from *N*-acetylglucosamine, its natural substrate, to glucose. The product of the reaction then is the free lactose (Figure 3B). A recent study reveals another feature of α-lactalbumin. It induces bovine β–1,4-galactosyltransferase to stimulate the transfer of *N*-acetylgalactosamine from its UDP derivative to free *N*-acetylglucosamine but not to glucose (Do, Do, and Cummings, 1995). It has also been reported to stimulate the activity of another, although distinct, transferase from snail that transfers GalNAc from UDP-GalNAc to GlcNAc as well as glucose residues (Neeleman and van den Eijnden, 1996). The significance of these findings remains to be clarified.

Under physiological conditions, galactosyltransferase stays bound to the membranes of the *trans* face of Golgi apparatus. During lactation there is a copious production of α-lactalbumin that passes through the lumen of Golgi. Here the two proteins form a complex in the presence of all the necessary substrates. It is this transient complex that synthesizes lactose. Both α-lactalbumin and lactose are then secreted. Hence, a continuous synthesis of α-lactalbumin is required during the entire period of lactation (Brew, 1970).

IV. COMPARISON OF THE DIFFERENT TYPES OF LYSOZYMES

A. Chicken and Goose Types

HEW lysozyme was the first of the lysozymes to be isolated in a pure enough form to be characterized properly. Its three-dimensional structure, incidentally the first enzyme for which this information was obtained, was determined in 1965 (Blake et al., 1965; Phillips, 1974). An enzyme with similar properties was found in tear, milk, and secretions or extracts from a variety of tissues as well as the egg whites of a number of avian species (Canfield et al., 1972; Hyslop et al., 1972; Schumacher, 1972; Acharya et al., 1989). A form of lysozyme with a different amino acid composition and properties was reported to be present in Embden goose egg white (Dianoux and Jollès, 1967; Canfield and McMurry, 1967). The purification of lysozyme from black swan egg white later yielded two distinct forms of the enzyme, only one of which cross-reacted with HEW anti-serum. The other had properties similar to the goose enzyme (Morgan and Arnheim, 1972). The comparative properties of the two types of enzymes are listed in Table 1. The enzyme in the egg white of black swan was later found to be similar to that in the egg whites belonging to the goose category (Simpson et al., 1980) and ostrich (Jollès, Périn, and Jollès, 1977).

It appears that the egg white of certain species of birds, hen, and Peking duck, for instance, contain only the C-type of lysozyme, while others, Embden goose and ostrich, for example, contain only the goose type. A third group of birds to which Canada goose, black swan, and duck belong contain both forms. Immunological studies have provided evidence that these two types of lysozymes are not two allelic forms of the same enzyme. They were also shown to be encoded in two distinct genes (Arnheim, 1972).

When the amino acid sequence of the goose type of lysozyme from black swan was determined and compared with HEW lysozyme, no sequence similarity was discernible at first (Simpson et al., 1980). Later, the sequence of the G-type of ostrich egg

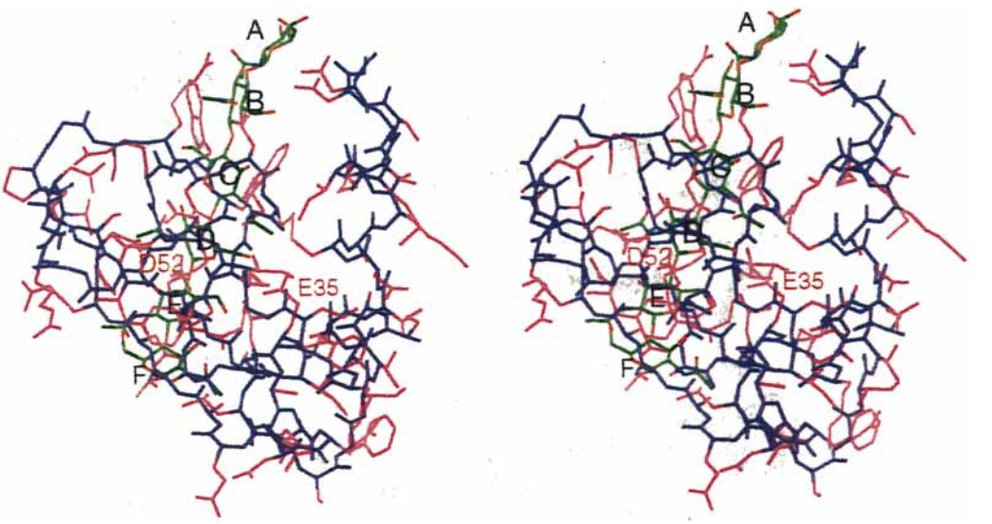

Plate 1. Stereodiagram depicting the binding of the hexasaccharide in the cleft that extends across the entire width of lysozyme. Biosym's InsightII program was used for generating the picture using 1lsz.pdb coordinate file. The hexasaccharide was built using the coordinates for the tetrasaccharide in 1lsz.pdb file and adding two more residues at the reducing end. The protein backbone chain is shown in purple, the side chains of the amino acid residues in red, and the hexasaccharide in green. The six binding sites A-F are labeled as are the catalytic residues Glu-35 and Asp-52. The other residues whose side chains interact with the bound substrate are listed in Table V.

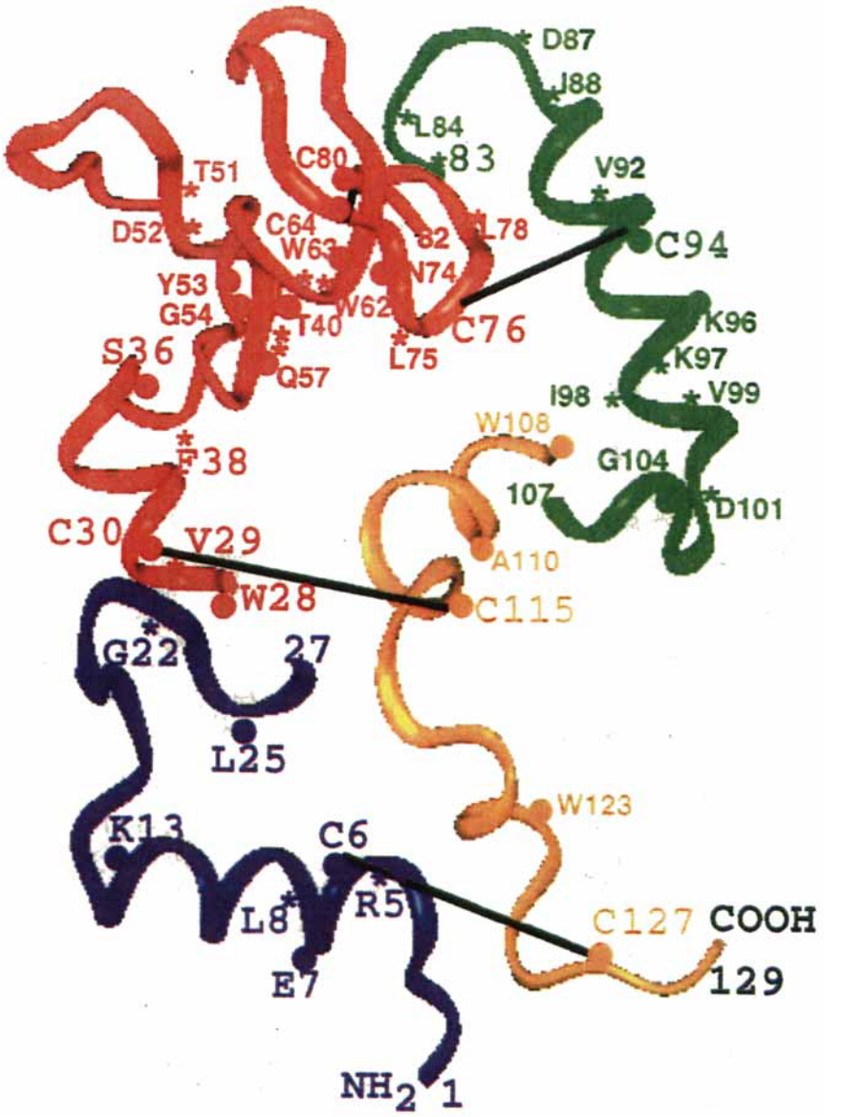

Plate 2. The similarities in the primary structures in the exon products of mixture of sixteen lysozymes and nine α-lactalbumins compiled in Table 1 (Acharya et al., 1989). Both are encoded in 4 exons represented by different colors. The four pairs of disulfide bonds they share are also shown. Twenty four amino acid residues, identified and circled are either invariant or varies in three or less number of species; twenty three residues, identified and indicated by asterisks are conserved substitutions. Six or seven of the 129 residues of lysozyme are deleted in α-lactalbumin. Homology is greatest in the sequences encoded in Exon 3, and least in that of Exon 4 (Table VI).

TABLE 1
Comparison of Some of the Properties of Lysozymes

Property	HEW C-type	COW foregut	GEW G-type	Phage T4
M.W. $\times 10^{-3}$	14.6	14.6	20.4	18.7
No. of amino acid residues	129	129	185	164
Cysteine	8	8	4	2
Free SH	0	0	0	2
Homology with lysozyme		Extensive	Weak	None
Spatially equivalent C_α to lysozyme C	–	92	78	
pH optimum	5–7 (broad)	5 (narrow)	6–7.5 (broad)	7.2–7.4
Catalytic residues	E-35; D52	E-35; D52	E73	E11; D20
Relative lytic activity with *M. lysodecticus*	1	—[a]	6	6
E. coli	—[a]	—[a]	6	250
$(NAG\text{-}NAM)_2$ hydrolysis	+	—[a]	–	–
Chitin	+	—[a]	–	—[a]
$(NAG)_4$ inhibition	+	+	–	—[a]

[a] Information not obtained.

lysozyme did show a slight resemblance to HEW lysozyme (Schoentgen, Jollès, and Jollès, 1982). Eight identical residues occupied the same position in a stretch of 20 residues, numbers 35 to 54 of HEW lysozyme and the corresponding stretch, numbers 73 to 92 in the ostrich enzyme (Figure 4A). There was also an observable resemblance in the sequences of residues 81 to 97 of HEW lysozyme and the corresponding sequence, residues 117 to 130 of the ostrich enzyme (Figure 4B). The latter had three deletions of amino acids. It was pointed out by Schoentgen, Jollès, and Jollès (1982) that the 20 amino acid stretch, residues 35 to 54, and the shorter stretch, residues 81 to 97, of HEW lysozyme are encoded within Exons 2 and 3, respectively (Jung et al., 1980). The catalytic amino acid residues, Glu-35 and Asp-52, which correspond to Glu-73 and Asp-90 in ostrich egg white lysozyme, are located in the first conserved region. Therefore, it seemed likely, that the other conserved amino acid residues in this region provide the environment conducive to the glycosidase activity. This sequence resemblance is seen to extend to the other goose type of lysozyme as well (Figure 4). The catalytically active glutamic acid residue occupies spatially identical positions in all the lysozymes (Weaver, Grütter, and Matthews, 1995). In residues 73 to 92 of HEW lysozyme, one of the homologous stretches forms a part of one side of the substrate-binding cleft and is believed to contribute some of the residues for substrate binding and for ensuring its specificity. This contention is strengthened by the similarity observed when the amino acids of these regions of all three goose type of lysozymes are aligned against the corresponding region of HEW lysozyme (Fig-

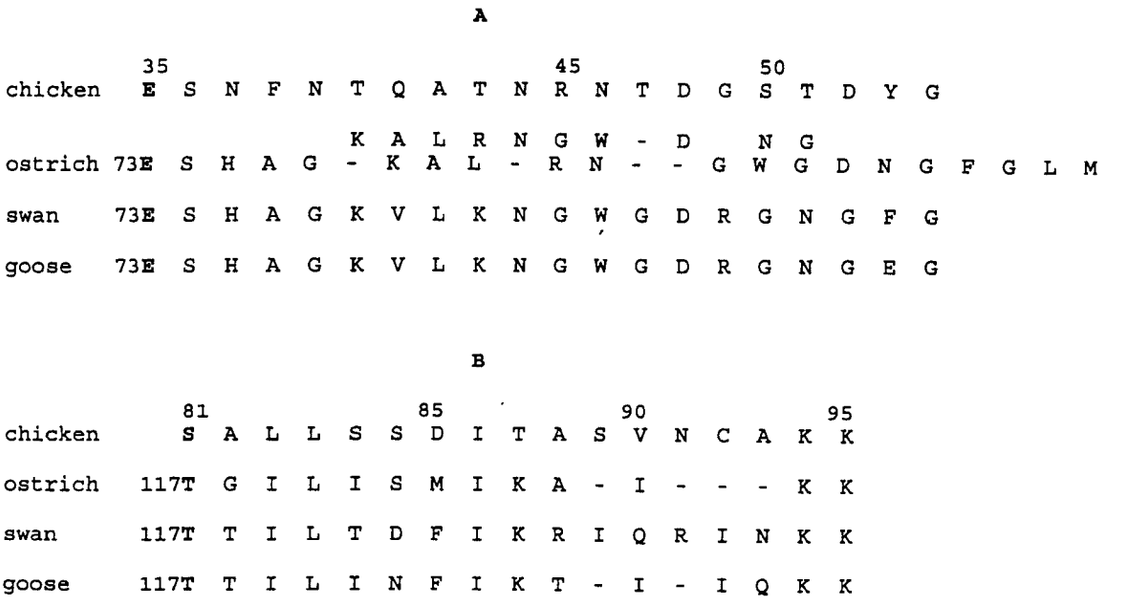

FIGURE 4. Homology between HEW lysozyme and the corresponding locations of the enzyme obtained from the egg white of ostrich, swan, and goose in two different segments, A and B. Realigning ostrich lysozyme residues shown above the main line provides a better homology. Data are taken from Schoentgen, Jollès and Jollès, 1982 (chicken and ostrich); Embden goose, Simpson and Morgan, 1983 (Embden goose), Simpson et al., 1980 (black swan).

ure 5). Glu-35 occupies identical positions in all of them. Asp-52 occupies the same position in ostrich enzyme if the deletion of four amino acid residues is taken into account. Swan and goose lysozymes do not have an aspartic residue at this position (Weaver, Grütter, and Matthews, 1995).

B. Stomach Lysozymes

The presence of a new form of lysozyme was reported in the stomach of animals that possessed cellulose fermenting fore-guts (Dobson, Prager, and Wilson, 1984; Jollès et al., 1984, 1990). This enzyme, while retaining extensive sequence identity with lysozyme C, differed from it in having a lower isoelectric point (6.2 to 7.7) compared with goat tear enzyme, which was above 9), a narrower pH profile at around 5, instead of 7.5, and a relative insensitivity to pepsin digestion (Table 1). The result of the comparison of the amino acid sequences of lysozyme of related species are consistent with the view that after the appearance of mammals, 70 to 80 million years ago, some evolved into species that possessed the ability to utilize the abundantly available cellulose of plants to meet their dietary requirements (Stewart and Wilson, 1987; Jollès et al., 1989; Irwin, Prager, and Wilson, 1992). Lacking cellulase among their digestive enzymes survival of these mammals depended on the development of a fore-gut and the establishment of a symbiotic relationship with specific microflora 60 million years ago. The enzyme in the microflora hydrolyzed the $\beta(1 \times 4)$ glycosidic bonds of glucose polymers and produced volatile fatty acids as their metabolic products that nour-

ished the tissues. More importantly, the microorganisms themselves became the source of valuable nutrients such as proteins, nucleic acids, and vitamins (Irwin and Wilson, 1989; Irwin and Wilson, 1990). In order to obtain these the microbial cells had to be lysed in a different chamber. The ability to accomplish this was acquired when a form of lysozyme evolved from the ancestral lysozyme gene following its duplication and adaptation to carry out its new role as a digestive enzyme. For the enzyme to act effectively in the environment of the true stomach it had to be less sensitive to low pH, be resistant to the action of pepsin, and also be available in adequate quantities.

The bovine stomach mucosa secretes at least three different but closely related lysozymes. They retain all but one of the 40 invariant amino acids of lysozyme C (Jollès et al., 1984). Other significant findings were (1) lysozyme present in all monogastric mammals contain the sequence, Asp101-Pro102 (human lysozyme numbering), the only peptide bond reported to be susceptible to acid hydrolysis at normal temperature. Proline in the lysozymes expressed in the bovine stomach is found to be deleted and the aspartic residue replaced with glutamate (Jollès et al., 1984). The effect of the Asp-Pro sequence on the stability of proteins has been questioned by Stewart and Wilson (1987) as the digestive lysozyme secreted by the fore-gut of leaf-eating langur retains this sequence. (2) Specific arginine residues are replaced with lysines, thereby lowering the arginine/lysine ratio. Lysyl peptide bonds are less sensitive than argininyl ones to tryptic action. Second, argininyl residues are naturally modified by diacetyl, one of the metabolic products present in the rumen, that reduces the activity of lysozyme (Stewart and Wilson, 1987). This is prevented by the substitution of arginine. (3) Cow genome contain as many as 10 copies of lysozyme genes (Irwin, Prager, and Wilson, 1987), at least four of which are known to be expressed in the stomach mucosa (Irwin and Wilson, 1989; Irwin, White, and Wilson, 1993). Amplification of the gene and the expression of multiple genes results in a 4- to 7-fold increase in lysozyme production (Irwin and Wilson, 1989). The genes expressed in the stomach are different from those expressed in the bovine kidney, mammary gland, and granulocytes (Ito et al., 1993; Steinhoff, Senft, and Seyfert, 1994) but are similar to those expressed in trachea (Takeuchi et al., 1993).

Lysozyme gene, which once served to develop a digestive function in ruminants, was recruited once again, 35 to 40 million years later, to perform the same role in leaf-eating monkeys. In this instance, the adaptation occurred without gene duplication following a course of convergent evolution through positive selection (Stewart, Schilling, and Wilson, 1987; Swanson, Irwin, and Wilson, 1991). Even though endowed with only a single gene the anterior lining of the stomach of the langur contains as much as 1 mg of lysozyme per gram of tissue, whereas the non-stomach tissues contain as little as 0.02 mg (Stewart, Schilling, and Wilson, 1987).

A recent report of a rigorous statistical analysis of the mutational substitution of nucleotide bases in the lysozyme genes of closely related primates provided strong evidence for the adaptive evolution of the gene that encodes the digestive lysozyme in leaf-eating colobine monkey (Messier and Stewart, 1997). Instances of such an adaptation at the molecular level are rare. This investigation involved a detailed comparison of synonymous and nonsynonymous mutations in the gene of the various closely related species. It revealed, among other significant observations, that two major episodes of Darwinian selection had occurred. In the first, nine amino acid substitutions had taken place in the course of the

evolution of the colobine lineage, five of which had occurred earlier independently in the lineage leading to the ruminant stomach lysozymes. The second episode of positive selection appears to have occurred in the lysozyme of species in the hominoid lineage. The significance of this development is at present unclear (Messier and Stewart, 1997).

A detailed description of the differences in the lysozymes present in the different tissues of the various species are beyond the scope of this review.

C. Phage T4 Lysozyme

The phage lysozyme has no discernible sequence identity with either HEW or goose type of lysozymes (Matthews et al., 1981; Grütter, Weaver, and Matthews, 1983), yet they show striking structural similarities. Some structural elements are common to all the three lysozymes, whereas other elements are common to either HEW and goose pair or goose and phage pair (Grütter, Weaver, and Matthews, 1995). These are shown in Figure 5. HEW lysozyme has been extended

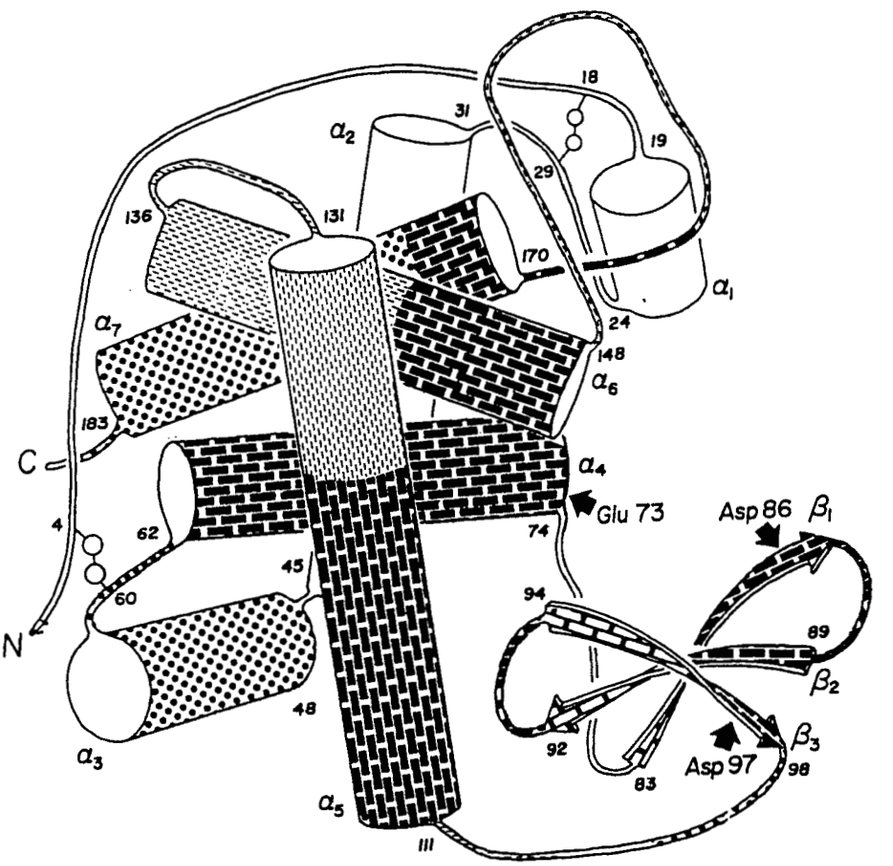

FIGURE 5. Schematic representation of the structural relationship of HEW, GEW, and T4 lysozymes showing common structural elements. Bricks indicate structures present in all three; dots, HEW and GEW; dashes, GEW and T4; open areas, GEW enzyme only. Asp 86 of GEW lysozyme seen in this diagram to be close to the catalytic Glu 73 is actually 10 Å away and hence is not involved in catalysis directly (Weaver, Grütter, and Matthews). (Reprinted with permission from Dr. Brian W. Matthews and Macmillan Journals.)

at the N-terminal in the goose enzyme and at the C-terminal in the phage enzyme, besides other short insertions (Grütter, Weaver, and Matthews, 1983). This is clearly seen by comparing Figure 5 with Plate 4. Besides the similarities in tertiary structures, the three enzymes show conspicuous likeness in their interactions with their respective substrates and the positions of the catalytic glutamate residues (Weaver, Grütter, and Matthews, 1995). In both HEW and T4 lysozymes aspartate, the other residue believed to be involved in catalysis also occupy similar positions with respect to the substrate. The goose enzyme, as mentioned before, lacks this catalytic residue. Because the interactions that bind the substrates to the enzymes involve the acetamido group of the substrate with the backbone atoms of the enzyme, the identity of the amino acids at the active site of the enzyme is apparently unimportant (Weaver, Grütter, and Matthews, 1995).

The possibility that the similarities in the structure and function of these lysozymes is the result of convergent evolution is discounted by the close structural similarities that exist between HEW and T4 enzymes. It has been proposed that all of the lysozymes originated from a common ancestor and diverged to an extent that no amino acid sequence homology was retained (Grütter, Weaver, and Matthews, 1983).

V. DIVERGENCE OF LYSOZYME *C* AND α-LACTALBUMIN

A. Comparison of the Global Properties

Following the discovery that the HEW lysozyme and bovine milk α-lactalbumin shared significant amino acid sequences, it seemed possible that they might also possess similar secondary and tertiary structures. Such a premise was supported by analogy with globins that possessed even less interspecies sequence homology and yet had strikingly close three-dimensional structures. Accordingly, Browne et al. (1969) made an effort to determine the structure of α-lactalbumin based on the assumption that all of its amino acid residues that are identical, or similar, and are located in comparable positions to those of lysozyme in the linear structure would also occupy comparable loci in the native folded states. A wire model of lysozyme molecule constructed to represent its backbone atoms (Blake et al., 1965) was modified stepwise, taking into account individual deletions and substitutions of the various amino acid residues of lysozyme to convert it to α-lactalbumin molecule. The resulting structure remained essentially unchanged even after the side chains of the substituted amino acids were put in place. Later this structure was further refined by Warme et al. (1974) by the computation of energy minima of successive sections of the molecule and finally of the entire molecule. The structure proposed also took into consideration the available experimental data gathered in a variety of physicochemical investigations on the distances between specific amino acid residues and of their reactivities. Experimental verification of the structure proposed was delayed by the difficulty in obtaining crystals of α-lactalbumin from the milk of any of the common species of mammals suitable for X-ray diffraction. Finally, baboon milk α-lactalbumin yielded crystals that were adequate for such studies (Aschaffenburg et al., 1979) and a low-resolution structure was determined (Smith et al., 1987). Soon a refined structure at high resolution was reported that was remarkably close to that predicted earlier from model building and theoretical considerations (Acharya et al.,

1989). The organization of the secondary structures and their encoding in the four exons are shown in Plates 2 and 3.* The structures of α-lactalbumin and lysozyme superposed over each other are shown in Plate 4**. Because the primary structure of baboon α-lactalbumin was not known at that time, there was some concern as to whether its amino acid sequence was indeed as different as were those of the other known α-lactalbumins from HEW lysozyme. This was soon dispelled when it was found that baboon α-lactalbumin was, as expected, closely similar to human α-lactalbumin, 112 of the 123 amino acid residues being either identical or conserved substitutions (Acharya et al., 1990; Acharya et al., 1991). It was later also established that the similarity extends to these proteins from other species, echidna, horse, and pigeon lysozymes on the one hand and human and baboon α-lactalbumins on the other (Acharya et al., 1994). These observations lend strong support to the theory that during the evolution and divergence of proteins, elements of their secondary and tertiary structures are retained more tenaciously than are the primary structures. The corollary is that only those mutations in the genes are sustained that allow the preservation of the stable secondary and tertiary structures of their protein products; the others are presumably selected out (McLachlan, 1987; Overington et al., 1990).

A comparison of the amino acid sequence of a number of lysozymes and α-lactalbumins are presented in Plate 2. The plate depicts the homology of the products of the four exons that code for both proteins. It is clear from the plate and the data presented in Table 2 that a substantial number of amino acids were altered during the divergent evolution of α-lactalbumins. The major difference in the properties of the two proteins are presented in Table 3. Typically, α-lactalbumins contain 123 amino acid residues, 6 shorter than lysozymes. A single exception found to date is rat α-lactalbumin, which is extended by a stretch of 17 residues at the C-terminal end (Dandekar and Qasba, 1981). This is the result of the mutation of the termination codon, TGA, at position 124 of the α-lactalbumin gene to GGA, which codes for glycine. Hence, the translation of the mRNA, when it occurs, continues until the next termination codon is encountered, which occurs at position 141.

An examination of the structures of the two proteins reveals that the cleft in the lysozyme molecule that divides it into two hemisphere-like lobes, and where its substrate is bound, is retained in α-lactalbumin (Plate 4). However, no lysozyme-like activity or a binding of its substrates or their analogs was detectable. A careful investigation designed to detect trace levels of the enzyme activity involving gravimetric and cell viability analysis of *Micrococcus luteus* cells did show such activity. Not surprisingly, the ratio of this activity in bovine α-lactalbumin was lower than that of HEW lysozyme by five orders of magnitude (White, Jr. et al., 1993). It is even uncertain whether even this low level is a reflection of lysozyme-like catalysis.

A second major difference is that all of the α-lactalbumins isolated so far contain 1 mol of tightly bound Ca-ion per mole of the protein. Binding constants have been reported to be as high as 10^7 under physiological conditions (Kronman, Sinha and Brew, 1981), which compares well with the values for many typical Ca^{2+}-binding proteins. Calcium ion was located in the region of the molecule that resembles the helix-loop-helix segment (Plates 4 and 5), the so-called E-F hand, of other calcium ion binding proteins such as calmodulin and parvalbumin (Stuart et al., 1986). In addition to this metal ion α-lactalbumin has

* Plates 2 and 3 appear after page 274.

** Plate 4 appears after page 274.

TABLE 2
Invariant and Conserved Amino Acid Residues in the Four Exons of Lysozymes and α-Lactalbumins and the Location of the Secondary Structures (2°) in Them

Amino acid residues	Exon 1 1–27			Exon 2 23–81			Exon 3 82–107			Exon 4 108–129		
	Inv	Cons	2°	Inv	Cons	2°	Inv	Cons	2°	Inv	Cons	2°
	C6	R5	α_A: 5–11	W28	V29	α_B: 25–36	C94	L83	$3_{10(3)}$:	W108		310(4):105–108
	E_7	L8	$3_{10(1)}$: 12–18	C30	F38		K96	L84	79–85	A110		α_D:109–113
	K_{13}	G22	$3_{10(2)}$: 19–23	S36	T51	β_1 42–45	G104	D87	α_C: 89–103	C115		$3_{10(5)}$: 119–123
	L25			T40	D52	β_2 50–53		I88			W123	
				Y53	I58			V92			C127	
Inv + Cons = 7 residues				G54	N59			K97				
% of residues = 26				Q57	R61			I98				
				W63	W62			V99				
				C64	L75			D101				
				N74	I78			M105				
				C76			Inv + Cons = 13 residues			Inv + Cons = 5 residues		
				C80			% of residues = 50			% of residues = 23		
				Inv + Cons = 22 residues								
				% of residues = 37								

Note: Eighteen lysozymes and 10 α-lactalbumins were examined. These were the sequences given in Table 1 (Acharya et al., 1989) and of cow stomach (Jollès et al., 1989), pig stomach (Jollès et al., 1990), lysozymes and α-lactalbumins from pig (Godovac-Zimmerman et al., 1990) and platypus (Shaw. et al., 1993). Of the conserved (cons.) amino acids, invariant (inv.) differed in less than three species, frequently in none. In the secondary structures α refers to α-helix, 3_{10} to 3_{10}-helix and β to the β-structure. The subscripts identify these. The amino acid residue numbers are those of HEW lysozyme.

TABLE 3
Comparison of the Overall Properties of Conventional Lysozyme and α-Lactalbumin

Properties	HEW LZ	Bovine α-LA
1. No. of amino acid residues	129	123
2. $(NAG\text{-}NAM)_3$ hydrolysis	Active	Absent
3. S-binding cleft	Present	Present
4. Substrate binding	$K_a \approx 10^4\ M^{-1}$	No binding
5. Ca^{2+} binding	No	Yes; 1 mol
6. Lactose synthesis in the presence of galactosyltransferase	None	Active
7. Complexation with galactosyltransferase	None	Active

Note: Some of the differences in the properties of a typical lysozyme and α-lactalbumin (α-LA) are presented. Because of the Ca^{2+} binding property of some lysozymes and differences in amino acids at corresponding positions, HEW lysozyme and cow milk α-actalbumin were chosen as representing the two different proteins.

been reported to bind a variety of other mono-, di-, and trivalent ions. Extensive investigation concerning these and the role they might play in lactose synthase and other activities of α-lactalbumins have been reviewed by Kronman (1989) and by McKenzie and White, Jr. (1991). Of the various metal ions, Zn^{2+} was deemed very important as its binding was believed to displace Ca^{2+} (Kronman, Sinha, and Brew, 1981). Subsequent investigation has made it clear that the binding of metal ions is an exceedingly complex process, there being numerous high- and low-affinity sites that had characteristics that are very sensitive to conformational changes in the protein (Kronman, 1989). Both Ca^{2+} and Zn^{2+} appeared to bind to the protein simultaneously under appropriate conditions (Musci and Berliner, 1985; Kronman, 1989). The presence of several Zn^{2+} binding sites was also indicated that were occupied successively, leading to increased susceptibility to tryptic and chymotryptic digestion as well as aggregation of the protein molecules (Permyakoff et al., 1991; Prestrelski, Byler, and Thompson, 1991). Crystals of human α-lactalbumin have been obtained that had both of these ions bound to it (Ren, Stuart, and Acharya, 1993). Zn^{2+} was located 17 Å away from Ca^{2+} toward the end of the cleft of the molecule corresponding to the catalytic site in lysozyme, a region that is not as rigid as is the Ca^{2+} binding site. Apo-α-lactalbumin formed on the removal of Ca^{2+} is in a "molten globule" state, a partially unfolded conformation of the protein. In Ca^{2+}-bound holoprotein, the native state, is much more stable thermally than is the apoprotein. Binding of Zn^{2+} did not change the geometry of the Ca^{2+} binding loop and caused only minor changes in the rest of the molecule (Ren, Stuart, and Acharya, 1993). It is significant that Zn^{2+}-bound α-lactalbumin crystals were obtained only under stringent conditions of crystallization. The role

Plate 3. Structure of α-lactalbumin that schematically indicates the main secondary structures associated with the four exon products (left). Exon product boundaries and disulfide bridges are also shown. At right the exploded view is presented. Cylinders represent α-helices and arrows β- structures.

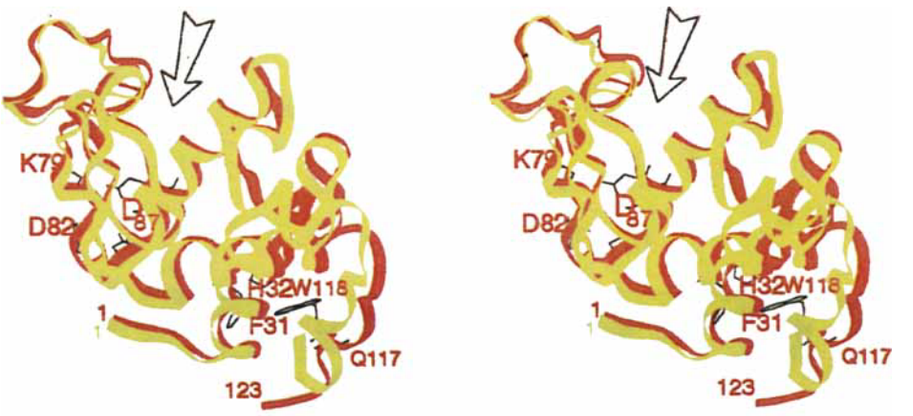

Plate 4. Stereodiagram showing the superposition of HEW lysozyme (yellow) and baboon α- lactalbumin (red) (data of Acharya et al., 1989). The coordinate files 1lsz.pdb and 1hml.pdb were used for generating the lysozyme and α-lactalbumin molecules, respectively.The arrow indicates the cleft in which the substrate of lysozyme is bound. The loop that binds Ca-ion in α-lactalbumin (K79 to D88) can be seen on the left side towards the middle and back of the molecule, and the hydrophobic cluster at the right.

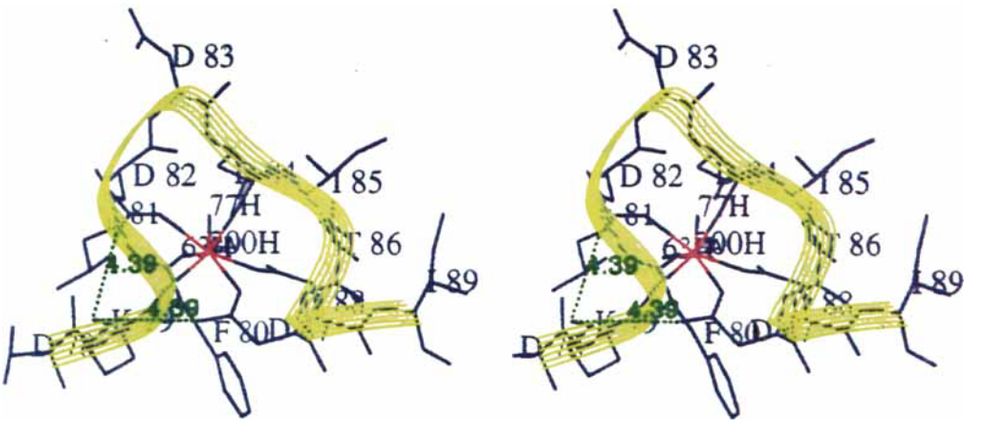

Plate 5. Ca-ion binding loop in baboon α-lactalbumin (1alc.pdb file). Orientations of some of the side chains (Table VII) are shown. The distances between the Ca-ion and its seven ligands (not shown, for the sake of clarity) are approximately 2.3 Å. The distances between ε-NH_2 of Lys-79 and the carboxyl atoms of both Asp-82 and Asp-87 are, as indicated, 4.39 Å. The possible relevance is discussed in the text.

that zinc ion plays in the action of α-lactalbumin is still unclear.

The third important feature that α-lactalbumin acquired during its evolution that distinguishes it from lysozyme is its ability to interact with galactosyltransferase (Table 3). In a partially purified state lactose synthase appeared to consist of a complex that on gel filtration separated into two protein components (Brodbeck and Ebner, 1966). Neither by itself catalyzed the synthesis of lactose from UDPGal and glucose. Investigations that followed established that one of them, the A protein, was identical to UDPGal-*N*-acetyl D-glucosamine galactosyltransferase (Brew, Vanaman, and Hill, 1968), whereas the other, the B-protein, was identical to α-lactalbumin (Brodbeck et al., 1967). Strangely, there appeared to be no definite stoichiometric relationship between the two purified proteins as far as enzyme activity was concerned (Brew, 1970), although a 1:1 association was detected in sedimentation experiments (Klee and Klee, 1972). Galactosyltransferase, which in the absence of α-lactalbumin, transfers the galactosyl residue from UDP-galactose to *N*-acetylglucosamine can also use glucose as the acceptor of galactose provided it is present in extremely high concentrations. The K_M for glucose was higher by three orders of magnitude in the absence of α-lactalbumin than in its presence (Klee and Klee, 1970; Klee and Klee, 1972; Hill and Brew, 1975). Increasing the ratios of α-lactalbumin to galactosyltranferase decreases the rate of the reaction using *N*-acetylglucosamine as galactose acceptor while increasing its reaction rate with glucose as the acceptor. Beyond a certain ratio α-lactalbumin becomes inhibitory to lactose synthesis (Brew, 1970). The kinetics of the transfer reactions in the presence of α-lactalbumin are very complex (Hill and Brew, 1975). Identification of the amino acid residues of galactosyltransferase and of α-lactalbumin that are involved in the association of the two proteins has been the aim of many investigations. A variety of spectroscopic and other physicochemical procedures have been employed for this purpose (McKenzie and White, Jr., 1991). Such investigations continue in many laboratories.

The differences between lysozyme and α-lactalbumin described briefly in the previous paragraphs are examined in greater detail in the following sections.

B. Gene Duplication and Divergence

The ancestry of lysozyme C family of proteins has been reported to go at least as far back as 600 million years (Myr). The basis for this estimate is the presence of lysozyme in three species of moths and the similarity of their amino acid sequences with lysozyme C of currently living species (Jollès et al., 1979; Prager and Wilson, 1988). This implied that the enzyme existed in living organisms even before the separation of vertebrates from invertebrates. It was possible to date some of the molecular events that occurred during the divergence of lysozyme and α-lactalbumin. This involved the construction of genealogical trees from the differences in amino acid and DNA sequences in the proteins and the application of the principle of parsimony. The procedure determines the most likely path from several that requires the minimum number of changes (mutations) to account for the differences between one species and another (Wilson, Carlson, and Whitey, 1977; Prager and Wilson, 1988). Uncertainties are inherent in such estimations because of the differences in the rates at which the third base of codon mutates from those of other two, as well as variations in the rates of evolution

of extant proteins at different stages. These were taken into consideration in the calculations. Phylogenetic relationship established earlier based on morphology and fossil ages also aided in the drawing of conclusions.

A comparison of the amino acid sequence of marsupial α-lactalbumin gene yielded an average figure of 261 ± 37 million years (Myr) (Shewale, Sinha, and Brew, 1984). It was concluded that following the duplication of the lysozyme gene, one of them retained the lysozyme characteristic function, while the other sustained independent mutations finally leading to the α-lactalbumin gene after the divergence of birds and mammals. Hayssen and Blackburn (1985) calculated a slightly earlier period, 300 Myr. Prager and Wilson (1988) estimated an even earlier period of gene duplication and suggested that it occurred at about the time mammals and birds split, over 300 Myr ago. More recent work (Dautigny et al., 1991), based on a comparison of the amino acid sequences of two trout lysozymes with those from moths and other species and of α-lactalbumins, concluded that the divergence of α-lactalbumins from lysozyme preceded the divergence of fishes from tetrapods. The age of gene duplication was estimated to be about 400 Myr. This estimate is supported by the observations of Grobler et al. (1994a).

A simple compilation of the main molecular events in the divergence of lysozymes and α-lactalbumins are presented in Figure 6. It is clear that α-lactalbumin gene originated long before its function as currently recognized was utilized. The age of placental mammals is accepted to be about 60 to 70 Myr. Mammary gland had developed and lactation originated long before. The presence of α-lactalbumin in the milk of marsupials and monotremes would suggest that the expression of α-lactalbumin

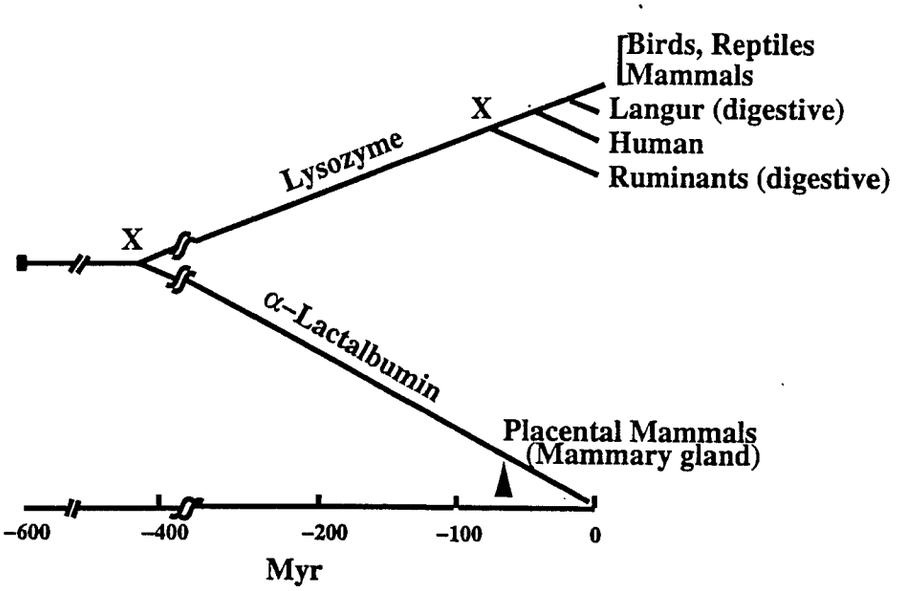

FIGURE 6. Evolutionary divergence of lysozymes and α-lactalbumins adapted from the reports quoted in the text. Phylogenetic split between birds and mammals occurred 300 Myr ago and the gene split resulting in α-lactalbumin occurred before that; ruminant stomach lysozyme genes originated about 60 Myr and langur lysozyme gene diverged from baboon's about 20 to 25 Myr ago. X indicates gene duplication. Placental mammals are estimated to have appeared around 60 to 70 Myr.

gene began over 200 Myr ago in their common ancestor, perhaps a therapsid reptile (Shewale, Sinha, and Brew, 1984; Shaw et al., 1993).

At least on two later occasions the lysozyme gene diverged again, once when the ruminants evolved and lysozyme was recruited to serve a role as a digestive enzyme. Gene duplications occurred then to produce multiple copies in order to produce the enzyme in the required quantities. These events are placed at about 20 Myr (Jollès et al., 1990; Irwin, Prager, and Wilson, 1992). Appropriate amino acid substitutions resulted again at a later time, 10 Myr, in the divergence of the digestive lysozyme in langur from baboon lysozyme. The human lysozyme lineage originated at an earlier time (Stewart, Schilling, and Wilson, 1987; Stewart and Wilson, 1987). Langur, human, and baboon have only one copy each of the lysozyme gene per haploid; a cow has 10, some of which are expressed in the stomach, while the others are expressed in non-stomach tissues (Jollès et al., 1990; Takeuchi et al., 1993; Ito et al., 1993). Sheep and deer also have 10 copies of lysozyme gene, while a goat only has seven (Irwin, Prager, and Wilson, 1992).

C. Acquisition of Calcium Ion Binding

α-Lactalbumin was discovered to be a Ca^{2+} binding metallo-protein by Hiraoka et al. (1980). No α-lactalbumin has yet been found that does not posses a strong Ca^{2+} binding site. Lysozymes, on the other hand, displayed no strong affinity for calcium ion until Nitta et al. (1987) reported that equine milk lysozyme had one bound calcium ion per mole just as did bovine α-lactalbumin. Human and HEW lysozyme did not bind Ca^{2+} stably enough to coelute with the protein. Later, donkey lysozyme was also shown to possess the strong Ca^{2+} binding site (Godovac-Zimmermann, Conti, and Napolitano, 1988) and so did the lysozyme from pigeon egg white (Nitta et al., 1988) and Echidna milk (Acharya et al., 1994). The Ca^{2+} binding lysozymes now belong to a category different from the non-metal binding conventional lysozyme C.

Initially, it was argued that α-lactalbumin acquired the Ca^{2+} binding characteristic during its evolution following lysozyme gene duplication. The alternative possibility that the ancestral lysozyme or its progenitor was a metalloprotein and that some of the current lysozymes lost the metal binding capacity was then ruled out (Stuart et al., 1986). Analysis of the relationship between DNA sequences of moth, trout, and some tetrapod lysozymes with mammalian α-lactalbumin and calcium binding lysozymes led Dautigny et al. (1991) to conclude that both the calcium binding proteins originated 400 Myr ago. Recently, two different forms of lysozymes were obtained from canine milk and spleen, respectively (Grobler et al., 1994a). The milk enzyme was of the nonconventional type and bound one Ca-ion, whereas the spleen enzyme was more similar to the avian and mammalian enzyme and did not bind Ca-ion. Therefore, tests were conducted on a variety of models for the evolution of the lysozyme family of proteins from a number of species. The two alternate routes considered for the calcium binding property were (1) the ancestral protein was a non Ca-ion binding one that later acquired this ability, and (2) Ca-ion binding was an early feature that was lost later in the conventional and insect lysozymes. The results of the investigation tended to favor the second possibility (Grobler et al., 1994a). Obviously, because examples known of calcium binding lysozymes being few in number, more extensive investigations are needed to establish this hypothesis.

Little data are available on the origin of the goose type of lysozyme or the phage lysozyme, undoubtedly because of scant identity in amino acid sequences. Goose lysozyme is mentioned as the product of a different genetic locus from that of lysozyme C (Wilson, Carlson, and White, 1977). Therefore, it is significant that there are conspicuous similarities in their crystal structures (Grütter, Weaver, and Mathews, 1983; Weaver et al., 1985). There are parts in the molecule that are common to all and then there are parts that are common to HEW and goose enzymes, while other parts are common to goose and T4 enzymes (Figure 5). The common ancestry of these enzymes is obvious from the similarities of the substrate binding cleft and the ability to superposition the common structural parts. However, amino acid sequences have altered to a point where the sequence similarities are hardly discernable. Alternately, the current enzymes are the products of convergent evolution, a possibility that at present appears to be less likely.

D. Gene Organization and Structural Homology

DNA sequences of the genes of chicken lysozyme (Jung et al., 1980), human lysozyme (Peters et al., 1989), and of six α-lactalbumins, rat (Qasba and Safaya, 1984), mouse (Vilotte and Soulier, 1992), cow (Vilotte et al., 1987), human (Hall et al., 1987), goat (Vilotte et al., 1991), and guinea-pig (Laird et al., 1988) have been determined. Each consists of four exons and three introns. The arrangement of these is shown schematically in Figure 7. The exon-intron boundaries are, as expected, precisely defined and conserved. The three exon-intron junctions split coding sequence of specific amino acids in exactly the same manner. Tryptophan coding TGG occurs at the junction between the first and second exons, the first base being at the 3′ end of Exon 1, the second and third at the 5′-end of Exon 2. One exception is the human α-lactalbumin in which the split codon is TTG, which codes for Leu-26, between Exon 1 and 2. The codon split between Exons 2 and 3 is that of Ala–82 in and of Lys-79 in α-lactalbumins. The third codon split between Exons 3 and 4 is also TGG, coding for Trp-108 in lysozyme and 104 in α-lactalbumins. In every instance the first two bases of this codon are located in Exon 3 and the third in Exon 4. This figure also shows that there is considerable difference between genes of the two proteins in the lengths of introns, goat and bovine lactalbumins being the most similar. The boundaries of all of the introns contain GT at the 5′ and followed by G more often than A. At the 3′ ends the sequence in all cases is CAG. It would appear the consensus splice signal is a trinucleotide. In many of the introns the similarities of the nucleotide residues extend even further. All the introns of goat and bovine α-lactalbumins have the sequence at GTGAGT at the 5′ end. It is interesting to note that although the amino acids flanking those that constitute the exon-intron junctions in each of the proteins are also conserved they are not necessarily the same ones in the two proteins. For instance, residue number 27 and 29 are invariant asparagine and conserved valine, respectively, in lysozymes. In contrast, in α-lactalbumins the corresponding position 25 is occupied by the invariant glutamic acid and position 27 by isoleucine, threonine, or valine. The flanking amino acids at the other junctions show similar differences between lysozymes and α-lactalbumins (Acharya et al., 1989).

An examination of the amino acids conserved in lysozymes and α-lactalbumins of a variety of species shows that, while there is less than 40% overall sequence identity

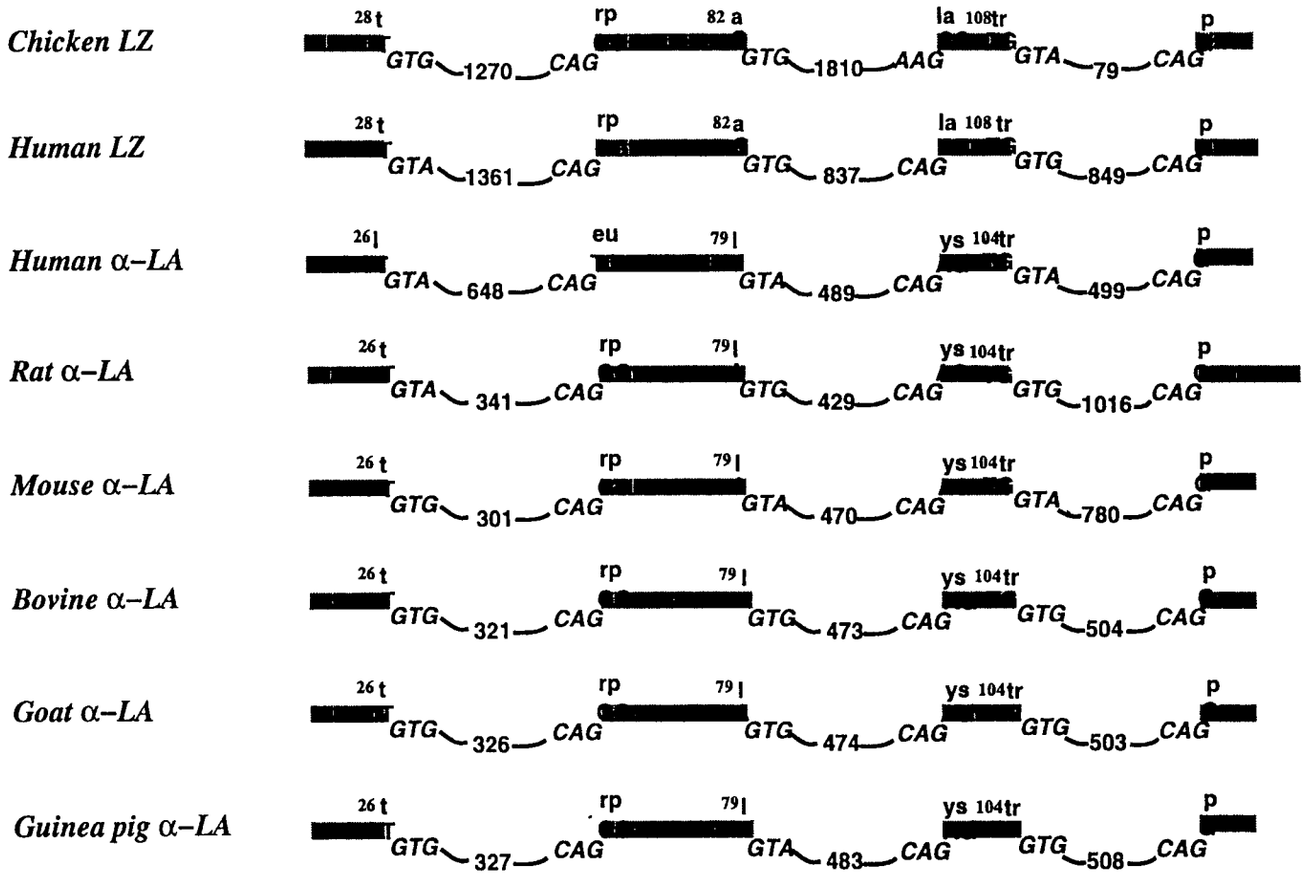

FIGURE 7. Organization of the exons and introns in lysozymes and α-lactalbumins. The solid bars represent the four exons and the wavy line below the introns. The bases at the junctions are also shown. The amino acids whose codons in the exons are split, the manner of the split, and the positions of the amino acids are indicated. The numbers within the intron loops are the number of the base pairs in them.

the conserved amino acids are not scattered uniformly throughout the molecules (Table 2 and Plate 2). Clearly, there are higher proportions of conserved amino acids in the segments that are coded in Exons 2 and 3 than in 1 and 4. 50% of the amino acids encoded in Exon 3 are conserved, while those in Exons 1 and 4 are only around 25.

Acharya et al. (1989) and Jung et al. (1980) have pointed out that Exon 1 encodes amino acids that constitute α-helix A and the short 3_{10}-helix residues. Exon 2 encodes most of the amino acids involved in the binding of the substrate to lysozyme and Glu 35 and Asp 52 that participate in its catalytic action. This region of the protein contains the α-helix B, the two anti-parallel β-pleated sheets, and the loop formed by the disulfide bridge between cyteines 61 and 77. The substrate is bound in the cleft, one ridge of which contains the β-sheet and the other forms a part of an α-helix and peptide segments that lack secondary structures.

Exon 3 encodes some of the substrate-binding amino acids and the loop in which the calcium ion is bound in α-lactalbumin. From the extent of homology of amino acids, this segment appears to be even more important for the maintenance of the structures of the two proteins.

The amino acid sequences encoded in Exon 4, the shortest segment, shows only 23% identity between lysozymes and α-lactalbumins (Table 2). However, of the 20 carboxyl terminal amino acids in α-lactalbumins, 17 (85%) are either invariant or conserved (Table 4). The conspicuous difference in Exon 4 of lysozymes and

TABLE 4
Amino Acids Found in the 20 C-Terminal Residues of α-Lactalbumins

Residue Number	Invariant	Conserved
104	Trp(Junction)(11)	
105		Leu(9); Glu(1); Lys(1)
106	Ala(11)	
107		His(10); Trp(1)
108		Lys(9); Gln(1); Glu(1)
110		Leu(8); Phe(2); Met(1)
111	Cys(11)	
112		Ser(8); Thr(1); Leu(1); Asn(1)
113		Glu(9); Asp(1); Ser(1)
114		Lys(8); Asn(1); Asp(1); deletion (1)
115		Leu(10); Asp(1)
116		Glu(6); Asp(4); Leu(1)
117		Gln(10); Asp(1)
118	Trp(11)	
119		Leu(5); Tyr(1); Val(1); Arg(3); Asn(1)
120	Cys(11)	
121		Glu(9); Lys(1); deletion(1)
122		Lys(7); Ala(1); Glu(1); deletion(2)
123		Leu(5); Gln(1); Pro(1); Trp(1) deletion(3)

Note: The above information compiled from the reports, the references of which are provided under Table 2. The numbers in parenthesis refer to the number of species in which each amino acid is found. The platypus C-terminal sequence has not been included in this compilation.

α-lactalbulmin led Qasba and Safaya (1984) and Vilotte et al. (1987) to propose a different origin for this exon in the two genes. The alternative explanation, a more rapid rate of mutations in this exon than in the other three was also considered by Qasba and Safaya (1984).

Investigation of the nature of gene organization has revealed other interesting features that the significance of which are not fully understood. 5′- and 3′-flanking regions of α-lactalbumin genes exhibit a high degree of sequence similarity. At the 5′ end it extends to –655 nucleotide from the transcription point in bovine α-lactalbumin gene. Within this there are two detectable conserved sequences besides the TATA and CAAT boxes. One occurs at about –140 to –110 in the human gene and consists of the consensus sequence, RGAAGRAAA(N)TGGACAGA-AATCAA (CG)-TTTCTA in which R stands for a purine and N for any nucleotide (Hall et al., 1987). This stretch has been subsequently found in guinea-pig (Laird et al., 1988), bovine, and goat genes (Vilotte et al., 1991; Vilotte and Soulier, 1992) at about the same location. The presence of a separate sequence in the genes of several caseins indicated either a hormone-sensitive regulatory role for it or one involved in the expression of milk protein genes. This stretch of conserved sequence, designated the milk box (Hall et al., 1987), occurs at about –280 and is present in all milk protein-encoding genes. An exception is κ-casein gene (Vilotte and Soulier, 1992). It is also not present in the lysozyme gene.

Qasba and Safaya (1984) had noticed two $(TG)_n$ repeats in intron-3 of rat α-lactalbumin gene. These sequences are believed to be involved in genetic recombination, rearrangement, or regulation of gene expression. As such, they were implicated in the unusual divergence of Exon 4 of α-lactalbumin from lysozyme. It was later shown (Meera, Ramesh, and Brahmachari, 1989) that the $(TG)_n$ repeats in this gene enabled the transition of DNA from the β-conformation to the left- handed Z-helices at varying levels of supercoiling to facilitate the processes involved.

Other sequences present in the gene include *alu* repeats in human α-lactalbumin intron-1, which is not present in rat (Hall et al., 1987). A second possible site was located at around –656 in the human gene. Peters et al. (1989) found one *alu* repeat in each intron and one in the non-coding region of Exon-4 of human lysozyme gene.

E. Absence of Enzyme Activity in α-Lactalbumin

The lysozyme molecule, superficially viewed, consists of two domains separated by a deep cleft. The larger of the two domains contains all the helices and comprises the N- and C-terminal residues, 1–36 and 89–129. The smaller domain is constituted by the intervening amino acid residues that form the β-sheet and a number of loops. These are encoded mostly in Exons 2 and 3, while the larger domain is encoded in Exon 1 and 4.

The close similarity of the three-dimensional structures of lysozyme and α-lactalbumin, whose genetic origins may go as far back 400 Myr (Figure 6) is an indication of the powerful drive to preseve the stable structure presumably formed then. The primary structure continued to be subjected to the inexorable changes. This implies that only those mutations endured that allowed the retention of the tertiary structure. The cleft being an intrinsic part of the structure of lysozyme, it was preserved in α-lactalbumin even though a substantial part, if not all, of it has no known functional role. The

evolution of α-lactalbumin followed crucial alterations in the primary structure of lysozyme that resulted in the loss of not only its enzyme activity but of even the capacity to bind the substrates or their analogs. The explanation is obvious when the amino acids in lysozyme involved in the binding of the substrates and their hydrolysis are examined with respect to those present in the corresponding positions in α-lactalbumins. Table 5 lists these amino acids. In addition, there are many water-mediated contacts between the atoms of the substrate and of the enzyme (Song et al., 1994). Of the 12 conserved amino acids reported to be directly involved in the binding of $(NAG\text{-}NAM)_3$ (Phillips et al., 1974), only five are found in α-lactalbumin from a significant number of species. Of these only Gln-59 is present in the α-lactalbumin of all species. Gln-57 and Trp-63 occur in all but one of the α-lactalbumins. Asn-46 is retained as Asn-44 in 9 of the 11 and Asp-107 as Asp-97 in all but 3. The rest of the amino acids have been replaced with residues that are substantially different. Residue number 34, which consist of aromatic side chains in lysozyme, is replaced with invariant His-32. The positions of the invariant Glu-35 in lysozyme that has an essential catalytic role and that is also involved in binding the substrate is occupied by an amino acid other than glutamate in all but one α-lactalbumin. The invariant lysozyme Asn-44 is found in

TABLE 5
The Amino Acids of HEW Lysozyme Interacting with $(NAG\text{-}NAM)_3$ and the Correspondng Residues in α-Lactalbumins

	Lysozymes		α-Lactalbumin
Residue no.	Majority of the amino acids	Interaction/ role	Corresponding number and the amino acid
34	Aromatic (17)	Peptide O	32: His(11)
35	Glu (17)	Peptide O; catalytic	33: Glu(1); Thr(5); others
37	Asn (10)	Amide N	35: Gly(11)
44	Asn (17)	Amide N	42: Val(10); Asn(1)
46	Asn (14)	Amide N	44: Asp(9); Tyr(1); Asp(1)
52	Asp (16)	Carbonyl O; catalytic	49: Glu(9); Ser(1); Phe(1)
57	Gln (16)	Amide O; peptidyl O	54: Gln(10); Asp(1)
59	Gln (16)	Peptide N	56: Gln(11)
62	Trp (13); Tyr (3)	Indole N	59: Ile(3); Leu(3); others
63	Trp (17)	Indole N	60: Trp(10); Phe(1)
101	Asp (13)	Carboxyl O	97: Asp(8); Glu(1); Ala(1); Leu(1)
107	Ala (17)	Peptide O	103: Tyr(9); Ala(1); His(1)
109	Val (15); Lys (2)	Peptide N	105: Leu(9); Lys(1); Glu(1)
114	Arg (7); His (7)	Guanidinium N; imidazole N	110: Leu(8); Phe(2); Met(1)

Note: The amino acids reported to interact are those listed by Phillips (1974) and Weaver, Grütter, and Matthews (1995). The data were taken from the sequences given in Table 1 (Acharya et al., 1989) and of the pig lysozyme (Jollès et al., 1989); α-lactalbumin sequences were from Acharya et al. (1989) and pig (Godovac-Zimmerman and Napolino, 1990) and platypus (Shaw et al., 1993). The data represents a total of 17 lysozymes and 11 α-lactalbumins. The numbers in parentheses are the number of species that contain the particular amino acid.

none of the α-lactalbumin, and neither are Asp-52, Trp-62, Ala-107, and Arg-114.

Determination of the crystal structure of baboon α-lactalbumin revealed that Tyr-103, which replaces Ala-107 of lysozymes, covers the cleft and blocks substrate binding at a site where the A and B sachharides of the hexamer substrate binds (Acharya et al., 1989). A similar situation is found in human α-lactalbumin (Acharya et al., 1990). The orientation of this tyrosine is such that only one internal water molecule remains here compared with three in lysozyme (Song et al., 1994). The lack of enzymatic activity is also due to the replacement of the crucial Glu-35 and of Asp-52, which aids in the process. Otherwise, a tetrameric or trimeric substrate might be expected to bind to α-lactalbumin at the binding sites C to F and undergo hydrolysis at the site between D and E.

Studies have been initiated in various laboratories to determine the role of individual amino acids in providing stability to these proteins and to evaluate the contribution of others that are functionally important by means of site-specific substitutions of selected amino acids or a stretch of these. Thr-40, Ile-55, and Ser-91 form part of cluster of amino acid in the interior of the lysozyme molecule just underneath the cleft. Their replacement with serine, valine, and threonine showed no overall structural changes (Wilson, Malcolm, and Matthews, 1992). Only the packing of the substituent side chains were altered and that reflected their accommodation in this region. They also had altered thermal stabilities. A hyperstable mutant of HEW lysozyme has been created by the incorporation of the following substitutions: H15L, A31V, I55L, S91T, D101S, and R114H. The cumulative effect on its thermostability was significally higher, T_m being + 10.5° and a bacteriolytic activity that was 2.5 times higher than that of the wild type (Shih and Kirsch, 1995). The mutant lysozyme D52S had a crystal structure similar to the wild type except for a slight displacement of a loop between the β-strands. It did bind the hexameric substrate but had less than 1% of the enzyme activity (Hadfield et al., 1994). Substitution of Trp-62 with tyrosine or phenylalanine decreased its affinity for the substrate, as reflected in a lower of K_m. However, it had increased catalytic activity (Maenaka et al., 1995).

Kumagai, Takeda, and Miura (1992) succeeded in creating hybrid proteins from goat α-lactalbumin gene in which its Exon 2 by itself or both Exon 2 and 3 were replaced with HEW lysozyme counterparts. A third mutant gene had its cleft blocking Tyr-103 replaced with alanine and Ile-59 with tryptophan, the natural acids in lysozyme. A fourth mutant had 10 substitutions, including that of Glu-49, which was changed back to the catalylytically helpful Asp-52 of lysozyme. The chimeric α-lactalbumin that contained the regions of Exon 2 and Exons 2 and 3 had now become catallytically active, although their activities were considerably lower than the wild-type protein. Strangely, the mutant that possessed most of the substrate binding as well as the two catalytic amino acids exhibited no enzyme activity. The results were indicative of the importance of additional amino acids, not recognized so far, for generating catalytic activity.

F. Calcium-Ion Binding Loop

Metal-ion binding by α-lactalbumin has attracted much attention ever since the presence in it of a strongly bound Ca-ion was reported by Hiraoka et al. (1980). This protein also showed evidence of interaction with a variety of other mono-, di-, and trivalent metal ions (Kronman, 1989; McKenzie

and White, 1991; Permyakov and Berliner, 1994), Of these, Ca^{2+} certainly remains bound under physiological conditions. It dissociates from the protein only at or below pH 4 when α-lactalbumin denatures (Kronman, 1989). Association of the other metal ions follow more complex patterns (Kronman, 1989; McKenzie and White, 1991; Permyakov and Berliner, 1994).

Calcium ion is held in place in α-lactalbumin by its coordination to seven oxygen atoms, three contributed by aspartyl carboxylates, two by peptidyl carbonyl groups, and two by different water molecules (Stuart et al., 1986). The arrangement of these ligands is that of a slightly distorted bipyramidyl pentagon with the Ca^{2+} at its center and the two peptidyl oxygen atoms at the pyramidyl apices. Ca-O distances of each of the ligands is between 2.3 Å and 2.5 Å. To obtain such an arrangement, apparently ideal for the coordination of Ca^{2+}, a segment of 10 contiguous amino acid residues, 79 to 88, bends to form an elbow or loop. This stretch of amino acids is flanked by two disulfide bridges, Cys 77 to Cys-61 and Cys-91 to Cys-73 (Plates 2 and 3). Cys-77 is located in a 3_{10} helix and Cys-91 in an α-helix (Acharya et al., 1989). Although the helix-loop-helix arrangement superficially resembles the E-F hand motif found in a number of Ca(II)-binding proteins, such as calmodulin, troponin C, and parvalbumin, it differs from it in a number of significant ways. (1) It is constituted by only 10 amino acid residues, two shorter than those of the typical Ca(II)-binding proteins. (2) It is flanked by two disulfide bridges, which make this segment much less flexible. (3) There is no amino acid sequence resemblence between them. (4) Finally, a conserved glycine residue found in position 6 of 90% of Ca(II)-binding protein is absent in the loop of α-lactalbumin (Acharya et al., 1989; Stuart et al., 1986). The geometry of the corresponding region in Ca-ion lacking lysozyme is similar except for the orientation of the side chains of some of the amino acids that still constitute the loop. These side chains point outward in lysozyme (Acharya et al., 1989).

The precise role of calcium ion bound to α-lactalbumin in lactose synthesis is unclear despite extensive investigation. Indisputably, it facilitates, and may even be indispensable, for the folding of the nascent or unfolded protein (Kronman, 1989). Renaturation of the completely denatured α-lactalbumin (U-state) has been reported to occur in two steps in a manner similar to the refolding of many other proteins. The first step, a very rapid process, involves the acquisition of the secondary structures to form a molten globule (A-state). In the second slower and rate-determining step, the various secondary structures coalesce with correct S-S pairing to form the native N-state (Kronman, 1989; Kuwajima, Mitani, and Sugai, 1990; Sugai and Ikeguchi, 1994). Calcium ion is involved in the latter process. Refolding of reduced, denatured apo-α-lactalbumin occurs at pH 8.0, or above, in the presence of a mixture of oxidized and reduced glutathione. When EGTA was present no more than 2% of the protein recovered its enzyme activity, a proportion that would be expected if random pairing of the eight SH groups had occurred. Analysis also indicated the presence of products containing mismatched disulfide bridges. In the presence of excess $CaCl_2$ reoxidation occurred rapidly and with correct SH pairing (Rao and Brew, 1989). Ewbank and Creighton (1993) have reported that reduced α-lactalbumin exists in various unstable conformations in equilibrium with each other and which undergo rapid disulfide-thiol exchanges. The native protein was regenerated only in the presence of calcium ion, which initially stabilized a three disulfide form from which the native state was derived rapidly

following the correct formation of the fourth disulfide.

The Ca(II) binding feature initially appeared to be one of the properties that distinguished α-lactalbumin from lysozymes. Following the identification of the Ca-ion ligands in bovine α-lactalbumin, when it was observed that three of the aspartyl residues were also present in equine lysozyme, Stuart et al. (1986) raised the possibility that equine lysozyme might also bind this metal ion. This prediction was soon shown to be true (Nitta et al., 1987). Later, pigeon egg white lysozyme was also found to bind one mol of Ca(II) (Nitta et al., 1980). Lysozyme from both sources exhibited an affinity for the metal ion similar to that of α-lactalbumins. The Ca-ion binding feature thus was no longer a distinguishing property of α-lactalbumin. Since then the list of calcium-ion binding lysoymes has grown. All of them possess highly conserved amino acid residues in the loop region. The data compiled from the various sources are presented in Table 3. The calcium-ion binding site of equine lysozyme had a geometry superimposable on bovine α-lactalbumin (Tsuge et al., 1992). This similarity was later reported to extend to echidna, guinea-pig, and goat lysozyme (Acharya et al., 1994; Pike, Brew and Acharya, 1996).

It can be seen from the data (Table 6) that besides the presence of the three aspartyl residues, 82, 87, and 88, which in α-lactalbumin provide the carboxyl oxygen atoms as ligands to the calcium ion, other amino acids in the region are also well conserved. Especially noteworthy is lysine-79, which is present in α-lactalbumin from all but 1 of 12 species. In the Ca-ion binding lysozymes, this amino acid is also found in all of the five species, as are the three aspartyl residues. These acids are generally absent in lysozymes that do not bind Ca-ion. Asp-84, the other peptidyl O-liganded residue, is found in 9 out of 12 α-lactalbumins. In the three remaining species it is substituted by glutamic acid in two and by asparagine in one. Calcium-ion binding lysozymes contain mostly asparagine. In contrast, this region in the conventional lysozyme shows much greater variability in sequences. Three residues numbered 80, 85, and 89 are conserved in both α-lactalbumin and lysozymes, whether the latter bind the metal-ion or not.

Rao and Brew (1989) and Kronman (1989) independently proposed that Ca(II) neutralizes the negative charges of some of the carboxyl groups of six closely placed aspartyl residues that would have otherwise hindered the folding of the molecule in this region and thereby prevented the correct pairing of the flanking SH-groups. This view is supported by the observed differences in the kinetics of folding of the unfolded α-lactalbumin and of the non-calcium binding lysozyme (Kronman, 1989; Sugai and Ikeguchi, 1994). A significant observation was that the geometry of the loop region of HEW lysozyme, despite the absence of Ca(II), is remarkably similar to that of α-lactalbumin (Acharya et al., 1989).

An examination was made of the loop region of α-lactalbumin using Quanta, a molecular modeling software from Polygene in Silicon Graphics work station IRIS 4D/70. The coordinates, obtained from Brookhaven Protein Data Bank, revealed that the side chain of Lys-79 pointed inwards, toward the Ca-ion and not outward as does Ala–82, the corresponding residue in lysozyme. Its ε-amino group is located at a distance of 4.49 Å from one of the two water molecules (No–63, according to the number assigned by Acharya et al. [1989]) and that is liganded to the calciumion. This amino group is also only 4.39 Å away from the carboxyl oxygen atoms of each of Asp-82 and Asp-87 (Plate 5).* Therefore, it is ideally situated to interact with these

* Plate 5 appears after page 274.

TABLE 6
Amino Acid Sequence of the Calcium-Binding and Flanking Regions of α-Lactalbumin and the Corresponded Regions of Lysozyme

α-Lactalbumin

Species	Residues 77	78	79*	80	81	82*	83	84*	85	86	87*	88*	89	90	91
Human	C	D	K	F	L	D	D	D	I	T	D	D	I	M	C
Guinea-pig	C	D	K	L	L	D	D	D	L	T	D	D	I	M	C
Cow	C	D	K	F	L	D	D	D	L	T	D	D	I	M	C
Rat	C	D	K	F	L	D	D	E	L	A	D	D	I	V	C
Mouse	C	D	K	L	L	D	D	E	L	D	D	D	I	A	C
Goat	C	D	K	F	L	D	D	D	L	T	D	D	I	V	C
Rabbit	C	E	N	F	L	D	D	N	L	T	D	D	V	K	C
Red Wallaby	Neck C	S	K	F	L	D	D	D	I	T	D	D	I	E	C
Camel	C	D	K	F	L	D	D	D	L	T	D	D	K	M	C
Horse	C	D	K	F	L	D	D	D	L	T	D	D	V	M	C
Pig	C	D	K	F	L	D	D	D	L	T	D	D	D	M	C
Platypus	C	S	K	L	L	D	D	D	I	L	D	D	I	K	C

Lysozyme Ca(II) binding lysozymes

	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94
Horse	C	S	K	L	L	D	E	N	I	D	D	D	I	S	C
Donkey	C	S	K	L	L	D	D	N	I	D	D	D	I	S	C
Pigeon	C	S	K	L	R	D	D	N	I	A	D	D	I	Q	C
Dog	C	S	K	F	L	D	D	N	I	D	D	D	I	A	C
Echidna	C	S	K	L	L	D	D	D	E	T	D	D	L	K	C

Non-calcium-binding lysozyme

Human	**C**	**S**	**A**	**L**	**L**	**Q**	**D**	**N**	**I**	**A**	**D**	**A**	**V**	**A**	**C**
Other Species	12C	11S 1N	10A 1K 1E	12L	9L 2M 1Q	7S 1E 1G 3Q	8S 2D 1N 1A	11D 1N	12I	10T 2A	6A 1D 1K 2E 1P 1Q	5A 4S 3T	10V 1I 1A	6N 3R 1Q 2R	12C

▲ Ca(II)-liganded residues. Residues 79 and 84 of α-Lactalbumin are liganded via carbonyl oxygen and 82, 87, and 88 via carboxyl oxygen, to Ca(II). Most of the data are derived from Table 1 (Acharya et al., 1989). In addition, the sources were α-lactalbumins from mouse (Villotte, Soulier, and Mercer, 1992); pig (Godovac- Zimmerman, Conti, and Napolitano, 1990); platypus (Shaw et al., 1993); lysozymes from dog (Grobler et al., 1994a), echidna (Teahan et al., 1991). The residue numbers are those of HEW lysozyme. Those of the others are adjusted to account for additions and deletions. In human lysozyme, for instance, this segment actually consists of residues 81 to 95 (Canfield, Collins, and Sobell, 1972; Peters et al., 1989).

carboxyl O-atoms as well as with one of the Ca(II)-liganded water molecules via other water molecules, as proposed by Ben-Naim, Ting, and Jernigan (1990). Such interactions, if they do occur, can be expected to further enhance the free energy of the binding of Ca(II) to the protein. These and some other possible ones are shown in Table 7.

An alternative suggestion is that the ε-NH_3^+ group of the Lys-79 situated close to the Ca(II)- binding site aids in further neutralizing the cluster of negative charges than Ca-ion alone (Haezebrouck et al., 1993). It is also to be expected that the other invariant and conserved residues, Asp-78, Phe-80, Leu-81, Asp-83, Asp-78, Leu-86, and Ile-89 (Table 6) would have some role in contributing toward the stability of this region of the molecule.

In its role as a modifier of galactosyltranferase, shifting the acceptor specificity of the enzyme, Ca^{2+}-lacking apo-lactalbumin was observed to be more effective than the Ca^{2+}-bound α-lactalbumin (Musci and Berliner, 1985). Although calcium ion stabilizes α-lactalbumin against thermal denaturation, Zn-ion induces a change in conformation that is closer to the less taut A-state in which this protein is believed to function as a component of the lactose synthase complex (Musci and Berliner, 1985). It has been proposed that the opposing actions of these two metal ions helps to fine tune the conformations of α-lactalbumin in order to modulate the activity of the

TABLE 7
Interactions of the Amino Acid Residues of the Calcium-Ion Binding Loop of Bovine α-Lactalbumin

			Interactions	
Residue	**Group**	**Orientation**	**Reported**	**Possible**
Asp-78[a]	Carboxylate	Outward		Solvent
Lys-79[b]	Carbonyl	Inward	Ca^{2+}	
	ε-Amino	Inward		H_2O—Ca^{2+}
Phe-80[a]	Phenyl	Outward		Ile-75
Leu-81[b]	Side Chain	Outward		Tyr-53
Asp-82[b]	Carboxylate	Inward	Ca^{2+}	
Asp-83[b]	Carboxylate	Outward		Solvent
Asp-84[a]	Carbonyl	Inward	Ca^{2+}	
	Carboxylate	Outward		Thr-86
Leu-85[a]	Side Chain	Outward		Ile-89
Thr-86[a]	Side Chain	Outward		Asp-84
Asp-87[b]	Carboxylate	Inward	Ca^{2+}	
Asp-88[b]	Carboxylate	Inward	Ca^{2+}	
Ile-89	Side Chain	Outward		Into the protein core

[a] Conserved amino acids.
[b] Invariant amino acids.

Note: The examination of the orientation of the side chains of the amino acid constituting the Ca^{2+} loop suggested that Phe-80, Leu-81, and Ile-89 had theirs pointing away from the loop toward the body of the protein; Asp-78, Asp 83 had theirs pointing toward the exterior. E-amino group pointed toward Ca^{2+} and was only 4.4 Å away from it.

galactosyltransferase (Musci and Berliner, 1985).

Recent investigations using mutants created by the substitution of specific amino acids have provided valuable information on the binding of Ca-ion by lysozymes. It had been observed earlier that the amino acid sequence of the loop of human lysozyme provides a partially formed Ca^{2+} binding site (Acharya et al., 1989). Even though it has only one (Asp-91) of the three carboxyl liganding aspartates, it does have the conserved Asp-87 of α-lactalbumin. The latter points outward, however, and hence should play no role in Ca-ion binding. (The numbering of the amino acids for human lysozyme in this discussion are the actual ones and differ from those in Table 6 by one.) The amide nitrogen of Gln-86 can replace one of the Ca^{2+}-liganding water molecule (Acharya et al., 1989). Kuroki et al. (1989) designed a mutant human lysozyme in which Gln-86 and Ala–92 were replaced with aspartate residues in order to obtain the three aspartyl residues that bind Ca^{2+} via the carboxyl groups (Figure 8). This D86/92 mutant now bound 1 mol of calcium ion with a binding constant of $5.0 \times 10^6\ M^{-1}$. The value for bovine α-lactalbumin is $4 \times 10^7\ M^{-1}$. Both apo- and Ca-ion bound holo D86/92 mutant were enzymatically active. The maximum activity of holo D86/92 lysozyme was at 80°, whereas the activities of apo-D86/92 mutant and the native enzyme were 65° to 70°, respectively. The Ca-ion bound mutant was more resistant to pronase digestion and had higher thermal stability. The denaturation temperatures were 80.3°, 76.5°, and 85.2°, respectively, for the wild type, apo-D86/92, and holo-D86/92. The binding of Ca(II) by the D86/92 mutant in the folded state was determined to be driven primarily by an increase in entropy, resulting from the release of Ca(II)-bound water molecules, as enthalpy changes were unfavorable (Kuroki et al., 1992).

Haezebrouck et al. (1993) created two mutants of human lysozyme, one in which Ala-92 was replaced with aspartate. This mutant (A92D) would thus contain two of the three Ca^{2+}-binding aspartyl carboxyl groups. The other mutant, M4, pictorially shown in Figure 8, contained three additional substitutions so that it had all five of the Ca^{2+}-binding ligands, Lys-83, Asp-86, Asp-88, Asp-91, and Asp-92 corresponding to Lys-79, Aspartyl 82, 84, 87, and 88 of α-lactalbumin (Table 6). The binding constants for the calcium ions were $2(\pm 1) \times 10^2\ M^{-1}$, $8 \times 10^3\ M^{-1}$, 1×10^7, and $1 \times 10^9\ M^{-1}$ for human wild-type lysozyme, A92D, M4, and bovine α-lactalbumin, respectively. These studies showed clearly that three aspartates, 86, 91, and 92 are necessary to create a strong calcium binding site; two were not sufficient. Further, the additional substitution of Ala–83 by Lys and Asn-88 by Asp, the carbonyl ligands of Ca-ion, doubles its affinity. A chimeric human lysozyme created by Pardon et al. (1995) by the replacement of nucleotides coding for 27 amino acid residues, comprising the calcium-ion binding loop as well as its flanking amino acids, with those found in bovine α-lactalbumin showed a Ca-ion binding constant of $2.5 \times 10^8\ M^{-1}$, one fourth of the K_A of α-lactalbumin. This chimeric molecule retained both of its catalytic residues, Glu-35 and Asp-52, and yet its enzymatic activity was only about 10% of that the native human lysozyme. This is indicative of the occurrence of subtle structural changes in the lysozyme molecule after the insertion of bovine α-lactalbumin sequences (Pardon et al., 1995).

G. Calcium Ion Binding and Divergence

The recognition of the strong Ca-ion binding site in α-lactalbumins and its absence in all the lysozymes then known

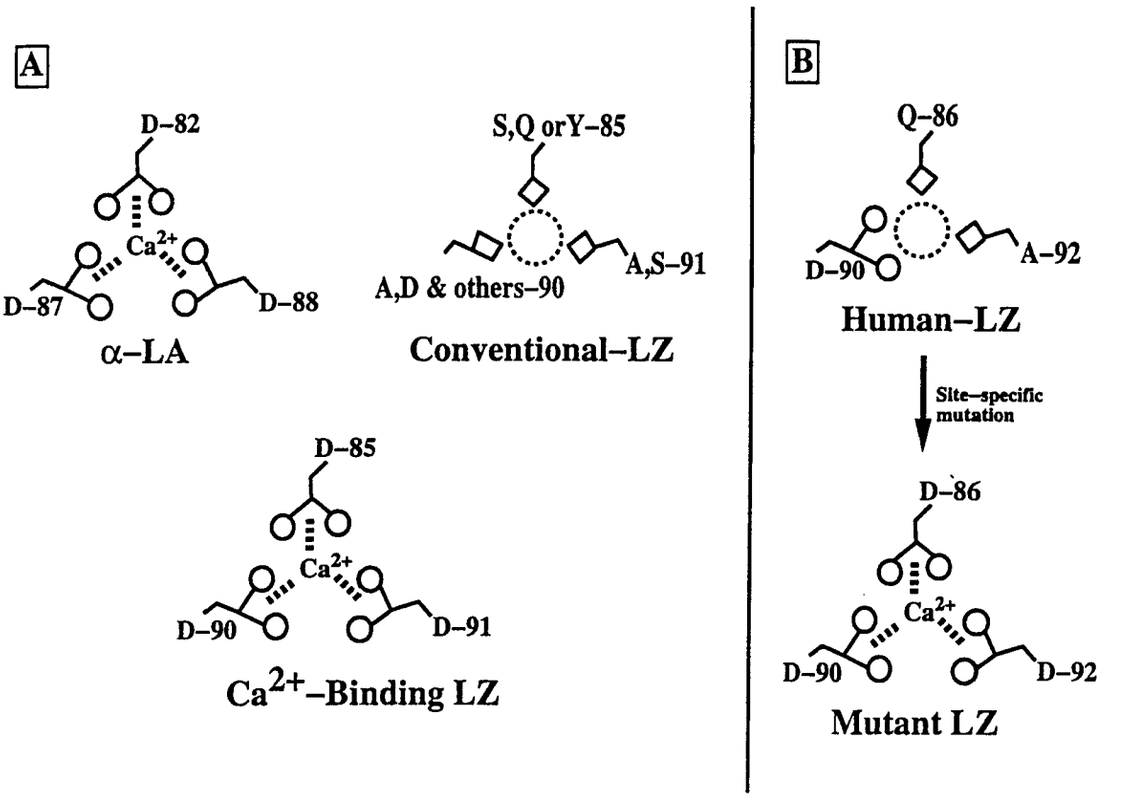

FIGURE 8. Schematic representation of Ca-ion binding in α-lactalbumin (α-LA), conventional and Ca^{2+}-binding nonconventional lysozymes (A). Two connected circles represent carboxyl groups and squares the side chains of other amino acid residues. They are identified and positions numbered. Not shown are the two carbonyl oxygen ligands and two water molecules that constitute the other ligands. Aspartyl residues of α-lactalbumin and the Ca^{2+}-binding lysozymes are absent in the conventional lysozyme. (B) The partially formed Ca-ion binding site of human lysozyme shows no affinity for Ca^{2+}. The substitution of Gln-86 and Ala–92 with Asp by mutagenesis converts it to a Ca- ion binding enzyme (Kuroki et al., 1989).

implied that this feature was an important factor in the divergence of these two proteins (Stuart et al., 1986). The differences between the larger EF hands of established calcium-ion binding proteins, including the lack of sequence homology with them, the dissimilarity in the geometry of the Ca^{2+}-coordination, a failure to find the pentagonal bipyramidal arrangement in any of the proteins, suggested an independent origin of the Ca-ion binding loop of α-lactalbumin (Stuart et al., 1986; Acharya et al., 1989). This view was further supported by the close similarity in the conformations of the backbone atoms of this region of bovine α-lactalbumin and of hen egg white lysozyme (Acharya et al., 1989). Based on these observations, Stuart et al. (1986) argued that the Ca-ion binding characteristic in α-lactalbumin arose during its divergence from lysozyme as a result of amino acid substitutions in this location, principally of the three aspartyl residues, 82, 87, and 88. The alternate possibility that their ancestral protein was a calcium ion binding one and that the current lysozymes lost this property was considered unlikely.

Prager and Wilson (1988) speculated on the possibility of the product of the duplicated lysozyme gene performing a role in

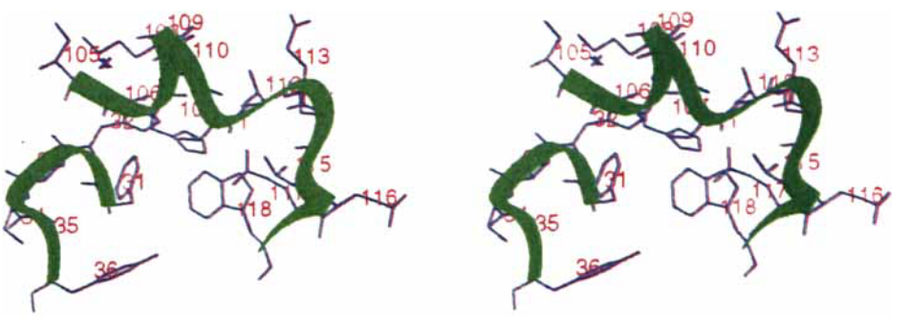

Plate 6. Arrangement of the amino acid residues of α-lactalbumin demonstrated to be involved in its interaction with galactosyltransferase. Phe-31, His-32, Gln-117 and Trp-118 constitute aromatic cluster I. The other amino acid residues, 105–110 are in the flexible C-terminal region.

ERRATUM

There is an error in the footnote for Plate 6 that appears in the article, "Molecular Divergence of Lysozymes and α-Lactalbumin," by Qasba and Kumar. The correct placement is following page 274. Please note this when referring to the plate in text. We apologize for the error.

calcium metabolism at some point during the long period intervening between gene duplication and the acquisition of its current role as a modifier of glycosyltransferase for lactose synthesis in mammals. Hayssen and Blackburn (1985) had proposed earlier that during the evolution of α-lactalbumin the protein product might have gone through an intermediate stage in which it possessed both functions. Echidna milk lysozyme is mentioned as such an evolutionary stage, more like α-lactalbumin than the conventional lysozyme (Hayssen and Blackburn, 1985; Acharya et al., 1994).

Nitta and Sugai (1989) conducted a comprehensive analysis of models of phylogenetic trees relating Ca^{2+}-binding lysozymes then available with conventional lysozymes C and α-lactalbumin. They concluded that the balance of different possibilities was in favor of the duplication of the ancestral lysozyme C gene sometime before the divergence of birds and mammals and that the duplicated gene acquired the ability to bind Ca-ion (Nitta and Sugai, 1989; Sugai and Ikeguchi, 1994). This accounted for the presence of Ca^{2+}-binding lysozyme in pigeon. At some point in the lineage leading to mammals another duplication of the Ca-binding lysozyme resulted in α-lactalbumin gene. The origin of the calcium-binding lysozyme was estimated to be about 300 Myr ago by Dautigny et al. (1991).

The belief in the acquisition of Ca-ion binding ability of α-lactalbumin and of some lysozymes lost ground when Grobler et al. (1994a) discovered two distinct types of lysozymes in dogs, one in milk and the other in spleen. Milk lysozyme contained 1 mol of Ca-ion per mol of protein. Spleen lysozyme, on the other hand, contained no bound Ca-ion. Sequence comparison revealed only 45% identity between them. Canine milk lysozyme was quite similar to the Ca-ion binding equine milk lysozyme, whereas spleen enzyme resembled the conventional mammalian lysozyme C. A preliminary investigation indicated that equine spleen also contained the conventional lysozyme but in much smaller amounts. Genes for both types are thus present and are expressed in each of these species. Therefore, these investigators question the divergence of α-lactalbumin from a single Ca-ion binding lysozyme gene proposed by Nitta and Sugai (1989). Instead, they suggest that Ca-ion binding was an ancient feature of the lysozyme family that slipped away from the line that led to the conventional lysozyme C. Nevertheless, it retained its molecular "memory" (Stuart et al., 1986). Hence, the metal-ion binding ability could be recovered simply by the introduction of the appropriate ligands. Further, amino acid sequences indicate that pigeon and echidna lysozymes arose from a gene line distinct from the one that produced mammalian Ca^{2+}-lysozymes (Grobler et al., 1994a).

The distribution of the Ca-ion binding feature of lysozymes in such diverse species as echidna, pigeon, dog, donkey, and horse raises important questions of evolutionary significance of this protein. If indeed this gene is much more widely distributed than currently recognized, which appears to be likely, the present uncertainty regarding the evolutionary relationship between its protein product, the conventional lysozyme, and α-lactalbumin will diminish only when structural data from a wider selection of species become available. Although our knowledge of the evolutionary history of phage lysozyme and lysozymes C and G is sparse, there is general agreement that α-lactalbumin diverged from lysozyme gene long before the appearance of mammals. The perplexing question is whether this gene remained silent during the intervening millions of years or it had some obscure functional role. If it was indeed silent another interesting question arises,

namely, what caused the persistence of this gene in a state that enabled it to encode a protein that largely retained its original three-dimensional structure.

H. Formation of a Complex Between α-Lactalbumin and Galactosyltransferase

The only feature remaining to be discussed that distinguishes α-lactalbumin from lysozyme is its ability to modify the reaction catalyzed by galactosyltransferase. A reconstruction of the evolution of this feature is even more elusive then the Ca-ion binding site of these two proteins.

The alteration induced in galactosyltransferase by α-lactalbumin, for example, the preference for the galactosyl acceptor, is an unusual, if not a unique, instance of a protein modifying the substrate specificity of an enzyme. Galactosytransferase belongs to a family of enzymes involved in the biosynthesis of the carbohydrate component of glycoproteins that perform a variety of essential biochemical roles (Masibay et al., 1993; Paulson and Colley, 1989). This enzyme under *in vitro* conditions transfers the galactosyl residue to *N*-acetyl-glucosamine to form *N*-acetyl-lactosamine. It shows no affinity for either lysozyme or α-lactalbumin per se. However, in the presence of its substrates, UDP-galactose, Mn^{2+}, either of the monosaccharides, *N*-acetyl-glucosamine or glucose, and at a relatively high concentration of α-lactalbumin it forms a 1:1 complex with α-lactalbumin, but not with lysozyme (Brew, 1970; Grobler et al., 1994b). There is an absolute requirement for a metal ion, Mn^{2+} being the most effective (Kronman, 1989; McKenzie and White, 1991). K_M of galactosyltransferase for glucose being high (2 m*M*), the transfer of galactose to glucose under physiological conditions is not important (Brew and Gobler, 1992). The situation in the presence of α-lactalbumin is quite different. UDP-galactose and Mn^{2+} together with glucose itself lower K_M for glucose by three orders of magnitude. Lactose synthesis then supersedes lactosamine synthesis (Richardson and Brew, 1979). When *N*-acetylglucosamine is provided as galactosyl acceptor α-lactalbumin reduces both K_M and V_{max} of the galactosyltransferase reaction so that α-lactalbumin becomes a strong inhibitor at high concentrations (Brew and Grobler, 1992). Although there is some disagreement on the kinetic mechanism followed by galactosyltransferase the one generally favored is the following. Mn-ion and UDP-galactose bind to the enzyme in that order, which is then followed by a random equilibrium binding of glucose and α-lactalbumin, their binding being synergistic. Each enhances the binding of the other (Brew and Grobler, 1992). Competitive inhibition by α-lactalbumin, when oligosaccharides are used as galactosyl acceptors, is taken as an indication that α-lactalbumin and oligosaccharide employ at least one common binding site near the catalytic center of the transferase. A model of lactose synthase, an adaptation of that proposed by Grobler et al. (1994b) is shown in Figure 9.

As mentioned earlier, α-lactalbumin modifies the reaction catalyzed by snail albumin transferase, UDP-GalNAc:GlcNAcβ-Rβ1→4-*N*-acetylgalactosaminyltransferase. However, it affects this enzyme differently from the manner it does bovine milk galactosyltransferase. It stimulates the transfer of GalNAc to GlcNAc and glucose (Neeleman and van den Eijnden, 1996). Curiously, the activity of a transferase that catalyzes the same reaction isolated from avian schistosome, *Trichobilharzia ocellata*, is not affected by α-lactalbumin. The molecular relationship between these enzymes is as yet not known. The association between

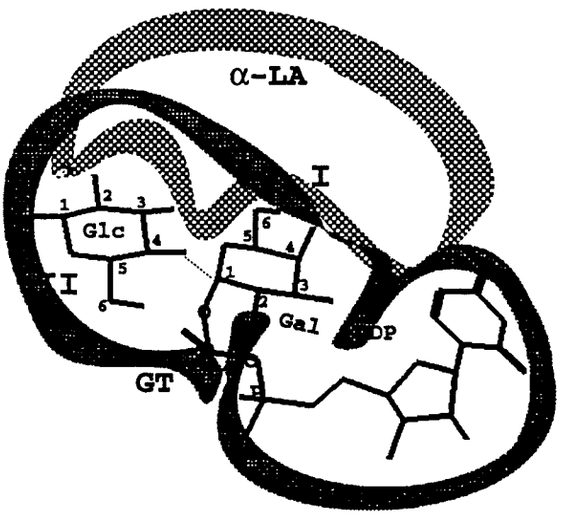

FIGURE 9. Schematic representation of lactose synthase complex adapted from the model proposed by Grobler et al., 1994b. The active site of galactosyltransferase is shown with bound UDP-galactose · Mn^{2+}. I represents the site where the galactose acceptor oligosaccharide would bind, which is occupied by α-lactalbumin. A binding site for glucose is then created out of the site, II (where *N*-acetylglucosamine can also be bound). Glucose is shown acting as a bridge between the two proteins. The orientation of the reacting sugar residues is presumptive. It accounts for the β-1,4-linkage in the products.

the two transferases, the only two proteins known, and α-lactalbumin raises questions of great importance regarding the evolution or divergence of α-lactalbumin.

Galactosyltransferase, normally a Golgi membrane protein, consists of four domains, a short N-terminal tail consisting of 23 to 24 amino acid residues extending into the cytoplasm, a trans-membrane signal anchor, an extended stem, and the catalytic domain residing in the lumen (Paulson and Colley, 1989; Masibay et al., 1993). The protein consists of about 400 amino acid residues with an approximate molecular weight of 50,000 (Yadav and Brew, 1990). Soluble forms of the enzyme, thought to be generated by the proteolysis of the catalytic component, is secreted in milk (Yadav and Brew, 1990). This constitutes the main source of the enzyme for the investigation of the formation of lactose synthase complex. It has been obtained also by the expression of a cDNA clones in *Escherichia coli* (Boeggeman et al., 1993; Wang et al., 1989). This has enabled the generation of specific mutants that are currently used to probe the nature of the complex (Grobler et al., 1994b; Pike, Brew, and Acharya, 1996).

Information about the nature of the catalytic site of galactosyltransferase is at present sparse. The catalytic activity resides in the C-terminal region because the deletion of 129 N-terminal amino acids results in little loss of enzymatic activities in the absence and presence of α-lactalbumin (Boeggeman et al., 1993). This work also indicated that Cys-134 is essential for the enzyme activities. It was shown earlier that this cysteine is involved in disulfide formation with cysteine-247. UDP-galactose has been shown to bind in a region between Lys-341 and Lys-351 (Yadav and Brew, 1990) and α-lactalbumin binds at a different site (Lee, Wong, and Wong, 1983). This site has been deduced to be located in the C-terminal half of the molecule, somewhere between residues 258 and 402 (Boeggeman, Balaji, and Qasba, 1995).

The retention of the substrate binding cleft of lysozyme C in α-lactalbumin without its ability to bind the substrates of the former was accounted for by the observation made by Acharya et al. (1989) that Tyr-103 of α-lactalbumin, which replaces Ala–107 of lysozyme, blocks the cleft where the A and B sachharide binding sites are located. Further, Glu-35, crucial for the enzyme activity of lysozyme, is substituted by Thr-33. An examination of the amino acid residues of lysozyme involved in the binding of the hexameric substrate reveals a few

important aspects of the divergence of α-lactalbumin. Of the various amino acids implicated by Phillips (1974) and Weaver, Grutter, and Matthews (1995) listed in Table 5 that are highly conserved in lysozyme, only five are found in significant numbers in α-lactalbumin. They are Glu-49, Gln-54, Gln-56, Trp-60, and Asp-97. The rest have been replaced by amino acids with side chains that are quite different in properties. For instance, catalytic Glu-35 is found in only one α-lactalbumin and Asp-52 in none. Significantly, many of the substituting amino acids, His-32, Gly-35, Val-42, Glu-49, Gln-54, Gln-56 and Leu-110, for instance, are now well conserved in α-lactalbumin. These substitutions, the blocking action of Tyr-103, and the absence of Glu-35 and Asp-52, undoubtedly caused the abolition of the catalytic activity of lysozyme. Of the 14 amino acids (Table 5), 10 are encoded in Exon 2, two in Exon 3, and two in Exon 4 (Table 2).

Equal in importance to the loss of enzyme activity in α-lactalbumin is the gain of a site(s) for interaction with galactosyltransferase and the creation of a glucose binding site in the lactose synthase complex. Some of the amino acids residues of α-lactalbumin involved have been identified by chemical modification of individual residues, crosslinking of the two proteins, and more recently by the employment of mutants in which specific amino acids were replaced with others of dissimilar properties (Richardson and Brew, 1980; Sinha and Brew 1981; Grobler et al., 1994b; Malinovskii, et al., 1996; Pike, Brew, and Acharya, 1996). The amino acids demonstrated to be involved are listed in Table 8, along with an indication of their role in the formation of a functional lactose synthase complex. These are by no means exclusive, judging from the presence of other conserved or invariant acids in α-lactalbumin (Table 5).

The effect of mutation of specific amino acids of α-lactalbumin has been determined on (1) the inhibition by α-lactalbumin of galactosyltransferase activity using *N,N*-diacetyl chitobiose as galactose acceptor, which provides a measure of α-lactalbumin-galactosyltransferase interaction, and (2) lactose synthesis using glucose as ac-

TABLE 8
Amino Acid Residues of α-Lactalbumin (α-LA) Demonstrated, and Suspected, to be Involved in Its Interaction with Galactosyltransferase (GT) To Form Lactose Synthase

Residue	Effect of chemical modification or mutagenic substitution
Phe-31	Replacement reduces both GT and glucose binding
His-32	Replacement reduces both GT and glucose binding; modification results in loss of activity
Ala–106	Replacement results in low interaction with GT only
His-107	Affects binding to GT only
Leu-110	Replacement reduces GT interaction only
Gln-117	Reduces binding to GT only
Trp-118	Reduces binding to GT only; chemical modification reduces activity suspected to be involved
Glu-25	Invariant in α-LA
Gly-35	Invariant in α-LA
Lys-74	Invariant in α-LA
Ile-95	Invariant in α-LA

ceptor, a measure of the creation of the glucose binding site. Various amino acid residues, on replacement, show different effects on complex formation. Some of them, Val-42 and Gln-54, for instance, corresponding to those involved in the substrate binding in lysozyme, following substitution, show little difference in either (Malinovskii et al., 1996). One of the cysteine pairs in α-lactalbumin is the disulfide formed between Cys-6 and Cys-120, which brings the two ends of the peptide chain close together. Replacement of Cys-6 with serine and the elimination of Cys-120 along with the rest of the three C-terminal residues had little effect on lactose synthesis. Elimination of residues 119 to 123 yielded a still functionally active product. In contrast, truncation of residues 117 to 123 yielded an inactive product, confirming the essentiality of Gln-117 and Trp-118 (Malinovskii et al., 1994). Clearly, one of the disulfide bridges, the one involving Cys-121, and five of the C-terminal amino acids are not required for the folding of α-lactalbumin or for its activity in lactose synthase complex.

The amino acids of α-lactalbumin involved in its association with galactosyltranferase appears to be assembled in three regions of its molecule. One, called aromatic cluster I, comprises Phe-31, His-32, Gln-117, and Trp-118 (Grobler et al., 1994b; Pike, Brew, and Acharya, 1996). Modification or substitution of any of them directly affects the functional characteristic of lactose synthase (Table 8). Of the residues involved, Phe-31 and His-32 are encoded in Exon 1, while Gln-117 and Trp 118 are encoded in Exon 4. The arrangement of these in the molecule is shown in Plate 6.* It can also be seen in Plate 4.

Aromatic cluster II, previously referred to as hydrophobic box (Acharya et al., 1990) comprises Trp-26, Phe-53, Trp-60, Tyr-103, and Trp-104 (Grobler et al., 1994b). Of these, Tyr-103, as mentioned earlier, straddles the substrate binding cleft. A mutant that had proline in its place exhibited reduced galactosyltransferase binding without affecting its K_M for glucose. Curiously, when Tyr-103 was replaced with alanine, the residue conserved in lysozyme, the mutant failed to fold properly. Mutants involving other amino acids were either not expressed in sufficient quantities or they failed to fold (Grobler et al., 1994b). Each of the acids belonging to this group is highly conserved in α-lactalbumin as well as lysozyme. Therefore, they are unlikely to have a major role in the formation of a functional lactose synthase. Accordingly, these as well as other residues that are conserved in both proteins (Plate 2) can be expected to account for the remarkable similarity of their structures.

The third region of α-lactalbumin that contains amino acids that are involved in its association with galactosyltransferase consist of residues 105 to 110 (Pike, Brew, and Acharya, 1996). Ala–106, His-107, and Leu-110 have major roles in complex formation with transferase and variable roles in generating glucose binding site in the complex (Table 8) (Malinovskii et al., 1996). Significantly, it is located close to the amino acids constituting aromatic cluster I (Plate 6). Aromatic clusters I and II are also near each other, the former being adjacent to one side of the cleft near which A and B sachharide residues bind in lysozyme (Brew and Grobler, 1992). Naturally, the possibility has been considered that the glucose binding site is generated here (Brew and Grobler, 1992). No clue to support this has been obtained so far.

The 20 C-terminal acids of α-lactalbumin coded in Exon 4 differ significantly from the products of the other exons in their homology with lysozyme (Table 2). Only five are common or conserved in both, a much smaller proportion than the products

* Plate 6 appears after page 290.

of the other exons. However, these are well conserved within α-lactalbumin from the different sources (Table 4). Crystal structures of this region in α-lactalbumin and lysozyme are also different (Harata and Muraki, 1992). In α-lactalbumin it exhibits a greater conformational flexibility than it does in lysozyme and the rest of its own molecule. Therefore, it is likely to play a major role in its interaction with galactosyltransferase (Pike, Brew, and Acharya, 1996). The need for a less rigid molten state form for its action has been proposed before (Kronman, 1989).

Other amino acid residues also not conserved in lysozyme are found to be conserved in α-lactalbumin. These can also be expected to have some role in α-lactalbumin functions. For instance, Glu-25, Gly-35, Glu-82, and Glu-83 are invariable, while Leu-115, Glu-116 are conserved besides the residues in calcium-ion binding loop. What the roles of these might be remains to be determined.

VI. IMPLICATIONS AND A GLIMPSE INTO THE FUTURE

Information gathered so far on the evolution of extant lysozymes and α-lactalbumin, presented in a condensed form in this review, raises many important questions. Some of them undoubtedly would be resolved soon with the employment of the powerful tools of molecular biology and the refined techniques available for the determination of protein structure and their stabilities. For example, the role of Ca-ion in α-lactalbumin, also present in lysozyme of certain species, can be expected to become clearer when α-lactalbumin mutants lacking Ca-ion-binding aspartates and other amino acids are investigated with respect to folding and biochemical activity. Replacement of α-lactalbumin gene with appropriate mutant genes is also likely to provide clues to the role of this ion.

The recruitment of lysozyme gene for adaptation, perhaps as long as 200 million years after its duplication, to serve a role in lactose synthesis is astonishing. It is clear from the results of the studies included in this review that in these two, as in other related proteins, the three-dimensional structure of protein molecules are conserved more tenaciously than are the amino acid sequences. Also, a superimposability of the structures need not reflect similarities of biochemical roles. What is it about the structural features of lysozyme that made it suitable for adaptation to interact with galactosyltransferase in the manner it does? Does the carbohydrate binding feature, the cleft in the molecule, or the flexible C-terminal amino acids separately or together constitute the basis for its choice? Would it be possible, as it was in the case of primate digestive lysozyme, to determine the course of the selective substitution of specific amino acids during the divergence α-lactalbumin from lysozymes? What was the role of nonfunctional α-lactalbumin gene that accounts for its persistence before the origin of mammals? Does it play a role in modifying the activities of glycosyltransferases in nonmammalian species?

The answers to these and other important questions may be a long time in coming. However, surprises have abounded in the realm of the origin of proteins. Some of the answers may be forthcoming sooner than expected.

ACKNOWLEDGMENTS

We express our sincere gratitude to Mrs. Shanta Kumar and Mrs. Kathleen Bayne for their help with the manuscript and

Drs. P. V. Balaji and Lakshman Iyer for their help with some of the figures.

REFERENCES

Acharya, K. R., Ren, J., Stuart, D. I., Phillips, D. C., and Fenna, R. E. 1991. Crystal structure of human α-lactalbumin at 1.7 Å resolution. *J. Mol. Biol.* **22**: 571–581.

Acharya, K. R., Stuart, D. I., Phillips, D. C., McKenzie, H. A., and Teahan, C. G. 1994. Models of the three-dimensional structures of echidna, horse and pigeon lysozymes: calcium-binding lysozymes and their relationship with α-lactalbumins. *J. Prot. Chem.* **13**: 569–584.

Acharya, K. R., Stuart, D. I., Phillips, D. C., and Scheraga, H. 1990. A critical evaluation of the predicted and X-ray structures of α-lactalbumin. *J. Prot. Chem.* **9**: 549–563.

Acharya, K. R., Stuart, D. I., Walker, N. P. C., Lewis, M., and Phillips, D. C. 1989. Refined structure of baboon α-lactalbumin at 1.7 Å resolution. *J. Mol. Biol.* **208**: 99–127.

Arnheim, N. 1972. Multiple genes for lysozyme. In *Lysozymes*. pp. 153–161. Osserman, E. F., Canfield, R. E., and Beychok, S., Eds., Academic Press, New York.

Aschaffenburg, R., Fenna, R. E., Phillips, D. C., Smith, S. G., Buss, D. H., Jenness, R., and Thompson, M. P. 1979. Crystallography of α-lactalbumin. III. Crystals of baboon milk α-lactalbumin. *J. Mol. Biol.* **127**: 135–137.

Bajaj, M. and Blundell, T. 1984. Evolution and the tertiary structure of proteins. *Ann. Rev. Biophys. Bioeng.* **13**: 453–492.

Ben-Naim, A., Ting, K. L., and Jernigan, R. L. 1990. Solvent effects on binding thermodynamics of biopolymers. *Biopolymers* **29**: 901–919.

Blake, C. C. F., Koenig, D. F., Mair, G. A., North, A. C. T., Phillips, D. C., and Sarma, V. R. 1965. Structure of hen egg-white lysozyme. A three-dimensional Fourier synthesis at 2 Å resolution. *Nature (London).* **206**: 252–761.

Boeggeman, E. E., Balaji, P. V., Sethi, N., Masibay, A. A., and Qasba, P. K. 1993. Expression of deletion constructs of bovine 47β-1,4-galactosyltransferase in *E. coli:* importance of Cys 134 for its activity. *Protein Eng.* **6**: 779–785.

Boeggeman, E. E., Balaji, P. V., and Qasba, P. K. 1995. Functional domains of bovine β-1,4-galactosyltransferase. *Glycoconjugate J.* **12**: 865–878.

Brew, K. 1970. Lactose synthetase: Evolutionary origin, structure and control. In: *Essays in Biochemistry*. Vol. 6. pp. 93–118. Campbell, P. N. and Dickens, E., Ed., Academic Press, New York.

Brew, K. and Grobbler, J. A. 1992. α-Lactalbumins. In *Advanced Dairy Chemistry*. pp. 191–229. Vol. 1. Fox, P., ed. Elsevier Press, London.

Brew, K., Vanaman, T. C., and Hill, R. L. 1967. Comparison of the amino acid sequence of bovine α-lactalbumin and hen's egg white lysozyme. *J. Biol. Chem.* **242**: 3747–3749.

Brew, K., Vanaman, T. C., and Hill, R. L. 1968. The role of α-lactalbumin and the A protein in lactose synthetase: A unique mechanism for the control of a biological reaction. *Proc. Natl. Acad. Sci. U.S.A.* **59**: 491–497.

Brodbeck, U. and Ebner, K. E. 1966. Resolution of a soluble lactose synthetase into two protein components and solubilization of microsomal lactose synthetase. *J. Biol. Chem.* **241**: 762–764.

Brodbeck, U., Denton, W. L., Tanahashi, N., and Ebner, K. E. 1967. The isolation and identification of the B protein of lactose synthetase as α-lactalbumin. *J. Biol. Chem.* **242**: 1391–1397.

Browne, W. J., North, A. C. T., Phillips, D. C., Brew, K., Vanaman, T. C., and Hill, R. L. 1969. A possible three-dimensional structure of bovine α-lactalbumin based on hen's egg-white lysozyme. *J. Mol. Biol.* **42**: 65–89.

Canfield, R. E., Collins, J. C., and Sobel, J. H. 1972. Human leukemia lysozyme. In: *Lysozymes*. pp.

63–70. Osserman, E. F., Canfield, R. E., and Baychock, S., Eds., Academic Press, New York.

Canfield, R. E. and McMurry, S. 1967. Purification and characterization of a lysozyme from goose egg white. *Biochem. Biophys. Res Comm.* **26**: 38–42.

Cerini, C., Kerjan, P., Astier, M., Gratecos, D., Mirande, M., and Sémériva, M. 1991. A component of the multisynthetase complex is a multifunctional aminoacyl-tRNA synthetase. *Embo. J.* **10**: 4267–4277.

Chothia, C. and Lesk, A. M. 1987. The evolution of protein structures. *Cold Spring Harbor Symp. Quant. Biol.* **LII**: 399–405.

Chothia, C. 1992. One thousand families for the molecular biologist. *Nature* **357**: 543–544.

Creighton, T. E. 1993. Evolutionary and genetic origins of protein sequences. In: *Proteins, Structure and Molecular Properties*. pp. 105–138. W.H. Freeman and Co., New York.

Dandekar, A. M. and Qasba, P. K. 1981. Rat α-lactalbumin has a 17-residue-long COOH-terminal hydrophobic extension as judged by sequence analysis of the cDNA clones. *Proc. Natl. Acad. Sci. U.S.A.* **78**: 4853–4857.

Dandekar, A. M., Robinson, E. A., Appella, E., and Qasba, P. K. 1982. Complete sequence analysis of cDNA clones encoding rat whey phosphoprotein: homology to a protease inhibitor. *Proc. Natl. Acad. Sci. U.S.A.* **79**: 3987–3991.

Dautigny, A., Prager, E. M., Danièle, P.-D., Jollès, J., Pakdel, F., Grinde, B., and Jollès, P. 1991. cDNA and amino acid sequences of rainbow trout (*Onchorhyneus mykiss*) lysozymes and their implications for the evolution of lysozyme and lactalbumin. *J. Mol. Evol.* **32**: 187–198.

Dianoux, A. and Jollés, P. 1967. Étude d'un lysozyme pauvre en cystine at en tryptophan le lysozyme de blanc d'oef d'oie. *Biochim. Biophys. Acta.* **133**: 472–479.

Do, K.-Y., Do, S.-I., and Cummings, R. D. 1995. α-Lactalbumin induces bovine milk β1,4-galactosyltransferase to utilize UDP-GalNAc. *J. Biol. Chem.* **270**: 18447–18451.

Dobson, C. M., Evans, P. A., and Radford, S. E. 1994. Understanding how proteins fold: the lysozyme story so far. *TIBS* **19**: 31–37.

Dobson, D. E., Prager, E. M., and Wilson, A. C. 1984. Stomach lysozymes of ruminants. I. Distribution and catalytic properties. *J. Biol. Chem.* **259**: 11607–11616.

Doolittle, R. F. 1985. The geneaology of some recent evolved vertebrate proteins. *TIBS* **10**: 233- 237.

Doolittle, R., F. 1990. New perspectives on evolution provided by protein sequences. *New Perspectives on Evolution. The Wistar Symposium Series* **4**: 165–173.

Ewbank, J. J. and Creighton, T. E. 1993. Structural Characterization of the disulfide folding intermediates of bovine α-lactalbumin. *Biochemistry* **32**: 3694–3707.

Fett, R. and Knippers, R. 1991. The primary structure of human glutaminyl-tRNA synthetase. *J. Biol. Chem.* **266**: 1448–1455.

Fleming, A. 1922. On a remarkable bacteriolytic element found in tissues and secretions. *Proc. Roy. Soc. London* **B93**: 306–317.

Fleming, A. 1974. Personal recollections of lysozyme and Fleming. In: *Lysozyme*. pp. XIII-XVII. Osserman, E. F., Canfield, R. E. and Beychock, S., Eds., Academic Press, New York.

Fothergill-Gilmore, L. A. 1986. The evolution of the glycolytic pathway. *TIBS* **11**: 47–51.

Freedman, R. B., Hirst, T. R. and Tuite, M. F. 1994. Protein disulfide isomerase: building bridges in protein folding. *TIBS* **19**: 331–336.

Gillespie, J. H. 1991. *Protein Evolution: Examples of Microadaptations in the Causes of Molecular Evolution*. pp. 3–40. Oxford University Press, Oxford.

Godovac-Zimmerman, J., Conti, A., and Napolitano, L. 1988. The primary structure of donkey

(*Equus asinus*) lysozyme contains the Ca(II) binding site of α-lactalbumin. *Biol. Chem. Hoppe-Seylor.* **369**: 1109–1115.

Godovac-Zimmerman, J., Conti, A., and Napolitano, L. 1990. The complete primary structure of α-lactalbumin isolated from pig (*sus scrofa*) milk. *Biol. Chem. Hoppe-Seyler.* **371**: 649–653.

Goodman, M., Czelusniak, J., Koop, B. F., Tagle, D. A., and Slightom, J. L. 1987. Globins: a case study in molecular phylogeny, In: *Evolution of a catalytic function. Cold Spring Harbor Symp. Quant. Biol.* **LII**: 875–890.

Grobler, J. A., Rao, K. R., Pervaiz, S. and Brew, K. 1994a. Sequences of two highly divergent canine type C lysozymes: Implications for the evolutionary origins of the lysozyme/ α-lactalbumin superfamily. *Arch. Biochem. Biophys.* **313**: 360–366.

Grobler, J. A., Wang, M., Pike, C. W., and Brew, K. 1994b. Study by mutagenesis of the roles of two aromatic clusters of α-lactalbumin in aspects of its action in the lactose synthase system. *J. Biol. Chem.* **269**: 5106–5114.

Grütter, M. G., Weaver, L. H., and Matthews, B. W. 1983. Goose lysozyme structure: an evolutionary link between hen and bacteriophage lysozymes? *Nature (London).* **303**: 828–830.

Hadfield, A. T., Harvey, D. J., Archer, D. B., MacKenzie, D. A., Jeenes, D. J., Radford, S. E., Lowe, G., Dobson, C. M., and Johnson, L. N. 1994. Crystal structure of the mutant D52S hen egg white lysozyme with an oligosaccharide product. *J. Mol. Biol.* **243**: 856–872.

Haezebrouck, P., De Baetselier, A., Joniau, M., Van Dael, H., Rosenberg, S., and Hanssens, I. 1993. Stability effects associated with the introduction of a partial and a complete Ca^{2+}-binding site into human lysozyme. *Protein Eng.* **6**: 643–649.

Hakansson, A., Zhivotovsky, B., Orrhenius, S., Sabharwal, H., and Svanborg, C. 1995. Apoptosis induced by a human milk protein. *Proc Natl. Acad. Sci. U.S.A.* **92**: 8064–8068.

Hall, L. and Campbell, P. N. 1986. α-Lactalbumin and related proteins: a versatile gene family with an interesting parentage. In: *Essays in Biochemistry.* Vol. 22. pp. 1–26. Marshall, R. D. and Tipton, K. Eds. Academic Press, New York.

Hall, L., Emery, C., Davies, S., Parker, D., and Craig, R. K. 1987. Organization and sequence of the human α-lactalbumin gene. *Biochem. J.* **242**: 735–742.

Harata, K. and Muraki, M. 1992. X-ray structural evidence for a local helix. loop transition in α-lactalbumin. *J. Biol. Chem.* **267:** 1419–1421.

Hartl, D. L. 1991. New Perspectives in the molecular evolution of genes and genomes. In: *New Perspectives of Evolution.* pp. 123–137. Warren, L. and Koproski, Eds., Wiley-Liss, New York.

Hayssen, V. and Blackburn, D. G. 1985. α-Lactalbumin and the origin of lactation. *Evolution* **39:** 1147–1149.

Higaki, J. N., Gibson, B. W., and Craik, C. S. 1987. Evolution of catalysis in the serine proteases. In: *Evolution of catalytic function. Cold Spring Harbor Symposia on Quant. Biol.* **LII:** 615–621.

Hill, R. L. and Brew, K. 1975. Lactose synthetase. In: *Advances in Enzymology.* **43**: 411–490. Interscience Publication, John Wiley, New York.

Hiraoka, Y., Segawa, T., Kuwajima, K., Sugai, S., and Murai, N. 1980. α-Lactalbumin: a calcium metalloprotein. *Biochem. Biophys. Res. Comm.* **95**: 1098–1104.

Holm, L. and Sander, C. 1996. Mapping the protein universe. *Science* **273:** 595–602.

Horwitz, J. 1992. α-Crystallin can function as a molecular chaperone. *Proc. Natl. Acad. Sci. U.S.A.* **89:** 10449–10453.

Hyslop, N. E., Jr., Kern, K. C., and Walker, W. A. 1972. Lysozyme in human colostrum and breast milk. In: *Lysozymes* pp. 449–462, Osserman, E. F., Canfield, R. E., and Beychock, S., Eds., Academic Press, New York.

Irwin, D. M., Prager, E. M., and Wilson, A. C. 1992. Evolutionary genetics of ruminant-lysozymes. *Animal Genetics* **23**: 193–202.

Irwin, D. M., White, R. T., and Wilson, A. C. 1993. Characterization of the cow stomach lysozyme genes: repetitive DNA and concerted evolution. *J. Mol. Evol.* **37**: 355–356.

Irwin, D. M. and Wilson, A. C. 1989. Multiple cDNA sequences and the evolution of bovine stomach lysozyme. *J. Biol. Chem.* **264**: 11387–11393.

Irwin, D. M., and Wilson, A. C. 1990. Concerted evolution of ruminant stomach lysozymes. Characterization of lysozyme cDNA clones from sheep and deer. *J. Biol. Chem.* **265**: 4944-4952.

Ito, Y., Yamada, H. Nakamura, M., Yoshikawa, A., Ueda, T., and Imoto, T. 1993. The primary structures and properties of non-stomach lysozymes of sheep and cow, and implication for functional divergence of lysozyme. *Eur. J. Biochem.* **213**: 649–658.

Jollès, J., Jollès, P., Bowman, B. H., Prager, E. M., Stewart, C.-B., and Wilson, A. C. 1989. Episodic evolution in the stomach lysozymes of ruminants. *J. Mol. Evol.* **28**: 528–535.

Jollès, J., Périn, J. P., and Jollès, P. 1977. The ostrich (*Struthio camelus*) egg white lysozyme. *Mol. Cell. Biochem.* **17**: 39–44.

Jollès, J., Prager, E. M., Alnemri, E. S., Jollès, P., Ibrahimi, I. M., and Wilson, A. C. 1990. Amino-acid sequences of stomach and non-stomach lysozymes of ruminants. *J. Mol. Evol.* **30**: 370–382.

Jollès, J., Schoentgen, F., Crozier, G., Crozier, L., and Jollès, P. 1979. Insect lysozymes from three species of Lepidoptera: their structural relatedness to the C (chicken) type lysozyme. *J. Mol. Evol.* **14**: 267–271.

Jollès, P., Schoentgen, F., Jollès, J., Dobson, D. E., Prager, E. M., and Wilson, A. C. 1984. Stomach lysozymes of ruminants. II. Amino acid sequence of cow lysozyme 2 and immunological comparisons with other lysozymes. *J. Biol. Chem.* **259**: 11617–11625.

Jörnvall, H., Höög, J., von Bahr-Lindström, H., and Vallee, B. L. 1987. Mammalian alcohol dehydrogenases of separate classes: intermediates between different enzymes and interclass isoenzymes. *Proc. Natl. Acad. Sci. U.S.A.* **84**: 2580–2584.

Jung, A., Sippel, A. E., Grez, M., and Schulz, G. 1980. Exons encode functional and structural units of chicken lysozyme. *Proc. Natl. Acad. Sci. U.S.A.* **77**: 5759–5763.

Keese, P. K. and Gibbs, A. 1992. Origin of genes: "Bing bang" or continuous creation. *Proc. Natl. Acad. Sci. U.S.A.* **89**: 9489–9493.

Klee, W. A. and Klee, C. B. 1970. The role of α-lactalbumin in lactose synthetase. *Biochem. Biophys. Res. Commun.* **39**: 833–841.

Klee, W. A. and Klee, C. B. 1972. The interaction of α-lactalbumin and the A protein of lactose synthetase. *J. Biol. Chem.* **247**: 2336–2344.

Koivu, J., Myllylä, R., Halaakoski, T., Pihlajaniemi, T., Tasanen, K., and Kivivikko, K. I. 1987. A single polypeptide acts both as the β subunit of prolyl 4-hydroxylase and as a protein disulfide-isomerase. *J. Biol. Chem.* **262**: 6447–6449.

Kornegay, J. R., Schilling, J. W., and Wilson, A. C. 1994. Molecular adaptation of a leaf-eating bird: stomach lysozyme of the hoatzin. *Mol. Biol. Evol.* **11**: 921–928.

Kronman, M. J. 1989. Metal ion binding and the molecular conformational properties of α-lactalbumin. *Crit. Rev. Biochem.* **24**: 565–667.

Kronman, M. J., Sinha, S., and Brew, K. 1981. Characteristics of the binding of calcium ion and other divalent metal ions to bovine α-lactalburmin. *J. Biol. Chem.* **256**: 8582–8587.

Kumagai I., Takeda, S., and Miura, K.-I. 1992. Functional conversion of the homologous proteins α-lactalbumin and lysozyme by exon exchange. *Proc. Natl. Acad. Sci. U.S.A.* **89**: 5887- 5891.

Kumar, S., Clarke, A. R., Hooper, M. L., Horne, D. S., Law, A. J. R., Leaver, J., Springbett, A., Stevenson, E., and Simons, J. P 1994. Milk composition and lactation of casein-deficient

mice. *Proc. Natl. Acad. Sci. U.S.A.* **91**: 6138–6142.

Kuroki, R., Kawakita, S., Nakamura, H., and Yutani, K. 1992. Entropic stabilization of a mutant lysozyme induced by calcium binding. *Proc. Natl. Acad. Sci. U.S.A.* **89**: 6803–6807.

Kuroki, R., Taniyama, Y., Seko, C., Nakamura, H., Kikuchi, M., and Ikehara, M. (1989). Design and creation of a Ca^{2+} binding site in human lysozyme to enhance structural stability. *Proc. Natl. Acad, Sci. U.S.A.* **86**: 6903–6907.

Kuroki, R., Weaver, L. H., and Matthews, B. W. 1993. A covalent enzyme-substrate intermediate with saccharide distortion in a mutant T4 lysozyme. *Science* **262**: 2030–2033.

Kuwajima, K., Mitani, M., and Sugai, S. 1989. Characterization of the critical state in protein folding. Effects of guanidine hydrochloride and specific Ca^{2+} binding on the folding kinetics of α-lactalbumin. *J. Mol. Biol.* **206**: 547–561.

Kuziora, M. A., Chalmers, Jr., J. H., Douglas, M. G., Hitzeman, R. A., Mattick, J. S., and Wakil, S. J. 1983. Molecular cloning of fatty acid synthetase genes from *Sachharomyces cerevisiae*. *J. Biol. Chem.* **258**: 11648–11653.

Laird, J. E., Jack, L., Hall, L., Boulton, P., Parker, D., and Craig, R. K. 1988. Structure and expression of the guinea-pig α-lactalbumin gene. *Biochem. J.* **254**: 85–94.

Lebioda, L. and Stec, B. 1988. Crystal structure of enolase indicates that enolase and pyruvate kinase evolved from a common ancestor. *Nature (London).* **333**: 683–686.

Lee, T. K., Wong, L.-J. C., and Wong, S. S. 1983. Photoaffinity labeling of lactose synthase with a UDP-galactose analog. *J. Biol. Chem.* **258**: 13166–13177.

Lonberg, N. and Gilbert, W. 1985. Intron-exon structure of the chicken pyruvate kinase gene. *Cell.* **40**: 81–90.

Maenaka, K., Matsushima, M., Song, H., Sunada, F., Watanabe, K., and Kumagai, I. 1995. Dissection of protein-carbohydrate interactions in mutant hen egg white lysozyme complexes and their hydrolytic activity. *J. Mol. Biol.* **247**: 281–293.

Malcolm, B. A., Rosenberg, S., Corey, M. J., Allen, J. S., de Baetselier, A., and Kirsch, J. F. 1989. Site-directed mutagenesis of the catalytic residues Asp-52 and Glu-35 of chicken egg white lysozyme. *Proc. Natl. Acad. Sci. U.S.A.* **86**: 133–137.

Malinovskii, V., Tian, J., Grobbler, J. A., and Brew, K. 1996. Functional site in α-lactalbumin encompasses a region corresponding to a subsite in lysozyme and parts of two adjacent flexible substructures. *Biochemistry* **35**: 9710–9715.

Masibay, A. S., Balaji, P. V., Boeggeman, E. E., and Qasba, R. K. 1993. Mutational analysis of the Golgi retention signal of bovine 3-1, 4-galactosyl transferase. *J. Biol. Chem.* **268**: 9908–9916.

Mason, D. Y., Jones, M., and Goodnow, C. C. 1992. Development of follicular localization of tolerant B lymphocytes in lysozyme/antilysozyme Ig M/Ig D transgenic mice. *Int. Immunol.* **4**: 163–175.

Matthews, B. W., Grütter, M. G., Anderson, W. F., and Remington, S. J. 1981. Common precursor of lysozymes of hen egg white and bacteriophage T4. *Nature (London).* **290**: 334–335.

McCarthy, A. D., and Hardie, D. G. 1984. Fatty acid synthetase — an example of protein evolution by gene fusion. *TIBS* **9**: 60–63.

McKenzie, H. A. 1996. α-Lactalbumins and lysozymes. In: *Lysozymes: Model Enzymes in Biochemistry and Biology*. pp. 365–409. Jollés, P., Ed., Birkhauser Verlag, Basel, Switzerland.

McKenzie, H. A. and White, F. H., Jr. 1991. Lysozyme and α-lactalbumin: structure, function, and interrelationships. In: *Advances in Protein Chemistry*. pp. 173–315. Anfinsen, C. B., Edsall, J. T., Richards, F. M., and Edenberg, D. S., Eds., Academic Press, New York.

McLachlan, A. D. 1987. Gene duplication and the origin of the repetitive protein structures. *Cold*

Spring Harbor Symp. Quant Biol. **LII**: 411–420.

Means, A. R., Putkey, J. A., and Epstein, P. 1988. Organization and evolution of genes from calmodulin and other calcium binding proteins. In: *Calmodulin: Molecular Aspects of Cellular Regulation.* pp. 17–33. Cohn, P. and Klee, C. B., Eds., Elsevier, Amsterdam.

Meera, G., Ramesh, N., and Brahmachari, S. K. 1989. Zintrons in rat α-lactalbumin gene. *FEBS. Lett.* **251**: 245–249.

Messer, M. and Nicholas, K. R. 1991. Biosynthesis of marsupial oligosaccharides: Characterization and developmental changes of two galactosyltransferases in lactating mammary glands of tammar wallaby, *Macropus eugenii*. *Biochim. Biophys. Acta* **1077**: 79–85.

Messier, W. and Stewart, C.-B. 1997. Episodic adaptive evolution of primate lysozymes. *Nature* **385**: 151–154.

Morgan, F. J. and Arnheim, N. 1972. Lysozyme of black swan, *Cygnus atratus*. In: *Lysozymes*, pp. 81–93. Osserman, E. F., Canfield, R. E., and Beychock, S., Eds., Academic Press, New York.

Musci, G. and Berliner, L. (1985). Physiological roles of zinc and calcium binding to α-lactalbumin in lactose biosynthesis. *Biochemistry* **24**: 6945–6948.

Nagel, G. M. and Doolittle, R. F. 1991. Evolution and relatedness in two aminoacyl-tRNA synthetase families. *Proc. Natl. Acad. Sci. U.S.A.* **88**: 8121–8125.

Neeleman, A. P. and van den Eijnden, D. H. 1996. α-Lactalbumin affects the acceptor specificity of *Lymnaea stagnalis* albumen gland UDP-GalNAc:GlcNAcβ-R β1 → *4*-N-acetylgalactosaminyltransferase: synthesis of GalNAcβ1 Æ 4Glc. *Proc. Natl. Acad. Sci. U.S.A.* **93**: 10111–10116.

Neidhart, D. J., Kenyon, G. L., Gerlt, J. A., and Petsko, G. A. 1990. Mandelate racemase and muconate lactonizing enzyme are mechanistically distinct and structurally homologous. *Nature* **347**: 692–694.

Newburg, D. S. and Neubauer, S. H. 1995. Carbohydrates in milks. In: *Handbook of Milk Composition.* pp. 273–349. Jensen, R. G., Ed., Academic Press, New York.

Nitta, K. and Sugai, S. 1989. The evolution of lysozyme and α-lactalbumin. *Eur. J. Biochem.* **182**: 111–118.

Nitta, K., Tsuge, H., Shimazaki, K., and Sugai, S. 1988. Calcium-binding lysozymes. *Biol. Chem. Hoppe-Seyler* **369**: 671–675.

Oftedal, O. T. and Iverson, S. J. 1995. Phylogenetic variation in the gross composition of milks. In: *Handbook of Milk Composition.* pp. 749–789. Jensen, R. G., Ed., Academic Press, New York.

Ohno, S. 1987. Early genes that were oligomeric repeats generated a number of divergent domains on their own. *Proc. Natl. Acad. Sci. U.S.A.* **84**: 6486–6490.

Osserman, E. F. 1972. Biological and Clinical studies. Introduction. In: *Lysozyme.* pp. 303–306. Osserman, E. F., Canfield, R. E., and Beychock, S., Eds., Academic Press, New York.

Overington, J., Johnson, M. S., Sali, A., and Blundell, T. L. 1990. Tertiary structural constraints on protein evolutionary diversity templates. Key residues and structure prediction. *Proc. R. Soc. Lond.* **B241**: 132–147.

Pancholi, V. and Fischetti, V. A. 1993. Glyceraldehyde-3-phosphate dehydrogenase on the surface group A streptococci is also on ADP-ribosylating enzyme. *Proc. Natl. Acad. Sci. U.S.A.* **90**: 8154–8158.

Pardon, E., Haezebrouck, P., De-Baetselier, A., Hooke, S. D., Fancourt, K. T., Desmet, J., Dobson, C. M., Van Dael, H., and Joniau, M. 1995. A $Ca^{(2+)}$-binding chimera of human lysozyme and bovine α-lactalbumin that can form a molten globule. *J. Biol. Chem.* **270**: 10514–10524.

Patel, M. S. and Roche, T. E. 1990. Molecular biology and biochemistry of pyruvate dehydrogenase complexes. *FASEB J.* **4**: 3224–3233.

Paulson, J. C. and Colley, K. J. 1989. Glycosyltransferases: structure, localization, and control of cell type-specific glycosylation. *J. Biol. Chem.* **264:** 17615–17618.

Pepys, M. B., Hawkins, P. N., Both, D. R., Vigushin, D. M., Tennent, G. A., Soutar, A. K., Totty, M., Nguyen, D., Blake, C. C. F., Terry, C. J., Feest, T. G., Zalin, A. M., and Hsuan, J. J. 1993. Human lysozyme gene mutations cause hereditary systemic amyloidosis. *Nature (London)* **362**: 553–557.

Permyakov, E. A. and Berliner, L. J. 1994. Co^{2+} binding to α-lactalbumin. *J. Prot. Chem.* **13**: 277–281.

Permyakov, E. A., Shnyrov, V. L., Kalinichenko, L. P., Kuchar, A., Reyzer, I. L., and Berliner, L. J. 1991. Binding of Zn(II) ions to α-lactalburmin. *J. Protein Chem.* **10**: 577–584.

Perutz, M. F. 1984. Species adaptation in a protein molecule. In *Advances in Protein Chemistry*. Vol. 36. pp. 213–244. Anfinsen, C. B., Edsall, J. T., and Richards, F. M., Eds., Academic Press, New York.

Peters, C. W., Kruse, U., Pollwein, R., Grzeschik, K.-H. and Sippel, A. E. 1989. The human lysozyme gene. Sequence organization and chromosomal localization. *Eur. J. Biochem.* **182**: 507–516.

Phillips, D. C. 1974. Crystallographic studies of lysozyme and its interactions with inhibitors and substrates. In: *Lysozyme*. pp. 9–30. Osserman, E. F., Canfield, R. E., and Beychock, S., Eds., Academic Press, New York.

Piatigorsky, J. and Wistow, G. 1991. The recruitment of crystallins: new functions precede gene duplication. *Science* **252**: 1078–1079.

Pike, A. C. W., Brew, K., and Acharya, K. R. 1996. Crystal structures of guinea-pig goat and bovine α-lactalbumin highlight the enhanced conformational flexibility of regions that are significant for its action in lactose synthase. *Structure* **4:** 691–703.

Post, C. B. and Karplus, M. 1986. Does lysozyme follow the lysozyme pathway? An alternative based on dynamic, structural and stereoelectric considerations. *J. Am. Chem. Soc.* **108**: 1317-1319.

Prager, E. M. and Wilson, A. C. 1988. Ancient origin of α-lactalbumin from lysozyme: analysis of DNA and amino acid sequences. *J. Mol. Evol.* **27**: 326–335.

Prestrelski, S. J., Byler, D. M., and Thompson, M. P. 1991. Effect of metal ion binding on the secondary structure of bovine α-lactalbumin as examined by infrared spectroscopy. *Biochemistry* **30**: 8797–8804.

Qasba, P. K. and Safaya, S. K. 1984. Similarities of the nucleotide sequences of rat α-lactalbumin and chicken lysozyme genes. *Nature (London)*. **308**: 377–380.

Ramasarma, T. 1994. One protein — many functions. *Curr. Sci. India.* **67**: 24–29.

Rao, K. R. and Brew, K. 1989. Calcium regulates folding and disulfide-bond formation in α-lactalbumin. *Biochem. Biophys. Res. Comm.* **163**: 1390–1396.

Ren, J., Stuart, D. I., and Acharya, K. R. 1993. α-Lactalbumin possesses a distinct zinc binding site. *J. Biol. Chem.* **268**: 19292–19298.

Richardson, R. H. and Brew, K. 1980. lactose synthase. An investigation of the interaction site of α-lactalbumin for galactosyltransferase by differential kinetic labeling. *J. Biol. Chem.* **255:** 3377–3385.

Richardson, J. S. and Richardson, D. C. 1989. The *de novo* design of protein structures. *TIBS* **14**: 304–309.

Sax, C. M. and Piatigorsky, J. 1994. Expression of α-crystallin/small heat shock protein/molecular chaperone genes in the lens and other tissues. *Advances in Enzymology*. Vol. 69. pp. 155–201. Meister, A., Ed., John Wiley & Sons, New York.

Schoentgen, F., Jollès, J., and Jollès, P. 1982. Complete aminoacid sequence of ostrich *(Struthis camelus)* egg white lysozyme, a goose type lysozyme. *Eur. J. Biochem.* **123**: 489–497.

Schumacher, G. F. B. 1972. Lysozyme in human genitalia. In: *Lysozymes*. pp. 427–447.

Osserman, E. F., Canfield, R. E., and Beychock, S., Eds. Academic Press, New York.

Shaw, D. C., Messer, M., Scrivener, A. M., Nicholas, K. R., and Griffiths. 1993. Isolation, partial characterization, and aminoacid sequence of α-lactalbumin from platypus (*Ornithothyneus anatinus*) milk. *Biochem. Biophys. Acta* **1161**: 177–186.

Shewale, J. G., Sinha, S. K., and Brew, K. 1984. Evolution of α-lactalbumins. The complete amino acid sequence of the α-lactalbumin from a marsupial (*Macropus rufogriseus*) and corrections to regions of sequences in bovine and goat α-lactalbumins. *J. Biol. Chem.* **259**: 4947–4956.

Shih, P. and Kirsch, J. F. 1995. Design and structural analysis of an engineered thermostable chicken lysozyme. *Prot. Sci.* **4**: 2063–2972.

Simpson, R. J., Begg, G. S., Dorow, D. S. and Morgan, F. J. 1980. Complete amino acid sequence of goose type lysozyme from egg white of black swan. *Biochemistry* **19**: 1814–1819.

Simpson, R. J. and Morgan, F. J. 1983. Complete amino acid sequence of Embden goose (*Anser anser*) egg white lysozyme. *Biochim. Biophys. Acta* **744**: 349–351.

Singh, R. and Green, M. R. 1993. Sequence-specific binding of transfer RNA by glyceraldehyde-3-phosphate dehydrogenase. *Science* **259**: 365–368.

Sinha, S. and Brew, K. 1981. A label selection procedure for determining the location of protein-protein interaction sites by cross-linking with bisimidoesters. Application to lactose synthase. *J. Biol. Chem.* **256:** 493–4204.

Sinnott, M. J. 1990. Catalytic mechanisms of enzymic glycosyl transfer. *Chem. Rev.* **90**: 1171-1202.

Smith, S. G., Lewis, M., Aschffenburg, R., Fenna, R. E., Wilson, I. A., Sundaralingam, M., Stuart, D. I., and Phillips, D. C. 1987. Crystallographic analysis of the three-dimensional structure of baboon α-lactalburmin at low resolution. Homology with lysozyme. *Biochem. J.* **242**: 353–360.

Song, H., Inaka, K., Maenaka, K., and Matsushima, M. 1994. Structural changes of active site cleft and different saccharide binding modes in human lysozyme co-crystallized with hexa-*N*-acetyl chitohexose at pH 4.0. *J. Mol. Biol.* **244**: 522–540.

Soulier, S., Mercier, J. C., Vilotte, J. L., Anderson, J., Clark, A. J., and Provot, C. 1989. The bovine and ovine genomes contain multiple sequences homologous to the α-lactalbumin-encoding genes. *Gene* **83**: 331–338.

Stacey, A., Schnieke, A., Kerr, M., Scott, A., McKee, C., Cottingham, I., Binus, B., Wilde, C., and Colman, A. 1995. Lactation is disrupted by α-lactalbumin deficiency and can be restored by human α-lactalbumin gene replacement. *Proc. Natl. Acad. Sci. U.S.A.* **92**: 2835–2839

Steinhoff, U. M., Senft, B., and Seyfert, H.-M. 1994. Lysozyme-encoding bovine cDNAs from neutrophil granulocytes and mammary gland are derived from a different gene than stomach lysozymes. *Gene* **143**: 271–276.

Stewart, C.-B., Schilling, J. W., and Wilson, A. C. 1987. Adaptive evolution in the stomach lysozymes of foregut fermenters. *Nature (London)* **330**: 401–404.

Stewart, C.-B. and Wilson, A. C. 1987. Sequence convergence and functional adaptation of stomach lysozymes from foregut fermenters. *Cold Spring Harbor Symp. Quant. Biol.* **LII**: 891–899.

Stinnakre, M. G., Villotte, J. L., Soulier, S., and Mercier, J. C. 1994. Creation and phenotypic analysis of α-lactalbumin-deficient mice. *Proc. Natl. Acad. Sci. U.S.A.* **91**: 6544–6548.

Strynadka, N. C. J. and James, M. N. G. 1991. Lysozyme revisited: crystallographic evidence for distortion of an N-acetylmuramic acid residue bound in site D. *J. Mol. Biol.* 1991 **220**: 401–424.

Stuart, D. I., Acharya, K. R., Walker, N. P. C., Smith, S. G., Lewis, M., and Phillips, D. C.

1986. α- Lactalbumin possesses a novel calcium binding loop. *Nature* **324**: 84–87.

Stumpf, P. 1984. Fatty acid biosynthesis in higher plants. In: *Fatty Acid Metabolism and Its Regulation* , pp. 155–179. Numa, S., Ed., Elsevier Science Amsterdam.

Südhoff, T. C., Goldstein, J. L., Brown, M. S., and Russel, D. W. 1985. The LDL receptor gene. A mosaic of exons shared with different proteins. *Science* **228**: 815–822.

Sugai, S. and Ikeguchi, M. 1994. Conformational comparison between α-lactalbumin and lysozyme. *Adv. Biophys.* **30**: 37–84.

Swanson, K. W., Irwin, D. M., and Wilson, A. C. 1991. Stomach lysozyme gene of langur monkey: Tests for convergence and positive selection. *J. Mol. Evol.* **33**: 418–425.

Takeuchi, K., Irwin, D. M., Gallup, M., Shinbrott, E., Kai, H., Stewart, C.-B. and Basbaum, C. 1993. *J. Biol. Chem.* **268**: 27440–27446.

Taylor, D. C., Cripps, A. W., and Clancy, R. L. 1995. A possible role for lysozyme in determining acute exacerbation in chronic bronchitis. *Clin. Exp. Immunol.* **102**: 406–416.

Teahan, C. G., McKenzie, H. A., Shaw, D. C., and Griffiths, M. 1991. The isolation and amino acid sequences of echidna (*Tachyglossus aculeatus*) milk lysozyme I and II. *Biochem. Int.* **24**: 85–95.

Tsuge, H., Ago, H., Noma, M., Nitta, K., Sugai, S., and Miyano, M. 1992. Crystallographic studies of a calcium binding lysozyme from equine milk at 2.5 Å resolution. *J. Biochem.* **111**: 141–143.

Tsugita, A. 1970. Phage lysozyme and other lytic enzymes. In *The Enzymes*, Vol. **5**, pp. 344-411. Boyer, P., Ed., Academic Press, New York.

Vilotte, J.-L., and Soulier, S. 1992. Isolation and characterization of the mouse α-lactalbumin-encoding gene: interspecies comparison, tissue- and stage-specific expression. *Gene* **119**: 287–292.

Vilotte, J. L., Soulier, S., and Mercier, J.-C. 1992. Sequence of the murine α-lactalbumin-encoding cDNA: interspecies comparison of the coding frame and deduced preprotein. *Gene* **112**: 251–255.

Vilotte, J.-L., Soulier, S., Mercier, J.-C., Gaye, P., Hue-Delahaie, D., and Furet, J.-P. 1987. Complete nucleotide sequence of bovine α-lactalbumin gene: comparison with rat counterpart. *Biochimie (Paris)* **69**: 609–620.

Vilotte, J. L., Soulier, S., Printz, C., and Mercier, J. C. 1991. Sequences of the goat α-lactalbumin-encoding gene: comparison with the bovine gene and evidence of related sequences in the goat genome. *Gene* **98**: 271–276.

Volbeda, A., Lahm, A., Sakiyama, F., and Suck, D. 1991. Crystal structure of *Penicilluim citrinum* P1 nuclease at 2.8 Å resolution. *EMBO J.* **10**: 1607–1618.

Wakil, S. J., 1989. Fatty acid synthase, a proficient multifunctional enzyme. *Biochemistry* **28**: 4523–4530.

Wang, M., Scott, W. A., Rao, K. R., Udey, J. Conney, G. E., and Brew, K. 1989. Recombinant bovine α-lactalbumin obtained by limited proteolysis of a fusion protein expressed at high levels in *Escherichia coli*. *J. Biol. Chem.* **264:** 21116–21121.

Warme, P. K., Momany, F. A., Rumball, S. V., Tuttle, R. W., and Scheraga, H. A. 1974. Computation of structures of homologous proteins. α-Lactalbumin from lysozyme. *Biochemistry* **13**: 768–782.

Weaver, L. H., Grütter, M. G., and Matthews, B. W. 1995. The refined structures of goose lysozyme and its complex with a bound trisaccharide show that the "goose type" lysozymes lack a catalytic aspartate. *J. Mol. Biol.* **245**: 54–68.

White, F. H., Jr., Balkwill, D. L., Meeter, D. A., and Merchant, K. K. 1993. Further studies on the lysozyme-like activity of α-lactalbumin: development of alternative methods of assay. *Anal. Biochem.* **212**: 263–268.

Wilson, A. C., Carlson, S. S., and White, T. J. 1977. Biochemical evolution. *Annu. Rev. Biochem.* **46**: 573–639.

Wilson, K. P., Malcolm, B. A., and Matthews, B. W. 1992. Structural and thermodynamic analysis of compensating mutations within the core of chicken egg white lysozyme. *J. Biol. Chem.* **267**: 10842–10849.

Wistow, G. 1993. Lens crystallins: gene recruitment and evolutionary dynamism. *TIBS* **18**: 301–306.

Wylie, D. C. and Vanaman, T. C. 1988. Structure and evolution of the calmodulin family of calcium regulating proteins. In: *Calmodulin*. pp. 1–15. Cohen, P. and Klee, C. B., Eds., Elsevier Science Publishers, Amsterdam.

Yadav, S. and Brew, K. 1990. Identification of UDP-galactose: *N*-acetylglucosamine 4-galactosyltranferase involved in UDP-galactose binding by differential labeling. *J. Biol. Chem.* **265:** 14163–14169.

Yadav, S. P. and Brew, K. 1991. Structure and function in galactosyltransferase. Sequence locations of α-lactalbumin binding site, thio groups and disulfide bond. *J. Biol. Chem.* **266:** 698–703.

Yang, D. C. H., Garcia, J. C., Johnson, Y. D., and Wahab, S. 1985. Multienzyme complexes of mammalian aminoacyl-tRNA synthetases. *Current Topics in Cellular Regulation* **26**: 325–335.

Zhang, Y., Liang, J.-Y., and Lipscomb, W. N. 1993. Structural similarities between fructose 1,6-biphosphatase and inositol monophosphatase. *Biochem. Biophys. Res. Comm.* **190**: 1080–1083.