CHAPTER 2

GENE RECRUITMENT: A NOVEL MECHANISM IN MOLECULAR EVOLUTION

MULTIFUNCTIONAL PROTEINS AND "GENE SHARING"

The greatest surprise in the study of crystallins has been that these long-lived structural proteins with apparently unique functional requirements in a highly specialized tissue may also be expressed in other tissues under quite different guises. Most strikingly the taxon-specific crystallins, those which are restricted to particular evolutionary lineages, have been found to be identical to enzymes.

In the recruitment of taxon-specific crystallins, genes encoding enzymes which have maintained their metabolic functions for hundreds of millions of years acquire greatly increased expression tissue—specifically in the lens. As a result, the protein product of the recruited gene becomes a major structural component of the lens. This protein now has two completely distinct functions; as a biological catalyst in many tissues, continuing its role as before, and as a bulk component of the refractive structure of the lens. There are many examples in which evolutionarily related proteins fulfill different functions. For example the β-chain of haptoglobin, an enzymatically inactive serum protein involved in hemoglobin binding, is structurally related to the serine protease superfamily which includes trypsin,[1] while the diverse members of the immunoglobulin superfamily have many specialized roles, as antibodies, cell surface receptors and cell adhesion molecules among other function.[2] The enzyme crystallins are different. They are not merely derived from enzymes but continue to serve that ancestral function while simultaneously serving as crystallins. The result is that one protein has two separate and distinct roles in the absence of gene

duplication. This protein multifunctionality is the key characteristic of the gene recruitment of crystallins. The term "gene sharing" has also been used to describe this phenomenon to emphasize that a single gene gives rise to two protein functions.[3]

Gene Recruitment of Crystallins: Modifying the Lens

There have actually been several distinct phases of crystallin gene recruitment in vertebrate lenses. The first occurred at the very earliest stages in the evolutionary development of the tissue. The lens first became a lens when it acquired a higher refractive index than the surrounding media. This was achieved by increasing the concentration of soluble proteins in the cytoplasm. As described in subsequent chapters this seems to have occurred through the recruitment of stress-related proteins which may have already had important roles in lens cells, probably involving interactions with cytoskeleton or other vulnerable and essential systems.[4-6] These initial recruitments involved increasing the expression of representatives of two protein superfamilies, the small heat shock protein/α-crystallin superfamily[4,7] and the βγ-crystallin superfamily.[4,8] Over time the recruited genes underwent multiplication and diversification to produce the proteins necessary for the high-refractive index lenses required for vision under water.

Since the ancestors of vertebrates evolved under water this environment had a profound effect on the emergence of vertebrate crystallins. With the cornea in direct contact with water, almost all the refractive power of a fish eye must come from the lens which consequently requires a high refractive index.[9] This is well illustrated by the extremely hard and dehydrated lenses in modern aquatic species which can have remarkably high concentrations of proteins (mainly crystallins) up to as much as 70% wet weight.[10,11]

Much later (about 350 million years ago) some vertebrates emerged from water into air. In this new environment a completely different kind of lens would have been needed. Rather than a high refractive index, rigid lens it would have been a great advantage to early terrestrial vertebrates to have a less myopic and more easily deformable lens capable of visualizing objects such as food and predators over a wide range of distances. This was achieved by modifying the protein content of a lens which had evolved under water. A major contributor to this process was a new episode of gene recruitment.

The proteins which make the largest contribution to the specialization of hard lenses are γ-crystallins.[10] As described in the following chapter, their structure, their biophysical properties and their phylogenetic and ontogenic distributions suggest that γ-crystallins, are needed to create or maintain a low-water content, high protein density cellular environment.[12-14] The most direct approach to softening a lens evolved under water would be to reduce the contribution made by the γ-crystallins and one way of doing this would be to recruit novel, additional crystallins with more normal hydration properties.[6]

Evolution has performed the experiment of softening the lens through gene recruitment of novel crystallins many times and the results of these experiments show us that the proteins best suited for this role are all globular metabolic enzymes.[6] Although at first sight this strict selectivity might seem surprising, on further reflection it becomes clear that many other classes of protein would be disqualified from recruitment as crystallins. Crystallins need to be highly stable structures capable of surviving without turnover for many years. They must retain solubility in a relatively high protein concentration environment while avoiding aggregation with other lens components to form light scattering centers. High concentrations of a particular protein should have no deleterious effects on cell metabolism, signaling pathways, transcription regulation, cytoskeletal structure or any of the other essential systems of the cell. For ease of recruitment, the genes for these new crystallins should already be expressed in the lens, or at least easy to induce.

Most of the proteins which satisfy these criteria turn out to be enzymes. One feature of enzymes which might be expected to pose problems following their recruitment as crystallins is their metabolic activity. High concentrations of some enzymes could disrupt the flow of metabolic pathways, sequestering substrate and co-factor molecules. Indeed, this may disqualify certain enzymes from suitability as crystallins. Those that succeed are generally involved in non-rate limiting reactions. In the case of at least one enzyme which acts as a crystallin in some species, α-enolase/τ-crystallin, overexpression in lenses and other tissues of transgenic mice has no evident deleterious effect.[15] Indeed it is possible that lens cells have a "metabolic compartment" separate from the bulk of the crystallins so that catalytic pathways are insulated from the overexpression of enzymes.

The other possible drawback of using an enzyme as a crystallin is its ability to sequester substrate and cofactor small molecules in the lens. However, this ability may actually be turned to advantage in certain circumstances.

Secondary Advantages: Protecting the Lens

The primary evolutionary pressure on terrestrial vertebrate lenses may have been to modify the optical properties of the lens. However secondary effects may also have been very important and these may even have been the dominant forces in several more recent phases of enzyme crystallin recruitment. These secondary advantages could take the form either of additional beneficial properties of a crystallin recruited primarily for refractive reasons or of the secondary recruitment of an additional crystallin for other purposes. These additional benefits probably include protective or stress related functions such as the filtering of harmful ultraviolet radiation, participating in maintenance of the osmotic balance of lens fibers or contributing to antioxidant mechanisms in the lens. Such properties have been proposed for ε-crystallin[16] and for other nicotinamide adenine dinucleotide cofactor

binding enzyme crystallins[17] which can sequester reduced cofactor in the lens. Since ε-crystallin was recruited into lenses which already had δ-crystallins[16] it is likely that properties of the enzyme other than its contribution to lens softening were important for its recruitment and maintenance in the lens. Reduced cofactors may act directly as UV filters or in redox reactions to protect oxidation of lens components.[16-18] The ability of recruited enzyme crystallins to sequester these cofactors in the lens may have conferred important benefits for some species such as those moving from an environment of dim light to one of full exposure to the sun.

Secondary advantages distinct from contributions to refractive index were probably also important for the original recruitments of α- and βγ-crystallins. These stress-related proteins may not only have helped to make the lens a lens but also may have enhanced the stability of other more vulnerable lens components, such as the cytoskeleton, the essential infrastructure of the elongated fiber cell, and transmembrane channel proteins. Indeed, the enzymes recruited as crystallins may even share with the stress protein crystallins some aspects of this role in associating with and possibly stabilizing cytoskeleton. Glycolytic enzymes in particular have been found to associate with various components of cytoskeleton, such as actin, and to form part of a "cytomatrix".[19-22]

WIDER IMPLICATIONS FOR MOLECULAR EVOLUTION

The gene recruitment of enzyme crystallins is a novel mechanism in molecular evolution but it is probably not unique to the lens. The peculiarities of structure, stability and evolutionary plasticity in the lens made the discovery of this form of gene recruitment possible, but the lessons learned are likely to have much wider applicability for understanding the origins of the duplicated multigene families and superfamilies which are such a dominant feature of the genomes of all organisms.[6,16,23]

A classical model in protein evolution proposes that in order for new protein functions to arise an existing gene must first undergo duplication[24,25] (Fig. 2.1). This is a purely random event. Following duplication, selection for the original function maintains one copy of the gene but the other gene is freed of selective pressures and consequently begins to experience sequence drift. During this period expression of the unselected gene is likely to cease since it has no selectable benefits for the cell and may indeed have deleterious effects. By chance this drifting pseudogene eventually acquires a useful new sequence. Also by chance it reacquires the ability to be expressed in a way which makes use of its new function. The problem faced by this classical model is that non-functional or non-essential genes are at high risk of elimination from the genome through deletion, insertion, rearrangement and loss of CG dinucleotides through the process known as "ripping."[26] The window of opportunity for the random acquisition

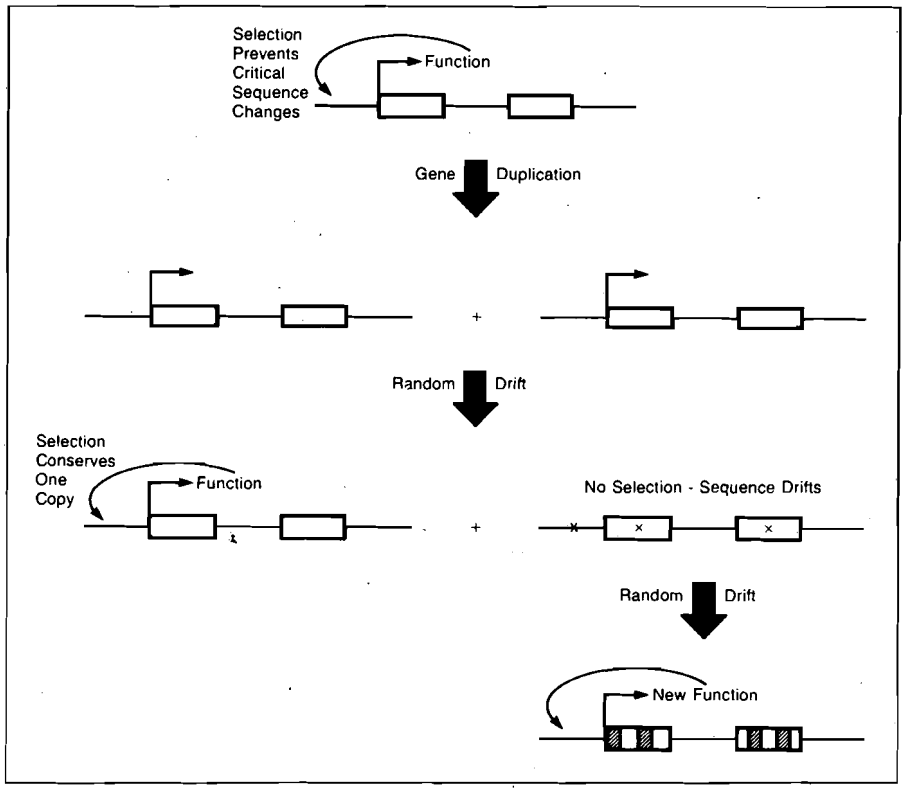

Fig. 2.1. The "classical" model for new protein functionalities and for gene duplication.

of a new role may be very short, perhaps too short for this process to be successful.

The gene recruitment of crystallins illustrates a powerful alternative strategy through protein multifunctionality or gene sharing (Fig. 2.2). In this scheme a new function arises in the product of a single gene. In the case of the crystallins an enzyme or stress protein acquires an additional structural role in the lens. The protein becomes subject to two sets of selective pressures but if neither of the functions impinges on the fitness of the other, protein multifunctionality of a single gene product may continue indefinitely. However, there may also be cases where adaptive conflicts occur. This is the situation when changes in the protein or its expression which are beneficial for one role actually begin to degrade its performance of the other role. Such adaptive conflict would provide a selective advantage for gene duplication and specialization. In this scheme, unlike the classical model, at no stage is a gene drifting without the protective effects of selection, and at each stage there is the potential for selective benefit.

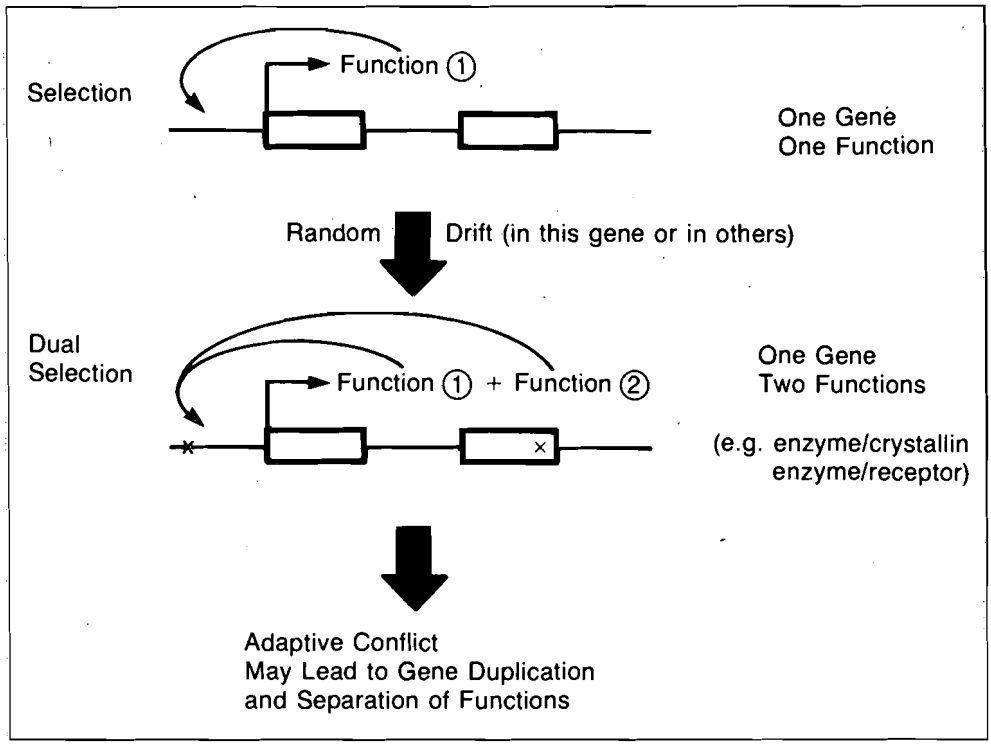

Fig. 2.2. The Gene Recruitment model for new protein functionalities and for gene duplication.

Other Paths to Multiple Functions

In the case of enzyme crystallins the recruitment to a new function occurs through tissue-specific modification of gene expression. However a protein could gain an additional function by other mechanisms. For example, neutral drift in amino acid sequences not essential for the main selected function of a protein could lead to the serendipitous acquisition of a new activity, such as a novel binding site. Something like this may explain the surprising discovery that the cytoplasmic glycolytic enzyme α-enolase has also been found as a plasminogen receptor exposed at the cell surface.[27] In this case the binding activity depends upon the C-terminal amino acid residue of the enzyme. If this role is indeed physiologically significant the enzyme may have two distinct functions.

Proteins could also acquire an additional function passively, through modifications to a second protein. For example, a serum protein might enhance its ability to bind to a cell surface if it happened to gain an additional binding site for a cell surface protein of previously unrelated function. The target protein would then find itself with a new function as a receptor for the serum protein without having under-

gone any sequence changes while the serum protein whose sequence had changed would still have only one function. Whatever the path followed, the result of recruitment is a single gene which encodes two or more protein functions. Evolution is inherently pragmatic and will make use of whatever functionalities or substrates are available. Thus there is every reason to expect that this sort of molecular opportunism may be quite widespread. Indeed, a number of well-characterized proteins have surprising "secret identities." These include the neurotrophic factor neuroleukin which is also the enzyme phosphohexose isomerase[28] and the protein which has been identified as protein disulfide isomerase, thyroid hormone binding protein, the β subunit of prolyl hydroxylase and the glycosylation site binding component of oligosaccharyl transferase.[29]

Gene Duplication

Although protein multifunctionality may be quite common, there are likely to be circumstances under which a gene and protein serving two masters may not be an evolutionarily stable condition. The requirements of the two roles may place contradictory pressures on the protein and give rise to adaptive conflict. In the gene recruitment model this conflict may be resolved by gene duplication, specialization and separation of function.

It is interesting to consider the possible applicability of this model to well-known examples of gene duplication outside the lens, such as the molecular evolution of digestive stomach lysozymes in ruminants.[30,31] Several different lineages of ruminant ungulates, monkeys and even a bird, the hoatzin,[32] have acquired multiple stomach lysozymes for the digestion of cellulose-metabolizing bacteria at acid pH. These enzymes seem to have been independently recruited from neutral pH lysozymes expressed in macrophages and elsewhere in defense against bacteria. In the classical model (Fig. 2.1) it would be assumed that a gene for a neutral pH lysozyme duplicated and that one copy drifted in sequence until it acquired both the attributes of protein sequence necessary for enzymatic function at low pH and expression in stomach.

However in the gene recruitment model an alternative scenario can be imagined (Fig. 2.2). For example, the promoter of a gene for a neutral pH lysozyme could have acquired an element which conferred additional expression in stomach while maintaining its original pattern of expression elsewhere. Since the enzyme was not adapted for this new low pH environment its activity would have been poor at best. However, so long as it was able to make some useful contribution to digestion there could have been enough selective advantage for this rudimentary gene recruitment to be maintained. Subsequently the new role of the enzyme could have been improved by selection for changes in protein sequence which enhances low pH stability and activity. However, at some point these beneficial changes for the digestive

role might have become sufficiently disadvantageous for the original neutral pH role of the multifunctional protein that an adaptive conflict was produced. At this stage a gene duplication of the gene for the enzyme would have had the great benefit of resolving the conflict and allowing rapid divergence in function and specialization of lysozymes for different functions and expression patterns.[31] One gene would have essentially reverted to the original role while the product of the other would have been free to acquire even more modifications to enhance its function in digestion.

Since the neutral and low pH roles of lysozymes in this hypothesis would have been so different the initial period of protein multifunctionality or gene sharing would necessarily have been brief and any sign of it would have been rapidly erased from the genome. In the lens, however, all the stages of this alternative model from gene recruitment through protein multifunctionality, adaptive conflict, gene duplication and subsequent specialization are illustrated by the varied examples of taxon-specific enzyme crystallins (Fig. 2.3). Indeed, presumably because of less serious problems of adaptive conflict in the lens, the enzyme crystallins frequently seem to be stable for very long periods at the initial one gene, two functions stage.

Additional Implications

In addition to their role in the evolution of new protein functions and in gene duplication, protein multifunctionality and gene recruitment have some other implications which may be worth considering. For instance, there is the conferring of a certain degree of economy in the number of genes required in an organism. Although this is unlikely to be of major importance in higher eukaryotes it could be of great significance in smaller genomes, particularly in viruses. From a different perspective, protein multifunctionality could have a great influence on the selective constraints experienced by a gene. In a protein which acquired a new function, sequences which were not well-constrained by the original function might come under more stringent selection while formerly well-conserved sequences might actually be forced to change to accommodate the new role. Clearly this could change the relative speeds of the "molecular clocks" for genes in different species. Indeed the recruitment of lactate dehydrydrogenase B (LDHB) as ε-crystallin in some birds but not others has been cited as the basis for varying clock rates for this enzyme in avian orders.[33]

Unexpected additional functions for proteins could also have implications for gene knock-out experiments and even for gene therapy. Targeting a gene or protein specifically for one known function might produce unexpected side-effects relating to a second hidden function. Similarly, taxon-specificity in such multifunctionality could contribute to some of the marked species differences seen in apparently homologous gene defects in different species.

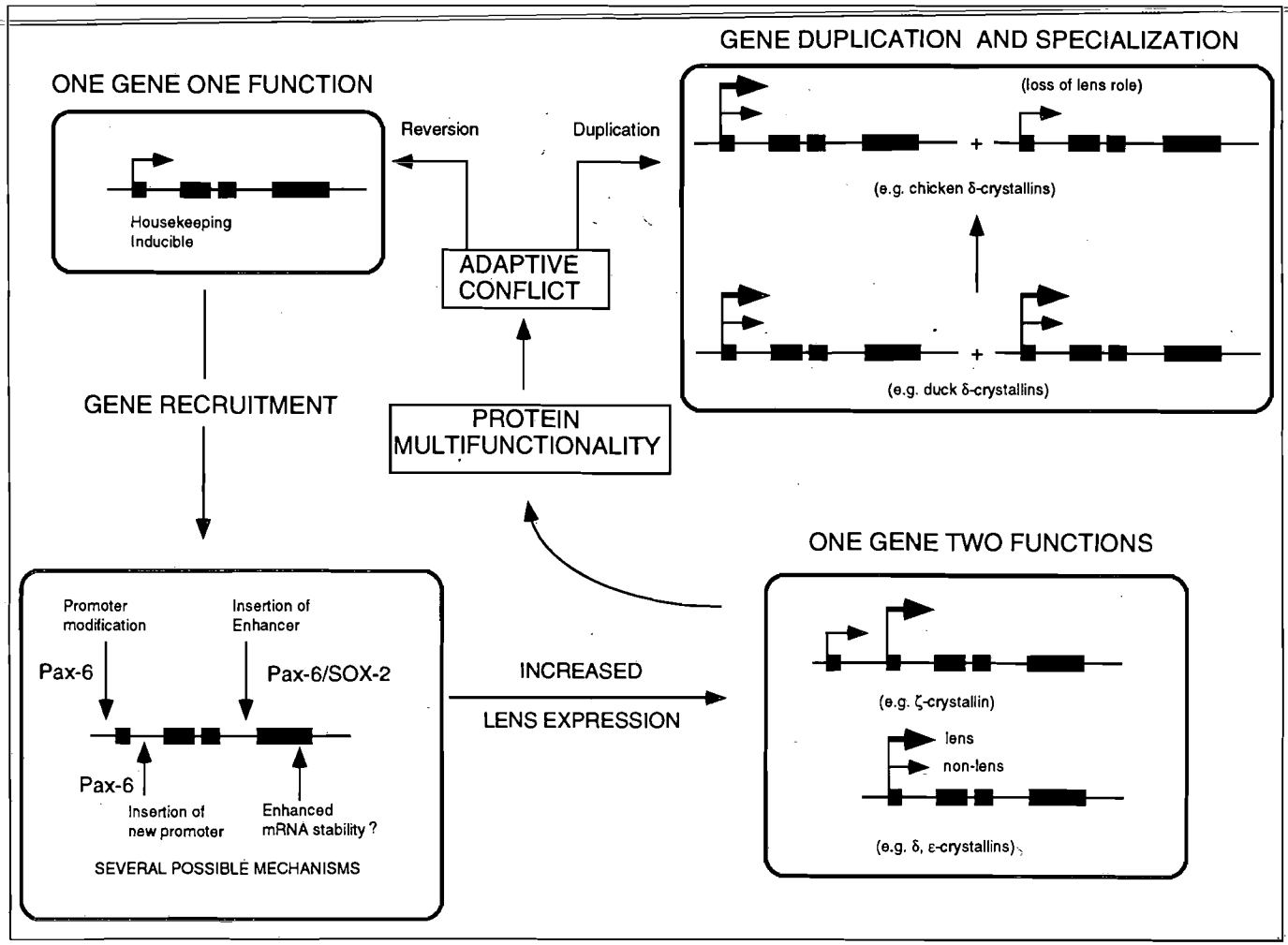

Fig. 2.3. Gene recruitment, protein multifunctionality and gene duplication in the crystallins.

REFERENCES

1. Kurosky A, Barnett DR, Lee TH et al. Covalent structure of human haptoglobin: a serine protease homolog. Proc Natl Acad Sci USA 1980; 77:3388-92.
2. Buck CA. Immunoglobulin superfamily: structure, function and relationship to other receptor molecules. Semin Cell Biol 1992;3:179-88.
3. Piatigorsky J, Wistow GJ. Enzyme/crystallins: Gene sharing as an evolutionary strategy. Cell 1989; 57:197-9.
4. Wistow G, Piatigorsky J. Lens crystallins: evolution and expression of proteins for a highly specialized tissue. Ann Rev Biochem 1988; 57:479-504.
5. de Jong WW, Hendriks W, Mulders JW, Bloemendal H. Evolution of eye lens crystallins: the stress connection. Trends Biochem Sci 1989; 14:365-8.
6. Wistow G. Lens crystallins: Gene recruitment and evolutionary dynamism. Trends Biochem Sci 1993; 18:301-6.
7. de Jong WW, Leunissen JAM, Voorter CEM. Evolution of the α-crystallin/small heat-shock protein family. Mol Biol Evol 1993; 10:103-26.
8. Lubsen NHM, Aarts HJ, Schoenmakers JGG. The evolution of lenticular proteins: the β- and γ-crystallin super gene family. Prog Biophys Mol Biol 1988; 51:47-76.
9. Fernald RD, Wright SE. Maintenance of optical quality during crystalline lens growth. Nature 1983; 301:618-20.
10. de Jong WW. Evolution of lens and crystallins. In: Bloemendal H, ed. Molecular and Cellular Biology of the Eye Lens. New York: Wiley-Interscience, 1981; 221-78.
11. Siezen RJ, Shaw DC. Physicochemical characterization of lens proteins of the squid *Nototodarus gouldi* and comparison with vertebrate crystallins. Biochim Biophys Acta 1982; 704:304-20.
12. Wistow G, Turnell B, Summers L et al. X-ray analysis of the eye lens protein γ-II crystallin at 1.9 A resolution. J Mol Biol 1983; 170:175-202.
13. Brakenhoff RH, Aarts HJ, Reek FH, Lubsen NH, Schoenmakers JG. Human γ-crystallin genes. A gene family on its way to extinction. J Mol Biol 1990; 216:519-32.
14. Siezen RJ, Wu E, Kaplan ED, Thomson JA, Benedek GB. Rat lens γ-crystallins. Characterization of the six gene product and their spatial and temporal distribution resulting from differential synthesis. J Mol Biol 1988; 199:475-90.
15. Kim RY, Wistow GJ. Expression of the duck α-enolase/τ-crystallin gene in transgenic mice. FASEB J 1993; 7:464-9.
16. Wistow GJ, Mulders JW, de Jong WW. The enzyme lactate dehydrogenase as a structural protein in avian and crocodilian lenses. Nature 1987; 326:622-4.
17. Rao PV, Zigler JS Jr. Extremely high levels of NADPH in guinea pig lens: correlation with ζ-crystallin concentration. Biochem Biophys Res Commun 1990; 167:1221-8.

18. Rao CM, Zigler JS Jr. Levels of reduced pyridine nucleotides and lens photodamage. Photochem Photobiol 1992; 56:523-8.
19. Clarke FM, Morton DJ, Stephan P, Wiedemann J. The functional duality of glycolytic enzymes: potential integrators of cytoplasmic structure and function. In: Ishikawa H, Hatano S, Sato H, eds. Cell Motility: Mechanism and Regulation (10th Yamada Conference). Tokyo: Tokyo University Press, 1985; 235-50.
20. Masters C. Cellular differentiation and the microcompartmentation of glycolysis. Mech Ageing Dev 1991; 61:11-22.
21. Masters C. Microenvironmental factors and the binding of glycolytic enzymes to contractile filaments. Int J Biochem 1992; 24:405-10.
22. Knull HR, Walsh JL. Association of glycolytic enzymes with the cytoskeleton. Curr Top Cell Regul 1992; 33:15-30.
23. Hughes AL. The evolution of functionally novel proteins after gene duplication. Proc R Soc Lond B Biol Sci 1994; 256:119-24.
24. Ohno S. Evolution By Gene Duplication. Heidelberg: Springer-Verlag, 1970.
25. Kimura M, Ohta T. On some principles governing molecular evolution. Proc Natl Acad Sci USA 1974; 71:2848-52.
26. Li WH, Graur D. Fundamentals of Molecular Evolution. Sunderland, MA: Sinauer Associates Inc., 1991.
27. Miles LA, Dahlberg CM, Plescia J, Felez J, Kato K, Plow EF. Role of cell-surface lysines in plasminogen binding to cells: identification of α-enolase as a candidate plasminogen receptor. Biochemistry 1991; 30:1682-91.
28. Chaput M, Claes V, Portetelle D et al. The neurotrophic factor neuroleukin is 90% homologous with phosphohexose isomerase. Nature 1988; 332:454-5.
29. Geetha-Habib M, Noiva R, Kaplan HA, Lennarz WJ. Glycosylation site binding protein, a component of oligosaccharyl transferase, is highly similar to three other 57 kd luminal proteins of the ER. Cell 1988; 54:1053-60.
30. Irwin DM, Prager EM, Wilson AC. Evolutionary genetics of ruminant lysozymes. Anim Genet 1992; 23:193-202.
31. Takeuchi K, Irwin DM, Gallup M et al. Multiple cDNA sequences of bovine tracheal lysozyme. J Biol Chem 1993; 268:27440-6.
32. Kornegay JR, Schilling JW, Wilson AC. Molecular adaptation of a leaf-eating bird: stomach lysozyme of the hoatzin. Mol Biol Evol 1994;11:921-8.
33. Wilson AC, Ochman H, Prager EM. Molecular time scale for evolution. Trends Biochem Sci 1987; 3:241-7.