Molecular Biology Intelligence Unit

Molecular Biology and Evolution of Crystallins: Gene Recruitment and Multifunctional Proteins in the Eye Lens

Graeme Wistow, Ph.D.

National Institutes of Health
National Eye Institute
Bethesda, Maryland, U.S.A.

R.G. Landes Company
Austin

Molecular Biology Intelligence Unit

MOLECULAR BIOLOGY AND EVOLUTION OF CRYSTALLINS: GENE RECRUITMENT AND MULTIFUNCTIONAL PROTEINS IN THE EYE LENS

R.G. LANDES COMPANY
Austin, Texas, U.S.A.

Submitted: June 1995
Published: August 1995

U.S. and Canada Copyright © 1995 R.G. Landes Company
All rights reserved. Printed in the U.S.A.

Please address all inquiries to the Publisher:
R.G. Landes Company, 909 Pine Street, Georgetown, Texas, U.S.A. 78626
or
P.O. Box 4858, Austin, Texas, U.S.A. 78765
Phone: 512/ 863 7762; FAX: 512/ 863 0081

U.S. and Canada ISBN 1-57059-299-3

International Copyright © 1995 Springer-Verlag, Heidelberg, Germany
All rights reserved.

International ISBN 3-540-60235-6

While the authors, editors and publisher believe that drug selection and dosage and the specifications and usage of equipment and devices, as set forth in this book, are in accord with current recommendations and practice at the time of publication, they make no warranty, expressed or implied, with respect to material described in this book. In view of the ongoing research, equipment development, changes in governmental regulations and the rapid accumulation of information relating to the biomedical sciences, the reader is urged to carefully review and evaluate the information provided herein.

Library of Congress Cataloging-in-Publication Data

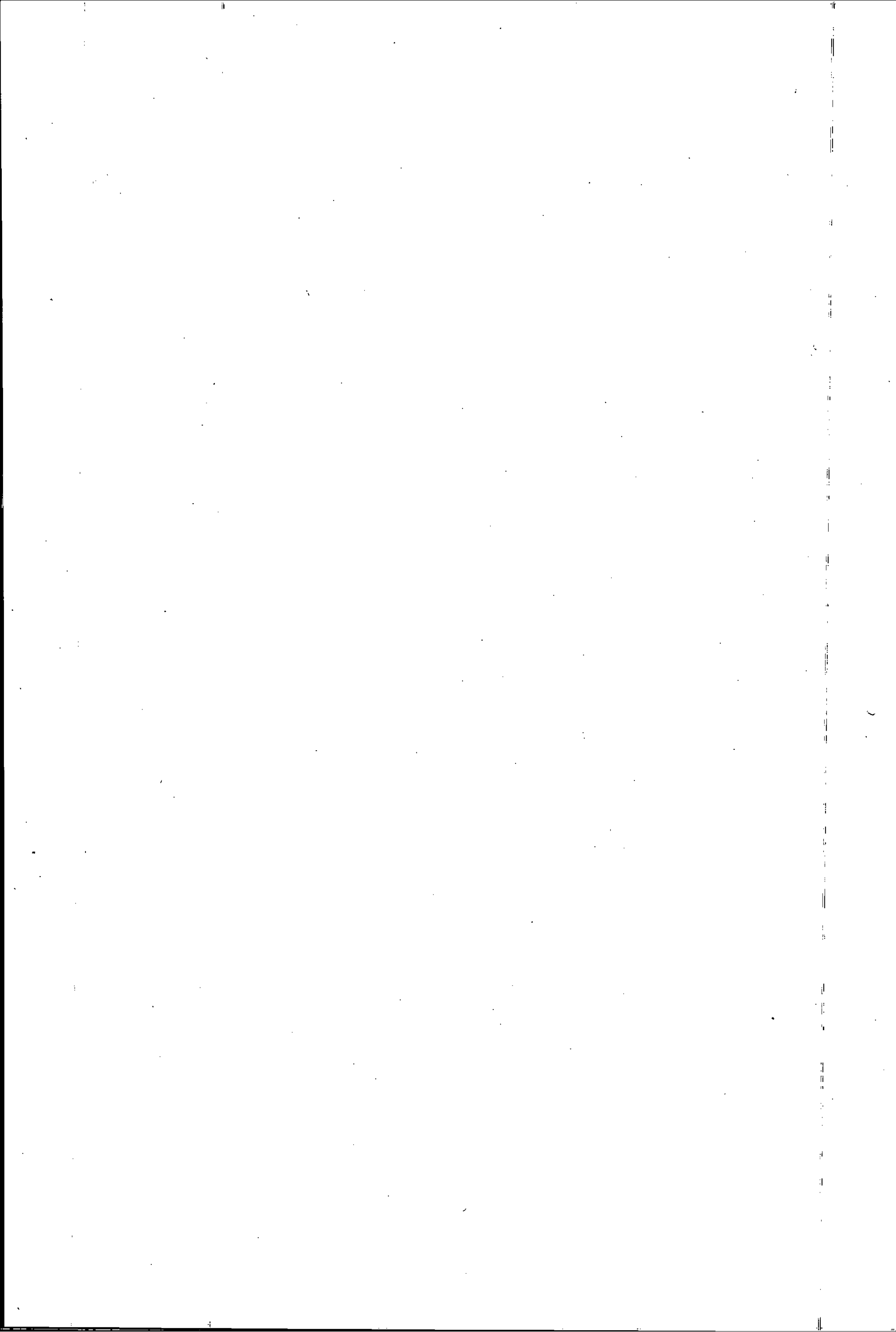
Wistow, Graeme J., 1953-
Molecular biology and evolution of crystallins : gene recruitment and multifunctional proteins in the eye lens / Graeme J. Wistow.
p. cm. — (Molecular biology intelligence unit)
Includes bibliographical references and index.
ISBN 1-57059-299-3 (alk. paper)
1. Crystalline lens—Molecular aspects. 2. Crystalline lens—Evolution. I. Title. II. Series.
[DNLM: 1. Lens, Crystalline—metabolism. 2. Lens, Crystalline—genetics. 3. Gene Expression. WW 260 W817m 1995]
QP478.W56 1995
591.1'823—dc20
DNLM/DLC 95-33050
for Library of Congress CIP

PUBLISHER'S NOTE

R.G. Landes Company publishes five book series: *Medical Intelligence Unit, Molecular Biology Intelligence Unit, Neuroscience Intelligence Unit, Tissue Engineering Intelligence Unit* and *Biotechnology Intelligence Unit.* The authors of our books are acknowledged leaders in their fields and the topics are unique. Almost without exception, no other similar books exist on these topics.

Our goal is to publish books in important and rapidly changing areas of medicine for sophisticated researchers and clinicians. To achieve this goal, we have accelerated our publishing program to conform to the fast pace in which information grows in biomedical science. Most of our books are published within 90 to 120 days of receipt of the manuscript. We would like to thank our readers for their continuing interest and welcome any comments or suggestions they may have for future books.

Deborah Muir Molsberry
Publications Director
R.G. Landes Company



CONTENTS

1. The Evolution of Eyes and Lenses 1
Light and Life 1
The Evolution of Eyes: Diversity and Success 4
Structure of the Vertebrate Eye 4
Structure and Development of Cellular Lenses 6
Pax-6 and Common Origins 13
Evolution Dynamism and the Vertebrate Lens: Environmental Pressures 17

2. Gene Recruitment: A Novel Mechanism in Molecular Evolution 25
Multifunctional Proteins and "Gene Sharing" 25
Wider Implications for Molecular Evolution 28

3. The Ubiquitous Crystallins: Stress Proteins Recruited to Lens 37
α-Crystallins: Members of the Small Heat Shock Protein Superfamily 37
The Small Heat Shock Protein Connection 39
Gene Structure 41
β- and γ-Crystallins: A Superfamily in the Vertebrate Lens 47
Protein Structure 52
A Wider Superfamily 59
Introns and Internal Duplications in β- and γ-Crystallin Gene Evolution 62
γ-Crystallins and the Evolution of the Lens 67

4. The Gene Recruitment of Enzymes as Crystallins 83
ε-Crystallin 83
δ-Crystallin 89
ρ-Crystallin 94
τ-Crystallin 94
Gecko Crystallins: A Response to Light? 97
Enzyme Crystallins in Mammals 99
η-Crystallin 101
ζ-Crystallin 103
λ-Crystallin 103
μ-Crystallin 104
Crystallins Elsewhere: Invertebrate Lenses, Light Organ Lens and Cornea 106
Other Crystallins 108
Recruitment Through Modified Gene Expression 109

5. **Crystallin Gene Expression: The Pax-6 Connection 123**
Pax-6, Eye Development and the Expression of Crystallins 124
Taxon-Specfic Crystallins 126
Pax-6 and Cataract 134
Other Taxon-Specific Crystallins 134
Ubiquitous Crystallins 138
Transgenics 151
Post-Transcriptional Control 153
Summary 154

6. **A Brief History of Lens and Crystallin Recruitment 165**

Index 169

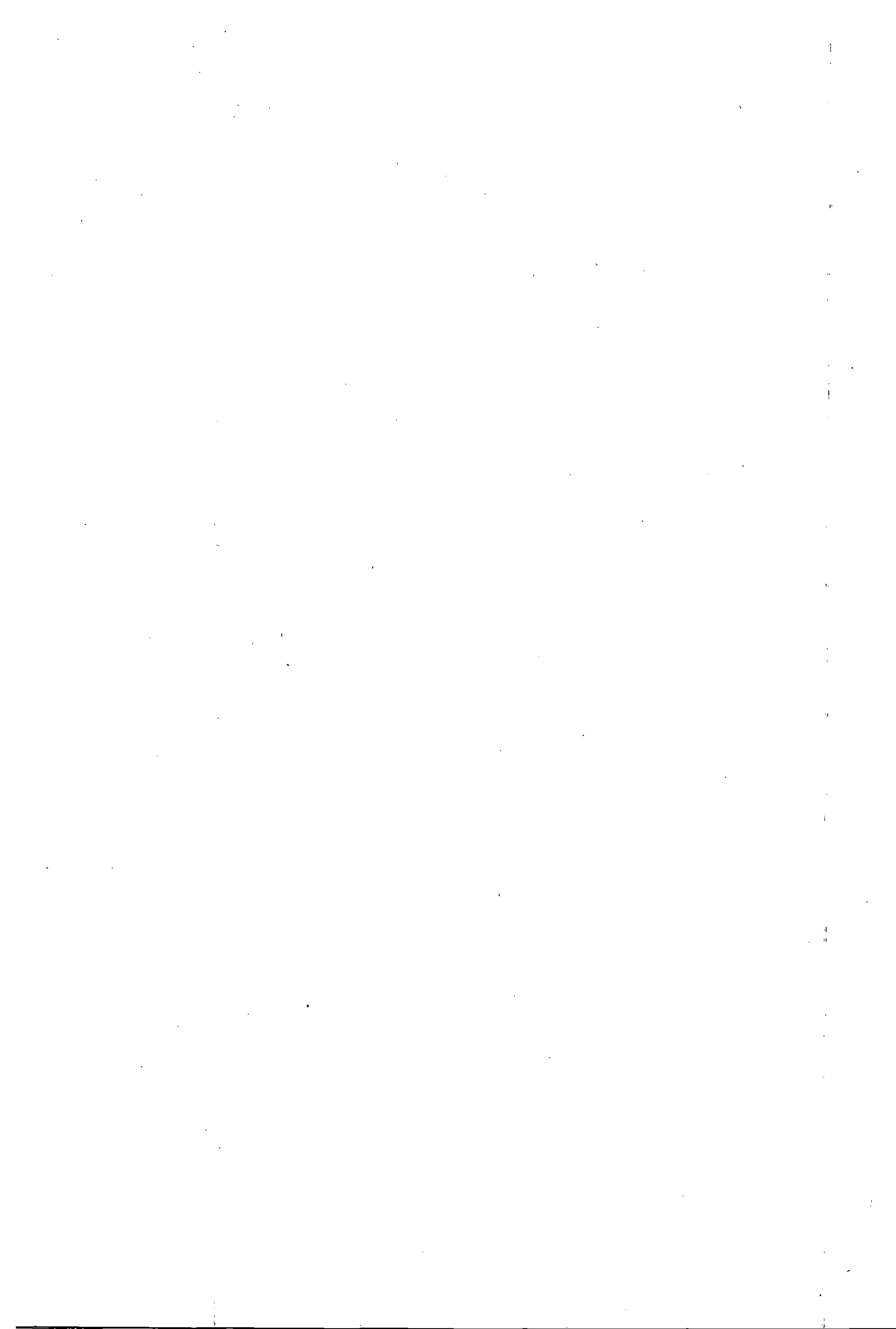
PREFACE

The eye lens is a remarkable transparent tissue with an important role in vision. Recent advances in our understanding of the lens in various species have revealed complex histories of molecular evolution and adaptation. In particular we have seen that the properties of the lens have always depended on the direct recruitment of existing proteins to new structural roles as crystallins. The lens turns out to be a particularly advantageous system for examination of evolutionary and developmental processes which may have wider significance beyond the lens itself.

Furthermore it has now been discovered that the molecular biology and development of the lens is intimately connected with ancient gene cascades which define "eye" in species from flies to mice. Thus we have seen a direct connection between essential tissue-determining genes such as *Pax-6* and *Sox-2* and the expression of crystallins in the lens.

This book describes our present view of the molecules of the lens in the context of the wider evolutionary history of the eye. Each chapter is intended to serve as review of specific areas; the evolution and development of the eye and the lens, the phenomenon of crystallin gene recruitment, the ubiquitous stress-related crystallins, the taxon-specific enzyme crystallins and the mechanisms of crystallin gene expression, followed by a closing summary.

Although this story has developed from studies of the lens it illustrates a number of important general processes in biology. It should therefore be of interest not only to those working directly on the eye but also to others involved in various fields of molecular and evolutionary biology.



CHAPTER 1

THE EVOLUTION OF EYES AND LENSES

Crystallins are the abundant, soluble structural proteins of cellular lenses in vertebrate and invertebrate eyes. The lens is a highly specialized tissue in a highly specialized organ. Its function is to control the refraction of light and image formation in the eye. Even though cellular lenses are relatively late and independently derived features of eyes, molecular studies of the lenses of vertebrates and invertebrates have revealed both a surprising diversity in composition and a surprising congruence in molecular mechanisms. These underlying similarities seem to be the result of a common evolutionary history shared by all metazoan eyes from a very early stage of organization.

Thus the evolutionary and developmental origins of crystallins, some of which are the results of quite recent recruitment events, are inextricably connected to the long evolutionary history of eyes and vision. Although the ability to sense and make sense of light and to modify behavior accordingly seems to be a complex and sophisticated behavior, one which is hard to mimic even with advanced technology, the origins of vision are surprisingly ancient in the history of life on earth.

LIGHT AND LIFE

Life is inherently opportunistic and inevitably it has made good use of the solar energy which penetrates the atmosphere as visible light. Green plants use chlorophyll-based photosynthetic systems in specialized organelles to harness the energy of light in chemical bonds[1] while some bacteria such as *Rhodopseudomonas* use the unrelated bacteriochlorophylls for a similar purpose.[1] Certain archaebacteria such as the halophile *Halobacterium halobium* (now *salinarium*) also use light as a source of energy. These prokaryotes possess an integral membrane protein called bacteriorhodopsin which consists of seven transmembrane α-helices arranged in a bundle.[2,3] A lysine residue in the seventh helix binds the chromophore retinaldehyde (retinal) through a Schiff's base linkage. When the bound retinal absorbs a photon it undergoes a

stereoisomerization from all-*trans* to 13-*cis*.[2,3] This triggers a conformational change in the protein and serves to activate a proton pump which provides an energy source for the cell.[3,4] A closely related protein, halorhodopsin, acts in a similar way as a light-powered chloride pump.[2] However archaebacteria, protists and animals go beyond using light as an energy source to exploit the broad spectrum, high frequency and directionality of light as the most information-rich of sensory media.

Remarkably it has been found that in addition to bacteriorhodopsin and halorhodopsin prokaryotic *Halobacteria* also possess two related proteins called SRI and II, where SR denotes sensory rhodopsin.[2,5,6] As their name suggests, these sensory proteins are required for a phototropic response. In eukaryotes, eyespots and eyes in species ranging from the unicellular photosynthetic protists *Chlamydomonas*[7] and *Euglena*[8] to complex multicellular animals also make use of the chromophore retinal bound to membrane proteins. In animals such as *Drosophila* and mammals, these proteins are known as opsins[9,10] (Fig. 1.1). Even in *Chlamydomonas* it seems likely that a similar retinal-binding, opsin is responsible for the response to light.[7] Like bacteriorhodopsin, opsins are integral membrane proteins with seven transmembrane α-helices in which retinal is bound via a Schiff's base to a lysine in the seventh helix. In animal opsins absorption of a photon causes a conformational change both in the bound chromophore retinal, usually from 11-*cis* to all-*trans*, and in the opsin itself. This change in structure initiates an amplifying cascade of signaling events, culminating in a release of neurotransmitters and a nerve impulse.[9-11] Depending on the organism involved this can lead to movement towards a candle flame or to understanding the written word.

The striking structural and functional similarities of the bacteriorhodopsin and opsin families could be the result of common descent from a single original "invention" of this protein motif at a very early stage in evolution. One might even speculate that this system for light absorption arrived in eukaryotic cells through a prokaryotic symbiont in a manner similar to the acquisition of organelles such as mitochondria and chloroplasts.[12] Unfortunately, there is far too little sequence similarity between bacteriorhodopsins and opsins to demonstrate homology.[13] By itself this does not eliminate the possibility of an evolutionary relationship since tertiary structure and functionality are often found to be well conserved even in the absence of obvious sequence similarity, as for example in the relationship between the 70 kDa heat shock proteins and actin.[14] However prokaryotic and eukaryotic opsins also differ in the stereoisoforms of retinal they bind. It is quite possible that in spite of their similarities the two families of proteins separately converged on the same structure since this seven-helical motif is thought to be extremely common in membrane protein receptors.[10,15,16]

While we may not be able to demonstrate common ancestry of bacteriorhodopsins and animal opsins there is still the real possibility that the visual pigment of the eukaryotic protist *Chlamydomonas* is related to animal opsins.[7] If true, this could place the root of eye evolution

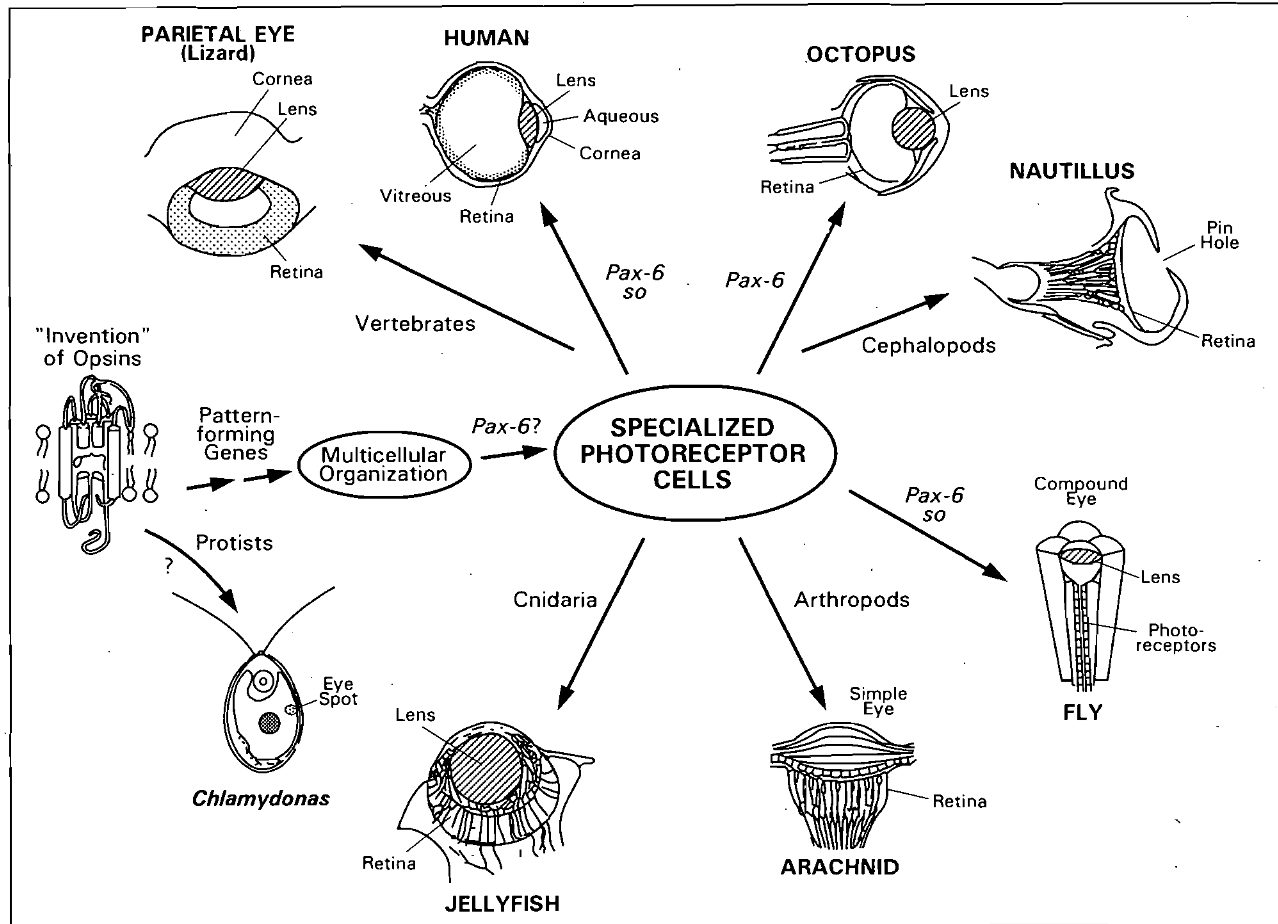


Fig. 1.1. The evolution of eyes. Diverse metazoan eyes may share a common origin and common molecular mechanisms of development involving Pax-6 *and other genes such as* sine oculis *(so).*[109] *All eyes and eyespots may share an even more ancient common origin in the evolutionary innovation of the opsin gene family. Figure is adapted from several sources including references 19, 27, 32, 110, 111.*

at least as far back as unicellular eukaryotes. Sequence data for protist opsins are eagerly awaited.

After the "invention" of opsin itself and of the mechanisms to couple its light receptor function to cellular responses the next big step in the evolution of eyes came with the arrival of multicellular organisms (Fig. 1.1). In metazoans it became possible to produce differentiated cells which could specialize in the production of opsins. These were the ancestors of the photoreceptor cells of animal eyes in which opsins are concentrated in arrays in plasma membranes. Photoreceptor cells occur in two major classes, the rhabdomeric photoreceptors composed of microvilli which are found in insect compound eyes and elsewhere or the ciliary photoreceptor cells typical of mammalian retinas.[9,11,17] These cells contain all the machinery of the visual cascade together with neural connections to transmit information to the rest of the organism. It now appears that the earliest achievement of specialized photoreceptor cells during multicellular organization may be ancestral to both rhabdomeric and ciliary photoreceptors since eyes from both lineages share the same fundamental molecular mechanisms of development and differentiation (Fig. 1.1).

THE EVOLUTION OF EYES: DIVERSITY AND SUCCESS

Even the most simple visual systems are a remarkable testimony to the power of natural selection and molecular evolution. Eyes are so useful that they are widespread in animals to the point that over 90% of living animal species have some kind of vision.[17-19] Some eyes, such as those in certain species of nematode, may be no more than light sensitive patches which may serve simply for up/down or light/dark orientation.[18,20] However, in many cases more of the information content of light is exploited by some kind of imaging system.

One of the simplest systems is found in the eye of the sea-going cephalopod mollusk *Nautilus*. The photoreceptor cells are arrayed in a curved retina much like that of the vertebrate eye.[17,20,21] Image formation is provided simply by a small hole in the front of the eye giving the form of a simple pinhole camera (Fig. 1.1). Light refracted through the pinhole can form a clear but dim image on the retina. Many other species, both invertebrates such as other mollusks, arachnids and jellyfish, and vertebrates (Fig. 1.1) have eyes in which light is concentrated and directed by a lens to give brighter images.[17,20,21] In some species, lenses may be used only as light concentrators, but elsewhere they are used for image formation, both gathering and focusing light, often correcting for the spherical and chromatic aberration which may afflict inorganic lenses as they do so.[17,22-24]

STRUCTURE OF THE VERTEBRATE EYE

The vertebrate eye (Fig. 1.2) is a spherical organ consisting of several transparent layers overlaying a photosensitive retina all contained

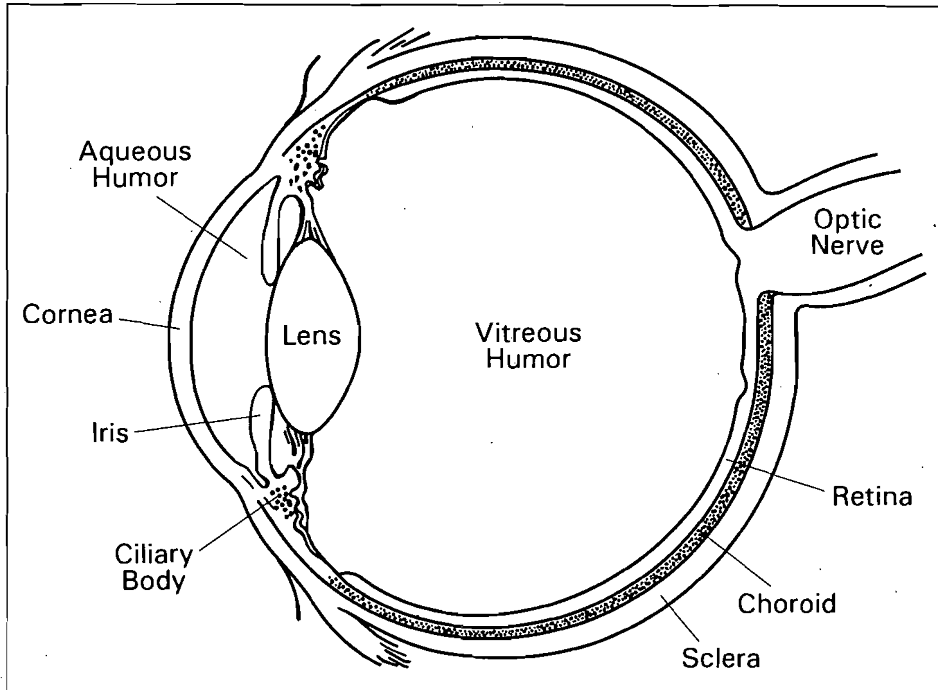


Fig. 1.2. The structure of the human eye.

in an opaque eyeball with a single anterior opening and a posterior connection to the brain via the optic nerve.[20,25,26] The first transparent layer, covering the anterior opening, is the cornea which is primarily an extracellular matrix of collagen overlaid by thin layers of endothelial and epithelial cells.[25,26] In many species the cornea provides a major part of the focusing power.[20] It also protects the interior of the eye from the external environment and absorbs much of the short wavelength ultra-violet radiation which could harm the sensitive interior. Both the physical curvature of the cornea and much of its nutrient supply derive from the aqueous humor, a clear fluid which fills the anterior chamber of the eye.[25,26] In the avascular anterior chamber the aqueous plays an essential role in transport of nutrients, growth factors and waste products for both the cornea and the anterior part of the lens.

The aperture of the eye, the pupil, is defined by a pigmented contractile tissue, the iris, which extends from the ciliary body.[25,26] Just behind the iris, suspended from the ciliary body by a system of ligaments, is the lens, a highly specialized cellular structure with various roles in producing a sharp visual image.[25,26] This image is projected onto the retina through the gelatinous vitreous body which fills the rest of the eyeball. Like the aqueous, the denser vitreous has nutritive

and other transport roles to play.[25,26] The retina consists of layers of nerve cells and the opsin-containing photoreceptors themselves.[25,26] Curiously, in the vertebrate eye these cells are arranged at the back of the retina so that light must first pass through the neural layers. Photoreceptor cells function without turnover throughout life. They continually produce membranous discs containing the opsins and other proteins of the visual cascade.[25,26] Particularly in the rod cells these form stacks of discs which, as they age, are shed and scavenged by the retinal pigment epithelium.[25,26] This light absorbing layer also serves to eliminate dazzle from internal reflection of unabsorbed light.

Other architectures for eyes, such as the use of directed bundles of photoreceptor cells perhaps coupled with wave guides rather than lenses have been exploited widely, most notably in the compound eyes of insects.[18,21] The eye of the scallop *Pecten* even makes use of a mirror instead of a lens for light gathering.[20] However this is unusual. Unlike astronomical telescopes in which the dominant form of optics is reflective, eyes are predominantly refractive.

Although we are most familiar with the idea of paired, symmetrically equivalent eyes many species have several sets of eyes, sometimes of different types. Many arthropods, such as *Drosophila*, have both compound eyes and small simple eyes with concentrating lenses.[18,21] Even vertebrates may have additional eyes. In lampreys, amphibians and some reptiles a small third (and sometimes fourth) eye, the parietal or median eye (Fig. 1.1), forms from a vesicle of neural ectoderm.[27] The posterior part of this vesicle develops photoreceptor cells to form a "retina" while the anterior part consists of a single layer of elongated, transparent cells, a "lens." The function of this eye is unknown although it seems likely that it has a role in setting diurnal rhythms. In birds and mammals the parietal eye has evolved into the pineal, the main source of the hormone melatonin.[28]

STRUCTURE AND DEVELOPMENT OF CELLULAR LENSES

A lens is basically a curved interface between regions of differing refractive indices. This kind of structure can be achieved in various ways. For a cellular lens, all that is required is for a monolayer of cells to elongate while constrained around the edges by contact with other cells, and for these cells to increase their protein concentration and hence their refractive index. Indeed, cellular lenses in both vertebrates and invertebrates are mainly composed of extremely elongated cells or, in the case of cephalopods, cell processes.[29-32] Discontinuities between adjacent cells are minimized and a uniform tissue consisting primarily of cytoplasm is formed.

The division into different cells allows the formation of a gradient of refractive index as protein content varies between layers of cells.[30,33] Generally this is used to increase the apparent convexity of the lens

and to increase its focusing power. Thus the center of the lens has a higher protein concentration and hence a higher refractive index than the cortical regions.[23,30,33] There may be a two fold difference in protein concentration between these two regions. The recent discovery that the lens in a chameleon acts as if it were concave rather than convex[34] suggests that this pattern may have been reversed to fit the peculiar optical requirements of this species.

Cell elongation, which is probably a largely osmotic process, is a key feature of lens development.[35-37] In vertebrate lenses it is the first recognizable stage in lens differentiation during embryogenesis (Fig. 1.3). A patch of epidermal ectoderm overlying neural ectoderm undergoes cell elongation to form the lens placode.[25,26,31,38] As development proceeds, there is a coordinate invagination of the lens placode and the underlying neural ectoderm. The lens placode pinches off from the surrounding ectoderm to form the lens vesicle. Even at this early stage, expression of the characteristic lens proteins, the crystallins can be detected.[39,40] The neural ectoderm layer goes on to form the eye cup which gives rise to the retina, ciliary body and other structures.

The lens vesicle consists of the original elongated cells of the placode, now at the posterior of the vesicle, and an anterior layer of undifferentiated, cuboidal cells. This arrangement allows for continual growth in the lens. The posterior, elongated cells undergo further elongation and differentiation becoming the primary fiber cells. They extend until they fill the lens vesicle and come into contact with the anterior layer. The anterior cells comprise the progenitors of the lens epithelium, a stem cell-like population which persists throughout life. While the central anterior epithelial cells remain rather quiescent more posterior cells migrate towards the lens equator where they enter a proliferative zone and go through mitosis (Fig. 1.4). At the lens equator cells undergo a dramatic terminal differentiation into enormously elongated new fiber cells. These secondary fiber cells overlay the primary fibers in concentric layers in a process which continues throughout life. The original embryonic primary fibers form the so-called lens nucleus. Later, fully mature secondary fibers may also constitute part of the nucleus, the densest region of the lens and a frequent locus for cataract formation in humans. Younger secondary fibers, particularly those which are still metabolically and synthetically active form the lens cortex.

Although cell elongation is essential for lens development, it is not the only determinant of the shape of the lens. During development vertebrate lenses acquire the lens capsule, essentially a basement membrane surrounding the lens, which helps to constrain its shape. While the epithelial cells are attached to the capsule the fiber cells detach during differentiation and eventually form contacts only with other fiber cells. This contact is mediated through numerous gap junctions[41] and probably also through an abundant lens fiber cell membrane protein called MIP26,[42] a member of a large family of water channel proteins.[43]

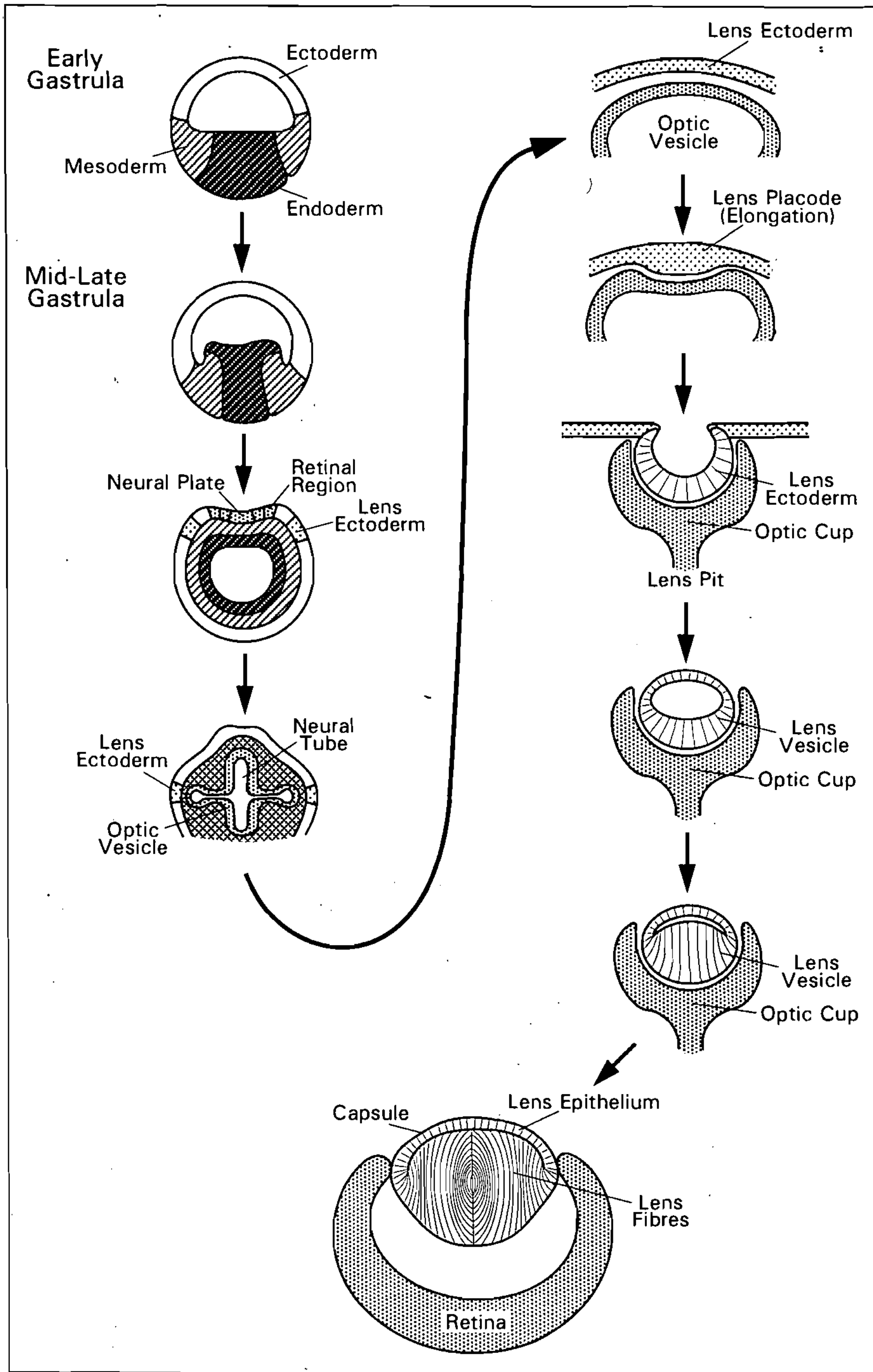


Fig. 1.3. The ontogeny of the vertebrate eye. This figure is a composite derived from a model of lens induction in Xenopus laevis [38] *shown on the left side, and later stages of eye and lens differentiation in the rat*[29] *shown on the right.*

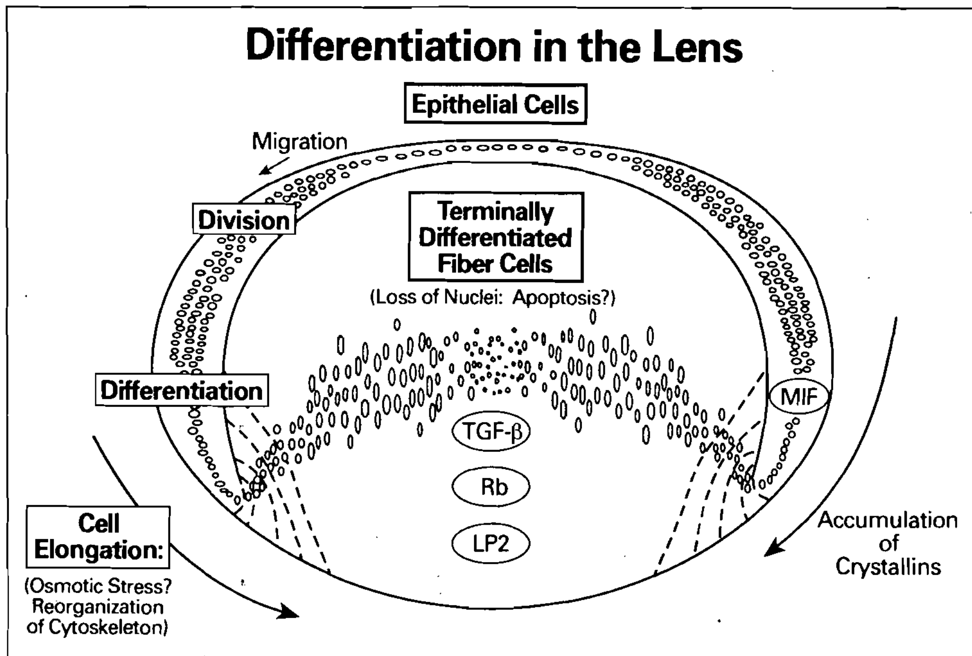


Fig. 1.4. Differentiation in a generalized vertebrate lens. Some non-crystallin molecular markers identified in various species are indicated; MIF,[59] TGFβ1,[58] Retinoblastoma protein (Rb)[112] and LP2, a lipid binding protein in lens.[113]

As fiber cells mature and are overlaid by younger layers their terminal differentiation continues and cellular organelles are lost. Nuclei, mitochondria and the structures of the endoplasmic reticulum are eliminated at a sharp division in the lens between one cell layer and the next.[44] From an optical standpoint the loss of organelles is usually interpreted as a loss of potential sites for light scattering along the optical axis of the lens. The loss of nuclei also eliminates any possibility of the fiber cells resuming proliferation, something which would certainly disrupt lens structure and transparency. However it is possible that the loss of nuclei is not an end in itself but is simply an inevitable part of the program of differentiation in this tissue. Lens cell differentiation has some intriguing parallels with programmed cell death mechanisms. Most notably the cell nuclei condense and the chromosomal DNA breaks down in a characteristic nucleosomal ladder.[45] Eventually the nuclei disappear into the cytoplasm. It has been pointed out that this also has some similarity to an abortive mitotic phase and the loss of nuclei may be the result of a failure of this phase to complete.[46]

Whatever the reason for nuclear breakdown, it is clear that mature fiber cells lose their ability to express genes or to synthesize new

protein and even decline in general metabolic capability.[47] Proteins in the mature fiber cells must survive without turnover throughout life. Lens proteins which are synthesized in the embryo and persist throughout life may thus be the oldest in the organism.

Growth Factors and Lens Differentiation

An unusual feature of the vertebrate lens is its relative isolation from other tissues. Although the lens may be vascularized during mammalian (but not avian) embryonic development the mature lens in all vertebrates is a completely avascular system.[47] This is necessary for transparency, but it means that nutrients and waste products must make their way to and from the lens from surrounding structures by diffusion. The closest structures are the ciliary body, to which the lens is attached through equatorial connections to the capsule, and the iris. Anteriorly lies the aqueous humor and the cornea. Posteriorly lies the vitreous and the retina. These surrounding structures communicate with the lens by means of diffusible growth factors and hormones and it is likely that the lens communicates with the rest of the eye in the same way. Growth of different tissues in the eye needs to be well coordinated and a deficiency in growth of one part, such as the lens, will lead to a general disruption of growth in the eye and microphthalmia.[48-50]

The compartmentalization of the eye allows the formation of gradients of growth factors across the lens. This may be a principal mechanism for control of lens differentiation. In the absence of growth factors, explanted rat lens epithelial cells remain quiescent but at increasing concentrations of acidic and basic fibroblast growth factors (aFGF, bFGF), the same cells in culture mimic the differentiation of the lens, they migrate, divide and finally increase in volume and synthesize differentiation-specific crystallins.[51-53] In this model high concentrations of aFGF and bFGF in the vitreous, lower concentrations in the posterior chamber (the space between the ciliary body and the lens), and low concentrations in the anterior chamber would be enough to regulate lens cell differentiation. This idea fits very well with some classic experiments in which lenses were reversed anterior to posterior in the embryonic chicken eye.[54] When this was done, epithelial cells which were now positioned posteriorly elongated, mimicking the differentiation of fiber cells. More recently, when a secretable form of aFGF was targeted to lens in transgenic mice, lens epithelial cells elongated and began to express differentiation-specific crystallins.[55]

Undoubtedly other growth factors are also involved in lens differentiation. These include IGF-1 which is important for lens cell growth in chicken,[56] platelet derived growth factor (PDGF) which when delivered in pulses maintains the transparency of rat lenses in organ culture[57] and TGFβ1 (transforming growth factor) which is localized in the fiber cells of mouse lens[58] (Fig. 1.4).

These growth factors may operate through a common path. A small protein expressed with moderate abundance in embryonic chick, mouse and human lens was found to be identical to a protein previously identified as macrophage migration inhibitory factor (MIF).[59] In lens the expression of this protein is associated with the equatorial region (Fig. 1.4). MIF is expressed in a delayed early response to mitogenic growth factors including bFGF, PDGF and TGFβ1 (which is mitogenic in NIH 3T3 fibroblast).[60,61] Antisense suppression of MIF in cultured cells blocks cell proliferation while constitutive expression allows cells to grow in the absence of serum.[61] MIF may be an essential part of the growth factor response in lens and other cells.

Evolution of a Cellular Lens

From Darwin onwards an explanation of the development of the multilayered vertebrate eye by step-wise selective processes has often been cited as one of the biggest challenges to classical evolutionary theory.[62] Recently a computer modeling exercise has demonstrated a possible selective path for the evolution of an eye in which each stage is a functional improvement over its predecessors.[63,64] This shows that the evolution of an eye superficially similar to those of vertebrates could occur rapidly, although since this treatment seems to consider zones of refractive index rather than discrete cells and tissues its ontogenic and phylogenetic implications for real eyes are not clear.

The evolution of a lens requires reasonable changes in structures of the eye with some benefit or lack of deleterious effect at each stage. A primitive ancestral stage in the evolution of the lens might have consisted of a single layer of elongated cells. This would resemble in some ways the so-called lens of the parietal or median eye present in many reptiles and amphibians[27] (although to what extent this structure acts as a lens is unknown). Such a structure could have served to protect the retina physically or from harmful radiation. It could also have begun to act as a concentrator of light.[20]

The size of a lens consisting of a single layer of cells is limited by the extent to which individual cells can elongate. Through the topological trick of forming the lens vesicle, the vertebrate eye lens is freed from these constraints and can grow throughout life adding new concentric layers of cells. It is by no means obvious how this trick was performed. It presumably occurred in one step perhaps through a single mutation in a gene controlling tissue-pattern formation or cell-cell recognition leading to the separation of lens cells from their surroundings. However convergent evolution has produced superficially very similar lenses in cephalopods, jellyfish and some other invertebrates although different developmental tricks have been used in different lineages. For example, in cephalopod lenses the concentric layers of fibers are not complete cells, instead they are cellular processes extending from a lentigenic region outside the lens proper.[65,66] Yet a very similar looking lens is the result.

Crystallins: Refraction and Transparency

The refractive power of the lens derives both from its curvature and its refractive index. The refractive index of the lens is largely a property of the crystallins, the soluble proteins of the "crystalline," or clear, lens which provide its bulk refractive structure.[47,67,68] Large amounts of crystallins accumulate in lens cells, particularly in the differentiated fiber cells (Fig. 1.4). Indeed, these proteins may account for as much as 80-90% of soluble protein in a highly proteinaceous tissue and, as such, are clearly structural proteins. Some dense, high refractive index lenses such as those of fish, rodents or squid, may have a protein content of 60% wet weight or more, although in many diurnal terrestrial species, particularly birds, the content is less than half this value.

Lens transparency is maintained by short-range order between crystallins.[69,70] Phase changes in the supramolecular structure of the lens, such as those which occur in cold cataract,[71] can lead to opacity by the creation of light scattering interfaces between zones of different refractive index. The intermolecular interaction of crystallins which define their supramolecular organization depend on the sequences and structures of the individual crystallins, as described in subsequent chapters.

Crystallins were originally characterized in a few domestic vertebrate species.[47,72] Size fractionation of native proteins revealed a few conspicuous size species which were named using the Greek alphabet, a convention which has been continued for vertebrate crystallins. In mammals, three classes were recognized, the α-, β- and γ-crystallins in descending order of native size, the α-crystallins being large aggregates, the β-crystallins dimers to octamers and the γ-crystallins monomers. In chickens the γ-crystallins were absent, apparently replaced by a different, multimeric crystallin which was named δ-crystallin. Subsequently many more species were examined and a new appreciation of crystallin diversity has emerged.

Crystallins may now be classified in two broad groups; the ubiquitous and the taxon-specific[37] (Table 1.1). The ubiquitous crystallins are represented in every vertebrate species examined suggesting that they reflect the composition of the ancestral vertebrate lens. They are the α-crystallins, consisting of two gene products, αA-crystallin and αB-crystallin; the β-crystallins which in mammals and birds are a family of six genes falling into two subgroups, βA and βB; and the γ-crystallins which in mammals form one family of highly similar genes expressed embryonically and neonatally and one more distantly related gene expressed in the cortical fibers of adult lenses. β- and γ-crystallins are related and may therefore share a common origin in the lens. A hypothetical ancestral vertebrate lens might have contained one α-crystallin and one ancestral β-crystallin.

In contrast the taxon-specific crystallins are major constituents of lenses only in defined evolutionary lineages (Table 1.1). They arose later than the ubiquitous crystallins as a result of discrete, indepen-

Table 1.1. Vertebrate crystallins

Ubiquitous: Stress-related	Taxon-specific: Enzymes	
α, β, γ	δ, ε, π	reptiles, birds
	ζ, η, λ, μ	mammals
	ρ	frogs
	τ	scattered distribution

dent recruitment events and were retained in descendant species.[37] The recruitment of taxon-specific crystallins provides an unusual opportunity to study events in the molecular evolution of species still at a stage of great diversity. Other systems may have enjoyed similar diversity in the distant past but this may have been obscured by subsequent extinctions and "rationalizations" of the pool of diversity.

Differential expression of the multiplicity of crystallins in each lens allows for establishment of smooth gradients of refractive index which enhance the optical properties of the lens, eliminating chromatic and spherical aberration.[22]

Crystallins and Stress

Although several different proteins serve as crystallins, many share the unifying of a connection to stress responses. There is a direct role for mammalian αB-crystallins in heat and osmotic shock[73-77] and other crystallins have more or less direct links with heat, osmotic, or various oxidative stresses[68,78-81] (Table 1.2). This may reflect the stressed condition of the lens itself. A tissue which undergoes enormous cell elongation, must continually maintain its balance of protein and water and be bathed in light for years on end. Crystallins seem to have been selected from proteins which are already expressed in the lens and which at high levels may have beneficial effects for lens stability. In particular several crystallins, including enzyme crystallins, may have a protective association with the cytoskeleton upon which the elongated fiber cells depend.[37,76] These properties of crystallins are discussed in later chapters.

PAX-6 AND COMMON ORIGINS

In spite of the enormous variety and evolutionary inventiveness of metazoan eyes and the high specialization of their component tissues, there is remarkable evidence that all these eyes and even derived tissues such as the vertebrate lens share a common origin. In vertebrates, flies, mollusks and worms the same molecular mechanisms are responsible for initiating the development of what otherwise appear to be such widely different eyes. Thus a gene responsible for normal eye development in man (*aniridia*)[82,83] and mouse (*small eye*)[84] turns out to be directly homologous to *eyeless* a gene essential for compound eye

Table 1.2. Crystallin connections

Heat Shock-
- αB is a sHSP in mammals
- Enolase (τ) is a HSP in yeast

Osmotic Stress-
- αB is induced by osmotic stress in mammals
- β and γ are related to proteins induced by dehydration
- ρ is related to aldose reductase, an osmotic stress protein
- Substrates of other enzyme crystallins may be osmolytes

Interaction with Cytoskeleton-
- α, β associate with cytoskeleton, plasma membrane
- LDH (ε), Enolase (τ), GAPDH (π) interact with "cytomatrix" in other cells

UV or oxidative stress-
- Several enzyme crystallins, ε, ζ, μ, ρ, π, bind NAD(P)H and sequester reduced and oxidized co-factors in lens where they could act as UV filters or as redox reagents

Chemical detoxification-
- Some enzymes used as crystallins are involved in detoxification: ζ is a quinone reductase; η is an aldehyde dehydrogenase; ρ is probably an aldo-keto reductase; squid SL1 is a glutathione reductase

development in *Drosophila*.[85] The product of these genes is a transcription factor, known in vertebrates as Pax-6, which belongs to a family of master-control factors whose expression determines the development of complex tissues.[86-89]

Indeed it seems that *Pax-6* may be the "master gene" for eye development in both mammals and flies (Fig. 1.1). When either *Drosophila* or mouse *Pax-6* is expressed ectopically in antenna, wing or leg of the fruit fly complete compound eyes develop in the targeted tissues.[90] This is a stunning demonstration of the existence of an ancient developmental control system which may be common to all eyes.

While ectopic expression of *Pax-6* could also be attempted in transgenic mouse embryos, mimicking the experiments in flies, the situation is likely to be more complex in vertebrates. After all, *Pax-6* is already expressed in other parts of the CNS[87] and even in pancreas.[91] Clearly the interplay of factors in different tissues can modulate the outcome of such developmental tinkering in less experimentally pliant organisms than *Drosophila*. There is also the interesting case of *Pax-6* in the eyeless nematode *C. elegans*. A homologue of *Pax-6*, vab-3/mab-18, has been found in this organism (A. Chisholm, personal communication).[92] It is expressed in sensory neurons and has an important role in correct formation of the "head."[93] *C. elegans* has no eyes but some other nematodes, such as *Mermis nigrescens*,[94] do have eyespots and phototaxis. It would not be surprising to find that *Pax-6* plays a key role in development of these structures. It seems likely that the

ancestors of *C. elegans* had eyes which were lost much like those of blind cave fish. What prevents eye formation in *C. elegans*? Has it lost target genes, such as those for opsins, or has it lost other factors downstream of *Pax-6* in the eye cascade?

Pax-6 operates at such a high level in the developmental cascade that it is upstream of all genes necessary to form an eye. Many of the genes which lie downstream of *Pax-6* must be different in flies and mice but the original trigger is the same for both and has been conserved for hundreds of millions of years. Other members of the cascade are also becoming known. *Sine oculis* is another homeodomain-encoding gene which is essential for eye development in *Drosophila* and it apparently has homologues in mammals which are also expressed in eye.[95] Other pattern-forming genes are also known to be expressed regionally in developing eyes of various species[96] including *Notch* of *Drosophila*[97] and numerous homeodomain-encoding genes such as *Msx-1* and *Msx-2* (formerly *Hox-7* and *Hox-8*)[98] and several *Hox* genes.[99]

This "eye cascade" must have evolved once early in metazoan evolution conferring the ability to respond to light. This gave the organisms which possessed it such advantages that their descendants came to dominate the animal world. As those descendant species radiated over hundreds of millions of years they continued to use the same ancestral molecular machinery even as the gross structure of their eyes diversified and adapted. Thus, although the common ancestor of octopus and human eyes is unlikely to have had a lens, its distant descendants both evolved superficially similar structures making use of some of the same common, inherited mechanisms.

This "master gene" role for *Pax-6* also helps explain the overlap in developmental origins and gene expression between eye and brain. In addition to eye, *Pax-6* is also expressed in the central nervous system,[87] particularly in the diencephalon[100] which is so closely related to eye in development. Its expression in various parts of the eye and brain is probably also responsible for the phenomenon of transdifferentiation among these tissues.

Eye and Brain

The eye has often been thought of as an offshoot of the brain (see ref. 67 for review). This seems logical enough since the optic cup which gives rise to the retina and ciliary body is derived from the neural ectoderm. However it has also been suggested that the eye came first.[101] Indeed it is striking to see that complex eyes are present in organisms, such as jellyfish, in which it is much harder to identify anything which could pass for a brain. This "eye-first" idea can be taken to its extreme if we entertain the possibility that the eyespots of protists share an evolutionary lineage with eyes of more complex organisms. If eyes came first, the brain could have developed as a center to process the information from the eye and to integrate it into behavioral responses.

As the image-forming and color-discriminating potential of the eye developed, so the brain developed further to make use of the information available. In this view the brain becomes a developmental extension of the retina rather than vice versa.

Whichever came first, the developmental unity of eye and brain are illustrated by the parietal eye of reptiles and amphibians[27] and its evolutionary descendent in birds and mammals, the pineal. In the parietal or median eye both a retina and a "lens" derive from the same neural ectoderm which gives rise to only the retina in the lateral eyes and which also gives rise to brain. Immunohistochemistry has suggested that the parietal eye lens shares molecular components with the lenses of the lateral eyes[102] but this has not yet been examined in detail at the molecular level. Thus the pineal which is regarded as part of the brain is descended from an eye similar in many ways to those with which we are familiar. Indeed, the chicken pineal expresses some of the same opsins as the retina of the lateral eyes.[103]

Transdifferentiation, Lens Regeneration and the Pax-6 Paradox

The close connection between differently derived parts of the eye and between eye and brain tissues is apparent in the remarkable phenomena of transdifferentiation and lens regeneration. In culture, cells from embryonic chicken adenohypophysis, iris and pigmented and neural retina can transdifferentiate to give rise to cell types resembling several differentiated tissue of the eye.[47,104-108] In particular all these systems can give rise to lens-like cells or lentoid bodies which express characteristic lens proteins. In many ways this is reminiscent of the derivation of the retina and lens of the parietal eye from neural ectoderm tissue. Lens can also be derived from other differentiated eye tissues in vivo. After lens removal in some species of newt, the dorsal and ventral iris, tissues of neural ectodermal origin, can regenerate a lens which expresses crystallins while in *Xenopus laevis* lens can regenerate from cornea.[47,65] Thus these differently derived tissues have the potential to follow the path of lens development even in the adult. However lens cells themselves are not capable of transdifferentiation into any other tissue. In this sense they are the lowest common denominator of differentiation potential in their developmental lineage.

The important role of *Pax-6* in all these tissues may be the basis for these phenomena. Recent work has shown that lens competence is a very early stage in development of the animal cap ectoderm and that earlier work which implied an inductive role for the optic cup was in error.[38] Furthermore, the earliest detection of Pax-6 expression in the chicken embryo is in a layer of cells which includes the presumptive lens.[100] Later, *Pax-6* is expressed in both lens and in neural ectoderm, including the developing retina and diencephalon.[100] Thus in vertebrates the lens appears to represent the minimal state of eye differen-

tiation. As other tissues lose expression of downstream components of the eye cascade they "revert" to the simpler level of the lens controlled by *Pax-6* and perhaps a few other high-level factors. Lens may thus be the minimal state of differentiation under control of *Pax-6* in vertebrates. It would be interesting to see whether pancreas, another site of *Pax-6* expression, could also transdifferentiate into "lens". The idea of lens being in some way the most fundamental outcome of *Pax-6* expression appears to be extremely paradoxical since the lens is by no means the fundamental tissue of the eye in any evolutionary or developmental sense. Yet, as will be described below, it turns out that expression of some crystallins, lens-specific or lens-preferred proteins, depends on binding of Pax-6 to the promoters or enhancers of their genes.

Crystallins would seem to be the final product of the cascade of gene expression necessary to form the lens. Similarly the vertebrate lens is probably the most recently evolved of eye structures. Why then are at least some crystallin genes under the control of the highest level eye "master gene"? The simplest explanation for the *Pax-6* paradox seems to be that the evolution of the lens necessitated use of transcription factors already expressed in the eye. As such Pax-6 fits the bill as much as any other factor in the cascade. At an earlier stage in evolution Pax-6 may have been principally involved in controlling other regulatory genes, such as those for other transcription factors, and may not have had a direct role in expression of eye-specific genes such as those encoding opsins. However, when the lens evolved, Pax-6 was co-opted or recruited to a new role in direct control of structural gene expression in the eye lens. Thus the gene recruitment of crystallins, described below, depended on the acquisition of Pax-6 binding sites and the consequent recruitment of *Pax-6* itself.

EVOLUTION DYNAMISM AND THE VERTEBRATE LENS: ENVIRONMENTAL PRESSURES

The eye provides a direct interface between the outside world and the internal world of perception and response. As an optical system, the properties of the eye are severely constrained by aspects of the external environment and the way of life of the organism. Thus the structure and light sensitivity of the retina differ according to whether the animal is diurnal, in which case it makes use of low-sensitivity, color-discriminating cone cell photoreceptors often with associated colored oil-drops,[20] or nocturnal, in which case it makes use of monochromatic vision through highly sensitive rod cell photoreceptors.[20] Similarly the refractive properties of the cornea and lens adapt to suit the needs of a fish, which requires a high refractive index to focus under water and has little use for vision at a distance, or of a hawk which needs a lower refractive index, accommodating lens to focus both at great distances in the air or close up in the nest.[20] The properties of

the cornea and lens may also evolve to filter harmful or dazzling radiation as species move from dim to bright light environments.

During the course of evolution vertebrate species have moved from water to land or from activity by day to activity by night. Perhaps more than any other organ this has placed unusual requirements on the eye to adapt and readapt its properties at both anatomical and molecular levels. Subsequent chapters will concentrate on describing the ubiquitous and taxon-specific crystallins, their structures and functions and the molecular biology of their gene recruitment and expression in the lens.

References

1. Scheer H, ed. Chlorophylls. Boca Raton: CRC Press, 1993.
2. Mukohata Y. Comparative studies on ion pumps of the bacterial rhodopsin family. Biophys Chem 1994; 50:191-201.
3. Krebs MP, Khorana HG. Mechanism of light-dependent proton translocation by bacteriorhodopsin. J Bacteriol 1993; 175:1555-60.
4. Chou KC. Conformational change during photocycle of bacteriorhodopsin and its proton-pumping mechanism. J Protein Chem 1993; 12:337-50.
5. Spudich JL. Protein-protein interaction converts a proton pump into a sensory receptor. Cell 1994; 79:747-50.
6. Spudich JL. Color sensing in the Archaea: a eukaryotic-like receptor coupled to a prokaryotic transducer. J Bacteriol 1993;175:7755-61.
7. Foster KW, Saranak J, Patel N, et al. A rhodopsin is the functional photoreceptor for phototaxis in the unicellular eukaryote *Chlamydomonas*. Nature 1984; 311:756-9.
8. James TW, Crescitelli F, Loew ER, McFarland WN. The eyespot of *Euglena gracilis*: a microspectrophotometric study. Vision Res 1992; 32:1583-91.
9. Nathans J. Rhodopsin: structure, function, and genetics. Biochemistry 1992; 31:4923-31.
10. Applebury ML. Relationships of G-protein-coupled receptors. A survey with the photoreceptor opsin subfamily. Soc Gen Physiol Ser 1994; 49:235-48.
11. Hurley JB. Transduction mechanisms of vertebrate and invertebrate photoreceptors. J Biol Chem 1994; 269:14329-32.
12. Margulis L. Symbiosis in Cell Evolution: microbial communities in the Archean and Proterozoic eons. 2nd ed. New York: Freeman, 1993.
13. Soppa J. Two hypotheses--one answer. Sequence comparison does not support an evolutionary link between halobacterial retinal proteins including bacteriorhodopsin and eukaryotic G-protein-coupled receptors. FEBS Lett 1994; 342:7-11.
14. Flaherty KM, McKay DB, Kabsch W, Holmes KC. Similarity of the three-dimensional structures of actin and the ATPase fragment of a 70-kDa heat shock cognate protein. Proc Natl Acad Sci USA 1991; 88:5041-5.
15. Strader CD, Fong TM, Tota MR, Underwood D, Dixon RA. Structure and function of G protein-coupled receptors. Annu Rev Biochem 1994; 63:101-32.

16. Larhammar D, Blomqvist AG, Wahlestedt C. The receptor revolution - multiplicity of G-protein-coupled receptors. Drug Des Discov 1993; 9:179-88.
17. Land MF, Fernald RD. The evolution of eyes. Annu Rev Neurosci 1992; 15:1-29.
18. Burr AH. Evolution of eyes and photoreceptor organelles in the lower phyla. In: Ali MA, ed. Photoreception and Vision in Invertebrates. New York: Plenum Press, 1984; 131-78. (NATO ASI Series A; 74).
19. Ali MA. Epilogue. In: Ali MA, ed. Photoreception and Vision in Invertebrates. New York: Plenum Press, 1984; 773-88. (NATO ASI Series A: Life Sciences; 74).
20. Walls GL. The Vertebrate Eye and Its Adaptive Radiation. facsimile of 1942 edition. New York, NY: Hafner, 1967.
21. Ali MA, ed. Photoreception and Vision in Invertebrates. New York: Plenum Press, 1984 (NATO ASI Series A: Life Sciences; 74).
22. Fernald RD, Wright SE. Maintenance of optical quality during crystalline lens growth. Nature 1983; 301:618-20.
23. Kroger RH, Campbell MC, Munger R, Fernald RD. Refractive index distribution and spherical aberration in the crystalline lens of the African cichlid fish *Haplochromis burtoni*. Vision Res 1994; 34:1815-22.
24. Kroger RH, Fernald RD. Regulation of eye growth in the African cichlid fish *Haplochromis burtoni*. Vision Res 1994; 34:1807-14.
25. Davson H. Physiology of the Eye. 5th ed. London: The Macmillan Press Ltd, 1990.
26. Berman ER. Biochemistry of the eye. 1st ed. New York: Plenum Press, 1991 (Blakemore C, ed. Perspectives in Vision Research).
27. Eakin RM. The Third Eye. Berkeley and Los Angeles: University of California Press, 1973.
28. Wiechmann AF. Melatonin: parallels in pineal gland and retina. Exp Eye Res 1986; 42:507-27.
29. McAvoy JW. Cell division, cell elongation and the co-ordination of crystallin expression during lens morphogenesis in the rat. J Embryol Exp Morphol 1978; 45:271-81.
30. Davson H. The Lens. In: Physiology of the Eye. 5th ed. London: Macmillan Press, 1990; 139-201.
31. Piatigorsky J. Lens differentiation in vertebrates. A review of cellular and molecular features. Differentiation 1981; 19:134-53.
32. Piatigorsky J, Horwitz J, Kuwabara T, Cutress CE. The cellular eye lens and crystallins of cubomedusan jellyfish. J Comp Physiol [A] 1989; 164:577-87.
33. Tardieu A, Veretout F, Krop B, Slingsby C. Protein interactions in the calf eye lens: interactions between β-crystallins are repulsive whereas in γ-crystallins they are attractive. Eur Biophys J 1992; 21:1-12.
34. Ott M, Schaeffel F. A negatively powered lens in the chameleon. Nature 1995; 373:692-4.
35. Beebe DC, Compart PJ, Johnson MC, Feagans DE, Feinberg RN. The mechanism of cell elongation during lens fiber cell differentiation. Dev Biol 1982; 92:54-9.

36. Beebe DC, Parmelee JT, Belcher KS. Volume regulation in lens epithelial cells and differentiating lens fiber cells. J Cell Physiol 1990; 143:455-9.
37. Wistow G. Lens crystallins: gene recruitment and evolutionary dynamism. Trends Biochem Sci 1993; 18:301-6.
38. Grainger RM. Embryonic lens induction: shedding light on vertebrate tissue determination. Trends Genet 1992; 8:349-55.
39. Zwaan J. The appearance of α-crystallin in relation to cell cycle phase in the embryonic mouse lens. Dev Biol 1983; 96:173-81.
40. Oguni M, Setogawa T, Hashimoto R, Tanaka O, Shinohara H, Kato K. Ontogeny of α-crystallin subunits in the lens of human and rat embryos. Cell Tissue Res 1994; 276:151-4.
41. Alcala J, Maisel H. Biochemistry of lens plasma membranes and cytoskeleton. In: Maisel H, ed. The Ocular Lens. New York: Marcel Dekker, Inc., 1985; 169-222.
42. Chepelinsky AB. The MIP Transmembrane Channel Gene Family. In: Peracchia C, ed. Handbook of Membrane Channels. San Diego: Academic Press, Inc., 1994; 413-32.
43. Chrispeels MJ, Agre P. Aquaporins: water channel proteins of plant and animal cells. Trends Biochem Sci 1994; 19:421-5.
44. Bassnett S, Beebe DC. Coincident loss of mitochondria and nuclei during lens fiber cell differentiation. Dev Dyn 1992; 194:85-93.
45. Appleby DW, Modak SP. DNA degradation in terminally differentiating lens fiber cells from chick embryos. Proc Natl Acad Sci USA 1977; 74:5579-83.
46. Gao CY, Bassnett S, Zelenka PS. Cyclin B, p34cdc2 and H1-kinase activity in terminally differentiating lens fiber cells. Dev Biol 1995; 169:185-94.
47. Harding JJ, Crabbe MJC. The lens: Development, proteins, metabolism and cataract. In: Davson H, ed. The Eye. v. 1B. New York: Academic Press, 1984; 207-492.
48. Kaur S, Key B, Stock J, McNeish JD, Akeson R, Potter SS. Targeted ablation of α-crystallin-synthesizing cells produces lens-deficient eyes in transgenic mice. Development 1989; 105:613-9.
49. Landel CP, Zhao J, Bok D, Evans GA. Lens-specific expression of recombinant ricin induces developmental defects in the eyes of transgenic mice. Genes Dev 1988; 2:1168-78.
50. Harrington L, Klintworth GK, Secor TE, Breitman ML. Developmental analysis of ocular morphogenesis in αA-crystallin/diphtheria toxin transgenic mice undergoing ablation of the lens. Dev Biol 1991; 148:508-16.
51. Chamberlain CG, McAvoy JW. Evidence that fibroblast growth factor promotes lens fibre differentiation. Curr Eye Res 1987; 6:1165-8.
52. Chamberlain CG, McAvoy JW. Induction of lens fibre differentiation by acidic and basic fibroblast growth factor (FGF). Growth Factors 1989; 1:125-34.
53. Peek R, McAvoy JW, Lubsen NH, Schoenmakers JG. Rise and fall of crystallin gene messenger levels during fibroblast growth factor induced terminal differentiation of lens cells. Dev Biol 1992; 152:152-60.

54. Coulombre J, Coulombre A. Lens development: fiber cell elongation and lens orientation. Science 1963; 142:1489-94.
55. Robinson ML, Overbeek PA, Verran DJ et al. Extracellular FGF-1 acts as a lens differentiation factor in transgenic mice. Development 1995; 121:505-14.
56. Beebe DC, Silver MH, Belcher KS, Van Wyk JJ, Svoboda ME, Zelenka PS. Lentropin, a protein that controls lens fiber formation, is related functionally and immunologically to the insulin-like growth factors. Proc Natl Acad Sci USA 1987; 84:2327-30.
57. Brewitt B, Clark JI. Growth and transparency in the lens, an epithelial tissue, stimulated by pulses of PDGF. Science 1988; 242:777-9.
58. Pelton RW, Saxena B, Jones M, Moses HL, Gold LI. Immunohistochemical localization of TGFβ1, TGFβ2, and TGFβ3 in the mouse embryo: expression patterns suggest multiple roles during embryonic development. J Cell Biol 1991; 115:1091-105.
59. Wistow GJ, Shaughnessy MP, Lee DC, Hodin J, Zelenka PS. A macrophage migration inhibitory factor is expressed in the differentiating cells of the eye lens. Proc Natl Acad Sci USA 1993; 90:1272-5.
60. Lanahan A, Williams JB, Sanders LK, Nathans D. Growth factor-induced delayed early response genes. Mol Cell Biol 1992; 12:3919-29.
61. Paralkar V, Wistow G. MIF: An essential intermediate for growth factor induced mitogenesis. FASEB J 1995; 9:A1332.
62. Darwin CR. In: Darwin F, ed. The Autobiography of Charles Darwin and Selected Letters. London: Murray, 1860.
63. Nilsson DE, Pelger S. A pessimistic estimate of the time required for an eye to evolve. Proc R Soc Lond B Biol Sci 1994; 256:53-8.
64. Dawkins R. Evolutionary biology. The eye in a twinkling. Nature 1994; 368:690-1.
65. McDevitt DS, Brahma SK. Ontogeny and Localization of the Crystallins in Eye Lens Development and Regeneration. In: McDevitt DS, ed. Cell Biology of the Eye. New York: Academic Press, Inc., 1982; 143-91.
66. West JA, Sivak JG, Pasternak J, Piatigorsky J. Immunolocalization of S-crystallins in the developing squid (*Loligo opalescens*) lens. Dev Dyn 1994; 199:85-92.
67. 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.
68. Wistow G, Piatigorsky J. Lens crystallins: evolution and expression of proteins for a highly specialized tissue. Ann Rev Biochem 1988; 57:479-504.
69. Benedek GB. Theory of transparency of the eye. Appl Optics 1971; 10:459-73.
70. Delaye M, Tardieu A. Short-range order of crystallin proteins accounts for eye lens transparency. Nature 1983; 302:415-7.
71. Siezen RJ, Fisch MR, Slingsby C, Benedek GB. Opacification of γ-crystallin solutions from calf lens in relation to cold cataract formation. Proc Natl Acad Sci USA 1985; 82:1701-5.

72. Bloemendal H. The Lens Proteins. In: Bloemendal H, ed. Molecular and Cellular Biology of the Eye Lens. New York: Wiley-Interscience, 1981; 1-47.
73. Klemenz R, Frohli E, Steiger RH, Schafer R, Aoyama A. αB-crystallin is a small heat shock protein. Proc Natl Acad Sci USA 1991; 88:3652-6.
74. Dasgupta S, Hohman TC, Carper D. Hypertonic stress induces αB-crystallin expression. Exp Eye Res 1992; 54:461-70.
75. Wistow G, Graham C. The duck gene for αB-crystallin shows evolutionary conservation of discrete promoter elements but lacks heat and osmotic stress response. Biochim Biophys Acta 1995: in press.
76. de Jong WW, Leunissen JAM, Voorter CEM. Evolution of the α-crystallin/small heat-shock protein family. Mol Biol Evol 1993; 10:103-26.
77. Sax C, Piatigorsky J. Expression of the α-crystallin/small heat shock protein/molecular chaperone genes in the lens and other tissues. In: Advances in Enzymology and Related Areas in Molecular Biology. v. 69. New York, NY: John Wiley & Sons Inc, 1994; 155-201.
78. de Jong WW, Hendriks W, Mulders JW, Bloemendal H. Evolution of eye lens crystallins: the stress connection. Trends Biochem Sci 1989; 14:365-8.
79. 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.
80. 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.
81. Rao CM, Zigler JS Jr. Levels of reduced pyridine nucleotides and lens photodamage. Photochem Photobiol 1992; 56:523-8.
82. Ton CC, Hirvonen H, Miwa H et al. Positional cloning and characterization of a paired box- and homeobox-containing gene from the aniridia region. Cell 1991; 67:1059-74.
83. Glaser T, Walton DS, Maas RL. Genomic structure, evolutionary conservation and aniridia mutations in the human PAX6 gene. Nat Genet 1992; 2:232-9.
84. Hill RE, Favor J, Hogan BL et al. Mouse small eye results from mutations in a paired-like homeobox-containing gene. Nature 1991; 354:522-5.
85. Quiring R, Walldorf U, Kloter U, Gehring WJ. Homology of the eyeless gene of *Drosophila* to the *Small eye* gene in mice and Aniridia in humans. Science 1994; 265:785-9.
86. Hill RE, Hanson IM. Molecular genetics of the Pax gene family. Curr Opinion Cell Biol 1992; 4:967-72.
87. Stoykova A, Gruss P. Roles of Pax-genes in developing and adult brain as suggested by expression patterns. J Neurosci 1994; 14:1395-412.
88. Walther C, Guenet JL, Simon D, et al. Pax: a murine multigene family of paired box-containing genes. Genomics 1991; 11:424-34.
89. Noll M. Evolution and role of Pax genes. Curr Opin Genet Dev 1993; 3:595-605.

90. Halder G, Callaerts P, Gehring WJ. Induction of ectopic eyes by targeted expression of the *eyeless* gene in *Drosophila*. Science 1995; 267:1788-92.
91. Turque N, Plaza S, Radvanyi F, Carriere C, Saule S. Pax-QNR/Pax-6, a paired box- and homeobox-containing gene expressed in neurons, is also expressed in pancreatic endocrine cells. Mol Endocrinol 1994; 8:929-38.
92. Zhang Y, Emmons SW. mab-18 contains a homeo domain similar to vertebrate pax and *Drosophila* paired box containing genes. Worm Breeder's Gazette 1993; 13:54.
93. Chisholm A, Horvitz B. The worm without a face: an update on vab-3. Worm Breeder's Gazette 1993; 13:48.
94. Burr AHJ, Babinszki CPF. Scanning motion, ocellar morphology and orientation mechanisms in the phototaxis of the nematode *Mermis nigrescens*. J Comp Physiol [A] 1990; 167:257-68.
95. Cheyette BN, Green PJ, Martin K, Garren H, Hartenstein V, Zipursky SL. The *Drosophila sine oculis* locus encodes a homeodomain-containing protein required for the development of the entire visual system. Neuron 1994; 12:977-96.
96. Beebe DC. Homeobox Genes and Vertebrate Eye Development. Invest Ophthalmol Vis Sci 1994; 35:2897-900.
97. Fortini ME, Rebay I, Caron LA, Artavanis-Tsakonas S. An activated Notch receptor blocks cell-fate commitment in the developing *Drosophila* eye. Nature 1993; 365:555-7.
98. Monaghan AP, Davidson DR, Sime C et al. The Msh-like homeobox genes define domains in the developing vertebrate eye. Development 1991; 112:1053-61.
99. Levine EM, Schechter N. Homeobox genes are expressed in the retina and brain of adult goldfish. Proc Natl Acad Sci USA 1993; 90:2729-33.
100. Li HS, Yang JM, Jacobson RD, Pasko D, Sundin O. Pax-6 is first expressed in a region of ectoderm anterior to the early neural plate: implications for stepwise determination of the lens. Dev Biol 1994; 162:181-94.
101. Polyak S. The Vertebrate Visual System. Chicago: University of Chicago Press, 1957.
102. McDevitt DS. Presence of lateral eye lens crystallins in the median eye of the American chameleon. Science 1972; 175:763-4.
103. Max M, McKinnon PJ, Seidenman KJ et al. Pineal opsin: a nonvisual opsin expressed in chick pineal. Science 1995; 267:1502-6.
104. Fedtsova NG. Induction of α- and β-crystallin synthesis in organ culture of the adenohypophyseal anlage of chickens is affected by 5-iododeoxyuridine and 5-bromodeoxyuridine. Ontogenez 1986; 17:396-401.
105. Fedtsova NG, Minina TA, Barabanov VM. Synthesis of a lens-specific antigen δ-crystallin in rudimentary chicken adenohypophysis. Biull Eksp Biol Med 1981; 92:314-6.
106. Yamada T. Transdifferentiation of Lens Cells and its Regulation. In: McDevitt DS, ed. Cell Biology of the Eye. New York: Academic Press, Inc., 1982; 193-242.
107. Eguchi G, Kodama R. Transdifferentiation. Curr Opin Cell Biol 1993; 5:1023-8.

108. Agata K, Kobayashi H, Itoh Y, Mochii M, Sawada K, Eguchi G. Genetic characterization of the multipotent dedifferentiated state of pigmented epithelial cells in vitro. Development 1993; 118:1025-30.
109. Zuker CS. On the evolution of eyes: would you like it simple or compound? Science 1994; 265:742-3.
110. Hargrave PA, McDowell JH, Feldmann RJ, Atkinson PH, Rao JK, Argos P. Rhodopsin's protein and carbohydrate structure: selected aspects. Vision Res 1984; 24:1487-99.
111. Crescitelli F, James TW, Erickson JM, Loew ER, McFarland WN. The eyespot of *Chlamydomonas reinhardtii*: a comparative microspectrophotometric study. Vision Res 1992; 32:1593-600.
112. Morgenbesser SD, Williams BO, Jacks T, DePinho RA. p53-dependent apoptosis produced by Rb-deficiency in the developing mouse lens. Nature 1994; 371:72-4.
113. Jaworski CJ, Wistow GJ. LP2; a member of the lipid/retinoid binding protein superfamily in bovine lens. Invest Ophthalmol Vis Sci 1994; 35:1706.

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

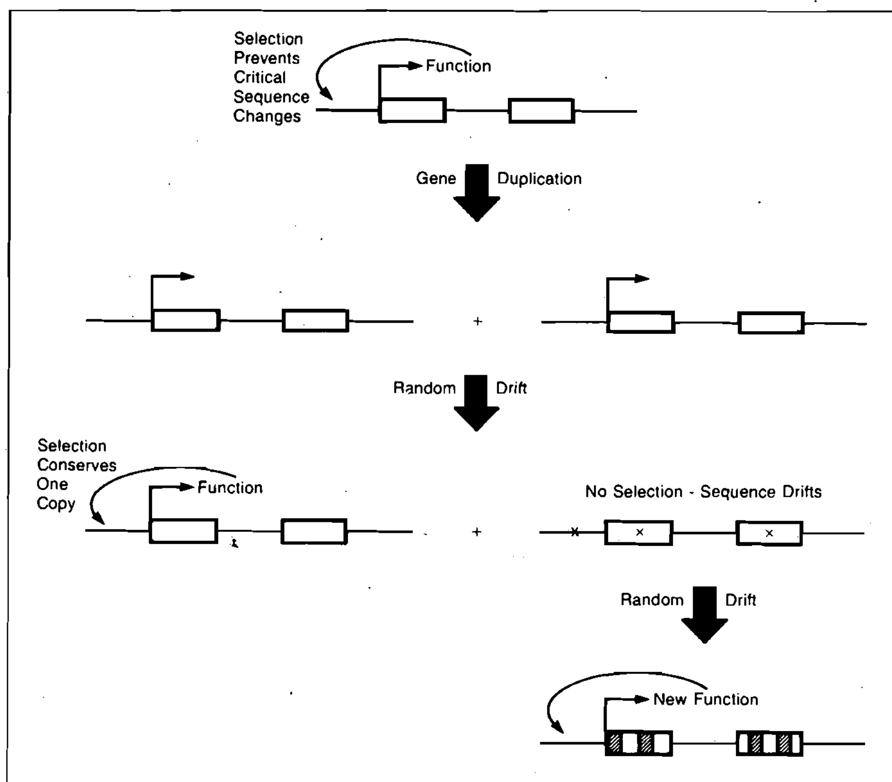


Fig. 2.1. The "classical" model for new protein functionalities and for gene duplication.

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.

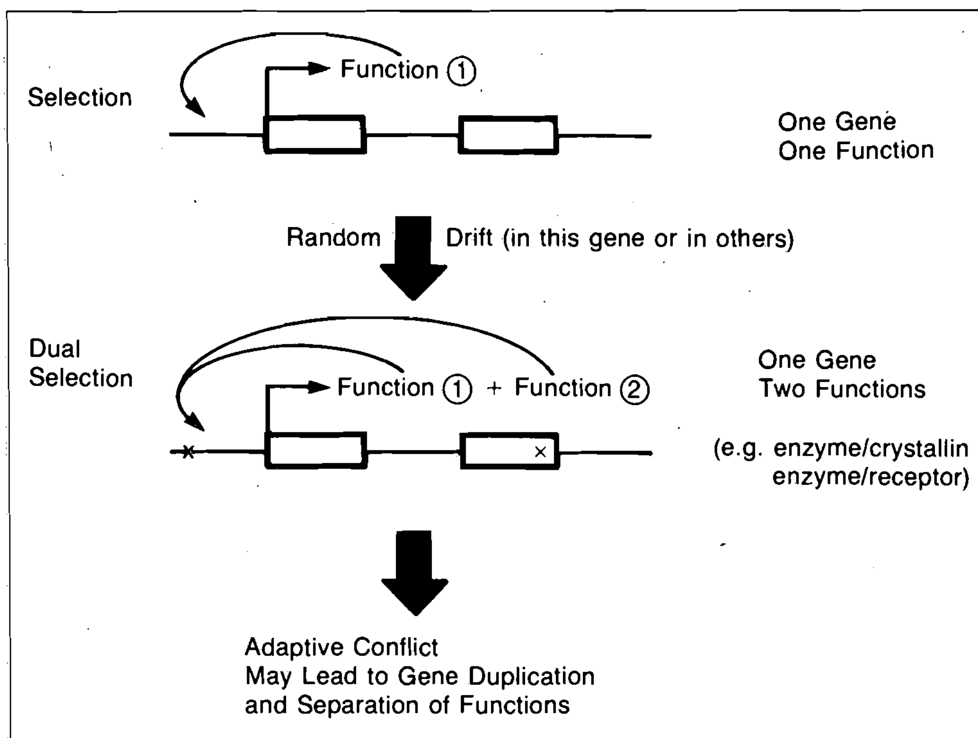


Fig. 2.2. The Gene Recruitment model for new protein functionalities and for gene duplication.

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.

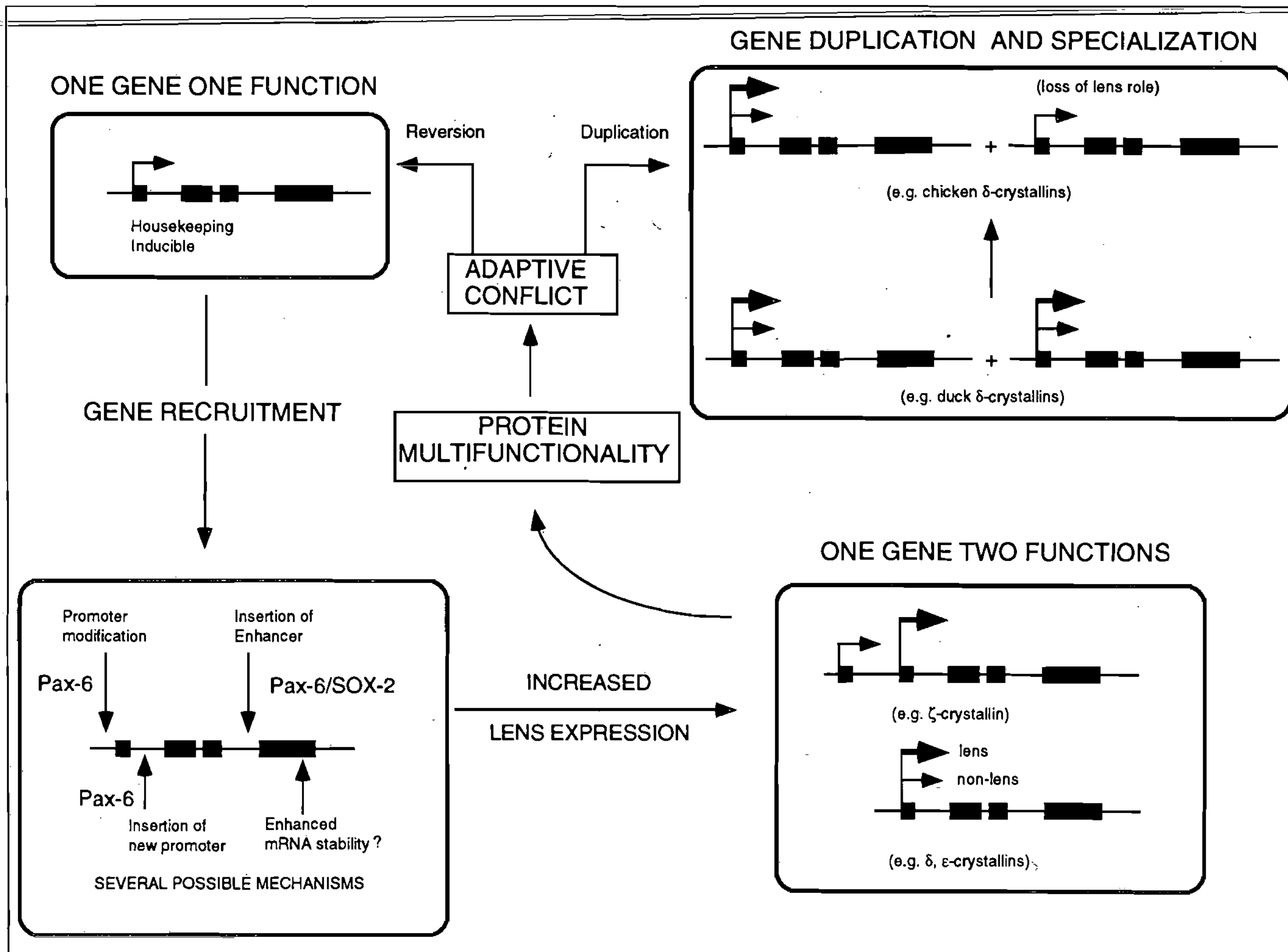
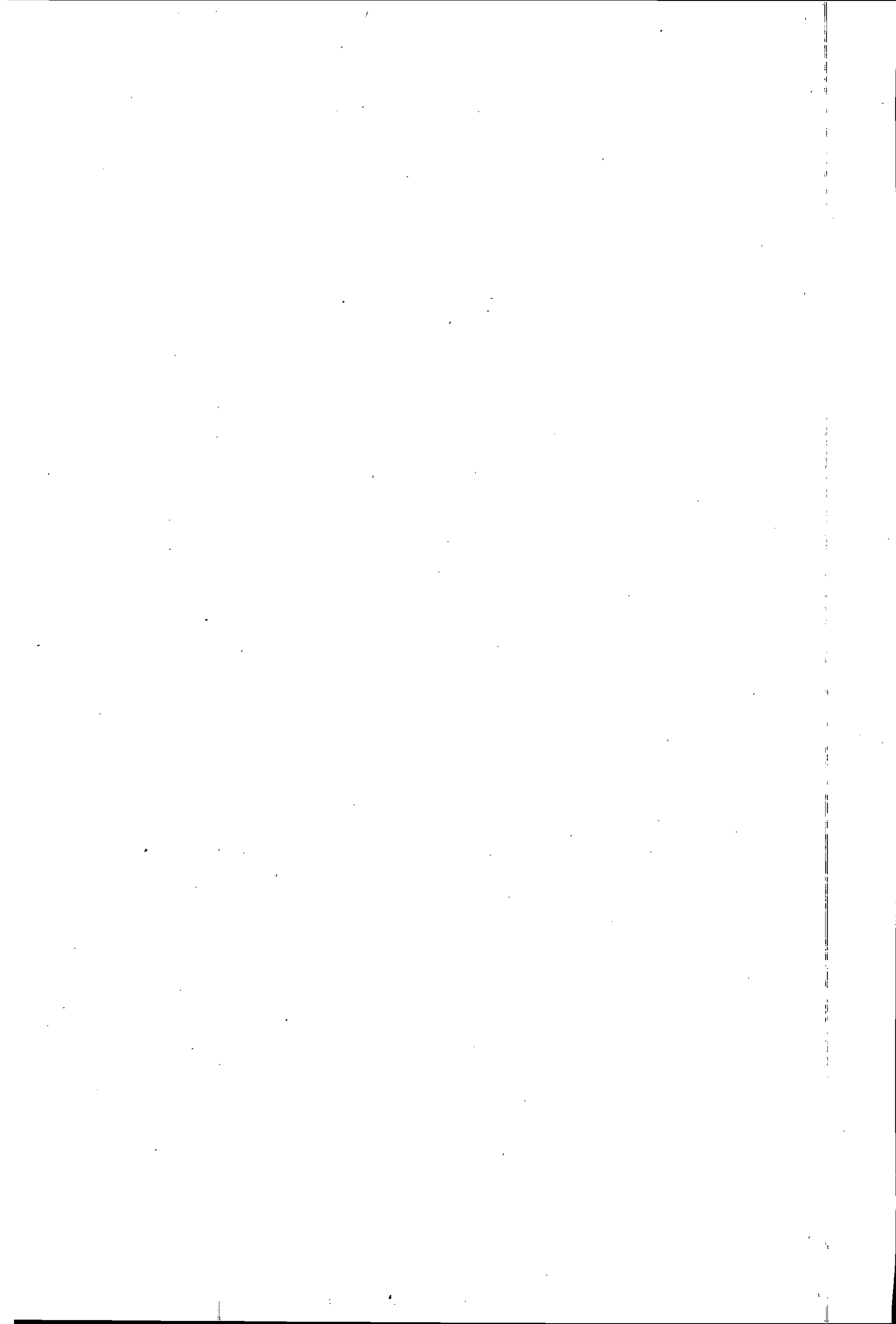


Fig. 2.3. Gene recruitment, protein multifunctionality and gene duplication in the crystallins.

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.



CHAPTER 3

THE UBIQUITOUS CRYSTALLINS: STRESS PROTEINS RECRUITED TO LENS

Three classes of crystallins, belonging to two protein superfamilies, are represented in the lenses of all vertebrates. These are the α, β and γ-crystallins. The ubiquity of these crystallins suggests that their recruitment occurred at a very early stage in the evolution of the lens. Furthermore all three classes of ubiquitous crystallins have at least some connection with various stress responses.

α-CRYSTALLINS: MEMBERS OF THE SMALL HEAT SHOCK PROTEIN SUPERFAMILY

α-Crystallins are present in the lenses of all vertebrates; indeed in many mammals they may be the major protein components of the lens.[1-3] Species as divergent as dogfish (*Squalus acanthias*) and mammals express the same two α-crystallin subunits, αA and αB.[4] αA-crystallin is usually the most abundant subunit and indeed may be the single most abundant gene product in the lens, although the αA/αB ratio varies considerably among species and αB seems to predominate in the dogfish lens.[1,2,4] Both subunits are about 20 kDa in size and have considerable sequence similarity; in chicken, αA[5] and αB[6] are 59% identical. This degree of conservation is reflected in the slow rate of sequence change in αA-crystallins[7] which has made them useful tools for molecular phylogenetic studies. αA-crystallin sequences form one of the most comprehensive databases for this purpose.[8-12]

Both α-crystallin gene products undergo a remarkable array of post-translational modifications including peptide bond cleavage,[1,2,13] cAMP dependent phosphorylation,[14-18] autophosphorylation,[19] deamidation,[20,21] transglutamination,[22] fatty acylation,[23] racemization[24] and cytoplasmic glycosylation.[25] The functional and structural consequences of these modifications are not clear although it seems likely that they could

have roles in modifying interactions of α-crystallins which each other and with other lens components.

Tertiary and Quaternary Structure

In spite of the growing interest in this family little is known of the tertiary structure of the protein subunits or their quaternary interactions. CD (circular dichroism) and ORD (optical rotary dispersion) spectroscopy both suggest that α-crystallin subunits have predominantly β-sheet structure with only a small content of α-helix.[26,27] ^{1}H NMR analysis has found that the C-terminal regions of bovine lens α-crystallin subunits are highly flexible and disordered.[28] However an x-ray structure determination is urgently needed.

One of the most notable features of α-crystallins is their propensity for aggregation. Large and variable aggregates form from populations of one subunit or from mixtures of both.[29,30] A number of different models for the quaternary structures for α-crystallins have been proposed in attempts to reconcile data from different experimental systems. Based on a variety of experiments which suggest that α-crystallin subunits form distinct populations with regard to accessibility, three-layer models of α-crystallins have been proposed.[31-33] From a structural point of view these models have the unsatisfactory feature of requiring identical subunits to occupy non-equivalent positions. In a different model the difference in hydrophobicity between the N-terminal and C-terminal regions of α-crystallins inspired the suggestion that α-crystallin subunits form protein micelles.[34]

Other experiments on the aggregation behavior of recombinant α-crystallin domains[35] led to the proposal that α-crystallin subunits assemble as tetramers or pseudo-tetramers on each face of either a rhombic dodecahedron or a cube,[36] a model in which all subunits occupy equivalent positions. Finally another model has been proposed based on the structure of molecular chaperones.[37] It seems very likely that α-crystallins have multiple modes of assembly depending on conditions, with dynamic interchange of subunits and easy transitions between alternative quaternary structures.[36,38,39]

Ontogeny

α-Crystallins are among the earliest crystallins to be expressed during vertebrate lens embryogenesis and continue to be expressed at high levels throughout life.[1,40-42] However the patterns of expression of the two genes appear to vary among species. For example, in human lens development αB-crystallin has been detected in the lens placode (see Fig. 1.2) while αA-crystallin was not observed until the lens vesicle stage.[42] In contrast, in the developing rat lens αA-crystallin appears first in the lens pit at embryonic stage E12,[42] similar to the timing of α-crystallin in mouse,[41] while αB-crystallin was not detected until stage E14 in elongating lens fibers.[42] The same authors also found differ-

ences in the distribution of the two α-crystallins in human and rat lenses such that αB-crystallin was found in both epithelial and fiber cells in human lens but was found only in fiber cells in rat lens. Interestingly, the work on developing mouse lens suggested that there was an association between the onset of crystallin synthesis and cell elongation.[41]

THE SMALL HEAT SHOCK PROTEIN CONNECTION

Although it was appreciated that crystallins must have arisen from molecular ancestors which predated the existence of the vertebrate lens,[2] it was also assumed that α- and other crystallins would be lens-specific structural proteins derived by gene duplication and specialization from those non-lens ancestors. The first indication of a non-lens relationship for crystallins came when the sequences of *Drosophila* small heat-shock proteins (sHSP) revealed surprisingly close similarities with αB-crystallin.[43]

The sHSP superfamily is ubiquitous in eukaryotes.[44-50] Distinct members of this superfamily are also found as egg antigens in *Schistosoma mansoni*[51,52] and muscle-associated proteins in the ascidian *Halocynthia roretzi*.[53] The superfamily is even represented in prokaryotes as surface antigens in *Mycobacterium leprae*[54] and *M.tuberculosis*,[55] as a ribosomal associated protein in *M.bovis*[56] and as stress induced proteins in *Escherichia coli*,[57] *Stigmatella aurantiaca*[58] and *Clostridium acetobutylicum*.[59] Recently a new vertebrate member has been added to this family, p20 a bovine protein with closer sequence similarity to α-crystallins than to any mammalian sHSP.[60] The large scale sequencing project for the nematode *Caenorhabditis elegans* has also uncovered a gene with closer similarity to the α-crystallin family than to other sHSP.[61]

The significance of the similarity between α-crystallins and the sHSP family became clearer when, following the discovery of non-lens expression of taxon-specific enzyme crystallins,[62] αB-crystallin too was detected outside the lens during difference library screening of hamster brains affected with the prion disease scrapie.[63] Indeed, this theme of elevated expression of αB-crystallin associated with neurological disorders has continued with the detection of high concentrations of αB-crystallin in Alexander's disease,[64] Lewy body disease[65] and the human prion associated Creutzfeldt-Jakob disease[66] among others.[3] Lower constitutive levels of αB-crystallin were also found in many adult tissues in both mammals and birds.[3,67-70]

The connection with the sHSP family was strengthened when it was discovered that mouse αB-crystallin was induced by heat-shock in mouse embryonic fibroblasts.[71] Thus in this species αB-crystallin itself is a sHSP. Like other heat shock proteins, αB-crystallin can also be induced by other insults such as osmotic stress,[72,73] ischemia[74] and by expression of some oncogenes.[75] However this stress response may be specific to mammals. In duck embryonic fibroblasts the endogenous

αB-crystallin gene is neither constitutively expressed nor induced by either heat or osmotic stress under conditions which elicit a stress response and HSP70 induction.[73] This correlates with an absence of canonical heat shock response elements in the promoter of the duck αB-crystallin gene.[73] It seems likely that both heat and osmotic stress responses in mammalian αB-crystallin genes are mediated through these elements and that αB-crystallin in the duck has lost this kind of inducibility. If the stress response is an ancestral feature of αB-crystallin predating its role in lens then it appears that duck αB-crystallin (and presumably its homologues in other birds) has become more specialized for lens expression and has given up a large part of the ancestral non-lens role.

There is no evidence for inducibility of αA-crystallin, but very low levels of the protein have been detected in non-lens tissues, particularly in spleen and thymus.[76] In general, however, it seems that αA-crystallin is much more specialized for lens-expression than αB.

Molecular Chaperones of the Cytoskeleton

The functions of heat shock proteins are beginning to become clear. The large heat shock proteins of the HSP70 family are ATPases with structural similarity to actin.[77] These proteins are ubiquitous and highly conserved from bacteria to vertebrates. They and other families of heat shock proteins such as the HSP60 class act as molecular chaperones guiding protein folding pathways and the transport and assembly of multisubunit complexes.[78] They also serve to prevent inappropriate interactions between partially assembled proteins.[78] Recent results suggest that α-crystallins and their sHSP relatives may lack a direct role in protein folding but may share some of the other aspects of molecular chaperones.

In contrast to HSP70, α-crystallins lack an identifiable ATPase domain, although there is some evidence for an association between α-crystallins and ATP.[19,79] However there is clear evidence that α-crystallins can prevent the insolubilization of proteins denatured by heat stress.[80] Unlike the HSP70 family, α-crystallins do not appear to have a direct role in protein folding or refolding since they are unable to protect or restore the enzyme activity of heat stressed enzymes.[3,80] This chaperone-like activity may be a reflection of a physiological role of α-crystallins in what might be described as an auxiliary chaperone role. α-Crystallins might participate in the assembly and disassembly of complex structures without a direct role in protein folding. This could be regarded as catalyzing a process which would otherwise have an unacceptably high activation energy due to the formation of incorrect interactions between constituent subunits of an unassembled structure.

A likely beneficiary of this chaperone effect of sHSP and α-crystallins may be the cytoskeleton. There is considerable evidence linking sHSP

and α-crystallins with management and protection of cytoskeleton in lens and in other cell types. For example, an actin depolymerizing protein of turkey gizzard was found to be a sHSP[81] while αB-crystallin itself can interact with actin,[74,82,83] desmin[82] and vimentin.[84] Chicken αB-crystallin has been associated with extensive cytomorphological remodeling in lens, notochord and myotome during embryogenesis.[85] In lens, α-crystallins participate in the ATP-dependent assembly and disassembly of vimentin and GFAP components of cytoskeleton during lens cell differentiation.[84] This role as a chaperone for cytoskeleton may be the basis for the increased thermotolerance conferred on cultured cells by overexpression of αB-crystallin.[86]

The Recruitment of a sHSP as a Crystallin

It seems likely that a sHSP was recruited as a crystallin in the primitive lens of an ancestor of vertebrates. This sHSP may have been associated with cytoskeleton in the elongating cells of the proto-lens. Shortly after recruitment this gene duplicated to produce the ancestors of the modern αA- and αB-crystallin genes. One gene was able to retain a general stress response role in many tissues in addition to its role as a crystallin. This gene became the αB-crystallin seen today. In contrast, the other gene specialized for the lens environment in both protein sequence and gene expression and substantially lost any non-lens role. This gene became the modern αA-crystallin. Indeed, although αA-crystallin and αB-crystallin are equivalent in many ways[3] there is evidence that αA-crystallin has some specialized properties which may have evolved specifically for its role in lens. In particular, some of the interactions between α-crystallins, lens cytoskeleton and membrane fractions seem to be specific to αA subunits.[87-91] This specialized interaction may be associated with the presence of the unusual "beaded filaments" found only in lens cytoskeleton.[92] It seems that the evolutionary specialization of α-crystallins for lens may be continuing since in birds the non-lens and stress role of αB-crystallin appear to have been lost.[73]

GENE STRUCTURE

Their ubiquity in modern species suggests that the genes for αA- and αB-crystallin must have arisen by duplication in a distant common ancestor of all vertebrates.[3] In man the two genes are found on different chromosomes, αA-crystallin (CRYA1) on chromosome 21[93-95] and αB-crystallin (CRYA2) on chromosome 11,[96] which is perhaps another reflection of the antiquity of their duplication. In spite of this ancient separation, αA- and αB-crystallins show a high degree of sequence similarity and identical gene structures. Genes for αA-crystallin have been cloned and at least partially sequenced from hamster,[97] mole rat,[98] mouse,[99,100] human[101,103] and chicken.[5] Genes for αB-crystallins have been cloned from hamster,[93] mouse,[67,104] rat,[105] human[106] and duck.[73] All these genes have the same general structure of three exons, with

the interesting exception that mammalian αA-crystallin genes also contain an alternatively spliced insert exon or pseudoexon in the first intron[97-99,101] (Fig. 3.1).

The Insert Exon of Mammalian αA-Crystallins

In spite of the close similarity in exon-intron structure of α-crystallin genes in mammals and birds and the high degree of conservation of sequence in α-crystallins, mammalian αA-crystallin genes have an interesting peculiarity. This was first discovered when a minor form of αA-crystallin in rat lens was found to contain an insertion of 23 amino acids in an otherwise perfectly conserved sequence.[107-109] Gene sequencing showed this insertion was due to alternative mRNA splicing of an insert exon in the first intron of rodent αA-crystallin genes.[97-99,110] Perhaps because of a slightly sub-optimal splice site recognition sequence (GC for GT) this alternative or insert exon was found to be spliced into mRNA in only about 10% of mature processed transcripts.[99,110] The function of the insert peptide is still unknown and the αA^{ins} subunit behaves much like αA-crystallin in assembly studies.[30]

The αA^{ins} subunit was originally thought to be a curiosity of the murine rodents, however its distribution has turned out to be more extensive (Fig. 3.2). By western blotting it has been detected in lens extracts of several other rodents, the European hedgehog (*Erinaceus europaeus*), elephant shrew (*Elephantulus rufescens*), both mega- and microchiropteran bats and in a marsupial, the tamar wallaby (*Macropus eugenii*).[102,111,112] It was not detected in the lenses of many other mammals or in any non-mammals. At the level of the genome, the alternative insert exon has also been detected by hybridization in elephant shrews, rodents, bats, marsupials and even the spiny anteater (*Echidna sp.*), a monotreme.[102] Again, there was no hybridization of alternative exon probes with birds, reptiles or amphibians[102] while gene sequencing of chicken αA-crystallin confirms the lack of a recognizable insert exon.[5] αA^{ins} may thus constitute a synapomorphy which distinguishes all mammalian lineages from other vertebrates.

The αA^{ins} subunit is not found in human or other primate lenses. Nevertheless, when the human gene for αA-crystallin was cloned a sequence very similar to the insert exon was found in the first intron.[101] Although closely matching the insert exon of rodent αA-crystallin genes this part of the human gene contains a number of non-silent sequence changes and a deletion which would cause a shift of reading frame if translated. Thus the insert-exon sequences in the human αA-crystallin gene constitute a pseudoexon which, like a pseudogene unprotected by selection, is undergoing accelerated sequence drift. DNA hybridization studies suggest that similar pseudoexons are also likely to be present in chimpanzee, Old World monkey and rabbit which, like humans, do not express the αA^{ins} subunit while in ungulates it appears that any sequences related to an ancestral insert exon have been lost.[102]

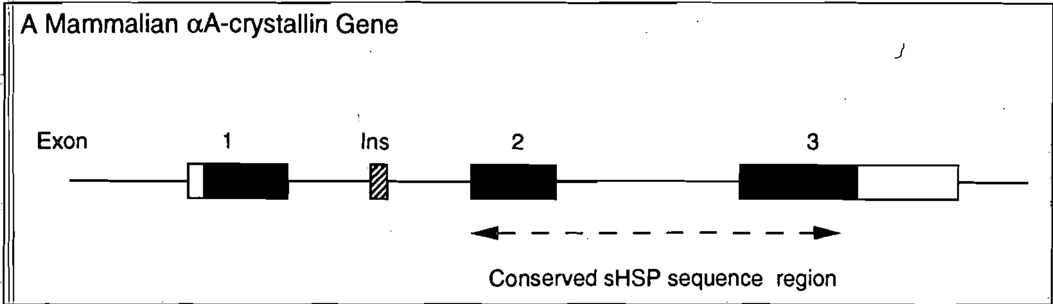


Fig. 3.1. Gene structure in the α-crystallins. The general structure of a mammalian αA-crystallin gene. Exons are shown as boxes, coding sequences are shaded. The alternative insert exon in intron 1 is peculiar to mammalian αA-crystallins. With this exception, αB-crystallin genes have the same general structure.

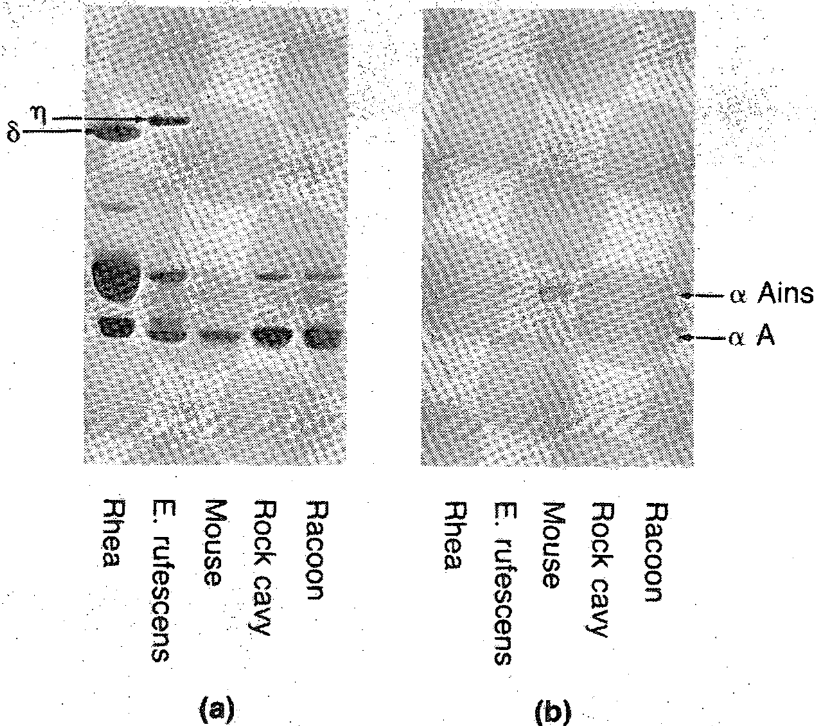


*Fig. 3.2. Detection of the αA^ins^ subunit in lenses of some mammals. Proteins were separated by SDS PAGE (left) and subjected to western blotting (right) with antisera to mouse αA-crystallin.[231] Lens extracts are from Rhea (*Rhea americana*) a bird; an elephant shrew (*Elephantulus rufescens*); mouse (*Mus musculus*); Rock cavy or moco (*Kerodon rupestris*); Raccoon (*Procyon lotor*). Migration positions of αA-crystallin and αA^{ins} subunits are marked. Two taxon specific crystallins, δ and η, are also marked.*

It thus appears that far from being a recent evolutionary experiment in murine rodents the insert exon and the encoded peptide have a much more ancient origin in mammalian evolution. Since its sequence is so well conserved it seems reasonable to suppose that αA^{ins} has a useful function when expressed. However, many modern placental species have abandoned its use altogether. It seems that this experiment in protein engineering answered a particular requirement of the lens in a common ancestor of monotreme, marsupial and placental mammals, a requirement which has been superseded by more recent evolutionary developments in many lineages. The story of αA^{ins} is another example of the dynamic molecular evolution of crystallins in the complex history of mammals.

Implications of Gene Structure for Protein Structure

The region of closest similarity in protein sequence between α-crystallins and other members of the sHSP superfamily corresponds to sequences coded by exons 2 and 3 of α-crystallin genes[3,113] (Fig. 3.1). This region of similarity is also that most highly conserved among all members of the superfamily, suggesting that it corresponds to an ancestral functional unit or structural domain. In p40 of *S.mansoni* there are actually two of these structural units in tandem.[4,51] In contrast, N-terminal protein sequences, which correspond to exon 1 of α-crystallins, are much more variable throughout the superfamily. In the α-crystallins themselves there is evidence for a two fold repeat in this region which contains limited patches of similarity with some vertebrate sHSP sequences.[4,113,114]

It has been proposed that the exons of α-crystallins encode discrete structural units.[113] The two fold repeat encoded by exon 1 in the α-crystallins suggests that this exon may have arisen by duplication and fusion of an ancestral motif. Intron 1 neatly separates the more variable N-terminal sequences from those which are universally conserved among members of the sHSP superfamily. Furthermore, when the sequences encoded by exons 2 and 3 are aligned according to intron positions another weak two fold repeat is observed. This repeat is also apparent in *Drosophila* sHSPs which themselves come from intronless genes.[113] This has been interpreted as reflecting the presence of two related structural motifs in the C-terminal domain of α-crystallins which are the result of an ancient duplication of a gene corresponding to one exon.

The observations of possible mapping between exons and protein sequence motifs led to the suggestion that α-crystallins and sHSPs have a bipartite structure consisting of a variable N-terminal domain which may or may not have its own internal symmetry, and a C-terminal domain (or pair of domains) with two fold symmetry.[113] The proposed C-terminal domain contains the most important conserved structural features of the sHSP family. This two-domain model was investigated

by cloning and expressing the putative domains of bovine αA-crystallin.[35] The C-terminal domain behaved as a well-folded discrete structure capable of forming dimers or tetramers while the more hydrophobic N-terminal domain tended to assemble in large aggregates. It was these results which led to the idea that α-crystallins may be able to form tetramers or pseudo tetramers which in turn assemble into larger aggregates through non-specific interactions of the more hydrophobic N-terminal domain.[36]

Like α-crystallins other sHSP are also synthesized as protein subunits of about 20 kDa size and aggregate to form large multimeric complexes very similar to those formed by α-crystallins.[115-117] Typically, for both sHSP and α-crystallins, these aggregates are spherical assemblies of 300-800 kDa, although a variety of complex sizes and shapes can be observed under different conditions.[3] Whatever the methods of subunit assembly, they are promiscuous enough to allow the formation of mixed assemblages of αA- and αB-crystallins with mammalian hsp27[116,118] and with the recently discovered p20 proteins.[60] The ability of these disparate proteins to multimerize is further evidence that assembly is primarily dictated by the evolutionarily conserved C-terminal domains.

Intron Positions in the sHSP/α-Crystallin Superfamily

The origins and functions of the introns which divide the coding sequences of most vertebrate genes have been the subjects of considerable speculation and controversy.[119-121] In many genes introns map very closely to structural divisions in protein sequences. This has led to the "introns early" hypothesis that introns were present from the earliest time when these genes were assembled and indeed that introns were responsible for the assembly of modern genes.[119,122] For others however the mapping of introns and protein motifs is a statistically insignificant "hopeful illusion." In this view introns were inserted randomly relatively late in the evolution of the eukaryotic genome.[121] An intermediate hypothesis has also been proposed which accepts the special positions of introns in many genes but suggests that they could still have been the result of "introns late" insertion.[123] This scheme envisages that intron insertion could have been directed by RNA structure retained from ancestral RNA "genes."

Although the tertiary structure of α-crystallin subunits is unknown, there are indications that the exons of α-crystallins do in fact correspond to structural motifs. Introns delineate regions of duplicated sequence and in particular they separate the sequences most highly conserved among α-crystallins and small heat shock proteins (sHSP) from more variable regions.[4,113] Thus exons 2 and 3 are hypothesized to correspond to a structural domain conserved in the superfamily[113] and this prediction has received support from recombinant studies expressing the isolated domains.[35]

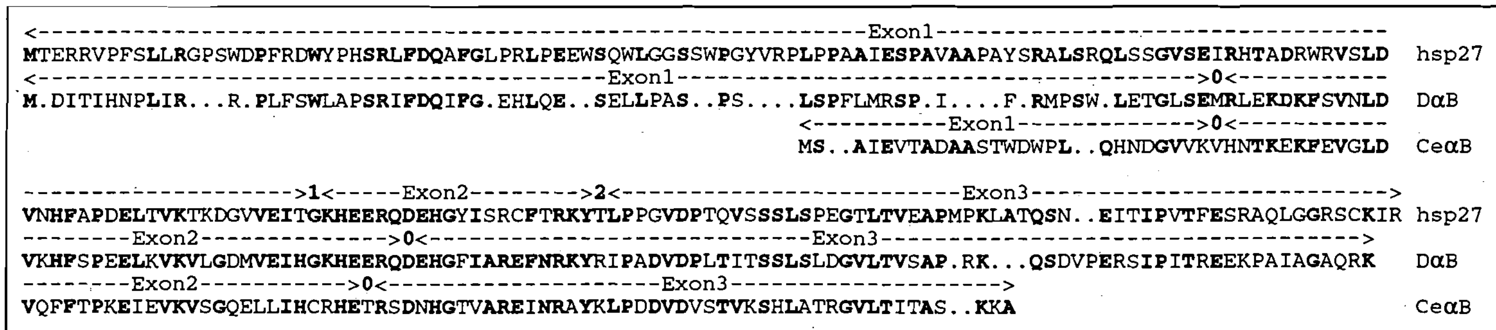


Fig. 3.3. Sequence alignment and relative intron positions of members of the sHSP/α-crystallin superfamily. Protein sequences are aligned and relative positions and phases of introns indicated. CeαB: predicted product of an "αB-like" C.elegans gene;[61] DαB: predicted product of duck αB-crystallin;[73] HSP27: predicted product of human HSP27 gene.[45]

Although the introns in α-crystallin genes seem to correspond to structural divisions they are not conserved in position or phase in the sHSP genes of mammals,[45,49,50] while several sHSP genes in yeast, plants and various vertebrates contain no introns at all.[46-48] It is not clear whether all these intronless genes are functional or may instead be processed pseudogenes.[50] However, some other sHSP family genes do share intron positions with α-crystallins. The relative position and splicing phase of the first intron of α-crystallin genes is exactly conserved in the hsp16 genes of *C.elegans*[44,113,124] and in the muscle-associated 29 kDa protein of *H.roretzi*.[53] Even greater similarity is found in an "αB-like" *C.elegans* gene discovered by large scale sequencing.[61] This gene has two introns both of which are in the same phase (0) as in α-crystallins (Fig. 3.3). The first is identical in position to intron 1 of α-crystallins and hsp16, while the second is displaced relative to α-crystallin intron 2 by only 3 codons.

In contrast the two introns of human hsp27 and mouse hsp25 genes do not closely correspond to the position of either α-crystallin intron and are in different phases (1 and 2 respectively)[45,49,50] (Fig. 3.3). This makes it highly unlikely that they are related to the α-crystallin introns by positional slippage. The 29 kDa protein gene has an intron at a relative position 6 codons downstream of intron 2 in α-crystallins.[53] This intron is in phase 2.

The following scenario for the evolution of gene structure in the sHSP/α-crystallin gene superfamily can be envisaged. The intron positions in α-crystallins may reflect the ancient assembly of a proto-sHSP from small fragments of RNA or DNA genetic material which in turn coded for protein structural motifs. This could have occurred either by intron-mediated assembly[119] or by subsequent intron insertion into RNA genes or transcripts at linker regions between ancestral folded RNA structural domains.[123] In an ancestor of the *C.elegans* hsp16 gene the equivalent of intron 2 was lost. In the gene for the 29 kDa protein such a lost intron may have been replaced by subsequent nearby insertion. In other lineages both introns were lost, perhaps via processed pseudogene intermediates. In some vertebrate sHSP genes there was then a reinsertion of introns which might have occurred directly at the DNA level with no potential for directed insertion. In this hypothesis the genes for α-crystallins and the αB-like sequence of *C.elegans* retained the ancestral structure mapping protein motifs and exons.

β- AND γ-CRYSTALLINS: A SUPERFAMILY IN THE VERTEBRATE LENS

In addition to the α-crystallins there are two other classes of ubiquitous crystallin which are represented in all vertebrate lenses. These are the β- and γ-crystallins. Several physico-chemical criteria led to the separate designation of these two classes.[1,125-127] The β-crystallins, with one interesting exception, were multimeric while the γ-crystallins were

monomers; the β-crystallins had blocked N-termini while those of γ-crystallins were free and the β- and γ-crystallins of mammals had different ranges of pI. Nevertheless, the β- and γ-crystallins have turned out to be more closely related than expected and belong to the same protein superfamily, the βγ-crystallins.[128-130] Furthermore, one original β-crystallin, βs-crystallin, is now redefined as a γ-crystallin.[131,132]

β-Crystallin Polypeptides

The first completely characterized set of β-crystallin polypeptides were those from bovine lens.[133] Several homologues from other species, including clear homologues in chicken, have also been sequenced.[6,129,130,134,135] All the β-crystallin polypeptides have sizes in the range 22-28 kDa. They are subclassified into βA (relatively acidic) and βB (relatively basic) subunits, although all β-crystallins have isoelectric points between 5.7 and 7.0.[1] In the bovine lens there are four βA polypeptides; βA1, A2, A3 and A4, and three βB polypeptides, βB1, B2 and B3.[130,133] The same nomenclature has generally been adopted for all β-crystallins in mammals and birds. Protein and nucleic acid sequencing has shown that βA1 and βA3 are products of the same gene and result from use of alternative initiator methionines.[133,136-138] Remarkably this feature of unknown utility is conserved in mammals, birds and frogs.[137-139]

All β-crystallin form dimers and higher aggregates of mixed composition.[1] The major β-crystallin, βB2, forms stable homodimers.[140] βA3/A1 is also able to form homodimers in vitro[141] but otherwise mixed multimers involving interactions between βA and βB subunits seem to be favored.[140] Dimers, trimers and perhaps tetramers of β-crystallin subunits form a β-crystallin fraction called βL (for low) in a native size range between about 40-100 kDa. In mammalian lenses larger aggregates of up to 200 kDa form the βH (for high) fraction which is dominated by octamers[142,143] containing up to 20% βB1 subunits and up to 35% βB2 subunits.[144]

Taxon-specific Differences

In spite of the conservation of clear homologues among the β-crystallins of mammals and birds there are some interesting differences in their properties which may relate to differences in lens structure and composition. This is illustrated in some taxon-specific variation in properties of βB1-crystallins. The βB1-crystallin of the chicken lens was originally called β35-crystallin for its apparent size in SDS PAGE.[145] This is considerably larger than the βB1-crystallins of mammals and larger than the 27 kDa deduced from its cDNA sequence.[134] Furthermore, the apparent size of this polypeptide even varies among birds, falling into two major mobility classes.[146] This anomalous behavior appears to be due to an essentially quantitative, post-translational modification involving some form of cytoplasmic glycosylation.[146] While there

is no evidence for any similar post-translational modification of β-crystallins in mammals, their βB1-crystallins have their own peculiarity. As discussed in detail below, mammalian βB1-crystallins alone among members of the βγ-crystallin superfamily have a unique sequence change in one of their characteristic structural motifs. Again, this may have altered the surface properties of the protein in a taxon-specific manner.

Other possible taxon-specific differences among β-crystallins of birds and mammals may be found in their phosphorylation patterns. In bovine lenses βB2-crystallin seems to be partially phosphorylated[147] while in chicken lens the only β-crystallin subject to this post-translational modification is βB3-crystallin.[148] Finally there may be taxon-specific differences in the relative abundance of certain β-crystallins. For example it has been suggested that the level of βA4-crystallin is lower in chicken than in bovine lens.[135]

The consequences of these differences is not known. However they may relate to the differences in protein content and "softness" of avian and mammalian lenses. In the bovine lens there is evidence that βB1-crystallin has an important role in the formation of βH.[149] The differences in βB1-crystallin among birds and mammals and perhaps even the differences in phosphorylation behavior of other βB-crystallin subunits may affect this aggregation into octamers such that bird lenses may not achieve these larger, more densely associated aggregates. This in turn may contribute to the greater hydration and softness of the bird lens.

γ-Crystallin Polypeptides

γ-Crystallins were originally characterized as the major components of a low molecular weight fraction in mammalian lenses.[1,150] In contrast to β-crystallins, they are strictly monomeric under in vitro conditions. Six γ-crystallins genes and their products were identified in the rat.[151,152] They were named γA-F and these systematic names have generally been applied to the γ-crystallins of other mammals, although not without certain problems (see below). This group of six γ-crystallins, constitute the originally defined γ-crystallin fraction of lens soluble proteins and they meet the original criterion of unblocked N-termini.

Although the best defined γ-crystallins are those found in mammals, multiple γ-crystallins have also been identified in the amphibians *Rana*[153] and *Xenopus*,[154] in the crocodilian *Caiman*[155] and in fish.[156,157] In all cases the multiple γ-crystallins are more similar within than among species, suggesting independent radiations or else highly constrained modifications to suit the requirements of each species. In the carp (*Cyprinus carpio*) γ-crystallins have remarkably high contents of methionine residues and this may be associated with the extremely high protein density of fish lenses.

In addition to the "classical" embryonic γ-crystallins there is another monomeric crystallin which is more widely distributed and better conserved in vertebrates and which was only recently redefined into this family. This is βs-crystallin, the protein which was formerly described as a monomeric β-crystallin.[1] By other criteria, such as the presence of a blocked N-terminus, this polypeptide seemed to be closer to the β-crystallins than to the γ-crystallins. However, protein, cDNA and gene sequencing has now shown that βs-, renamed γs-crystallin, is actually a more distantly related member of the γ-crystallins.[131,132] γs-Crystallin is expressed in fish, mammals and birds and is well conserved in sequence.[131,158-160] Indeed, while there are no clearly homologous relationships between any other fish or amphibian γ-crystallins and those of mammals, γs-crystallin is 70% identical between carp and bovine lens.[131,158] γ-Crystallins were thought to be absent from birds, however the reclassification of βs-crystallin, which is expressed in birds has altered this perception and allowed γ-crystallins the claim of ubiquity in vertebrates.[161,162] Even so, the absence of the classical γ-crystallins from bird lenses is striking and probably highly significant for the properties of the soft avian lens.

Expression Patterns

The expression patterns of β- and γ-crystallins in the developing vertebrate lens reveal something of their roles in building and maintaining the optical properties of the tissue. Both classes are expressed predominantly in the fiber cells. The γ-crystallins and certain β-crystallins such as βB1-crystallin are particularly strict in their preference for these terminally differentiated cells.[163-169] Although β-crystallins are expressed from early developmental stages in the lens, their expression continues and rises after birth so that the highest concentrations are usually found in the lens cortex. However, the pattern of expression varies among the individual β-crystallins. In contrast, the main group of mammalian γ-crystallins are expressed predominantly during embryogenesis, although some continue to be expressed for a time after birth, and there is species variability such that γD is the last to be active in human lens while in rat the last is γB.[152,165,166,170]

The result is that the highest concentrations of γA-F-crystallins are found in the central regions of the lens, the so-called lens nucleus which represents the embryonic lens. As the expression of the lens nuclear γ-crystallins declines γs-crystallin, the outlying member of the family, replaces them.[167] This led to its early identification as "cattle γ" a γ-crystallin peculiar to the mature, as opposed to the embryonic, bovine lens.[171] Thus γs-crystallin is found in the younger, more hydrated cortical fiber cells while the main group of mammalian γ-crystallins is associated specifically with the densest, highest refractive regions of the lens.

In the distribution of γ-crystallins, ontogeny seems to match phylogeny. Among species, the highest abundance of γ-crystallins is found in the hard lenses of fish as well as in nocturnal, burrowing rodents such as rats and mice. Only the cortical γs-crystallin is found in birds[162] and in human lenses only two out six γ-crystallin genes produce significant amounts of protein.[170,172,173] Embryonic γ-crystallins appear to have a specialized role in maintaining the stability of a low-water, high-protein concentration environment. In contrast, γs-crystallin and some of the β-crystallins help to maintain an environment of lower protein concentration. The result is a lens with a refractive index which declines from center to periphery. This increases the apparent convexivity of the lens and can also eliminate the spherical and chromatic aberration which afflicts lenses of uniform substance.[174]

This specialization of γ-crystallins seems to be borne out by the lack of evidence for any functional role for them outside the lens. Although sensitive RNA protection methods have detected non-lens expression of γ-crystallin family genes in *Xenopus* larvae, there has so far been no detection of γ-crystallin protein outside the lens.[154] In contrast, there is recent evidence that some β-crystallin proteins are expressed

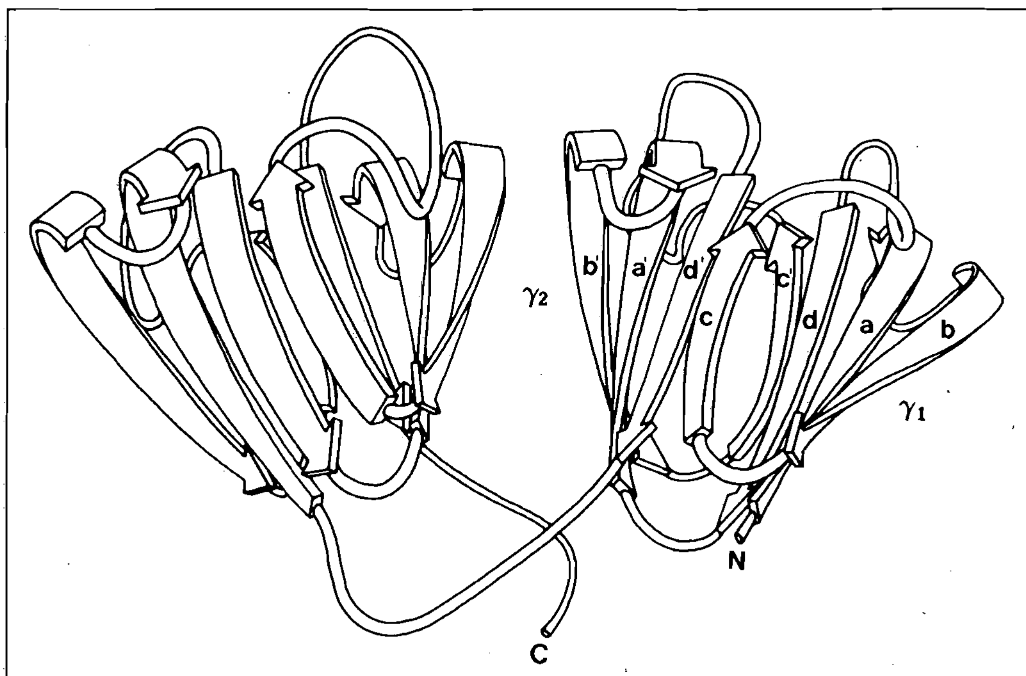


Fig. 3.4. The structure of bovine γB-crystallin[176-179] illustrated as a ribbon tracing of the polypeptide chain. (A modification of a drawing by Jane Richardson.)[206] The first two motifs of the N-terminal domain are shown as γ1 and γ2 with their β-strands lettered a-d and a'-d' respectively.

in non-lens tissues of newly hatched chickens, especially in retina, brain and kidney.[175] Their function in these tissues is not yet known.

PROTEIN STRUCTURE

Although little is known about the non-lens or ancestral function of the βγ-crystallin superfamily, a great deal is known about the three dimensional structure of the proteins and the gene organization of its members. This knowledge suggests a rich and complex evolutionary history. Since structure is so important in an understanding of this superfamily it will be described in some detail.

X-ray crystallography has determined the tertiary structures of several β- and γ-crystallins in detail. The first example solved, which has become the archetype of the class, was bovine γII-crystallin[176-179] (which now has the systematic name γB-crystallin) (Fig. 3.4). Very similar structures have also been determined for bovine γIIIb (γD)[180,181] and γIVa (equivalent to rat γE or mouse γF) crystallins.[182] Bovine γB-crystallin has a remarkably symmetrical structure of two domains, each containing a further two fold repeat of a characteristic structural motif. This γ-crystallin motif fits the structural category of a modified "greek-key." This same pattern is seen in βB2-crystallins for which an x-ray structure analysis is now also available.[183,184] Each motif bears a distinctive sequence signature which is required for correct folding of the polypeptide chain.

The structural motif of the βγ-crystallin superfamily consists of four antiparallel β-strands, *a-d*[177] (Fig. 3.4). The βγ motif fold is achieved as follows. Using the numbering scheme for motif 1 of bovine γB-crystallin as reference (Fig. 3.5), the first pair of strands, *a-b*, form a β-hairpin with an unusual turn between the strands which folds back onto the hairpin, burying some side chains. This is the principal distinguishing feature of this superfamily and requires certain key amino-acid residues. At position 6, on strand *a*, is an aromatic, usually tyrosine, which is buried by the folded back loop. At position 11 on the bend itself is another aromatic whose side chain also contributes to the interaction with the hairpin. At position 13 is an absolutely conserved glycine residue. As the polypeptide backbone negotiates the bend it adopts several positions of unusual Φ/ψ angles, this puts residue 13 in such a position that there is no room to accommodate a side chain, hence the requirement for glycine. After the *a-b* hairpin, strand *c* bends away from the plane occupied by the other three strands, crossing over to the other side of the protein domain. This strand is usually short and leads into the most variable part of the motif, a long *c-d* loop which crosses back over the "top" of the motif to the β-sheet in which the *a,b* and *d* strands of the motif lie.

As it bends back into the *d* strand which is hydrogen bonded to strand *a* in a β-sheet, the polypeptide backbone again encounters the unusual structure formed between strands *a* and *b* which actually bur-

ies the first residue of strand *d*, at position 34. There is only limited space for the side chain of this residue. Furthermore the side chain is in a position to hydrogen bond to the peptide backbone amide of residue 11 which would otherwise be unbonded and shielded from water, an energetically unfavorable situation. Only a serine residue can fit both requirements for size and hydrogen bonding and indeed serine is almost always found at this position. However, glycine and alanine could also fit into the pocket although they would be unable to form hydrogen bonds. It turns out that evolution has already performed this structural experiment in the fourth motif of mammalian βB1-crystallins.[133,185] These polypeptides, unlike their homologue in chicken[134] have alanine rather than serine at position 34 (Fig. 3.5). The structural consequences of this substitution are not clear. It seems probable that it would lead to looser packing of the folded hairpin. Since this change has been conserved among mammals it may modify the surface properties and intermolecular interactions of βB1-crystallin to fit the supramolecular organization of mammalian lenses.

The complete folding pathway results in a supersecondary structural motif with a three stranded β-sheet, *b-a-d*, and a lone β-strand *c*. This structure leaves exposed most of the hydrophobic side chains and it seems unlikely that it could exist in isolation. A complete globular domain is formed by assembly of two motifs around a pseudo-twofold axis such that the lone strand *c* of one motif interacts with strand *d* of the other. This forms a single globular domain, a wedged shaped structure of two β-sheets with strand patterns *b-a-d-c'* with a compact hydrophobic core between the two β-sheets (Fig. 3.4). Each β- and γ-crystallin monomer contains two of these domains.

Monomers and Dimers

The two domains of γB-crystallin are highly stable structures which are able to fold independently.[186] In the monomeric γ-crystallins, the two of these two-motif domains in the same molecule associate about another pseudo-twofold axis, joined by a bent connecting peptide so that symmetry related motifs 2 and 4 interact[176,177] (Fig. 3.6). In spite of their independent folding pathways, interaction between the two domains seems to be necessary for maximal stability.[187] Most of the residues of a γ-crystallin polypeptide are contained in the globular domains. Apart from the residues of the connecting peptide, the only other residues outside the globular domains are in the short, exposed extension at the C-terminus which is available for intermolecular interactions, perhaps acting as a "spacer" as suggested for β-crystallin extensions.[188] At the N-terminus in contrast there is no extension. The sequence begins with the first residue of strand *a*, a conserved glycine, and its amino terminus does not protrude from the surface of the N-terminal domain. This presumably explains the lack of N-terminal acetylation in γ-crystallins since their N-termini provide no target for aminopeptidase activity.

```
                                                                                  METQTVQQELESLPTTK (-2/-1) [N-extension] βA3

                                             MSQPAAKASATAAVNPGPDGKGKAGPPPGPAPGSGPAPAPAPAPAQPAPAAKAELPPGSYK (-1/1) [N-extension] βB1
                                                                                 MAQTNPMPG.....SVGPWK (-1/1) [N-extension] βA3
                                                                                                   GK (-1/1)               γB
                                                                                                   <-

     LVVFEQENFQGRRVEFSGECLNLGDRGFERVRSIIVTSGP   (1/2)  WVAFEQSNFRGEMFVLEKGEYPRWDTWSSS..YRSDRLMSFRPIKM   (2/3) [N-domain]   βB1
     ITIYDQENFQGKRMEFTSSCPNVSERNFDNVRSLKVECGA   (1/2)  WVGYEHTSFCGQQFVLERGEYPRWDAWSGSNAYHIERLMSFRPICSA (2/3) [N-domain]   βA3
     ITFYEDRGFQGHCYECSSDCPNL.QPYFSRCNSIRVDSGC----------WMLYERPNYQGHQYFLRRGDYPDYQQWMGFN....DSIRSCRLIPQ  (2/3) [N-domain]   γB
     -a-->       <-b->  <c>             <-d-> <-       a-->       <-b->   <c>               <-d->
                 Motif 1                                          Motif 2

DAQEHKLCLFEGANFKGNTMEIQEDDVPSLWVYGFCDR.VGSVRVSSGT (3/4) WVGYQYPGYRGYQYLLE....PGDFRHWNEWG..AFQPQMQAVRRLRD      [C-domain]   βB1
NHKESKITIFEKENFIGRQWEI.CDDYPSLQAMGWPNNEVGSMKIQCGA (3/4) WVCYQYPGYRGYQYILECDHHGGDYKHWREWGSHAQTSQIQSIRRIQQ      [C-domain]   βA3
HTGTFRMRIYERDDFRGQMSEI.TDDCPSLQDR.FHLTEVHSLNVLEGS-------WVLYEMPSYRGRQYLL....RPGEYRRYLDWG..AMNAKVGSLRRVMD      [C-domain]   γB
   <-a-->       <-b->   <c>               <-d-> <-       a-->       <-b->      <c>                <-d->
<-->            Motif 3                                             Motif 4
connecting
peptide

                                                                       RQWHREGCFPVLAAEPPK      [C-extension] βB1
                                                                       FY                      [C-extension] γB
```

Fig. 3.5. Sequence alignment of β- and γ-crystallins. Three bovine protein sequences taken from the databases, βB1-crystallin, βA3/A1-crystallin and γB-crystallin. Underlines show the two alternative translation starts for βA3- and βA1-crystallins. Sequences are divided to illustrate both exonic and protein structures. Intron positions and phases are indicated, taken from several gene sequences in GenBank (see also ref. 129). Introns labeled according to the scheme in Figure 3.9. Key residues in the motif signature are shown in bold type. The extent of β-strands is shown with arrows.

β-crystallins are different. They can form homo- and heterodimers, and they do this by a unique mechanism revealed by x-ray analysis of bovine βB2-crystallin.[183,184] The basis of the intermolecular interaction is essentially the same as the interdomain interaction in a γ-crystallin monomer except that this time the two domains come from different molecules. Each subunit has a very similar domain structure to γ-crystallins. However, instead of a bent connecting peptide and contact between two domains in the same molecule as in γ-crystallin, the connecting peptide of βB2-crystallin adopts an extended conformation (Fig. 3.6) which allows both domains to interact with another subunit through motif 2/motif 4' contacts across a two fold axis (Fig. 3.7).

β-Crystallins also differ from γ-crystallins in having long, relatively unstructured extensions. The βB family have extensions at both the N- and C-termini (Fig. 3.5). In contrast βA subunits have only an N-terminal extension while their C-termini are actually shorter than those in γ-crystallins. It has been hypothesized that either the terminal extensions or the connecting peptides provide the basis for the different conformations of β- and γ-crystallins.[183,189] This question has been

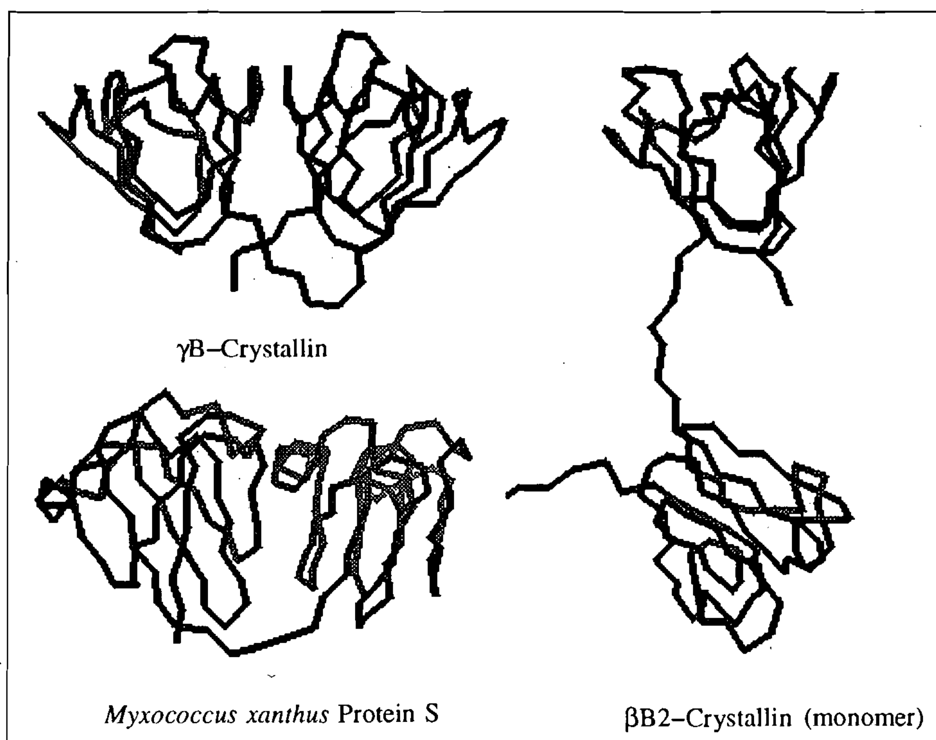


Fig. 3.6. Chain traces of three polypeptides of the βγ-crystallin superfamily. Coordinates were taken from the Protein Data Bank and displayed using the program Quanta on a Silicon Graphics workstation. (See text for references.)

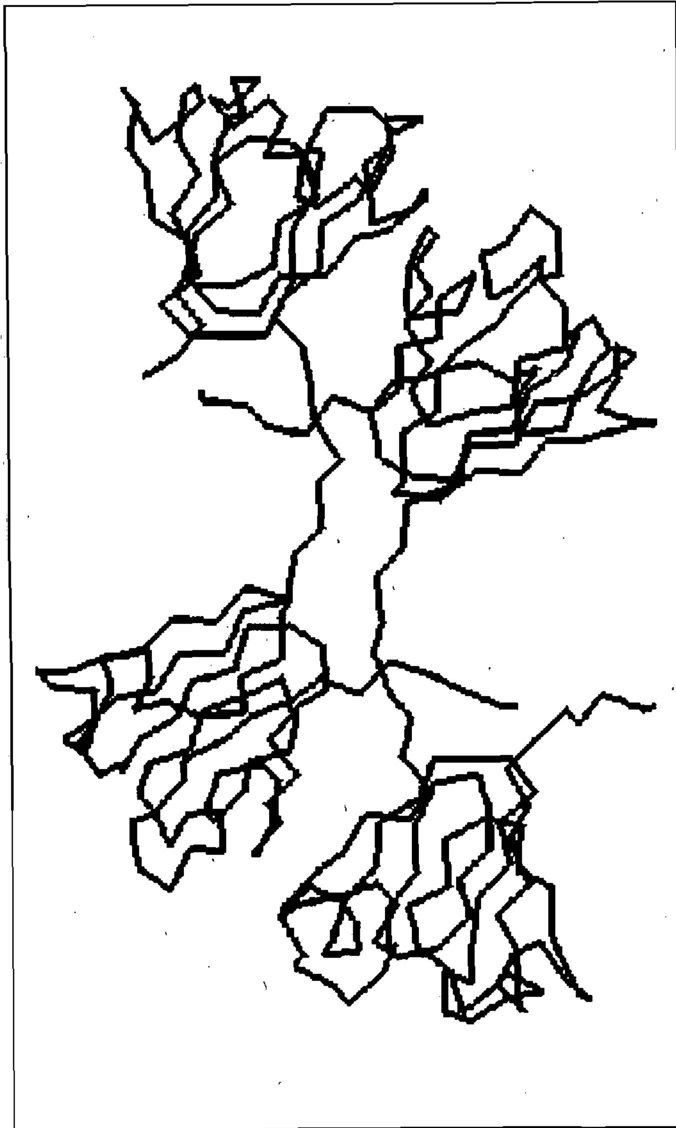


Fig. 3.7. Chain tracing of a bovine βB2-crystallin dimer.

addressed by numerous studies involving proteolytic modification of β-crystallins,[190] x-ray analysis,[183,184,188] ^{1}H NMR spectroscopy[144,191] and synthesis of chimeric recombinant crystallins.[141,187,192-194]

Early model-building[189] and proteolysis studies[190] had suggested that the N-terminal arm of βB2-crystallin was important for dimerization. However more recent studies have contradicted this idea. NMR and x-ray structure analyses[144,183,191] find that the extensions of βB2-crystallin are highly flexible and unstructured. Recombinant protein experiments

also show that dimers of βB2 or βA3/A1 can form in the absence of either or both extensions.[192,194] However x-ray structure analysis does suggest that at least part of the N-terminal extension may be involved in formation of βB2 tetramers.[188] There is less consensus about the role of the connecting peptide. Transposition of the βB2 connecting peptide into γB yields a monomeric protein[187] while similar transposition of the γB connecting peptide into mouse βA3 does not affect the ability of the recombinant protein to dimerize.[193] These results suggest that monomer or dimer formation is independent of the connecting peptide. However when the γB connecting peptide was transposed into βB2, the recombinant protein was found to be monomeric.[196] These apparently contradictory results await resolution.

Form for Function: The Role in Lens

Many of the structural features of β- and γ-crystallins can be interpreted in terms of the functional requirements of their role as crystallins.

γ-Crystallins need to be highly stable proteins capable of existing in the highest protein concentration regions of the lens throughout life. Aiding them in this role, they have compact globular domains with few protruding loops exposed for proteolysis. Bovine γB-crystallin is extremely symmetrical, yet the domains achieve an extra degree of close packing between the two four-stranded β-sheets of each domain through adoption of a slight asymmetry.[177] This allows rows of hydrophobic residues from opposing β-sheets to interdigitate, something which a perfectly symmetrical structure could not do.

Another somewhat unusual feature of γ-crystallins is their aminoacid composition. The core residues of γ-crystallins are rich in aromatic and sulfur-containing residues. The high content of cysteine residues seems paradoxical for a lens protein since oxidation is regarded as a major threat and potential cause of cataract.[1] This suggests that the sulfur atoms of γ-crystallins are present for important structural and functional reasons. One possibility which has been suggested[177] is that the d-orbitals of these atoms together with π-orbitals of aromatic residues are involved in formation of molecular bonds, shared networks of electrons binding the hydrophobic core. Some of the cysteine residues are actually exposed on the surface where they might be involved in similar interactions with other molecules in the densely packed regions of the lens. Interestingly, in fish γ-crystallins, which achieve even higher concentrations than those of mammals, the sulfur content, particularly of methionines, is even higher.[156,157] This may also reflect a use of polarizable electrons for non-hydrophobic contacts in low water concentrations.

X-ray analysis of bovine γB-crystallin under somewhat oxidizing conditions also revealed another possible useful function of some cysteine residues. Under these conditions, the protein forms a disulfide

bond between a cysteine in strand *c* and on in the *c-d* loop.[177] This tightens up a bend, but otherwise does not disrupt the structure. Thus this protein has the ability to absorb oxidizing potential harmlessly without forming any intermolecular bonds. This could provide an additional protection for the lens from oxidizing insult.

A final feature of amino acid composition is the very low lysine content of γ-crystallins (2 residues in bovine γB) compared to a considerably higher level of the other basic residue, arginine (20 in γB). Lysine is usually one of the most common residues in animal proteins. Its diminished content in γ-crystallins may help to reduce the potential for various kinds of post-translational modification which have been postulated to lead to cataract, such as carbamylation.[1]

The surface properties of γ-crystallins are key to its interactions with water and with other proteins. Bovine γB-crystallin has about half of its surface polar and charged side chains involved in intramolecular ion pairs or hydrogen bonds,[176,177] accounting for one third of the total molecular surface.[195] This is an unusually high fraction. It reduces the potential for binding shells of surrounding water molecules and may contribute to the usefulness of γ-crystallins in a relatively dehydrated environment. Because of their surface features, γ-crystallins may be less susceptible to certain kinds of phase separation at the high concentrations found in lens. Since they do not bind as much water as other proteins, their presence allows the maintenance of a low-water concentration and hence a high refractive index.

The surface properties of γ-crystallins may explain some unusual aspects of their biophysical behavior. Although they maintain their solubility at high concentrations, three of the mammalian γ-crystallins, γD, γE and γF whose genes form a consecutive group in the gene cluster, have a temperature sensitive phase separation.[152,173,196] This so-called "cold cataract" occurs when a solution of crystallins is cooled and is reversible. It suggests that the affinity of these proteins for a more hydrated environment is not strong and that energetically they are close to favoring a phase with more protein-protein interactions. In fact there is apparently an overall attraction among γ-crystallins while in contrast α- and β-crystallins have repulsive interactions.[197] Indeed, recent NMR studies suggest that at the concentrations of the lens nucleus, γ-crystallins may not be truly monomeric but may enter a state of "macromolecular crowding."[198] This is presumably essential for achieving the high protein concentrations and hence high refractive index of the lens nucleus without precipitation.

Chemical modification studies suggest that exposed cysteine residues may be involved in the phase separation which occurs in cold cataract and hence in attractive interactions among γ-crystallins.[199] Interestingly, in humans two of the four "cold cataract" crystallins, γE and γF, are pseudogenes.[170,200] Their inactivation probably contributes to the softening of the human lens relative to those of rodents in which all six γ-crystallins are expressed.

γ-Crystallins also have another odd behavior. On simple column separations they tend to elute as if they were smaller than half their actual molecular size.[201] The smaller than expected apparent size of the molecule in solution might be due to the presence of a more loosely held shell of water molecules than would be typical for other soluble proteins. Another possibility is that the two domains move rather independently in solution. This is something which could be answered by NMR solution structure analysis of the kind which have been performed for Protein S of *Myxococcus xanthus*.[202,203]

β-Crystallins have many similarities with γ-crystallins in domain structure.[183] However they have fewer surface ion-pairs and hydrogen bonds consistent with a role less specialized for a low-water environment. They are able to form various sizes of aggregates by forming dimers and higher multimers with subunits arranged in antiparallel "bunches."[188] This permits flexibility in supramolecular structure according to the requirements of different regions of the lens.

What is the function of the β-crystallin extensions? They may be involved in higher quaternary structures through direct protein interactions.[188] However it has also been suggested that these exposed polypeptides may act as "spacers" in the supramolecular structure of the lens.[188] Another possibility is that they have a principally entropic role in maintaining the solubility of large β-crystallin aggregates. The free extensions seem to be highly mobile in solution. This freedom would be lost in a precipitated phase. Thus the extensions, spinning like propellers on the outside of β-crystallin multimers, may contribute to an entropic component of the free energy of solution, opposing superaggregation, phase changes and opacity. Indeed, proteolytic cleavage of the N-termini of β-crystallins reduces their solubility.[204] In the rat lens this cleavage occurs as part of a normal maturation process which seems to contribute to the hardening of the lens nucleus and to its high refractive index.[204]

A WIDER SUPERFAMILY

The characteristic structural signature in the protein sequence of βγ-crystallin superfamily members (Fig. 3.5) has allowed the detection of non-lens members whose sequences are in the databases. Two of these proteins are expressed during spore or cyst formation in response to stress in certain micro-organisms.

The sequence of Protein S of the bacterium *Myxococcus xanthus* suggested the presence of a four motif, two domain structure similar to those of β- and γ-crystallins.[205] This was recently confirmed by a solution NMR analysis which revealed a structure remarkably similar to that of γ-crystallin[202,203] (Fig. 3.6). There were two notable differences. First, the long *c-d* loops of motifs 1 and 3 in Protein S form regular α-helices, reminiscent of but much more ordered than the irregular helices of motifs 2 and 4 in γ-crystallins. Second, the interdomain contact in the bacterial protein resembles neither that of a γ- or

β-crystallin. It is unsymmetrical, involving contact between motifs 2 and 3 while that in γ-crystallin is highly symmetrical and involves motifs 2 and 4. This difference may be due to the calcium binding properties of Protein S.

Spherulin 3a is expressed in the eukaryotic slime mold, *Physarum polycephalum*.[206] It too bears the signature required for folding of the βγ-crystallin superfamily motif. However, uniquely for known members of this superfamily, instead of four motifs and two domains spherulin 3a has only two motifs and one domain. It thus illustrates an earlier stage in molecular evolution than related proteins, predating the second internal duplication of sequence. Intriguingly, although it is only a one-domain structure, spherulin 3a has a fairly long N-terminal extension suggestive of those found in β-crystallins. By analogy, this might be expected to participate in higher order quaternary interactions.

The functions of these two proteins are not known. However, unlike more closely related species, prokaryotic *M.xanthus* and eukaryotic *P.polycephalum* share an unusual stress response. Both of them sporulate in response to osmotic stress induced by polyols, sugar alcohols. Since both organisms also share expression of members of the βγ-crystallin superfamily when undergoing spore or cyst formation it is possible that these two phenomena are related and that the ancestral role of the βγ-crystallin superfamily may have been to act as osmotic stress proteins.[62,206]

Until recently, no non-crystallin members of this superfamily were known in vertebrates. However, a candidate for this class has now been proposed.[207] An epidermis differentiation-specific protein (EDSP) in embryos of *Cynops pyrrhogaster*, an amphibian, has been cloned.[208] The N-terminal half of EDSP contains four, somewhat divergent, βγ-like motifs which suggest a two-domain, four motif structure like that of a γ-crystallin.[207] At the protein sequence level, EDSP is more similar to β- and γ-crystallins (28-30% identical) than to Protein S (25%).

Evolution of the Superfamily

The remarkable internal symmetry of each member of the βγ-crystallin superfamily suggests that they are derived from an extremely distant ancestor corresponding to a single structural motif. Such an isolated structure would probably be unable to satisfy hydrogen bonding requirements or to exclude water from its hydrophobic residues. It could achieve a stable conformation only by forming a homodimer. However the x-ray analysis of γB-crystallin shows that a heterodimeric association of two similar but different motifs, as in a γ-crystallin domain, produces a closer packing of core residues than would be possible in a perfectly symmetrical homodimer.[177] The added stability of such a heterodimeric structure could have been attained by duplication of the ancestral "motif gene" followed by sequence divergence to yield two isolated motifs, A and B (Fig. 3.8). These would have been required to fold together. The efficiency of the folding of each motif

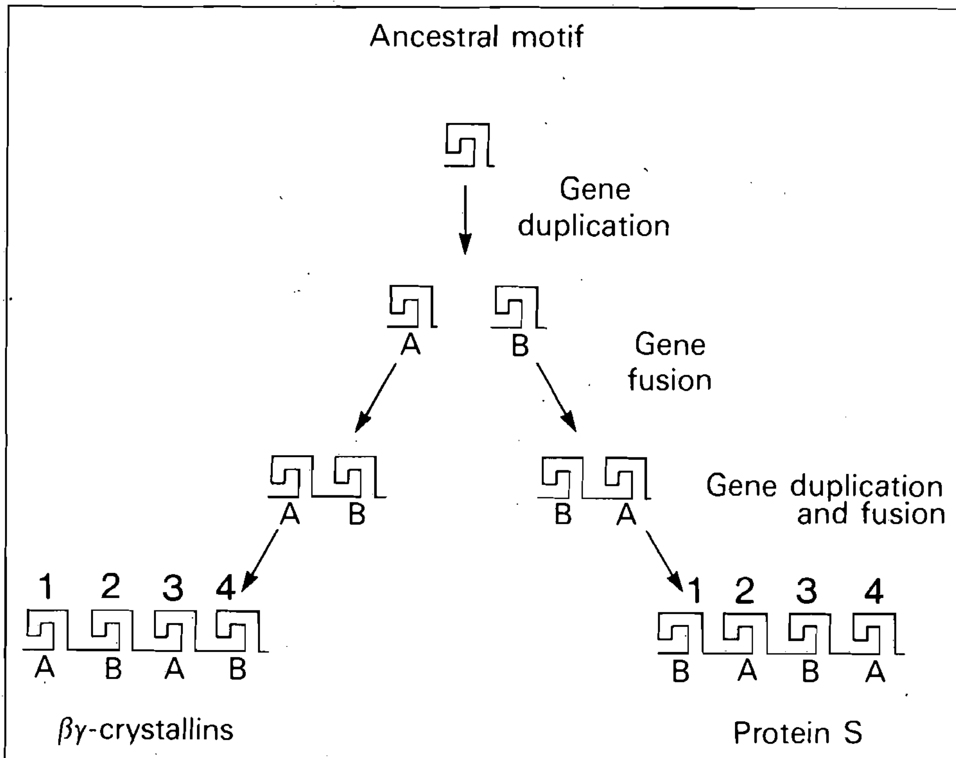


Fig. 3.8. Internal duplications in the evolution of the βγ-crystallin superfamily.

and their dimerization into a stable globular structure would have clearly been enhanced by fusion of the two proto-genes. The resultant fused gene would have coded for a one-domain protein with motif pattern AB, similar to one domain of a γ-crystallin or to spherulin 3a. At this stage there would have been no obvious mechanism to determine the order of fusion of the two motifs, so that AB and BA would have been equally likely.

Further duplication would have created the ABAB pattern seen in β- and γ-crystallins in which motifs 1 and 3 are both A-type while motifs 2 and 4 are B-type. In the crystallins, the B-type motif generally has a slightly longer *c-d* connecting loop which tends to adopt an irregular α-helical conformation. *Cynops* EDSP seems to lie on the same evolutionary pathway as the crystallins and also has the ABAB motif pattern.[207] In *M.xanthus* Protein S, one pair of motifs also exhibits slightly longer *c-d* loops which in this protein adopt the form of regular α-helices. However, these are motifs 1 and 3, not 2 and 4. This gives Protein S a BABA pattern of motifs,[202,205] suggesting an independent history of duplication and fusion starting from a very early stage.

INTRONS AND INTERNAL DUPLICATIONS IN β- AND γ-CRYSTALLIN GENE EVOLUTION

Further support for the idea that there were multiple independent gene duplication and fusion events in the evolution of this superfamily comes from the gene structure of the β- and γ-crystallins themselves.[129]

In vertebrates, the genes for the related families of β- and γ-crystallins exhibit striking examples of mapping between repeated protein structural motifs and exons. However this mapping is coupled with a clear history of internal duplication and it is by no means easy to discern the path taken in the assembly of each gene family. As we have seen, the members of the protein superfamily have a repeated structure of conserved motifs. Two motifs assemble into one domain and in most cases two domains are linked by a connecting peptide. N- and C-terminal peptide extensions may also be present. In β-crystallins, each of the four conserved structural motifs is encoded by a separate exon while one or two additional exons encode the N-terminal extension (Fig. 3.9). In γ-crystallin genes, in contrast, instead of each motif being encoded in separate exons the motifs are encoded in pairs so that one exon corresponds to a complete domain of two motifs.

For comparative purposes the exons and introns of β- and γ-crystallin genes can be labeled according to the motif structure of the proteins encoded so that homologous exons have similar names (Fig. 3.9). In this scheme the numbering of exons reflects the encoded motifs. Thus the third exon of a β-crystallin codes for motif 1 and can be designated exon 1. In γ-crystallins the second exon codes for motifs 1 and

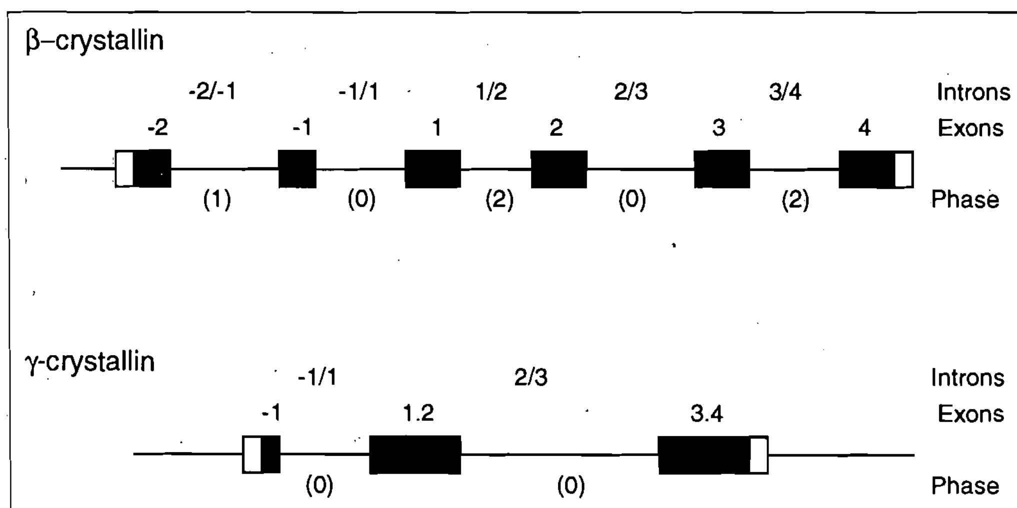


Fig. 3.9. Schematic gene structures for typical β- and γ-crystallins. Exons are shown as boxes. Coding sequences are shaded. In some β-crystallins the first exon is non-coding. Introns and exons are numbered to show similar mapping relative to protein motifs. Intron phases are shown in parentheses.

2 and can be designated exon 1.2. Introns can then be designated according to the flanking exons so that the interdomain intron of both β- and γ-crystallins is intron 2/3. Upstream exons have negative numbers. The first three codons of γ-crystallin motif 1 are contained in exon -1 which has some similarity to both exons -2 and -1, which encode the N-terminal extensions of β-crystallin genes.[129]

In β-crystallins the introns which delineate motifs form two pairs (Figs. 3.5, 3.9). Introns -1/1 and 2/3 are in phase 0. Intron -1/1 divides the N-terminal extension sequences from those of motif 1 while 2/3 divides motif 2 sequences from the connecting peptide. In spite of these differences there are intriguing similarities in the positions of these introns relative to the repeated motif structure (Fig. 3.5). The introns forming the other pair, 1/2 and 3/4, fall between motifs in each domain. Their positions relative to the motif structure are close to that of the other pair but not identical (Fig. 3.5) and they have phase 2. Introns -1/1 and 2/3 in γ-crystallins correspond precisely to those in β-crystallins. The γ-crystallin genes lack introns 1/2 and 3/4.

These two families of genes clearly share common ancestry. The "introns early" model would predict that the original common ancestor gene would have arisen by the intron-mediated assembly of two "motif genes." Duplication of this gene gave rise to the β-crystallins such that introns 1/2 and 3/4 of β-crystallins represent the original intron. γ-Crystallins would have arisen by independent duplication of a copy of the two-motif gene in which the ancestral intron was lost.[129] The general similarity of introns -1/1 and 2/3 could reflect shared ancestry such that one represents a duplication of the other resulting from an unequal crossover between two copies of the same gene. However this occurred, it would have been necessary to generate the extra sequence of the connecting peptide at the N-terminal end of motif 3.

All two-domain members of the superfamily, including β- and γ-crystallins, Protein S of *M.xanthus* and EDSP of *Cynops* have connecting peptides of 4 or 5 residues between domains. From a structural view it seems unlikely that a protein of this family could have two domains and no connecting peptide. Thus when the last round of duplication occurred the connecting peptide must have been present immediately. It could have arisen from an N- or C-terminal peptide of a one-domain precursor, but again it is not simple to reconcile the precise conservation of intron 2/3 in both β- and γ-crystallin families with a model of independent duplication. In fact the simplest explanation for the identical positioning of this intron is that it was already present in a common ancestor of both families. Otherwise this intron could be the result of directed insertion at a special position perhaps delineated by conserved RNA structure.[123]

Spherulin 3a of the eukaryote *P.polycephalum* is a one-domain member of the βγ-crystallin superfamily resembling an ancestral stage before the last internal duplication which gave rise to the crystallins.

The gene for spherulin 3a lacks introns. If introns are ancestral in this family, they were lost in spherulin 3a and in Protein S, partially lost in γ-crystallins and retained in β-crystallins. Alternatively, introns were not present ancestrally but were progressively added to vertebrate crystallin genes. As more members of this superfamily are uncovered, it will be interesting to see what patterns of introns their genes possess. Possibly the ancestral genes of β- and γ-crystallins were intronless even at the four motif stage. Introns may have been gained by directed insertion and then became duplicated within a gene through gene conversion mechanisms. The two families may have diverged prior to the last round of intron insertion and propagation in the ancestral β-crystallin gene.

The additional 5' exon -2 and intron -2/-1 of β-crystallins could have arisen by a variety of mechanisms. Some similarity has been noted between first and second exons of β-crystallins and between these sequences and the first exons of γ-crystallin genes.[129] This has led to the suggestion that there was a duplication of the first exon in β-crystallins.[129] In βA3/A1 both duplicated initiator methionines were retained while in βB1-crystallin the first initiator was lost causing the first exon to become non-coding.[129] However intron -2/-1 is in phase 1 so it seems unlikely that it represents a duplication of intron -1/1.

Gene Multiplication in the Evolution of β- and γ-Crystallins

Distinct families of β- and γ-crystallin are found throughout the vertebrates, from fish and (probably) lampreys to amphibians, reptiles, birds and mammals. This suggests that the molecular lineages of the two families were well established at a very early stage in lens evolution. Possibly both families pre-date the vertebrate lens and were recruited together. However a more parsimonious hypothesis is that an ancestral β-crystallin was recruited as one of the original crystallins. Subsequent gene duplications and divergence give rise to multiple β-crystallins and to the more lens-specialized γ-crystallins.

Homologous βA- and βB-crystallin genes are found in both birds and mammals. In the human genome the genes for βB2, βB3, βA4 and a pseudogene for βB2-crystallin genes are closely linked on chromosome 22.[209-212] βB2, βB3 and βBA4 are linked on mouse chromosome 5[213] while βB2 and βB3 are known to be linked in rat.[209] In chicken βB1 and βA4 are arranged head to head only 2 kb apart.[135] However human βA3/A1-crystallin is located on chromosome 17.[214,215] This suggests that there might at one time have been a single β-crystallin cluster which is now beginning to disperse.

βA3/A1- and βA4-crystallins have been observed in amphibians[39,139] and hybridization studies suggest the existence of βA- and βB-crystallin related sequences in all classes of vertebrates.[216] Thus the βA- and βB-crystallin families also seem to have had a very early origin. In the course of surveying fish for possible taxon-specific crystallins, we have

obtained several tryptic peptides of β-crystallin subunits from the lens of a teleost fish, the surf perch (unpublished). All the peptides clearly belong to the βB family. However it is less easy to assign these peptides to specific family members since individual peptides share some sequence hallmarks of different mammalian/avian βB-crystallins. This preliminary analysis confirms the presence of βB-crystallins in fish but raises the possibility that this family underwent independent radiation in fish and in terrestrial vertebrates.

ANCESTRAL β-CRYSTALLINS?

It is possible that one of the existing β-crystallins is directly homologous to the ancestral β-crystallin which gave rise to the whole β-crystallin family and perhaps even to the γ-crystallins. Several features of βB2-crystallin suggest that it has a special structural and evolutionary significance. βB2-crystallin is the most highly conserved of the βB-crystallin subunits.[39] In those vertebrate lenses which have been examined it is the major β-crystallin subunit[1,2] and it plays a role in organizing other β-crystallin subunits.[140] Unlike other β-crystallins, βB2 shows a strong propensity for formation of highly stable dimers[140,183] and can therefore exist in a stable form in the absence of other β-crystallin subunits. Indeed, βB2-crystallin has the highest thermal stability of any β-crystallin.[217,218] Furthermore, βB2 appears to be the β-crystallin subunit with the highest non-lens expression and is the first crystallin to be detected at elevated level in transdifferentiation of chicken neural retina.[175] In most of these "special" categories βB2 is closely followed by βA3 which also happens to be even more highly conserved in amino acid sequence.[39,219] Could one of these proteins be the modern homologue of the pre-lens ancestor of this family?

γ-CRYSTALLINS

γ-Crystallins present a more complicated picture. The single γ-crystallin of the adult mammalian lens, γs-crystallin is well conserved among mammals and fish.[131,158] In contrast the six embryonic γ-crystallins, γA-F, are highly conserved in mammals,[129] absent from birds[162,220] and apparently non-homologous to γ-crystallins in amphibians and fish.[153,156] The six embryonic genes of mammals form a tight cluster on a single chromosome, 2 in man,[22-223] 1 in mouse[224] and 9 in rat.[151,225] Their products are highly similar to each other, ranging from about 75% to 98% identical. Analysis of the rat γ-crystallin gene cluster has yielded evidence of gene conversion.[226] Nevertheless, clearly homologous genes are present in different mammals. The distantly related but evolutionarily conserved γs-crystallin is not linked to the γ-crystallin cluster and in the human genome is located on chromosome 3.[132]

It seems likely that all modern placental mammals have six homologous embryonic γ-crystallin genes although since only in rat, mouse and human have all six genes been characterized,[129,227] this view is subject

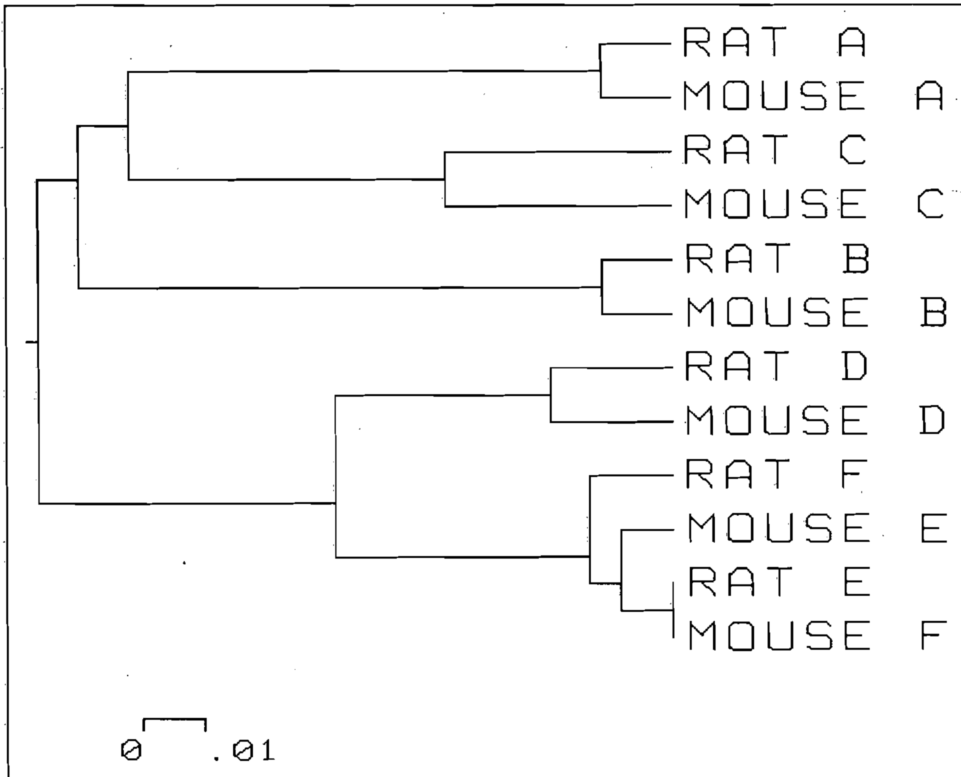


Fig. 3.10. Cladogram comparing protein sequences for mouse and rat γ-crystallins. Sequences taken from Swissprot database. Tree was constructed using the UPGMA p-distance option in the program MEGA.[232] Scale represents 1% difference. (Figure kindly provided by Dr. Cynthia Jaworski). Note that the high-phase separation temperature cryoproteins γD,E and F (152) are on the same branch and that rat γE and mouse γF are identical.

to revision. In rat and mouse four γ-crystallins, γA-D, are clearly homologous as judged by both gene and protein sequence[151,227] (Fig. 3.10). This suggests that these four genes were established prior to the separation of murine rodents 10-20 million years ago. γE- and γF-crystallins are a different story. These two proteins are the most similar pair of γ-crystallins, 98% identical in rat and 96% identical in mouse. However, designating homologues for this pair of crystallins between rat and mouse is not straightforward. When application of a rationalized nomenclature for γ-crystallins was attempted it seemed clear that the protein coded by the cDNA originally named mouse γ2-crystallin was identical to that coded by rat γE-crystallin.[128] However, when gene sequences and relative positions in the γ-crystallin gene cluster were compared it appeared that the mouse γ2-crystallin gene was actually more equivalent to the gene for rat γF-crystallin.[129] Consequently, since it was felt that genomic organization should have precedence over protein

sequence, the genes now known as rat γE and mouse γF-crystallin produce an identical protein (Fig. 3.10).

There are two possible reasons for this paradox. The first is that gene conversion has shuffled the identities of the γ-crystallins. The second explanation is that a prototype γE/F crystallin was present in an ancestor of rats and mice. Just before these species diverged the gene duplicated and untranslated sequences diverged rapidly establishing γE and γF genes with essentially identical protein products. After divergence of the species, the protein products of the two daughter genes diverged in sequence more slowly in response to particular selective pressures. These pressures led to certain coordinated sequence changes in one protein, equivalent to the sequence of rat γE-crystallin. However, in mouse these changes were produced in the γF-crystallin gene, giving the remarkable result of an identical protein produced by a different gene in a closely related species.

The human genome also contains six complete genes for γ-crystallins.[170] In protein sequence, the products of γA-D genes again seem to be homologues of the equivalent rodent genes, suggesting that these proteins and their genes predate the radiation of modern placental mammals. Humans also have two pseudogenes, superficially equivalent to the γE and γF genes of rodents.[170,200] Thus there may have been six γ-crystallin genes in the common ancestor of primates and of rodents. The human genome also contains a fragmentary γ-crystallin gene sequence and it has been suggested that this represents a lost copy of one of the γE/F genes, possibly the true homologue of rat γF, which has been deleted in primates.[170] The remaining gene may have then duplicated again to restore the complement of genes to six. However the history of these genes is by no means unambiguous. Insertion of repetitive elements, gene conversion, sequence drift in pseudogenes and perhaps varying pressure on expressed genes in species whose lenses are as different as humans and rats have all had their effect.[129]

γ-CRYSTALLINS AND THE EVOLUTION OF THE LENS

The γ-crystallins of mammals are much more similar to each other than are the multiple γ-crystallins of a frog, which generally show the same degree of conservation and therefore the same apparent age as β-crystallins.[129] Thus the mammalian γ-crystallins appear to be a much younger family than either β-crystallins or the γ-crystallins of an amphibian. Part of this similarity may be due to gene conversion resulting from the tight clustering of γ-crystallin genes in mammals. However, both the clustering and the conservation of sequence could also be the result of a relatively recent re-invention of γ-crystallins, perhaps a series of duplications of a single gene, the most recent of which gave rise to the γE/γF pair.

In this hypothesis, γ-crystallin gene expression in the distant reptilian ancestors of mammals would have declined just as it did in the

ancestors of birds and for the same reasons. Later in evolution however, the ancestors of modern placental mammals seem to have abandoned the diurnal habit which evidently favors soft lenses and instead adopted a nocturnal, burrowing habit similar to that of mice and rats today. This idea is supported by certain features of the mammalian visual system. In particular, most mammals are essentially color blind, lacking cone cell photoreceptors required for color discrimination and instead relying on the rod cells which are specialized for dim-light vision.[228] Indeed, all placental mammals lack the colored oil-drop bodies which are associated with cone cells in the retinas of fish, reptiles and birds and even in non-placental mammals such as marsupials.[228,229]

Due to the low photosensitivity of cones, color vision requires bright light. It therefore has no value to animals which are active in dim light and in such species there is no selective pressure for its retention. In the same way, animals living as mice do now have no need of a soft, accommodating lens. As rats and mice demonstrate, a harder, spherical lens is more useful, perhaps because its short focal length permits detailed examination of objects at close quarters.

Nocturnal, burrowing ancestors of placental mammals may have lost both their color vision and their soft, diurnal lenses. They may have essentially re-invented the hard lenses of their distant aquatic ancestors by increasing the expression of γ-crystallins. If γ-crystallins were already on the path to elimination, their revival could have easily been achieved by multiplication of a surviving gene to regenerate a family of γ-crystallins forming a tightly linked cluster of recently duplicated genes. Later, when descendent species became diurnal or partially diurnal once more, this dynamic evolutionary process would have reversed yet again, recapitulating the softening of the lens. Thus, while all six γ-crystallin genes are expressed at high levels in rat, only two out of six γ-crystallin genes, γC and γD, are expressed at significant levels in human lens. Two other human γ-crystallin genes, γE and γF, are pseudogenes while γA and γB are expressed at very low levels.[170,172,173] As a result the soft human lens has no more than one third the γ-crystallin content of the hard rodent lens.[152,230]

Modulation of γ-crystallin content seems to be one of the key mechanisms for modifying the properties of the vertebrate lens. This strategy for molecular engineering of the optical properties of the lens has also been hypothesized to be the underlying reason for the recruitment of taxon-specific enzyme crystallins.[62] The introduction of these new proteins into the lens may serve to dilute or replace the γ-crystallins and hence contributes to a lower protein concentration in the lens of terrestrial vertebrates.

References

1. Harding JJ, Crabbe MJC. The lens: Development, proteins, metabolism and cataract. In: Davson H, ed. The Eye. v. 1B. New York: Academic

Press, 1984; 207-492.

2. 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.
3. de Jong WW, Leunissen JAM, Voorter CEM. Evolution of the α-crystallin/small heat-shock protein family. Mol Biol Evol 1993; 10:103-26.
4. de Jong WW, Leunissen JAM, Leenen PJM, Zweers A, Versteeg M. Dogfish α-crystallin sequences: comparison with small heat shock proteins and *Schistosoma* egg antigen. J Biol Chem 1988; 263:5141-9.
5. Thompson MA, Hawkins JW, Piatigorsky J. Complete nucleotide sequence of the chicken αA-crystallin gene and its 5' flanking region. Gene 1987; 56:173-84.
6. Sawada K, Agata K, Eguchi G. Crystallin gene expression in the process of lentoidogenesis in cultures of chicken lens epithelial cells. Exp Eye Res 1992; 55:879-87.
7. de Jong WW, van Amelsvoort AM, van der Ouderaa FJ, Bloemendal H. Slow rate of evolution of αA chain of α-crystallin. Nature [New Biol] 1973; 246:233-6.
8. de Jong WW, Zweers A, Versteeg M, Nuy-Terwindt EC. Primary structures of the α-crystallin A chains of twenty-eight mammalian species, chicken and frog. Eur J Biochem 1984; 141:131-40.
9. de Jong WW, Leunissen JAM, Wistow GJ. Eye lens crystallins and the phylogeny of placental orders: Evidence for a macroscelid-paenungulate clade? In: Szalay FS, Novacek MJ, McKenna MC, eds. Mammal Phylogeny: Placentals. New York, NY: Springer Verlag, 1993; 5-12.
10. de Jong WW. Superordinal affinities of Rodentia studied by sequence analysis of eye lens protein. In: Luckett WP, Hartenberger J-L, eds. Evolutionary relationships among rodents: a multidisciplinary analysis. New York: Plenum Publishing Corp., 1985; 211-26.
11. de Jong WW, Goodman M. Mammalian phylogeny studied by sequence analysis of the eye lens protein α-crystallin. Sonderdruck Z Saugetierkunde 1982; 47:257-76.
12. Stapel SO, Leunissen JA, Versteeg M, Wattel J, de Jong WW. Ratites as oldest offshoot of avian stem - evidence from α-crystallin A sequences. Nature 1984; 311:257-9.
13. Voorter CE, de Haard-Hoekman WA, van den Oetelaar PJ, Bloemendal H, de Jong WW. Spontaneous peptide bond cleavage in aging α-crystallin through succinimide intermediate. J Biol Chem 1988; 263:19020-3.
14. Chiesa R, Gawinowicz-Kolks MA, Kleiman NJ, Spector A. The phosphorylation sites of the B2 chain of bovine α-crystalli. Biochem Biophys Res Commun 1987; 144:1340-7.
15. Chiesa R, Gawinowicz-Kolks MA, Spector A. The phosphorylation of the primary gene products of α-crystallin. J Biol Chem 1987; 262:1438-41.
16. Voorter CEM, Mulders JWM, Bloemendal H, de Jong WW. Some aspects of the phosphorylation of α-crystallin A. Eur J Biochem 1986; 160:203-10.

17. Spector A, Chiesa R, Sredy J, Garner W. cAMP-dependent phosphorylation of bovine lens α-crystallin. Proc Natl Acad Sci USA 1985; 82:4712-6.
18. Voorter CE, de Haard-Hoekman WA, Roersma ES, Meyer HE, Bloemendal H, de Jong WW. The in vivo phosphorylation sites of bovine αB-crystallin. FEBS Lett 1989; 259:50-2.
19. Kantorow M, Piatigorsky J. α-Crystallin/small heat shock protein has autokinase activity. Proc Natl Acad Sci USA 1994; 91:3112-6.
20. de Jong WW, Mulders JW, Voorter CE, Berbers GA, Hoekman WA, Bloemendal H. Post-translational modifications of eye lens crystallins: crosslink phosphorylation and deamidation. Adv Exp Med Biol 1988; 231:95-108.
21. Voorter CE, Roersma ES, Bloemendal H, de Jong WW. Age-dependent deamidation of chicken αA-crystallin. FEBS Lett 1987; 221:249-52.
22. Groenen PJ, Bloemendal H, de Jong WW. The carboxy-terminal lysine of αB-crystallin is an amine-donor substrate for tissue transglutaminase. Eur J Biochem 1992; 205:671-4.
23. Manenti S, Dunia I, Benedetti EL. Fatty acid acylation of lens fiber plasma membrane proteins. MP26 and α-crystallin are palmitoylated. FEBS Lett 1990; 262:356-8.
24. Groenen PJ, van den Ijssel PR, Voorter CE, Bloemendal H, de Jong WW. Site-specific racemization in aging αA-crystallin. FEBS Lett 1990; 269:109-12.
25. Roquemore EP, Dell A, Morris HR, et al. Vertebrate lens α-crystallins are modified by O-linked N-acetylglucosamine. J Biol Chem 1992; 267:555-63.
26. Li LK, Spector A. The optical rotatory dispersion and circular dichroism of calf lens α-crystallin. J Biol Chem 1967; 242:3234-6.
27. Li LK, Spector A. Circular dichroism and optical rotatory dispersion of the aggregates of purified polypeptides of α-crystallin. Exp Eye Res 1974; 19:49-57.
28. Carver JA, Aquilina JA, Truscott RJ, Ralston GB. Identification by 1H NMR spectroscopy of flexible C-terminal extensions in bovine lens α-crystallin. FEBS Lett 1992; 311:143-9.
29. Thomson JA, Augusteyn RC. On the structure of α-crystallin: construction of hybrid molecules and homopolymers. Biochim Biophys Acta 1989; 994:246-52.
30. Hendriks W, Weetink H, Voorter CEM, Sanders J, Bloemendal H, de Jong WW. The alternative splicing product αAins-crystallin is structurally equivalent to αA and αB subunits in the rat α-crystallin aggregate. Biochim Biophys Acta 1990; 1037:58-65.
31. Bindels JG, Siezen RJ, Hoenders HJ. A model for the architecture of α-crystallin. Ophthal Res 1979; 11:441-52.
32. Tardieu A, Laporte D, Licinio P, Krop B, Delaye M. Calf lens α-crystallin quaternary structure. A three-layer tetrahedral model. J Mol Biol 1986; 192:711-24.
33. Walsh MT, Sen AC, Chakrabarti B. Micellar subunit assembly in a three layer model of oligomeric α-crystallin. J Biol Chem 1991; 266:20079-84.

34. Augusteyn RC, Koretz JF. A possible structure for α-crystallin. FEBS Lett 1987; 222:1-5.
35. Merck KB, de Haard-Hoekman WA, Oude Essink BB, Bloemendal H, de Jong WW. Expression and aggregation of recombinant αA-crystallin and its two domains. Biochim Biophys Acta 1992; 1130:267-76.
36. Wistow G. Possible tetramer-based quaternary structures for α-crystallins and small heat-shock proteins. Exp Eye Res 1993; 56:729-32.
37. Carver JA, Aquilina JA, Truscott RJ. A possible chaperone-like quaternary structure for α-crystallin. Exp Eye Res 1994; 59:231-4.
38. van den Oetelaar PJ, van Someren PF, Thomson JA, Siezen RJ, Hoenders HJ. A dynamic quaternary structure of bovine α-crystallin as indicated from intermolecular exchange of subunits. Biochemistry 1990; 29:3488-93.
39. de Jong WW, Lubsen NH, Kraft HJ. Molecular Evolution of the Eye Lens. In: Chader G, Osbourne N, eds. Progress in Retinal and Eye Research. v. 13. : Elsevier Science Ltd, 1994; 391-442.
40. McAvoy JW. Cell division, cell elongation and distribution of α- β- and γ-crystallins in the rat lens. J Embryol Exp Morphol 1978; 44:149-65.
41. Zwaan J. The appearance of α-crystallin in relation to cell cycle phase in the embryonic mouse lens. Dev Biol 1983; 96:173-81.
42. Oguni M, Setogawa T, Hashimoto R, Tanaka O, Shinohara H, Kato K. Ontogeny of α-crystallin subunits in the lens of human and rat embryos. Cell Tissue Res 1994; 276:151-4.
43. Ingolia TD, Craig EA. Four small *Drosophila* heat shock proteins are related to each other and to mammalian α-crystallin. Proc Natl Acad Sci USA 1982; 79:2360-4.
44. Russnak RH, Candido EP. Locus encoding a family of small heat shock genes in *Caenorhabditis elegans*: two genes duplicated to form a 3.8-kilobase inverted repeat. Mol Cell Biol 1985; 5:1268-78.
45. Hickey E, Brandon SE, Potter R, Stein G, Stein J, Weber LA. Sequence and organization of genes encoding the human 27 kDa heat shock protein [published erratum appears in Nucleic Acids Res 1986 Oct 24; 14(20):8230]. Nucleic Acids Res 1986; 14:4127-45.
46. Raschke E, Baumann G, Schoffl F. Nucleotide sequence analysis of soybean small heat shock protein genes belonging to two different multigene families. J Mol Biol 1988; 199:549-57.
47. Susek RE, Lindquist SL. hsp26 of *Saccharomyces cerevisiae* is related to the superfamily of small heat shock proteins but is without a demonstrable function. Mol Cell Biol 1989; 9:5265-71.
48. Krone PH, Snow A, Ali A, Pasternak JJ, Heikkila JJ. Comparison of regulatory and structural regions of the *Xenopus laevis* small heat-shock protein-encoding gene family. Gene 1992; 110:159-66.
49. Gaestel M, Gotthardt R, Muller T. Structure and organisation of a murine gene encoding small heat-shock protein Hsp25. Gene 1993; 128:279-83.
50. Frohli E, Aoyama A, Klemenz R. Cloning of the mouse hsp25 gene and an extremely conserved hsp25 pseudogene. Gene 1993; 128:273-7.

51. Nene V, Dunne DW, Johnson KS, Taylor DW, Cordingley JS. Sequence and expression of a major egg antigen from *Schistosoma mansoni*: Homologies to heat shock proteins and α-crystallins. Mol Biochem Parasitol 1986; 21:179-88.
52. Cao M, Chao H, Doughty BL. Cloning of a cDNA encoding an egg antigen homologue from *Schistosoma mansoni*. Mol Biochem Parasitol 1993; 58:169-71.
53. Takagi T, Yasunaga H, Nakamura A. Structure of 29-kDa protein from ascidian (*Halocynthia roretzi*) body wall muscle. J Biochem (Tokyo) 1993; 113:321-6.
54. Nerland AH, Mustafa AS, Sweetser D, Godal T, Young RA. A protein antigen of *Mycobacterium leprae* is related to a family of small heat shock proteins. J Bacteriol 1988; 170:5919-21.
55. Verbon A, Hartskeerl RA, Schuitema A, Kolk AH, Young DB, Lathigra R. The 14,000-molecular-weight antigen of *Mycobacterium tuberculosis* is related to the α-crystallin family of low-molecular-weight heat shock proteins. J Bacteriol 1992; 174:1352-9.
56. Tantimavanich S, Nagai S, Nomaguchi H, Kinomoto M, Ohara N, Yamada T. Immunological properties of ribosomal proteins from *Mycobacterium bovis* BCG. Infect Immun 1993; 61:4005-7.
57. Allen SP, Polazzi JO, Gierse JK, Easton AM. Two novel heat shock genes encoding proteins produced in response to heterologous protein expression in *Escherichia coli*. J Bacteriol 1992; 174:6938-47.
58. Heidelbach M, Skladny H, Schairer HU. Heat shock and development induce synthesis of a low-molecular-weight stress-responsive protein in the myxobacterium *Stigmatella aurantiaca*. J Bacteriol 1993; 175:7479-82.
59. Sauer U, Durre P. Sequence and molecular characterization of a DNA region encoding a small heat shock protein of *Clostridium acetobutylicum*. J Bacteriol 1993; 175:3394-400.
60. Kato K, Goto S, Inaguma Y, Hasegawa K, Morishita R, Asano T. Purification and characterization of a 20-kDa protein that is highly homologous to αB crystallin. J Biol Chem 1994; 269:15302-9.
61. Wilson R, Ainscough R, Anderson K et al. 2.2 Mb of contiguous nucleotide sequence from chromosome III of *C. elegans*. Nature 1994; 368:32-8.
62. Wistow G. Lens crystallins: Gene recruitment and evolutionary dynamism. Trends Biochem Sci 1993; 18:301-6.
63. Duguid JR, Rohwer RG, Seed B. Isolation of cDNAs of scrapie-modulated RNAs by subtractive hybridization of a cDNA library. Proc Natl Acad Sci USA 1988; 85:5738-42.
64. Iwaki T, Kume-Iwaki A, Liem RK, Goldman JE. αB-crystallin is expressed in non-lenticular tissues and accumulates in Alexander's disease brain. Cell 1989; 57:71-8.
65. Lowe J, Landon M, Pike I, Spendlove I, McDermott H, Mayer RJ. Dementia with β-amyloid deposition: Involvement of αB-crystallin supports two main diseases. Lancet 1990; 336:515-6.
66. Renkawek K, de Jong WW, Merck KB, Frenken CW, van Workum FP,

Bosman GJ. αB-crystallin is present in reactive glia in Creutzfeldt-Jakob disease. Acta Neuropathol (Berl) 1992; 83:324-7.

67. Dubin RA, Wawrousek EF, Piatigorsky J. Expression of the murine αB-crystallin gene is not restricted to the lens. Mol Cell Biol 1989; 9:1083-91.
68. Iwaki T, Kume-Iwaki A, Goldman JE. Cellular distribution of αB-crystallin in non-lenticular tissues. J Histochem Cytochem 1990; 38:31-9.
69. Lee DC, Kim RY, Wistow GJ. An avian αB-crystallin. Non lens expression and sequence similarities with both small (hsp27) and large (HSP70) heat shock proteins. J Mol Biol 1993; 232:1221-6.
70. Kato K, Shinohara H, Kurobe N, Inaguma Y, Shimizu K, Ohshima K. Tissue distribution and developmental profiles of immunoreactive αB crystallin in the rat determined with a sensitive immunoassay system. Biochim Biophys Acta 1991; 1074:201-8.
71. Klemenz R, Frohli E, Steiger RH, Schafer R, Aoyama A. αB-crystallin is a small heat shock protein. Proc Natl Acad Sci USA 1991; 88:3652-6.
72. Dasgupta S, Hohman TC, Carper D. Hypertonic stress induces αB-crystallin expression. Exp Eye Res 1992; 54:461-70.
73. Wistow G, Graham C. The duck gene for αB-crystallin shows evolutionary conservation of discrete promoter elements but lacks heat and osmotic stress response. Biochim Biophys Acta 1995:in press.
74. Chiesi M, Longoni S, Limbruno U. Cardiac α-crystallin. III. Involvement during heart ischemia. Mol Cell Biochem 1990; 97:129-36.
75. Klemenz R, Frohli E, Aoyama A et al. αB-Crystallin accumulation is a specific response to Ha-ras and v-mos oncogene expression in mouse NIH 3T3 fibroblasts. Mol Cell Biol 1991; 11:803-12.
76. Kato K, Shinohara H, Kurobe N, Goto S, Inaguma Y, Ohshima K. Immunoreactive αA crystallin in rat non-lenticular tissues detected with a sensitive immunoassay method. Biochim Biophys Acta 1991; 1080:173-80.
77. Flaherty KM, McKay DB, Kabsch W, Holmes KC. Similarity of the three-dimensional structures of actin and the ATPase fragment of a 70-kDa heat shock cognate protein. Proc Natl Acad Sci USA 1991; 88:5041-5.
78. Hartl FU, Hlodan R, Langer T. Molecular chaperones in protein folding: the art of avoiding sticky situations. Trends Biochem Sci 1994; 19:20-5.
79. Reddy MC, Palmisano DV, Groth-Vasselli B, Farnsworth PN. 31P NMR studies of the ATP/α-crystallin complex: functional implications. Biochem Biophys Res Commun 1992; 189:1578-84.
80. Horwitz J. α-Crystallin can function as a molecular chaperone. Proc Natl Acad Sci USA 1992; 89:10449-53.
81. Miron T, Vancompernolle K, Vandekerckhove J, Wilchek M, Geiger B. A 25-kD inhibitor of actin polymerization is a low molecular mass heat shock protein. J Cell Biol 1991; 114:255-61.
82. Bennardini F, Wrzosek A, Chiesi M. αB-crystallin in cardiac tissue. Association with actin and desmin filaments. Circ Res 1992; 71:288-94.
83. Gopalakrishnan S, Takemoto L. Binding of actin to lens α-crystallins. Curr Eye Res 1992; 11:929-33.

84. Nicholl ID, Quinlan RA. Chaperone activity of α-crystallins modulates intermediate filament assembly. EMBO J 1994; 13:945-53.
85. Scotting P, McDermott H, Mayer RJ. Ubiquitin-protein conjugates and αB crystallin are selectively present in cells undergoing major cytomorphological reorganisation in early chicken embryos. FEBS Lett 1991; 285:75-9.
86. Aoyama A, Frohli E, Schafer R, Klemenz R. αB-crystallin expression in mouse NIH 3T3 fibroblasts: glucocorticoid responsiveness and involvement in thermal protection. Mol Cell Biol 1993; 13:1824-35.
87. FitzGerald PG, Graham D. Ultrastructural localization of αA-crystallin to the bovine lens fiber cell cytoskeleton. Curr Eye Res 1991; 10:417-36.
88. Mulders JWM, Stokkermans J, Leunissen JAM, Benedetti EL, Bloemendal H, de Jong WW. Interaction of α-crystallin with lens plasma membranes. Eur J Biochem 1985; 152:721-8.
89. Rameakers FCS, Selten-Versteegen AME, Bloemendal H. Interaction of newly synthesized α-crystallin with isolated lens plasma membranes. Biochim Biophys Acta 1980; 596:57-63.
90. Fleschner CR, Cenedella RJ. Examination of a lens 'native' plasma membrane fraction and its associated crystallins. Curr Eye Res 1992; 11:739-52.
91. Ifeanyi F, Takemoto L. Specificity of α-crystallin binding to the lens membrane. Curr Eye Res 1990; 9:259-65.
92. Alcala J, Maisel H. Biochemistry of lens plasma membranes and cytoskeleton. In: Maisel H, ed. The Ocular Lens. New York: Marcel Dekker, Inc., 1985; 169-222.
93. Quax-Jeuken Y, Quax W, van Rens G, Khan PM, Bloemendal H. Complete structure of the αB-crystallin gene: conservation of exon-intron distribution in the two non-linked α-crystallin genes. Proc Natl Acad Sci USA 1985; 82:5819-23.
94. Hawkins JW, Van Keuren ML, Piatigorsky J, Law ML, Patterson D, Kao F-T. Confirmation of assignment of the human α1-crystallin gene (CRYA1) to chromosome 21 with regional localization to q22.3. Hum Genet 1987; 76:375-80.
95. Petersen MB, Jaworski CJ, Lewis JG, Antonarakis SE. PvuII and XhoI/EcoRV polymorphisms adjacent to the αA-crystallin (CRYA1) gene on human chromosome 21. Nucleic Acids Res 1990; 18:4300.
96. Ngo JT, Klisak I, Dubin RA, et al. Assignment of the αB-crystallin to human chromosome 11. Genomics 1989; 5:665-9.
97. van den Heuvel R, Hendriks W, Quax W, Bloemendal H. Complete structure of the hamster αA-crystallin gene. Reflection of an evolutionary history by means of exon shuffling. J Mol Biol 1985; 185:273-84.
98. Hendriks W, Leunissen J, Nevo E, Bloemendal H, de Jong WW. The lens protein αA-crystallin of the blind mole rat, *Spalax ehrenbergi*: Evolutionary change and functional constraints. Proc Natl Acad Sci USA 1987; 84:5320-4.
99. King CR, Piatigorsky J. Alternative RNA splicing of the murine αA-crystallin gene: protein-coding information within an intron. Cell 1983; 32:707-12.

100. Jaworski CJ, Chepelinsky AB, Piatigorsky J. The αA-crystallin gene: conserved features of the 5'-flanking regions in human, mouse, and chicken. J Mol Evol 1991; 33:495-505.
101. Jaworski CJ, Piatigorsky J. A pseudo-exon in the functional human αA-crystallin gene. Nature 1989; 337:752-4.
102. Jaworski CJ. The Human αA-crystallin Gene [Dissertation]. The George Washington University, 1992; 161.
103. Jaworski CJ. A re-assessment of mammalian αA-crystallin sequences using DNA sequencing: implications for anthropoid affinities of tarsier. J Mol Evol 1995; in press.
104. Frederikse PH, Dubin RA, Haynes JI, Piatigorsky J. Structure and alternate tissue-preferred transcription initiation of the mouse αB-crystallin/small heat-shock protein gene. Nucleic Acids Res 1994; 22:5686-94.
105. Srinivasan AN, Bhat SP. Complete structure and expression of the rat αB-crystallin gene. DNA Cell Biol 1994; 13:651-61.
106. Dubin RA, Ally AH, Chung S, Piatigorsky J. Human αB-crystallin gene and preferential promoter function in lens. Genomics 1990; 7:594-601.
107. Cohen LH, Westerhuis LW, Smits DP, Bloemendal H. Two structurally closely related polypeptides encoded by 14-S mRNA isolated from rat lens. Eur J Biochem 1978; 89:251-8.
108. Cohen LH, Westerhuis LW, de Jong WW, Bloemendal H. Rat α-crystallin A chain with an insertion of 22 residues. Eur J Biochem 1978; 89:259-66.
109. de Jong WW, Cohen LH, Leunissen JA, Zweers A. Internally elongated rodent α-crystallin A chain: resulting from incomplete RNA splicing? Biochem Biophys Res Commun 1980; 96:648-55.
110. King CR, Piatigorsky J. Alternative splicing of αA-crystallin RNA. Structural and quantitative analyses of the mRNAs for the αA2- and αAins-crystallin polypeptides. J Biol Chem 1984; 259:1822-6.
111. Hendriks W, Sanders J, de Leij L, Ramaekers F, Bloemendal H, de Jong WW. Monoclonal antibodies reveal evolutionary conservation of alternative splicing of the αA-crystallin primary transcript. Eur J Biochem 1988; 174:133-7.
112. Jaworski CJ, Kim HS, Piatigorsky J, Wistow G. The evolution of αAins-crystallin in mammals. Invest Ophthalmol Vis Sci 1990; 31:507.
113. Wistow G. Domain structure and evolution in α-crystallins and small heat-shock proteins. FEBS Lett 1985; 181:1-6.
114. Barker WC, Ketcham LK, Dayhoff MO. A comprehensive examination of protein sequences for evidence of internal gene duplication. J Mol Evol 1978; 10:265-81.
115. Arrigo AP, Suhan JP, Welch WJ. Dynamic changes in the structure and intracellular locale of the mammalian low-molecular-weight heat shock protein. Mol Cell Biol 1988; 8:5059-71.
116. Zantema A, Verlaan-De Vries M, Maasdam D, Bol S, van der Eb A. Heat shock protein 27 and αB-crystallin can form a complex, which dissociates by heat shock. J Biol Chem 1992; 267:12936-41.
117. Chen Q, Osteryoung K, Vierling E. A 21-kDa chloroplast heat shock protein assembles into high molecular weight complexes in vivo and in

Organelle. J Biol Chem 1994; 269:13216-23.
118. Kato K, Shinohara H, Goto S, Inaguma Y, Morishita R, Asano T. Copurification of small heat shock protein with αB-crystallin from human skeletal muscle. J Biol Chem 1992; 267:7718-25.
119. Gilbert W, Marchionni M, McKnight G. On the antiquity of introns. Cell 1986; 46:151-3.
120. Darnell JE, Doolittle WF. Speculations on the early course of evolution. Proc Natl Acad Sci USA 1986; 83:1271-5.
121. Palmer JD, Logsdon J JR. The recent origins of introns. Curr Opin Genet Dev 1991;1:470-477.
122. Rabaey M, Segers J. Changes in the polypeptide composition of the bovine corneal epithelium during development. In: Trevor-Roper PD, ed. Vith Congress of the European Society of Ophthalmology. London: Academic Press, 1981; 41-4.
123. Wistow G. Protein structure and introns. Nature 1993; 364:107.
124. Russnak RH, Jones D, Candido EP. Cloning and analysis of cDNA sequences coding for two 16 kilodalton heat shock proteins (hsps) in *Caenorhabditis elegans*: homology with the small hsps of *Drosophila*. Nucleic Acids Res 1983; 11:3187-205.
125. Harding JJ, Dilley KJ. Structural proteins of the mammalian lens: A review with emphasis on changes in development, aging and cataract. Exp Eye Res 1976; 22:1-73.
126. Bloemendal H, ed. Molecular and Cellular Biology of the Eye Lens. New York: Wiley, 1981.
127. Berman ER. Biochemistry of the eye. 1st ed. New York: Plenum Press, 1991 (Blakemore C, ed. Perspectives in Vision Research).
128. Wistow G, Piatigorsky J. Lens crystallins: evolution and expression of proteins for a highly specialized tissue. Ann Rev Biochem 1988; 57:479-504.
129. 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.
130. van Rens GL, de Jong WW, Bloemendal H. A superfamily in the mammalian eye lens: the β/γ-crystallins. Mol Biol Rep 1992; 16:1-10.
131. Quax-Jeuken Y, Driessen H, Leunissen J, Quax W, de Jong W, Bloemendal H. βs-Crystallin: structure and evolution of a distinct member of the βγ-superfamily. EMBO J 1985; 4:2597-602.
132. van Rens GL, Raats JM, Driessen HP et al. Structure of the bovine eye lens γs-crystallin gene (formerly βs). Gene 1989; 78:225-33.
133. Berbers GA, Hoekman WA, Bloemendal H, de Jong WW, Kleinschmidt T, Braunitzer G. Homology between the primary structures of the major bovine β-crystallin chains. Eur J Biochem 1984; 139:467-79.
134. Hejtmancik JF, Thompson MA, Wistow G, Piatigorsky J. cDNA and deduced protein sequence for the βB1-crystallin polypeptide of the chicken lens. Conservation of the PAPA sequence. J Biol Chem 1986; 261:982-7.
135. Duncan MK, Haynes JI, Piatigorsky J. The βA4-crystallin gene is tightly

linked to the βB1-crystallin gene in chickens. Gene 1995; in press.

136. Quax-Jeuken Y, Janssen C, Quax W, van den Heuvel R, Bloemendal H. Bovine β-crystallin complementary DNA clones. Alternating proline/alanine sequence of βB1 subunit originates from a repetitive DNA sequence. J Mol Biol 1984; 180:457-72.
137. Peterson CA, Piatigorsky J. Preferential conservation of the globular domains of the βA3/A1-crystallin polypeptide of the chicken eye lens. Gene 1986; 45:139-47.
138. Hogg D, Tsui L-C, Gorin M, Breitman ML. Characterization of the human β-crystallin gene HuβA3/A1 reveals ancestral relationships among the βγ-crystallin superfamily. J Biol Chem 1986; 261:12420-7.
139. Luchin SV, Zinovieva RD, Tomarev SI et al. Frog lens βA1-crystallin: the nucleotide sequence of the cloned cDNA and computer graphics modelling of the three-dimensional structure. Biochim Biophys Acta 1987; 916:163-71.
140. Slingsby C, Bateman OA. Quaternary interactions in eye lens β-crystallins: basic and acidic subunits of β-crystallins favor heterologous association. Biochemistry 1990;29:6592-9.
141. Hope JN, Chen HC, Hejtmancik JF. βA3/A1-crystallin association: role of the N-terminal arm. Protein Eng 1994; 7:445-51.
142. Bindels JG, Koppers A, Hoenders HJ. Structural aspects of bovine β-crystallins: physical characterization including dissociation-association behavior. Exp Eye Res 1981; 33:333-43.
143. Berbers GA, Boerman OC, Bloemendal H, de Jong WW. Primary gene products of bovine β-crystallin and reassociation behavior of its aggregates. Eur J Biochem 1982; 128:495-502.
144. Cooper PG, Carver JA, Truscott RJ. ^{1}H-NMR spectroscopy of bovine lens β-crystallin. The role of the βB2-crystallin C-terminal extension in aggregation. Eur J Biochem 1993; 213:321-8.
145. Ostrer H, Piatigorsky J. β-Crystallins of the adult chicken lens: relatedness of the polypeptides and their aggregates. Exp Eye Res 1980; 30:679-89.
146. Wistow G, Roquemore E, Kim HS. Anomalous behavior of βB1-crystallin subunits from avian lenses. Curr Eye Res 1991; 10:313-9.
147. Kleiman NJ, Chiesa R, Kolks MA, Spector A. Phosphorylation of β-crystallin B2 (βBp) in the bovine lens. J Biol Chem 1988; 263:14978-83.
148. Voorter CE, Bloemendal H, de Jong WW. In vitro and in vivo phosphorylation of chicken βB3-crystallin. Curr Eye Res 1989; 8:459-65.
149. Bloemendal H. Biosynthesis of lens crystallins. Chap. 5. In: Bloemendal H, ed. Molecular and cellular biolgy of the eye lens. New York: Wiley-Interscience, 1981; 189-220.
150. Bloemendal H. The Lens Proteins. In: Bloemendal H, ed. Molecular and Cellular Biology of the Eye Lens. New York: Wiley-Interscience, 1981; 1-47.
151. den Dunnen JT, van Neck JW, Cremers FP, Lubsen NH, Schoenmakers JG. Nucleotide sequence of the rat γ-crystallin gene region and compari-

son with an orthologous human region. Gene 1989; 78:201-13.
152. 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.
153. Tomarev SI, Zinovieva RD, Chalovka P, Krayev AS, Skryabin KG, Gause GG Jr. Multiple genes coding for the frog eye lens γ-crystallins. Gene 1984; 27:301-8.
154. Smolich BD, Tarkington SK, Saha MS, Grainger RM. *Xenopus* γ-crystallin gene expression: Evidence that the γ-crystallin gene family is transcribed in lens and non-lens tissues. Mol Cell Biol 1994; 14:1355-63.
155. Chiou S-H, Chang W-P, Lo C-H, Chen S-W. Sequence comparison of γ-crystallins from the reptilian and other vertebrate species. FEBS Lett 1987; 221:134-8.
156. Chang T, Jiang YJ, Chiou SH, Chang WC. Carp γ-crystallins with high methionine content: cloning and sequencing of the complementary DNA. Biochim Biophys Acta 1988; 951:226-9.
157. Croft LR. Amino and carboxy terminal sequence of γ-crystallin, from haddock lens. Biochim Biophys Acta 1973; 295:174-7.
158. Chang T, Chang WC. Cloning and sequencing of a carp βs-crystallin cDNA. Biochim Biophys Acta 1987; 910:89-92.
159. Zarina S, Abbasi A, Zaidi ZH. Primary structure of βs-crystallin from human lens. Biochem J 1992; 287:375-81.
160. Zarina S, Slingsby C, Jaenicke R, Zaidi ZH, Driessen H, Srinivasan N. Three-dimensional model and quaternary structure of the human eye lens protein γs-crystallin based on β-and γ-crystallin X-ray coordinates and ultracentrifugation. Protein Sci 1994; 3:1840-6.
161. McDevitt DS, Croft LR. On the existence of γ-crystallin in the bird lens. Exp Eye Res 1977; 25:473-81.
162. van Rens GL, de Jong WW, Bloemendal H. One member of the γ-crystallin gene family, γs, is expressed in birds. Exp Eye Res 1991; 53:135-8.
163. Ostrer H, Beebe DC, Piatigorsky J. β-crystallin mRNAs: differential distribution in the developing chicken lens. Dev Biol 1981; 86:403-8.
164. Hejtmancik JF, Beebe DC, Ostrer H, Piatigorsky J. δ- and β-crystallin mRNA levels in the embryonic and posthatched chicken lens: temporal and spatial changes during development. Develop Biol 1985; 109:72-81.
165. Murer-Orlando M, Paterson RC, Lok S, Tsui LC, Breitman ML. Differential regulation of γ-crystallin genes during mouse lens development. Dev Biol 1987; 119:260-7.
166. van Leen RW, van Roozendaal KEP, Lubsen NH, Schoenmakers JGG. Differential expression of crystallin genes during development of the rat eye lens. Dev Biol 1987; 120:457-64.
167. Aarts HJM, Lubsen NH, Schoenmakers JGG. Crystallin gene expression during rat lens development. Eur J Biochem 1989; 183:31-6.
168. Voorter CE, de Haard-Hoekman WA, Hermans MM, Bloemendal H, de

Jong WW. Differential synthesis of crystallins in the developing rat eye lens. Exp Eye Res 1990; 50:429-37.

169. Goring DR, Breitman ML, Tsui LC. Temporal regulation of six crystallin transcripts during mouse lens development. Exp Eye Res 1992; 54:785-95.
170. 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.
171. Kabasawa I, Tsunematsu Y, Barber GW, Kinoshita JH. Low molecular weight proteins of the bovine lenses. Exp Eye Res 1977; 24:437-48.
172. Russell P, Meakin SO, Hohman TC, Tsui LC, Breitman ML. Relationship between proteins encode by three human γ-crystallin genes and distinct polypeptides in the eye lens. Mol Cell Biol 1987; 7:3320-3.
173. Siezen RJ, Thomson JA, Kaplan ED, Benedek GB. Human lens γ-crystallins: isolation, identification, and characterization of the expressed gene products. Proc Natl Acad Sci USA 1987; 84:6088-92.
174. Fernald RD, Wright SE. Maintenance of optical quality during crystalline lens growth. Nature 1983; 301:618-20.
175. Head MW, Peter A, Clayton RM. Evidence for the extralenticular expression of members of the β-crystallin gene family in the chick and a comparison with δ-crystallin during differentiation and transdifferentiation. Differentiation 1991; 48:147-56.
176. Blundell T, Lindley P, Miller L et al. The molecular structure and stability of the eye lens: x-ray analysis of γ-crystallin II. Nature 1981; 289:771-7.
177. 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.
178. Summers L, Wistow G, Narebor M et al. X-ray Structure of the Lens Specific Proteins: The Crystallins. In: Hearn MTW, ed. Peptide and Protein Reviews. v. 3. New York: Marcel Dekker, 1984:147-68.
179. Najmudin S, Nalini V, Driessen HPC et al. Structure of the bovine eye lens protein γB(γII)-crystallin at 1.47A. Acta Cryst 1993; D49:223-33.
180. Chirgadze YN, Sergeev YV, Fomenkova NP, Oreshin VD. Polypeptide chain pathway in γ-crystallin IIIb from calf lens a resolution. FEBS Lett 1981; 131:81-4.
181. Chirgadze Y, Nevskaya N, Vernoslova E et al. Crystal structure of calf eye lens γ-crystallin IIIb at 2.5 A resolution: its relation to function. Exp Eye Res 1991; 53:295-304.
182. White HE, Driessen HP, Slingsby C, Moss DS, Lindley PF. Packing interactions in the eye-lens. Structural analysis, internal symmetry and lattice interactions of bovine γIVa-crystallin. J Mol Biol 1989; 207:217-35.
183. Bax B, Lapatto R, Nalini V et al. X-ray analysis of βB2-crystallin and evolution of oligomeric lens proteins. Nature 1990; 347:776-80.
184. Lapatto R, Nalini V, Bax B et al. High resolution structure of an oligomeric eye lens β-crystallin. Loops, arches, linkers and interfaces in βB2 dimer compared to a monomeric γ-crystallin. J Mol Biol 1991; 222:1067-83.
185. den Dunnen JT, Moormann RJ, Lubsen NH, Schoenmakers JGG. Intron insertions and deletions in the β/γ-crystallin gene family: The rat

βB1 gene. Proc Natl Acad Sci USA 1986; 83:2855-9.
186. Rudolph R, Siebendritt R, Nesslauer G, Sharma AK, Jaenicke R. Folding of an all-β protein: independent domain folding in γII-crystallin from calf eye lens. Proc Natl Acad Sci USA 1990; 87:4625-9.
187. Mayr EM, Jaenicke R, Glockshuber R. Domain interactions and connecting peptides in lens crystallins. J Mol Biol 1994; 235:84-8.
188. Nalini V, Bax B, Driessen H, Moss DS, Lindley PF, Slingsby C. Close packing of an oligomeric eye lens β-crystallin induces loss of symmetry and ordering of sequence extensions. J Mol Biol 1994; 236:1250-8.
189. Wistow G, Slingsby C, Blundell T, Driessen H, de Jong W, Bloemendal H. Eye-lens proteins: the three-dimensional structure of β-crystallin predicted from monomeric γ-crystallin. FEBS Lett 1981; 133:9-16.
190. Berbers GA, Brans AM, Hoekman WA, Slingsby C, Bloemendal H, de Jong WW. Aggregation behavior of the bovine β-crystallin Bp chain studied limited proteolysis. Biochim Biophys Acta 1983; 748:213-9.
191. Carver JA, Cooper PG, Truscott RJ. ^{1}H-NMR spectroscopy of βB2-crystallin from bovine eye lens. Conformation of the N- and C-terminal extensions. Eur J Biochem 1993; 213:313-20.
192. Kroone RC, Elliott GS, Ferszt A, Slingsby C, Lubsen NH, Schoenmakers JGG. The role of the sequence extensions in β-crystallin assembly. Protein Engng 1994; 7:1395-9.
193. Hope JN, Chen HC, Hejtmancik JF. Aggregation of βA3-crystallin is independent of the specific sequence of the domain connecting peptide. J Biol Chem 1994; 269:21141-5.
194. Trinkl S, Glockshuber R, Jaenicke R. Dimerization of βB2-crystallin: the role of the linker peptide and the N- and C-terminal extensions. Protein Sci 1994; 3:1392-400.
195. Sergeev YV, Chirgadze YN, Mylvaganam SE, Driessen H, Slingsby C, Blundell TL. Surface interactions of γ-crystallins in the crystal medium in relation to their association in the eye lens. Proteins 1988; 4:137-47.
196. Blundell TL, Lindley PF, Miller LR et al. Interactions of γ-crystallin in relation to eye lens transparency. Lens Res 1983; 1:109-31.
197. Tardieu A, Veretout F, Krop B, Slingsby C. Protein interactions in the calf eye lens: interactions between β-crystallins are repulsive whereas in γ-crystallins they are attractive. Eur Biophys J 1992; 21:1-12.
198. Stevens A, Wang SX, Caines GH, Schleich T. ^{13}C-NMR off-resonance rotating frame spin-lattice relaxation studies of bovine lens γ-crystallin self association: effect of 'macromolecular crowding'. Biochim Biophys Acta 1995; 1246:82-90.
199. Pande J, Berland C, Broide M, Ogun O, Melhuish J, Benedek G. Suppression of phase separation in solutions of bovine γIV-crystallin by polar modification of the sulfur-containing amino acids. Proc Natl Acad Sci USA 1991; 88:4916-20.
200. Meakin SO, Breitman ML, Tsui LC. Structural and evolutionary relationships among five members of the γ-crystallin gene family. Mol Cell Biol 1985; 5:1408-14.
201. Bhat SP, Spector A. A simple method to characterize γ-crystallin synthe-

sized in vitro. Exp Eye Res 1984; 39:317-23.
202. Bagby S, Harvey TS, Eagle SG, Inouye S, Ikura M. NMR-derived three-dimensional solution structure of protein S complexed with calcium. Structure 1994; 2:107-22.
203. Bagby S, Harvey TS, Kay LE, Eagle SG, Inouye S, Ikura M. Unusual helix-containing greek keys in development-specific Ca2+-binding Protein S. ^{1}H, ^{15}N and ^{13}C assignments and secondary structure determined with the use of multidimensional double and triple resonance heteronuclear NMR spectroscopy. Biochemistry 1994; 33:2409-21.
204. David LL, Shearer TR, Shih M. Sequence analysis of lens β-crystallins suggests involvement of calpain in cataract formation. J Biol Chem 1993; 268:1937-40.
205. Wistow G, Summers L, Blundell T. *Myxococcus xanthus* spore coat protein S may have a similar structure to vertebrate lens βγ-crystallins. Nature 1985; 315:771-3.
206. Wistow G. Evolution of a protein superfamily: relationships between vertebrate lens crystallins and micro-organism dormancy proteins. J Mol Evol 1990; 30:140-5.
207. Wistow G, Jaworski J, Rao PV. A non-lens member of the βγ-crystallin superfamily in a vertebrate, the amphibian *Cynops*. Exp Eye Res 1995:in press.
208. Takabatake T, Takahashi TC, Takeshima K, Takata K. Protein synthesis during neural and epidermal differentiation in *Cynops* embryo. Develop Growth Differ 1991; 33:277-82.
209. Aarts HJ, Den Dunnen JT, Lubsen NH, Schoenmakers JG. Linkage between the βB2 and βB3 crystallin genes in man and rat: a remnant of an ancient β-crystallin gene cluster. Gene 1987; 59:127-35.
210. Hogg D, Gorin MB, Heinzmann C et al. Nucleotide sequence for the cDNA of the bovine βB2 crystallin and assignment of the orthologous human locus to chromosome 22. Curr Eye Res 1987; 6:1335-42.
211. Hulsebos TJ, Bijlsma EK, Geurts van Kessel AH, Brakenhoff RH, Westerveld A. Direct assignment of the human βB2 and βB3 crystallin genes to 22q11.2——q12: markers for neurofibromatosis 2. Cytogenet Cell Genet 1991; 56:171-5.
212. van Rens GL, Geurts van Kessel AH, Bloemendal H. Localization of the βA4-crystallin gene (CRYBA4) on human chromosome 22 in the region q11.2—›q13.1. Cytogenet Cell Genet 1992; 61:180-3.
213. Hulsebos TJM, Jenkins NA, Gilbert DJ, Copeland NG. The β crystallin genes on human chromosome 22 define a new region of homology with mouse chromosome 5. Genomics 1995; 25:574-6.
214. Law ML, Cai GY, Hartz J et al. Localization of a β-crystallin gene, Human βA3/A1 (gene symbol: CRYB1), to the long arm of human chromosome 17. Cytogenet Cell Genet 1986; 42:202-7.
215. Sparkes RS, Mohandas T, Heinzmann C, Gorin MB, Zollman S, Horwitz J. Assignment of a human β-crystallin gene to 17cen-q23. Hum Genet 1986; 74:133-6.

216. van Rens GL, Hol FA, de Jong WW, Bloemendal H. Presence of hybridizing DNA sequences homologous to bovine acidic and basic β-crystallins in all classes of vertebrates. J Mol Evol 1991; 33:457-63.
217. McFall-Ngai M, Horwitz J, Ding LL, Lacey L. Age-dependent changes in the heat-stable crystallin, βBp, of the human lens. Curr Eye Res 1986; 5:387-94.
218. Maiti M, Kono M, Chakrabarti B. Heat-induced changes in the conformation of α- and β-crystallin: unique thermal stability of α-crystallin. FEBS Lett 1988; 236:109-14.
219. Aarts HJM, Jacobs EH, van Willigen G, Lubsen NH, Schoenmakers JG. Different evolution rates within the lens-specific β-crystallin family. J Mol Evol 1989; 28:313-21.
220. Treton JA, Jones RE, King CR, Piatigorsky J. Evidence against γ-crystallin DNA or RNA sequences in the chick. Exp Eye Res 1984; 39:513-22.
221. den Dunnen JT, Jongbloed RJ, Geurts van Kessel AH, Schoenmakers JG. Human lens γ-crystallin sequences are located in the p12-qter region of chromosome 2. Hum Genet 1985; 70:217-21.
222. Lubsen NH, Renwick JH, Tsui LC, Breitman ML, Schoenmakers JG. A locus for a human hereditary cataract is closely linked to the γ-crystallin gene family. Proc Natl Acad Sci USA 1987; 84:489-92.
223. Shiloh Y, Donlon T, Bruns G, Breitman ML, Tsui LC. Assignment of the human γ-crystallin gene cluster (CRYG) to the arm of chromosome 2, region q33-36. Hum Genet 1986; 73:17-9.
224. Skow LC, Donner ME, Huang SM et al. Mapping of mouse γ-crystallin genes on chromosome 1. Biochem Genet 1988; 26:557-70.
225. den Dunnen JT, Szpirer J, Levan G, Islam Q, Schoenmakers JG. All six rat γ-crystallin genes are located on chromosome 9. Exp Eye Res 1987; 45:747-50.
226. den Dunnen JT, Moormann RJM, Lubsen NH, Schoenmakers JGG. Concerted and divergent evolution within the rat γ-crystallin gene family. J Mol Biol 1986; 189:37-46.
227. Graw J, Liebstein A, Pietrowski D, Schmitt-John T, Werner T. Genomic sequences of murine γB- and γC-crystallin-encoding genes: promoter analysis and complete evolutionary pattern of mouse, rat and human γ-crystallins. Gene 1993; 136:145-56.
228. Walls GL. The Vertebrate Eye and Its Adaptive Radiation. facsimile of 1942 edition. New York, NY: Hafner, 1967.
229. Walls GL. The reptilian retina. I. A new concept of visual cell evolution. Amer J Ophthal 1934; 17:892-915.
230. Thomson JA, Augusteyn RC. Ontogeny of human lens crystallins. Exp Eye Res 1985; 40:393-410.
231. Wistow G. Lens crystallins: a model system for gene recruitment. In: Zimmer EA, White TJ, Cann RL, Wilson AC, eds. Molecular Evolution. Producing the Biochemical Data. v. 224. San Diego, CA: Academic Press, 1993:563-75. (Methods in Enzymology).
232. Kumar S, Tamura K, Nei M. MEGA: Molecular Evolutionary Genetics Analysis software for microcomputers. Comput Appl Biosci 1994; 10:189-91.

CHAPTER 4

The Gene Recruitment of Enzymes as Crystallins

Taxon-specific enzyme crystallins are found in all reptiles and birds which have been examined, in amphibians of the genus *Rana* and in several species of mammal (Fig. 4.1). One atypical enzyme crystallin is also found in certain fish, lamprey, turtles and some other reptiles and birds. With this last exception, all the taxon-specific crystallins can be localized to specific lineages and their recruitment can be attributed to a single event in the ancestry of each lineage. Taxon-specific crystallins may modify the properties of the lens either through "diluting" the effects of γ-crystallins which contribute to hard, high-refractive index lenses, or through secondary protective effects such as UV filtration or contributions to protecting against oxidative or other stresses.

ε-CRYSTALLIN

Our understanding of crystallin gene recruitment and the occurrence of enzymes as taxon-specific crystallins began with ε-crystallin and in many ways this protein remains an archetype. It was discovered as a major component of the lenses in many birds and in all the crocodilians examined.[1] ε-Crystallin can be extremely abundant. In the lens of a hummingbird (*Calypte ana*) it was found to make up more than 40% of total soluble protein[2] (Fig. 4.2). In many other species, particularly seabirds, water fowl and others which hunt their food in bright light, ε-crystallin is 10-25% of total soluble protein.[3] The biggest surprise about ε-crystallin came when the protein from duck (*Anas platyrhynchos*) lens was partially sequenced and was found to be very closely related to lactate dehydrogenase B (LDHB), the heart muscle isoform of the glycolytic enzyme.[1,3] Peptide sequences of purified duck heart LDHB and duck lens ε-crystallins were identical except for age-related deamidation of two asparagine residues in the much older lens protein and ε-crystallin was found to have LDH activity.[3] Subsequently the identity of LDHB and ε-crystallin was confirmed by cDNA and genomic cloning which showed that the heart and lens proteins are indeed the products of the same single gene.[4,5]

Ubiquitous Stress-Protein Crystallins	*Related* or **Identical**	Activity	Cofactor	Species
αA, αB	**small heat shock protein** (mammalian αB)	Solubilize heat-stressed protein	none	all vertebrates
βA1-4, βB1-3 γA-F, γS	*M.xanthus Protein S* *P.polycephalum spherulin 3a* *EDSP of Cynops*	osmotic stress?	none	all vertebrates: γA-F not in birds
Taxon-specific Enzyme Crystallins				
δ1,δ2	**argininosuccinate lyase** (ASL)	δ2 has ASL activity	none	most birds, reptiles
ε	**lactate dehydrogenase B**	LDH activity	NADH	many birds, crocodiles
ζ	**NADPH:quinone oxidoreductase**	quinone reductase activity	NADPH	hystricomorphs, camels
η	**cytosolic aldehyde dehydrogenase**	retinal dehydrogenase	NAD+	elephant shrews
λ	*hydroxyacyl-CoA dehydrogenase*	?	NADH ?	rabbits, hares
μ	*ornithine cyclodeaminase* *glutamyl-tRNA reductase*	?	NADPH	some marsupials
π	**GAPDH**	GAPDH activity	NAD	geckos (*Phelsuma*)
ρ	*aldo-keto reductases*	?	NADPH	frogs (*Rana*)
τ	**α-enolase**	low enolase activity	none	some reptiles, birds several other species
Invertebrates				
S	glutathione-S transferase	weak GST activity in some subunits	GSH	octopus, squid
Ω	*ALDH*	no ALDH activity detected	NAD+?	octopus
J	?	?	none	jellyfish

Fig. 4.1. The crystallins.

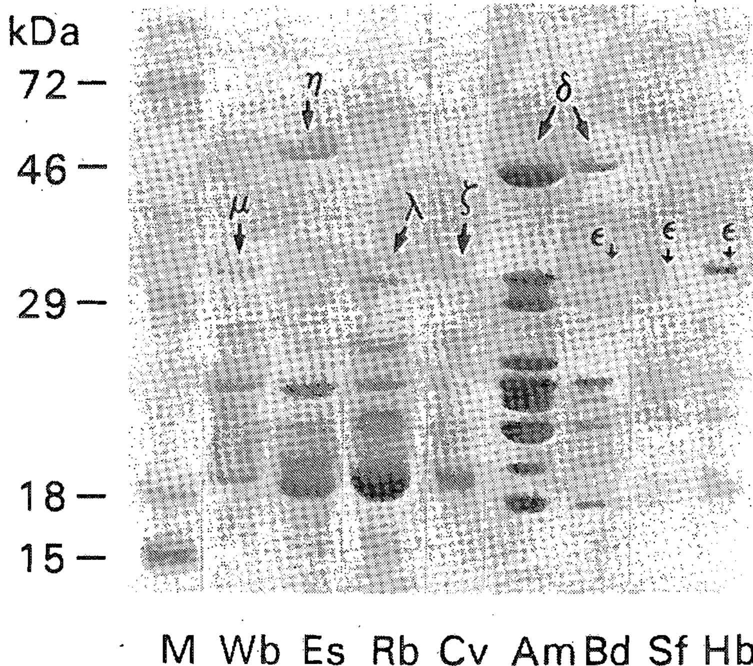


*Fig. 4.2. Taxon-specificity in crystallins. SDS PAGE[23] of lens extracts from some mammals and birds. Some major taxon-specific crystallins are indicated. M: size markers; Wb: tamar wallaby (*Macropus eugenii*); Es: elephant shrew (*Elephantulus rufescens*); Rb: rabbit (*Oryctolagus cuniculus*); Cv: rock cavy (*Kerodon rupestris*); Am: american merganser (*Mergus merganser*); Bd: black duck (*Anas rubripes*); Sf: chimney swift (*Chaetura pelagica*); Hb: Anna's hummingbird (*Calypte anna*).*

This was a remarkable and unexpected discovery. Instead of being a specialized lens structural protein, LDHB/ε-crystallin is the product of hundreds of millions of years of evolution as a glycolytic enzyme. It has become a crystallin by direct gene recruitment without prior gene duplication. This means that the same protein produced by a single gene is performing as a crystallin in the lens while still maintaining its normal pre-recruitment role as an enzyme in other tissues. Thus, in addition to the selective forces acting on this protein in its role as an enzyme, it now experiences another set of pressures from the new role in lens.

Sequence Changes and Adaptive Conflict

The effect of such pressures are apparent in LDHB/ε-crystallin itself. Although both enzyme and crystallin are identical in the same organism, sequence comparison with the LDHB polypeptides of other species reveal some unusual changes. In particular, two amino acid residues, Asn 114 and Phe 118, which are conserved in both LDHA

and LDHB sequences from species throughout the vertebrates, are changed to glycines in most species which have recruited the enzyme as a crystallin.[1,3,4] These two residues are close together on the surface of the protein, lying on the same side of an α-helix which runs across the top of the active site cleft in the LDH subunit (Fig. 4.3). The phenylalanine reside in particular forms an exposed, hydrophobic bump on each of the four subunits of the LDH tetramer. In duck LDHB/ε-crystallin, the substitution of glycines at positions 114 and 118 creates instead a flat patch on each subunit. This evidently has no beneficial effect on enzyme activity since it has never appeared in other vertebrate LDH sequences. Instead it must be due to the second role in lens. Given the critical aspects of protein-protein interactions in the lens it is likely that the modification serves to remove a potential site for protein aggregation.

This substitution of Phe 118 is found in many birds and in the crocodilian sequences which have been examined. Since one gene encodes both enzyme and crystallin, in those species which have recruited ε-crystallin, the LDHB enzyme in heart muscle also has this unusual sequence change. In most species which use ε-crystallin it appears that the modified LDHB functions well enough that these changes do not exert a significant evolutionary burden. However, there are two species of bird which have not followed this path. Both the chimney swift

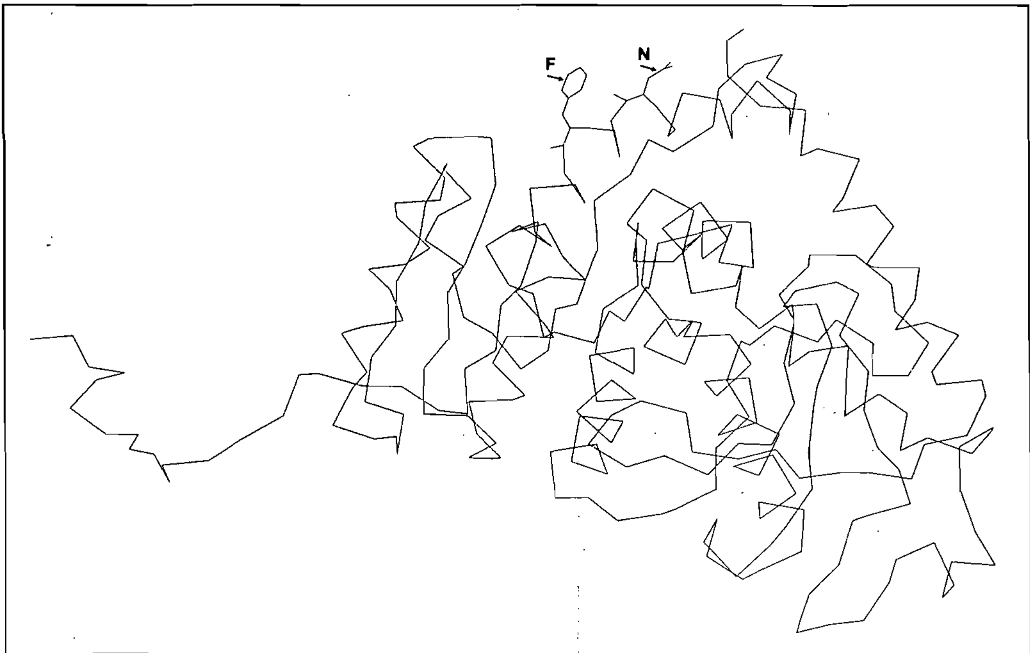


Fig. 4.3. Sequence modifications in an enzyme recruited as a crystallin. The exposed positions of Asn 114 (N) and Phe 118 (F) shown on a backbone trace of an LDH subunit.[150]

(*Chaetura pelagica*) and Ana's hummingbird (*Calypte ana*) have high levels of ε-crystallin in their lenses. These two species, although superficially very different, belong to the same order, Apodiformes.[6] They also share the characteristic of very energy-intensive life styles. Swifts spend a large part of their life on the wing at high speed in pursuit of insects while hummingbirds maintain an extremely high metabolic rate as they hover to collect nectar from flowers. When the ε-crystallins from the lenses of these two species were sequenced it was found that the Asn 114/Phe 118 combination of residues was still present.[2] Conceivably, in these species even a small decrease in the function of LDHB as an enzyme might have measurable consequences for the animal. As a result there may have been particular pressure for LDHB sequence to have been maintained to optimize its enzymatic role.

However, if the sequence changes in other species are indeed a response to selective pressures in the lens, then in swifts and hummingbirds there must have been a different response to accommodate these pressures. This may well have been the case, for both these species have reduced the content of another crystallin in their lens. In the case of the swift, δ-crystallin is completely absent as a detectable

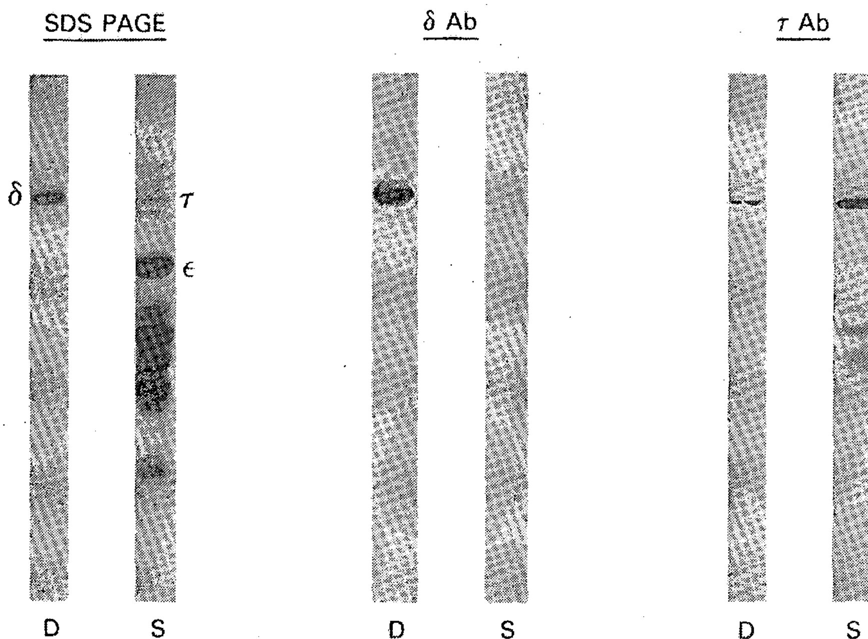


*Fig. 4.4. The swift (*Chaetura pelagica*) lacks δ-crystallin. SDS PAGE*[23] *and western blots of lens extracts from embryonic duck (*Anas platyrhynchos*) (as control) and from adult swift.*

crystallin (Fig. 4.4), while in the hummingbird it has been reduced to much lower levels than in other birds[2] (Fig. 4.2). One possible interpretation of this coincidence is that unmodified LDHB may interact with abundant δ-crystallin, perhaps through Asn114/Phe118, in a way which is detrimental to the lens. This can be overcome by a sequence modification to LDHB/ε-crystallin. However, if that solution is not advantageous an alternative would be to remove or reduce the δ-crystallin. A third solution which may also have been employed by many species of bird, is the reversion or loss of the ε-crystallin phenotype by LDHB. Many modern birds do not have abundant ε-crystallin in the lens. Since ε-crystallin is also present in crocodiles it was probably first recruited in a common ancestor of all archosaurs, including crocodiles, birds and dinosaurs. This implies that avian species which do not use this crystallin must have lost its expression after the divergence of birds and crocodiles.

This raises some interesting questions about the recruitment of ε-crystallin. It arose in lenses which already had δ-crystallin and which therefore had already adapted for the terrestrial environment. It must therefore have conferred some important secondary benefit to lens in order to have been retained by so many descendent species for so long. If the sequence modifications to LDHB were important they must have been made early on. This condition would have been retained in crocodiles, ducks and other birds to the present. Other birds, like chicken which lacks ε-crystallin,[3] must have lost ε-crystallin expression and undergone a reversion in their LDHB sequence, suggesting that there is indeed a real advantage to the function of LDHB in having the Asn 114/Phe 118 sequence. In contrast, as the ancestors of swifts and hummingbirds increased their energy budgets they too underwent a reversion in LDHB sequence but this was compensated by an alternative strategy which preserved high expression of ε-crystallin but reduced or removed δ-crystallin.

Protective Roles for Enzyme Crystallins

What special benefit could ε-crystallin bring to a lens? One suggestion is that although high levels of LDHB are not needed in lens for the purposes of glycolysis, the selective value of the protein comes from its ability to bind its cofactor NADH.[3] The reduced form of NAD+ (nicotinamide adenine dinucleotide) absorbs strongly in the near ultra violet at around 340 nm. Unlike those of mammals, bird retinas contain cone cell photoreceptors which have a peak sensitivity of 370 nm[7] allowing birds to see in the near UV. While this is undoubtedly useful under many circumstances it could also cause problems. Shorter wavelengths are scattered more efficiently by dust particles in the air, which is why the sky appears to be blue to our eyes. The blue-end of the spectrum thus contributes disproportionately to glare in bright light. For birds hunting insects against a brightly lit sky UV glare could be

a problem. It would also be a problem for birds and even crocodiles looking down through water for prey. UV glare would reflect off the surface while there would be little transmitted UV in the images from under the surface. Under these conditions it might be advantageous to filter out some of the UV. This could be achieved by sequestering NADH in the lens through binding to LDHB/ε-crystallin.

Strikingly, all the birds which use LDHB/ε-crystallin seem to fit this profile. They consist mainly of water hunters such as herons, gannets and gulls and bright light feeders such as swifts and hummingbirds.[1-3] Most birds which lack ε-crystallin, such as sparrows, chickens and owls feed on the ground, in low light or, in the case of penguins, entirely under water. Indeed, there is a fascinating correlation between the presence of ε-crystallin and an enigmatic feature of the avian retina, the pecten.[8] This is a folded, conical structure which protrudes from the back of the eye toward the lens. Unlike the rest of the avian retina it is vascularized and it is thought that it acts as a means of delivering nutrition to the eye. Interestingly, the degree of involution and hence the surface area of the pecten is higher in just the kind of bright light feeders which have ε-crystallin in their lenses.[8] Could one function of the pecten be to deliver nutrients to the lens to increase its content of NADH? Unfortunately there is no direct evidence to support this idea and the design of experiments to investigate it further has been daunting. However such a mechanism would allow birds dynamic control of the UV absorption of their lenses in response to environmental conditions.

δ-CRYSTALLIN

If ε-crystallin is the archetype for gene recruitment, gene sharing and the secondary benefits of recruitment, δ-crystallin has the same significance as a new model for gene duplication.

Although it was through ε-crystallin that the realization of the nature of taxon-specific enzyme crystallin came about, the first example of this class was already known though unrecognized as such. δ-Crystallin was observed as the first and most abundant of the soluble proteins of the developing chicken (*Gallus gallus*) lens.[9-11] Since these lenses also lack the γ-crystallins (as originally defined), δ-crystallin was seen a replacement for γ-crystallins in birds. In fact δ-crystallin is probably the most widespread and one of the oldest taxon-specific crystallins. Almost all birds, with the exception of swifts,[2] and probably all the reptiles which have been examined have abundant δ-crystallin (Fig. 4.2). It must therefore have been recruited at a very early stage in the reptile/bird lineage. At first it probably served to dilute the γ-crystallins which were present in reptilian lenses and are still present in at least some species today. Later, in the birds, δ-crystallin completely replaced the embryonic γ-crystallins and by itself formed as much as 90% of the soluble protein of the central, nuclear regions of the lens. The

difference in protein content and properties between the lens of a fish and that of a bird are striking. Bird lenses are among the softest and most hydrated known with typical protein contents of as little as 20%.[10] They often exhibit remarkable powers of accommodation and contribute to the unmatched visual acuity of birds.

δ-Crystallin is very different from γ-crystallin. It forms tetramers of 50 kDa subunits, giving it a native molecular size of 200 kDa, similar to that of the βH fraction. Unlike α-, β- and γ-crystallins which have predominantly β-sheet conformation, δ-crystallin has a high content of α-helix.[12,13] Only one form of δ-crystallin was ever cloned from chicken lens, but two similar genes, designated δ1 and δ2 were found closely linked in the chicken genome.[14-17] The predicted amino acid sequences of the products of these two genes showed 91% identity, and there was further strong similarity in non-coding sequences such as introns and untranslated regions of gene transcripts. δ1-Crystallin was the gene expressed at high levels in lens. δ2-Crystallin was named for its similarity to δ1 although there was no evidence that it actually served as a crystallin and its expression in lens was much lower than that of δ1-crystallin.[18]

When the first sequences for chicken δ-crystallin were produced no similarity to other proteins was noted. Coincidentally the sequence for yeast argininosuccinate lyase (ASL) was published at the same time.[19] However, at this early stage in the development of sequence databases the two sequences were not compared. It was not until the cDNA for human ASL was determined and entered into a database that the close similarity between chicken δ-crystallins and ASL was realized.[20-23] ASL is usually associated with the urea cycle in mammals but is also expressed in non-ureotelic tissues such as those of the eye,[24] and may be involved in various pathways such as nitric oxide synthesis.[25] Sequence comparisons showed that in fact it was δ2-crystallin which was most similar to the ASL enzymes of human and yeast.[20] Southern blot hybridization of chicken genomic DNA showed that the two δ-crystallin genes were the only ASL-like sequences present.[26] This suggested that δ2-crystallin was actually chicken ASL. But what about δ1-crystallin?

ASL enzyme activity was higher in embryonic chicken lens, which contains more than 80% δ-crystallin, than in other tissues[26] but this activity was very low compared to that of the purified human enzyme, showing that the chicken lens crystallin was not a fully active ASL enzyme. From these results it seemed that there had been a duplication of the ASL gene in birds. While one gene maintained the function of an enzyme, the other diverged and specialized as a crystallin. Since enzymatic activity was not essential for this new, structural role, δ1-crystallin lost its ancestral ASL activity. Thus chicken δ-crystallin appeared to follow a different model for taxon-specific crystallins from that exemplified by LDHB/ε-crystallin, a model without protein multifunctionality. Although the term "gene sharing" was first used in

connection with these proteins, the chicken δ-crystallins, unlike duck LDHB/ε-crystallin, have separate, specialized functions encoded by separate genes. However, chicken δ-crystallins do not tell the whole story of this family.

Previous work comparing the δ-crystallins of chicken with those of the duck (*Anas platyrhynchos*) had suggested that more isoforms of δ-crystallin were present in the duck.[27] Given the presence of two genes, this raised the possibility that in duck both genes were being expressed as crystallins. The first indication that this was correct came when crude duck lens extract was measured for ASL activity and was found to possess fully a quarter the activity of the purified human enzyme.[26] Thus in duck, unlike chicken, ASL/δ2-crystallin is indeed a multifunctional, taxon-specific enzyme crystallin. Later, full length cDNAs for both δ1- and δ2-crystallin were cloned from duck lens and both were found to be highly abundant.[28,29] Subsequently both duck[30] and chicken[31] δ1- and δ2-crystallins were expressed in different systems and for both species it was confirmed that δ2-crystallin is an active ASL while δ1-crystallin has no detectable activity.

ASL/δ2-crystallin is an enzyme crystallin in ducks, geese and swans, closely related members of the Anseriformes.[26,28,32] Most other birds which have been examined, such as pigeon (*Columba livia*)[33] have very low levels of lens ASL activity like the chicken. However, the ostrich (*Struthio camelus*), a ratite, also has high levels of ASL activity in its lens suggesting that in this flightless bird ASL/δ2-crystallin is expressed as a crystallin.[34] Since ducks and ostriches are so distantly related this suggests that the condition of two active δ-crystallin genes in the lens is ancestral. Following the model of LDHB/ε-crystallin and several other examples of taxon-specific enzyme crystallins it is very likely that ASL was recruited as a crystallin in an ancient ancestor of reptiles. Initially, like most other enzyme crystallins, one gene would have served two functions, with low level expression of the enzyme in various tissues and very high level expression in the lens.

The sequence changes forced on LDHB by the secondary role as ε-crystallin show how competing selective pressures can act on a bifunctional recruited gene. This can set up an adaptive conflict in which changes beneficial for one role are deleterious for the other. Under these circumstances there is selective advantage in gene duplication and specialization. This is what probably occurred at some point in the evolution of δ-crystallin. Duplication allowed one gene to adapt to whatever extra requirements the lens environment dictated. However it is clear that ASL itself can still function as a crystallin although it may require the presence of the more specialized δ1-crystallin to do so.

Although expression of ASL/δ2-crystallin can continue in the presence of δ1-crystallin it is evidently not required. Thus expression of ASL/δ2-crystallin in lens may be lost over time. In the same way non-lens expression of δ1-crystallin is also non essential and it too would be

expected to decline with time. Indeed this is essentially what seems to have happened in most birds which use only δ1-crystallin as a structural protein in the lens. Outside the lens, levels of mRNA for both δ1 and δ2 are generally very low but that for ASL/δ2-crystallin predominates, especially with age.[35,36] With the exception of another chicken eye tissue, the cornea, there is no evidence for expression of δ1-crystallin protein in non-lens tissues.[37] However if δ1-crystallin was expressed, its subunits would be capable of forming mixed tetramers with the enzymatically active ASL/δ2.[30] Indeed, when recombinant chicken δ1-crystallin was expressed in cultured mouse cells endogenous ASL activity was actually reduced, presumably by formation of mixed tetramers with lower activity.[31] Whether there is any benefit from such a non-lens role for δ1-crystallin in birds is unknown and perhaps unlikely. Certainly mammals are able to regulate ASL activity perfectly well by other means.

One prediction of the scenario presented here for the evolution of δ-crystallin is that at one time ancestors of birds had only one ASL/δ-crystallin gene. Recently we have obtained peptide sequences from the δ-crystallin of a reptile, the tuatara, a survivor of an ancient group, the sphenodonts. All the peptides of tuatara δ-crystallin examined seem to come from one sequence which more closely resembles δ2-crystallin and human ASL than it does δ1-crystallin. This is at least suggestive that this reptile expresses ASL as a crystallin and may not have acquired the lens-specialized δ1-crystallin. DNA analysis is now needed to investigate the number of ASL/δ-crystallin-like genes in this animal.

His89/Gln89

The lack of ASL activity in δ1-crystallin is rather surprising, considering that in the duck, δ1- and δ2-crystallins are 94% identical in amino acid sequence.[28] In both chicken and duck the two genes are closely linked in the same orientation and separated by only 4-4.5 kb.[16,29] This seems to have made them prone to gene conversion, and for both species the two δ-crystallins are more similar to each other than either is to its homologue in the other species. However, the enzymatically inactive δ1-crystallins in chicken and duck conserve at least one key amino acid difference from δ2-crystallins and from other ASL sequences. While active enzymes have histidine at position 89 this is replaced by glutamine in δ1-crystallins. Mutation of His89 to glutamine in human ASL reduces enzyme activity by over 90%.[38] Since activity is not eliminated, this residue is probably not involved in the central reaction of catalysis but has some important associated role. Surprisingly, a cDNA clone for a pigeon δ-crystallin was found to code for His89 and to be more similar to δ2-crystallins than to δ1-crystallins.[33] Since pigeon lens extract and total δ-crystallin has low ASL activity the authors concluded that His89 does not have an important enzy-

matic role. However since the clone was derived by PCR there is no evidence that it corresponds to the majority of expressed protein in pigeon lens and may in fact represent a low abundance δ2-crystallin.

Since the His89/Gln89 sequence change in chicken and duck δ1-crystallins is maintained in spite of gene conversion it may be under strong selection for some aspect of crystallin function. It seems unlikely that the selection is specifically for the loss of enzyme activity, since high levels of ASL activity in lens are not per se a problem for several species. The structures of turkey δ1-crystallin[13] and a modified form of duck δ2-crystallin[39] have now been determined by x-ray crystallography. Both reveal highly symmetrical tetramers arranged to form remarkable core bundles of 20 α-helices. These structures also show that His89 is not part of the active site of the enzyme although it is exposed on the surface nearby. Perhaps this residue has an accessory role in facilitating binding or release of substrates or products or a role in a conformational change necessary for function. Since the residue is surface exposed its change from histidine to glutamine may optimize or stabilize an interaction beneficial to its role as crystallin, such as a specific interaction with lens cytoskeleton. Alternatively this change might help maintain a more stable, less flexible structure for δ-crystallin. As a by product, this change contributes to a loss of enzymatic function but this is probably not the important, selected outcome.

Elements of Neutrality

δ-Crystallins also serve to illustrate another important point about enzyme crystallins in general. That is that these proteins are not recruited for their enzyme activity. Bird lenses do not need high levels of ASL activity but neither are such high levels of activity harmful. Birds do not even have a specific structural requirement for δ-crystallin. As shown by the chimney swift, δ-crystallin gene expression in lens can be entirely eliminated, perhaps as a response to adaptive conflict.[2] Its place in the swift lens is apparently taken quite adequately by LDHB/ε-crystallin while in the barn swallow (*Hirundo rustica*), a bird with rather similar hunting methods to those of swifts, δ-crystallin is abundant as normal in most other birds and reptiles. In other words the requirement is for a structural protein, not an enzyme, but even at this level the choice of a particular protein is at least partially neutral.

Intron-Slippage in Duck ASL/δ2-Crystallin

While the origins of introns are controversial,[40-42] it seems to be accepted that they may contribute to protein and genome evolution in many ways, including acting as sites for insertion of new coding sequences into existing stable structures. In duck ASL/δ2-crystallin additional protein sequence has been gained through a recent splice-site slippage. A single base change (GT->GC) in the splice site recognition

sequence has led to use of a cryptic site six nucleotides further into the intron.[28,29] This causes the in frame insertion of two amino acids into the N-terminal region of the enzyme sequence. This insertion is unique in the ASL family in eukaryotes.[28] However it has evidently not had any major deleterious effect on enzyme activity since duck δ2-crystallin is an active ASL.[30] Whether, on the other hand, the insertion has any beneficial effects for the role of the protein in the lens is not known, although it is interesting that it creates an RGD tripeptide, a cell attachment motif.[43] In many cases, homologous genes of very distantly related species have introns in similar but non identical positions. This could reflect either independent insertion of introns at susceptible regions or slippage of the whole intron. Duck ASL/δ2-crystallin illustrates the first stage in such slippage. A reciprocal slip at the 3' end of the same intron could restore the number of amino acid residues but move the intron in a 3' direction.

ρ-CRYSTALLIN

ρ-Crystallin, the only taxon-specific crystallin identified so far in amphibians, is found in the lenses of frogs of the genus *Rana*. Although its enzymatic specificity is not known, ρ-crystallin binds NADPH (nicotinamide adenine dinucleotide phosphote) and belongs to an aldo-keto reductase superfamily which includes aldehyde and aldose reductase, prostaglandin F synthase and several detoxification enzymes.[44-47] The x-ray structure analysis of aldose reductase shows that this superfamily uses a structural motif consisting of eight parallel βα units which was first observed in triose phosphate isomerase (TIM) and now known to be very common among enzymes.[48]

τ-CRYSTALLIN

α-Enolase/τ-crystallin is in many ways atypical of enzyme crystallins. However it also serves to illustrate some general processes of crystallin gene recruitment. Unlike every other known example of enzyme crystallins, α-enolase/τ-crystallin has a rather widespread and patchy distribution. Furthermore while all other enzyme crystallins in vertebrates are associated with terrestrial species it is the only one known to be prominently expressed in some aquatic species. It was first discovered as the 48K protein in sea lamprey (*Petromyzon marinus*), where it was estimated to make up 13% of total lens protein, and as τ-crystallin in a turtle (*Pseudemys scripta*) where it was also very abundant[49,50] (Fig. 4.5). A survey of other species detected 48K/τ-crystallin in numerous species, including several reptiles, birds and fish such the air-breathing gar (*Lepisosteus oculatus*).[49] Its abundance varied among species, such that it was easily detected in domestic duck lenses but was barely detectable in chicken lens. At this time it was not thought be present in mammalian lenses.

Protein sequencing showed that turtle τ-crystallin was probably identical to α-enolase,[20] and this assignment allowed the interpretation

τ-crystallin/α-enolase in Birds and Reptiles

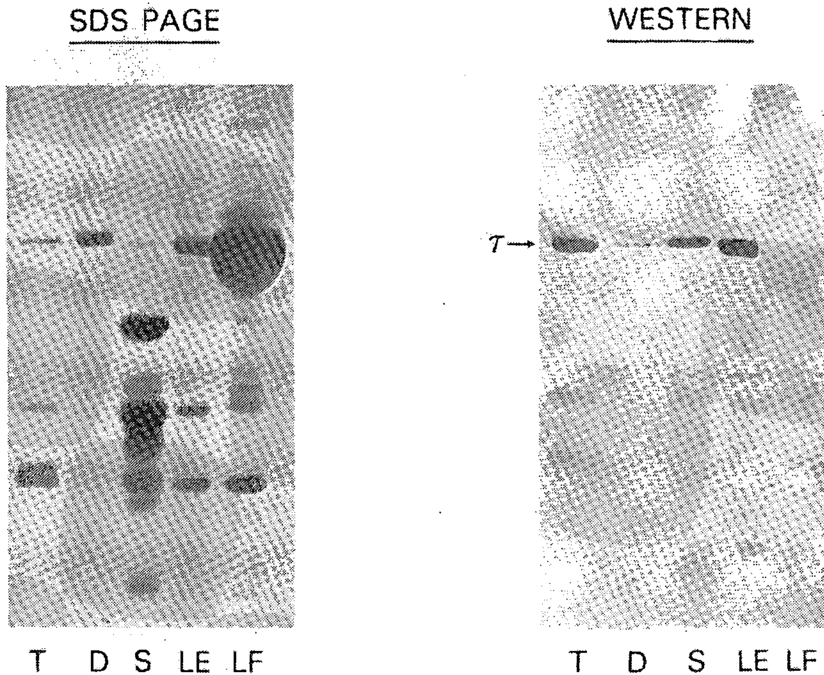


*Fig. 4.5. α-enolase/τ-crystallin (τ). SDS PAGE of lens extracts[23] (left) and western blot with anti-lamprey τ-crystallin antiserum (right). T: turtle (*Pseudemys scripta*); D: embryonic duck (*Anas platyrhynchos*); Sw: chimney swift (Chaetura pelagica); LE: embryonic chicken (*Gallus gallus*) lens epithelial cells; LF: embryonic chicken lens fiber cells.*

of peptide composition data from the lamprey protein which was also seen to match α-enolase.[51] Purified turtle τ-crystallin was shown to have enolase activity, albeit at a low level.[20,51] Since the previous survey had identified fairly high levels of τ-crystallin in duck lens, full-length cDNA for τ-crystallin was cloned from this source. Duck lens τ-crystallin was shown to be the product of a single gene and both τ-crystallin and human α-enolase cDNA probes hybridized to identical band in southern blots of duck genomic DNA.[51]

But is α-enolase/τ-crystallin a crystallin? In domestic duck, mRNA for α-enolase is more abundant in embryonic lens than in liver.[51] However the very high levels of expression in the lens seem to be a rather transient feature of embryogenesis. Analysis of the gene promoter for τ-crystallin found it to be highly active in all cultured cells with no tissue preference.[52] When the entire duck gene was expressed in transgenic

mice elevated levels of α-enolase were found in all tissues.[53] This experiment illustrated the point that lenses are capable of acquiring large increases (about seven fold in this case) in concentration of an enzyme in one step without serious problems. Indeed in the transgenic mice the duck transgene increased the level of α-enolase to close to parity with some individual β-crystallin subunits.[53] However in terms of expression patterns the duck gene behaved in adult mice in a very similar way to the endogenous α-enolase with no preference for lens.

These observations may be explained by the discovery that α-enolase is expressed at high levels in vivo in many stem cell populations, such as corneal limbus.[54] The function, if any, of high levels of α-enolase in stem cells is not known. It is possible that enolases may play structural roles in addition to their role in glycolysis. For example, γ-enolase has been found to be associated with the centrosome in HeLa cells[55] while several glycolytic enzymes are believed to form a cytomatrix with cytoskeletal proteins.[56-58]

Among the stem cells with high levels of α-enolase are lens epithelia (Fig. 4.5). α-Enolase/τ-crystallin has been measured at 9% of total protein in adult chicken epithelia and 12% in adult duck,[59] and it is likely that levels are higher in embryonic tissue. α-Enolase is at much lower levels in differentiated fiber cells (Fig. 4.5). It seems that α-enolase is an enzyme which is necessarily expressed at high levels in lens epithelia. In a small lens such as that of the lamprey, where epithelium constitutes a large fraction of total lens mass, α-enolase will accordingly achieve high overall levels. In other species such as mammals α-enolase is a prominent abundant enzyme but in a larger lens may not exceed .1% of total protein.[60]

Thus in many species α-enolase is intermediate between the low levels of many enzymes and the high levels of crystallins. Even then, it achieves concentrations in epithelial cells which are certainly in the structural range. Furthermore, overexpression of α-enolase does not seem to have any harmful effects in transgenic mice.[53] Thus *a priori* this enzyme is a good candidate for recruitment to even higher level expression in the lens. Indeed, there are clearly examples in which α-enolase unambiguously achieves the level of a crystallin. In turtle lens, for example, α-enolase/τ-crystallin is 46% of total protein in the epithelial cells and 6.5% in the nucleus giving an overall abundance of about 10%.[50,59] Immunohistochemistry shows high levels of α-enolase/τ-crystallin throughout the embryonic lens, including fiber cells.[51] The occurrence of τ-crystallin in diverse taxa may either reflect its "normal" high levels in epithelial cells or independent parallel recruitment of the same suitable gene in different lineages.

Intron Positions and Protein Structure

The gene for α-enolase/τ-crystallin from the domestic duck (*Anas platyrhynchos*)[52] and the homologous human α-enolase gene[61] have the

same pattern of intron positions. The structure of enolase from yeast has been determined by x-ray crystallography[62,63] and the sequence of enolase is sufficiently well conserved that a model of duck α-enolase can be built on the yeast enzymes' coordinates (unpublished). Enolase has two domains, an N-terminal domain of about 150 residues linked by a connecting peptide to a C-terminal domain of about 280 residues. The C-terminal domain was first thought to conform to the 8-fold β-strand/α-helix supersecondary structure repeat, $(\beta\alpha)_8$, typified by triose phosphate isomerase (TIM).[48,62] More detailed analysis showed that enolase actually has a different folding topology, $\beta\beta\alpha\alpha(\beta\alpha)_6$.[63] The TIM structural family, including other enzymes such as pyruvate kinase, has been used to illustrate the idea that all exons represent discrete structural elements and that genes were assembled from such units by intron-mediated mechanisms.[64] This was supported by the observation that introns in TIM map near the ends of β-strands or α-helices and not in the middle.

Enolase seems to contradict this model. First it has achieved a very similar tertiary structure through a different folding pattern and no significant sequence similarity. Second, introns mapping to the TIM barrel-like C-terminal domain of enolase do not neatly delineate secondary structures and three map to the sequences within α-helices (Fig. 4.6). These observations suggest that the βα barrel may simply be a thermodynamically stable structure available to many protein sequences regardless of their evolutionary origins.

However, the N-terminal domain of enolase is different. In this region introns fall neatly between supersecondary structures and one intron exactly corresponds to the join between the two domains (Fig. 4.6). Thus this domain could very well reflect the structures of ancestral motifs. Since enolase is present in prokaryotes[65] where it lacks introns the question again arises of whether the vertebrate introns were inserted or whether ancient introns were deleted in bacteria. It is possible that the N-terminal domain is descended from a well-structured RNA which directed intron insertion accordingly while the insertion into the C-terminal domain was essentially random.

GECKO CRYSTALLINS: A RESPONSE TO LIGHT?

The acquisition and loss of taxon-specific crystallins has been associated with changes such as moving from diurnal to nocturnal habits.[66] This recruitment could be driven either by the primary pressures of modifying the optical properties of the lens or by secondary pressures such as UV or oxidative stress. This has prompted an examination of geckos, lizards which adopt several different habits and which in some lineages may have moved from diurnality to nocturnality and back again.[8,67] Indeed, in two genera of diurnal geckos taxon-specific crystallins were found while none were apparent in nocturnal species.[68,69] In both *Phelsuma* and *Lepidodactylus* taxon-specific subunits of about

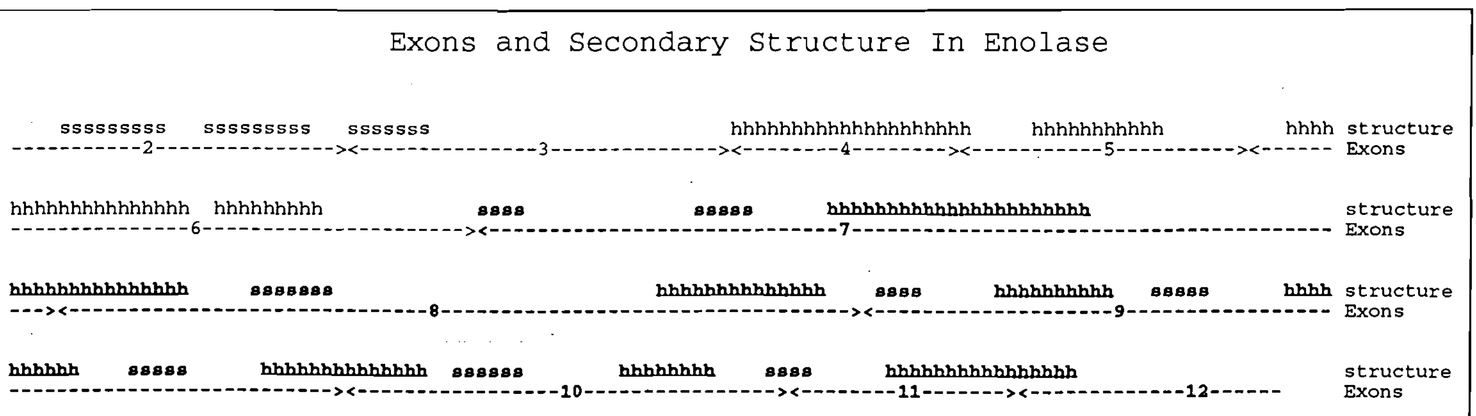


Fig. 4.6. Exons and secondary structure in enolase. Secondary structures of yeast enolase and a model of duck α-enolase/τ-crystallin are shown. (s) Residues in β-strands; (h) residues in α-helices. One character represents one residue. The relative positions of exons in the duck gene are shown by arrows.[52]
Exons 2-6 correspond to the N-terminal domain of the protein, exons 7-12 (bold face) correspond to the C-terminal αβ barrel.

36 kDa were observed, similar in size to several known crystallins. The protein in ***Phelsuma*** has been named π-crystallin and identified as the enzyme glyceraldehyde 3-phosphate dehydrogenase (GAPDH).[69] Interestingly GAPDH is expressed at moderately abundant but sub-crystallin levels in mammals. The protein in ***Lepidodactylus*** is also being characterized[68] and is different from π-crystallin (de Jong and Röll, unpublished).

These results suggest the independent recruitment of taxon-specific enzyme crystallins in response to a species moving into a brighter light environment. Since geckos are reptiles which already have high levels of "lens-softening" δ-crystallins, the new crystallins must be conferring secondary advantages, probably in protection against UV as was suggested for ε-crystallin and some other taxon-specific crystallins.[3,70] Interestingly, GAPDH, like LDH and enolase, is a glycolytic enzyme which has been shown to associate tightly with actin and other components of cytoskeleton.[56-58] These enzymes may be able to play a role in stabilization of cytoskeleton in lens fiber cells in addition to any other functions.

ENZYME CRYSTALLINS IN MAMMALS

The reptile/bird lineage presents a fairly uniform picture of the evolutionary processes by which lens composition is modified (Fig. 4.7). At some very early stage ASL/δ-crystallin was recruited and the adaptation of the lens to the terrestrial, diurnal environment began. Later, at least in birds, embryonic γ-crystallin expression was lost and in the archosaurs LDHB was recruited as ε-crystallin while diurnal geckos have also independently recruited different taxon-specific crystallins.[68,69] However the overall similarity in the retention of δ-crystallin in all descendants of early reptiles is remarkable.

Mammals too are descended from a major group of early terrestrial vertebrates, the synapsids, the so-called "mammal-like reptiles." This lineage includes the pelycosaurs which formed one of the dominant group of land animals in the early permian, the therapsids of the later permian and the cynodonts of the early triassic.[71] If it was useful for the reptile/bird line to modify their lenses with a recruited enzyme crystallin it seems reasonable to expect that the same applied to the ancestors of mammals. However, most modern placental mammals do not express taxon-specific crystallins and are instead limited to the ancient α- and β-crystallins and a group of very highly conserved γ-crystallins whose genes are closely clustered on one chromosome. Where taxon-specific crystallins do exist in mammals they are rather tightly limited in phylogenetic distribution. If the ancestors of mammals indeed had a taxon-specific crystallin it must have been lost by most descendent species which, in at least some cases, have independently recapitulated the process of gene recruitment (Fig. 4.7). Candidates for the role of ancestral synapsid taxon-specific crystallin are δ-crystallin itself which

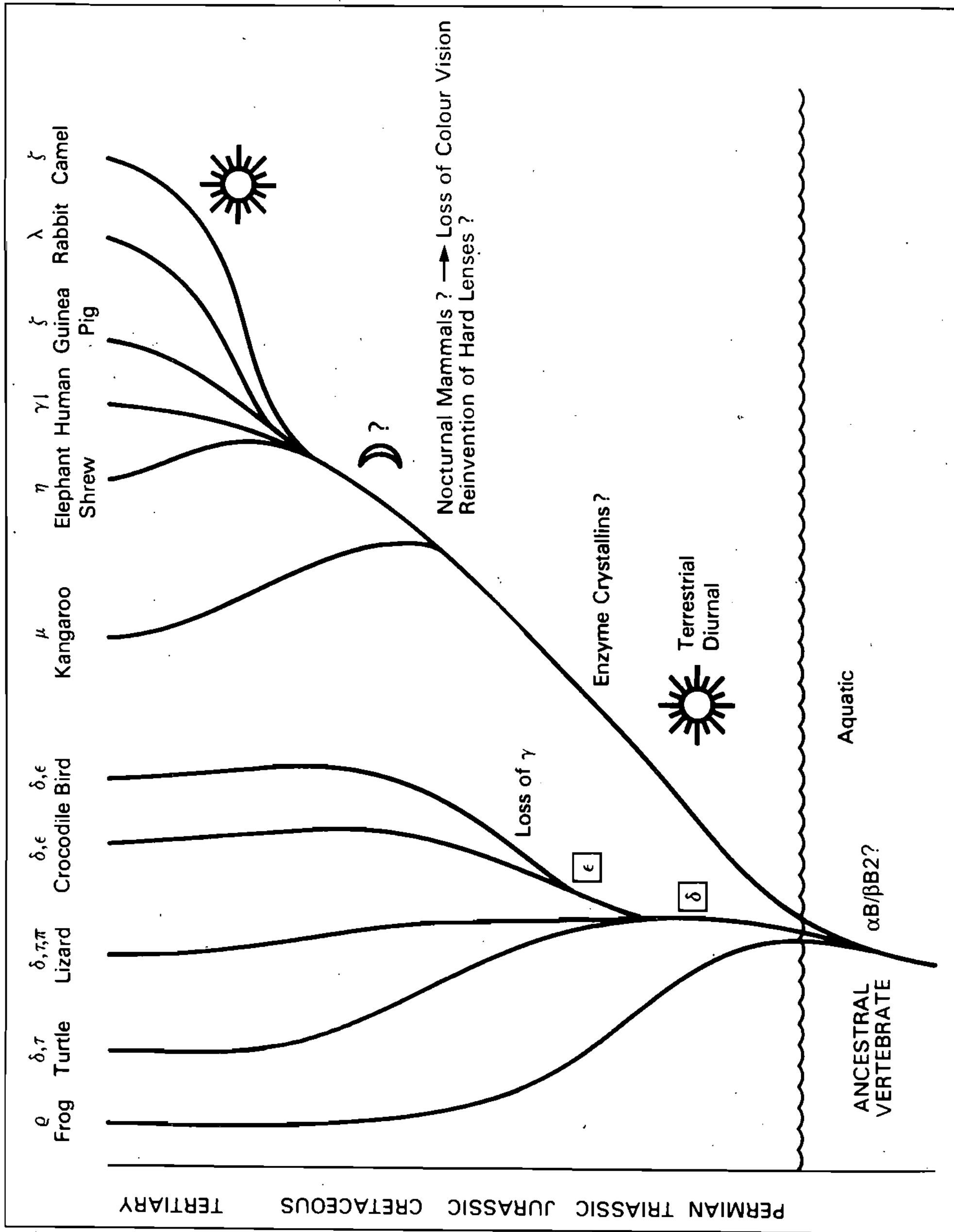


Fig. 4.7. An evolutionary tree for vertebrates illustrating the distribution of taxon-specific crystallins. Up-dated from Wistow G, Trends Biochem Sci 1993; 18:301-306.

might have been recruited in a common ancestor of both major reptilian lineages, μ-crystallin which has a wide but patchy distribution in marsupials, η-crystallin which is found in an ancient diurnal group of placentals and ζ-crystallin which is found in two separate lineages of modern placentals.

η-CRYSTALLIN

The most abundant taxon-specific crystallin in placental mammals is η-crystallin[60] which is probably the single major protein component in the lenses of macroscelids, elephant shrews (Fig. 4.2). These small mammals are active, diurnal insectivores. They are generally regarded as being relatively primitive and have even served as an illustration in a popular text for what some early mammalian ancestors might have been like.[72] Most experts have placed macroscelids in phylogenetic groups such as the Insectivora, which would include shrews, or the Glires, which in most classifications includes rodents and lagomorphs.[73] However, protein sequence data from the αA-crystallin of *Elephantulus rufescens*, the rufous elephant shrew, show identity with those of the hyrax.[60,74] This would place the macroscelids in an early offshoot of the placental mammal family tree which appropriately also includes elephants as part of a group called the paenungulates.[74] Whatever their affinities, elephant shrews are not shrews.[75] They have large eyes and soft lenses.

η-Crystallin accounts for about a quarter of total lens protein in both *E.rufescens* and *E.edwardi* while in a species of another genus, *Macroscelides proboscideus*, it contributes about 10% of total protein.[60] Whether this represents generic or individual difference is not known. In the one example of *E.rufescens* which was examined, η-crystallin seemed to have largely supplanted γ-crystallins in a manner reminiscent of δ-crystallin in birds.[60] In *M.proboscideus*, with its lower content of η-crystallin, γ-crystallins could be detected. Again, individual differences may be significant and the loss of γ-crystallins in the specimen of *E.rufescens* may have been due to aging effects in an old animal.

Partial protein sequence of η-crystallin and immunochemical reactivity suggested possible identity with cytoplasmic aldehyde dehydrogenase (ALDH1).[60] The very close sequence similarity of η-crystallin to ALDH1 has now been confirmed by cloning η-crystallin from the lenses of *E.edwardi* and *M.proboscideus*.[76] Sequence data show that η-crystallins clearly group with ALDH1 of other vertebrates and that all the residues required for ALDH enzymatic function[77] are conserved.

The cytoplasmic enzyme ALDH1 has very low activity towards the soluble aldehydes which are good substrates for related enzymes such as the mitochondrial ALDH2. However, ALDH1 is widely expressed and highly conserved suggesting that it has an important function.[77,78] It now appears that at least one role for this enzyme is as retinaldehyde (retinal) dehydrogenase,[79,80] an activity which converts the aldehyde retinal to retinoic acid, an important activator of gene expression and a po-

tent morphogen in development.[81,82] In the mouse eye, ALDH1 expression is an early marker for development in the retina with preferential expression in dorsal retina.[79,83] Indeed, it has been suggested that ALDH-derived retinoic acid might have a role in inducing *Pax-6* expression itself in the determination of the eye as an organ.[83]

When recombinant *E.edwardi* η-crystallin was tested for ALDH activity it was found to be enzymatically active using 11-*cis*-retinal as substrate.[76] Together with its close sequence similarity to ALDH1 this enzyme activity shows that η-crystallin is indeed an ALDH1. However, when non-lens expression of η-crystallin was examined there was a surprise. PCR analysis of *M.proboscideus* liver detected mRNA for both η-crystallin and a second, more abundant ALDH1 in liver.[76] Further PCR analysis of elephant shrew tissues suggests that η-crystallin is the predominant form of ALDH1 in retina and iris as well as in lens. Thus η-crystallin may have been recruited directly from an eye-preferred ALDH1 isoform. Alternatively η-crystallin may be the second known example, after δ-crystallin, in which the gene recruitment of a taxon-specific enzyme crystallin is associated with gene duplication. This idea is supported by cladistic analysis which suggests that the separation of the two genes for ALDH1 occurred at an early stage in the evolution of elephant shrews but probably after their lineage had split off from those of most other placental mammals.[76]

The predominance of η-crystallin in eye tissues explains why this taxon-specific crystallin retains its retinal dehydrogenase activity while serving as a crystallin. ALDH activity is essential for normal development in many tissues. In lens itself ALDH1 is expressed from early stages[79] and is present at reasonably high sub-crystallin levels in lenses of many species.[60] Retinoic acid receptors have been implicated in expression of γ-crystallin genes[84,85] while overexpression of retinoid binding proteins in lens causes developmental defects.[86,87] High level expression of a hypothetical inactive η-crystallin might sequester the active enzyme in low activity heterotetramers and essentially eliminate ALDH activity with possibly serious consequences for lens development. These problems would be avoided through the retention of ALDH activity by η-crystallin even as it acquired a structural role in lens.

Like some other enzyme crystallins, η-crystallin binds a nicotinamide adenine dinucleotide cofactor. However in this case the preferred cofactor is NAD+, the oxidized rather than the reduced form, although it is not yet known whether levels of both NAD+ and NADH are elevated in elephant shrew lenses. The recruitment of η-crystallin may have been selected through its modification of the optical properties of the lens rather for a secondary role in protection against oxidative stress. If a protective role does exist it may be against the toxic effects of aldehydes rather than against more generalized oxidative stress.

ζ-CRYSTALLIN

It was essential for our recognition of the enzyme crystallins that some of the first to be discovered were well known enzymes which had already entered the sequence databases.[23] Others however were discovered first as crystallins and only later defined as enzymes. One such is ζ-crystallin which was first discovered in guinea pig (*Cavia porcellus*) lens where it accounts for 7-10% of total soluble protein.[88] A congenital cataract in this species was found to be associated with loss of a 35 kDa crystallin subunit.[88,89] This turned out to be a taxon-specific crystallin belonging to the alcohol dehydrogenase (ADH) superfamily.[90,91] However, unlike ADH, ζ-crystallin binds NADPH rather than NADH, suggesting a role as a reductase.[92,93] After testing a number of possible substrates it was found that ζ-crystallin is a novel NADPH:quinone oxidoreductase.[94] Like LDHB/ε-crystallin, ζ-crystallin is the product of a single gene which is also expressed in other tissues at lower levels.

ζ-Crystallin was also detectable at crystallin levels in some related South American hystricomorph rodents including degu (*Octodon degus*) and rock cavy or moco (*Kerodon rupestris*) but not in coypu (*Myocasor coypu*) or in other rodents.[93] Then most surprisingly it was also found in camelids, Old World camels and New World llamas.[88,93,95,96] Both hystricomorphs and camelids have their origins in South America. However they are so distantly related phylogenetically that the presence of ζ-crystallin in both groups must have been due to independent recruitment. Recent gene sequencing for guinea pig and llama ζ-crystallins confirms the independence of the recruitments, although there are some remarkable parallels in mechanisms used (see ref. 97 and below).

Independent recruitment of the same enzyme as a crystallin should not perhaps have been totally unexpected since the pool of suitable enzymes is obviously finite and some may be easier to recruit than others. Indeed, as discussed below, parallel recruitment in widely divergent species may have occurred several times in the case of α-enolase/τ-crystallin.[49]

λ-CRYSTALLIN

λ-Crystallin seems to be another previously unknown enzyme. It has been found only in rabbit (*Oryctolagus cuniculus*) and hare (*Lepus europaeus*).[98] λ-Crystallin is distantly related to hydroxyacyl- and hydroxybutyryl-CoA dehydrogenases and conserves the consensus NADH binding site, however its activity, if any, has not yet been determined.[98] From a phylogenetic standpoint, it is interesting that λ-crystallin was not seen in pika (*Ochotona princeps*) which is classified as lagomorphs with rabbits and hares. Pikas also differ from the other lagomorphs in that they express αA^{ins}, the product of alternative splicing of αA-crystallin while rabbits do not.[99]

μ-CRYSTALLIN

μ-Crystallin was first observed as a 35 kDa subunit accounting for up to a quarter of total soluble protein in the lenses of some Australian marsupial mammals[60] (Fig. 4.2). So far it has been seen in all macropods (kangaroos, tree kangaroos and wallabies) which have been examined (unpublished). It was also identified in the lens of the only dasyurid examined, a carnivorous quoll, Dasyurops maculata.[60] It was not detectable in several other species, including various possums and wombat (unpublished) but a low level was seen in the sugar glider. It was not detectable by western blot in the only New World marsupial examined, the Virginia opossum (*Didelphis virginiana*).[60] Its pattern of occurrence suggests an early recruitment in a common ancestor of Australian marsupials perhaps 120million years ago and subsequent loss in several species. However more marsupial species, particularly from South America, need to be examined.

In contrast to placental mammals, marsupials retain in their retinas the cone cells with oil-drops found in diurnal reptiles and other vertebrate species.[8,67] This suggests that the ancestors of marsupials may not have experienced the nocturnal, burrowing phase which may have been an important part of placental mammal evolution (Fig. 4.7). Consequently marsupials may have retained other features from the eyes of their diurnal reptilian ancestors. If so, μ-crystallin may have an even more ancient origin than currently indicated and may be a candidate for the hypothetical ancestral enzyme crystallin of the reptilian ancestors of all mammals.

In grey kangaroo (*Macropus fuliginosus*) μ-crystallin is the product of a single gene which is expressed at high levels in lens and at lower levels in retina and brain, presumably in an enzymatic role.[100] Peptide sequences of μ-crystallin from kangaroo and quoll lenses could not convincingly demonstrate a relationship with any known proteins. However when kangaroo μ-crystallin was cloned it was found to be significantly similar (over 30% identity in predicted amino acid sequence) to ornithine cyclodeaminases (OCD) of the bacterium *Agrobacterium tumefaciens*.[100] The similarity was not apparent from peptide sequences because these all came from the more hydrophilic N- and C-terminal regions rather than the more hydrophobic central regions of the sequence which contain the highest similarity among OCDs and therefore presumably contain the active site.

OCD is an unusual enzyme involved in metabolism of opines, amino acid derivatives produced when *A. tumefaciens* invades plant cells.[101-103] OCD catalyses the conversion of ornithine directly to proline in the presence of NAD+.[102,103] In the more familiar biochemical pathways of standard textbooks this conversion requires two enzymes and passes through a glutamate semialdehyde intermediate. The mechanism of the OCD reaction is not known but it seems likely that it would involve a similar intermediate. Like OCD μ-crystallin binds a nicotinamide

adenine dinucleotide cofactor but in contrast to OCD it is NADPH rather than NAD+,[104] suggestive of a role as a reductase. Indeed both μ-crystallin and OCDs are related to another family of enzymes which are reductases, the glutamyl-tRNA reductases (GluTR) (Segovia and Wistow, in preparation). These unusual enzymes convert glutamyl-tRNA to a glutamate semialdehyde[105] using NADPH as cofactor. At this stage it seems reasonable to hypothesize that all three families of proteins are enzymes involved in unusual amino-acid metabolism and that they may share common affinities for derivatives of glutamate such as glutamate semialdehydes. This raises some interesting possibilities for the function of μ-crystallin.

Although μ-crystallin was discovered in marsupial lenses, it is conserved and expressed in other tissues in other vertebrates. Human μ-crystallin has been cloned from retina and brain and is over 80% identical to the kangaroo sequence. Northern blots of human mRNA detect μ-crystallin in neural, muscle and kidney tissue.[100] However, immunohistochemical detection of the protein shows a more restricted pattern. Although low levels are apparent in the epithelial and equatorial regions of human and rat lenses and may also be present as a general background in most nervous tissue, the highest levels of protein immunoreactivity are found specifically in the outer segments of the photoreceptors of the retina in human, rat and chicken (ref. 104 and in preparation). Indeed μ-crystallin is one of the earliest markers for photoreceptor development in embryogenesis. In this regard it is intriguing that another crystallin, η-crystallin, is derived from ALDH1 which is an even earlier marker for retinal development.[79]

The enzymatic function of μ-crystallin has still not been determined. All attempts to catalyze conversion of ornithine or proline with any nicotine dinucleotide cofactor have failed. Furthermore, although the protein, which appears to be a dimer, can easily be isolated from lens extracts using its ability to bind to blue sepharose it has proved very difficult to maintain its solubility (unpublished), an unexpected problem for a crystallin. However we do know that this protein is related to enzymes which metabolize derivatives of glutamate and that it is expressed in photoreceptors. It so happens that glutamate is the neurotransmitter of the photoreceptors[106] and that these cells are also highly susceptible to glutamate and ornithine toxicity.[78,107,108] Indeed, in the disease gyrate atrophy a systemic deficiency in ornithine metabolism is manifest as a rather specific syndrome of photoreceptor loss in the retina.[107,108] Perhaps μ-crystallin participates in metabolism of glutamate and as such plays an unanticipated role in the normal function of photoreceptors? Indeed, the human gene for μ-crystallin maps close to a breakpoint on chromosome 16 associated with cataract and micro-phthalmia.[109,110] Human and kangaroo genes for μ-crystallin have now been cloned and are being analyzed (unpublished).

Since gene expression in the retina may involve transcription factors which are shared by lens, such as Pax-6, μ-crystallin may have been a facile target for gene recruitment as a crystallin. The gene for μ-crystallin may have required only minor modifications to promoter sequences to allow high expression in lens.

CRYSTALLINS ELSEWHERE: INVERTEBRATE LENSES, LIGHT ORGAN LENS AND CORNEA

S-CRYSTALLINS

A variety of lenses are also found in the exceedingly diverse eyes of invertebrates. Most cephalopod mollusks such as octopus, squid and cuttlefish have eyes which are superficially very similar to those of vertebrates[8,111,112] (see Fig. 1.1). They have cellular lenses which seem to have arisen by convergent evolution to produce a similar solution to the problem of refining the optical properties of the eye. Like the lenses of fish, cephalopod lenses are extremely hard and relatively dehydrated, with protein content up to 70% wet weight.[113] Again like vertebrate lenses, most of the protein in these lenses consists of highly abundant soluble proteins with subunit sizes between 20 and 30 kDa which have been called S-(for squid) crystallins.[113] Although they are not related to any vertebrate crystallins S-crystallins share with taxon-specific crystallins a derivation from enzymes, in this case glutathione S-transferases (GST).[20,114] Unlike most examples of vertebrate enzyme crystallins, S-crystallins are generally lens-specific. They are also ubiquitous in cephalopod lenses, like α-, β- and γ-crystallins in vertebrates, rather than taxon-specific. S-crystallins are encoded by large gene families, again reminiscent of the multiple β- and γ-crystallins. Furthermore, most S-crystallins lack detectable enzymatic activity, although at least one squid protein does have some GST activity.[115]

It appears that the cephalopod lens is very ancient. This has allowed a greater degree of specialization for lens than is seen in vertebrates. The common ancestor of modern cephalopods must have recruited a GST as its original crystallin. Possibly this detoxification enzyme was already being expressed abundantly in the ocean-exposed eye in a protective role. Subsequently there was gene multiplication and specialization to produce more than a dozen lens-specific crystallins in the modern squid. Indeed, the specialization is such that the major GST enzyme of the squid is not expressed at high levels in lens. The antiquity of the cephalopod lens is further emphasized by the presence of other specialized components, such as a lens-specific tubulin.[116]

Sequence analysis of cephalopod S-crystallins shows they have the same exon-intron structure as mammalian GSTs.[117] This similarity, together with the degree of sequence similarity seen in various genes is consistent with the idea that these mollusks are more closely related to vertebrates than are other invertebrates such as arthropods.

ALDH: A Crystallin for All Seasons?

Although the predominant S-crystallins are ubiquitous in cephalopod lenses there is also at least one taxon-specific crystallin in this group. In addition to the S-crystallins, the octopus lens also contains a fairly prominent 59 kDa subunit which was named Ω-crystallin.[118] Surprisingly, when it was cloned, Ω-crystallin was found to be related to class 1 and 2 aldehyde dehydrogenases of vertebrates[119,120] and therefore distantly related to η-crystallin in elephant shrews.[60,76] However, unlike η-crystallin, Ω-crystallin has not conserved all the residues thought to be essential for enzymatic function and lacks detectable ALDH activity.[119,120] Again, it seems that the recruitment of this protein occurred sufficiently long ago that specialization has occurred.

Remarkably another member of the ALDH superfamily is also expressed abundantly in another lens-like tissue of the squid. The light organ of some squid emits light for signaling or camouflage purposes. Light is produced by luminescent bacteria and is diffused through a translucent "lens" derived from muscle tissue. A major soluble protein in this tissue has been named L-crystallin.[121] It too belongs to the ALDH superfamily, although it is not closely related to either η- or Ω-crystallins. This is another example in which soluble structural proteins of a functional lens have been recruited from an available enzyme.

The involvement of the ALDH superfamily in transparent tissues does not stop with the squid light organ. The vertebrate cornea consists mainly of a stroma of aligned bundles of collagen.[122] This is maintained by thin layers of cells on each surface, the outer epithelium exposed to the air and the endothelium which contacts the aqueous of the eye. Damage to these cells can lead to osmotic swelling of the stroma and opacity. Analysis of bovine corneal epithelium found a single major soluble protein component, BCP54.[123] Analysis of chicken cornea showed that this protein was not abundant,[37,124] showing that corneal proteins too may be taxon-specific. When BCP54 was characterized[125,126] it proved to be identical to ALDH III, the so-called tumor-inducible ALDH which is about 30% identical to the class I and II enzymes.[127] Surprisingly, the major site of constitutive expression of this inducible detoxification enzyme seems to be mammalian cornea.

Does this mean that ALDH superfamily members are in some way inherently transparent proteins? After all, the name "transparentin" was once proposed for BCP54.[128] In fact there is no reason to believe that these enzymes are "more transparent" than other proteins. Transparency depends on a medium lacking light absorbance and light scattering.[129,130] Absorbance is due to chromophores while scattering is due mainly to irregularly distributed objects or discontinuities which have sizes on the order of the wavelength of incident light. Since proteins are small it is more reasonable to talk of transparent solutions than transparent proteins.

So why are ALDH superfamily members so frequently recruited to transparent tissues? All these tissues are subject to osmotic stress, either in the swelling process which creates them or in maintaining their structure. Perhaps ALDH enzymes have some involvement in responses to such stress? In any case, it is likely that for some functional reason, these enzymes are easy to recruit under the conditions which give rise to transparent tissues. With their role as detoxification enzymes they may also play a protective role against toxic aldehydes which might result from oxidative insult, particularly to membranes. At least in terrestrial species they may also have a role in filtering UV radiation[125] and in this part of the spectrum they may actually be inherently opaque proteins.

What is clear is that overexpressed aldehyde dehydrogenases are not essential for lens or for cornea. Their recruitment is taxon-specific and many other species survive with only normal enzymatic levels of these proteins. BCP54/ALDHIII is only prominent in mammalian cornea while in chicken no single protein dominates to the same extent.[124] Instead several enzymes are quite abundant, including α-enolase[37] which other data show is localized to corneal limbus in mouse.[54] The most interesting discovery in chicken cornea was that enzymatically inactive δ1-crystallin, the major component of chicken lens is present at detectable levels.[37] There are two possible reasons why this lens-specialized protein might be abundant in cornea. The most likely is that cornea and lens share certain transcription factors, such as Pax-6.[131-136] Expression of a crystallin in the cornea could thus be due to overlap in transcriptional specificity of related tissues. Alternatively, whatever role δ1-crystallin has specialized to perform in lens, which might be something like stabilization of a particular kind of cytoskeleton or a role in osmoregulation, could also be beneficial for cornea. Again such a role cannot be essential for transparency since mammalian corneas survive quite well without expression of this protein.

OTHER CRYSTALLINS

Cellular lenses are also found in hydromedusan and cubomedusan jellyfish. The lens of the cnidarian *Cladonema radiatum* contains two major soluble proteins of 40 kDa and 70 kDa subunit size,[137] while certain cubomedusan jellyfish contain crystallins of 20 kDa and 35 kDa subunit size which have been named J-crystallins.[138] Sequence analysis of J1-crystallins, including cloning three intronless genes,[139] revealed no obvious identity with known proteins, although there is a weak but intriguing similarity to a region of the HSP60 family of molecular chaperones (Wistow, unpublished).

Other species of invertebrate have various kinds of acellular lenses. Some use inorganic materials while others use secreted structural proteins. Since crystallins have always been thought of as soluble proteins, these invertebrate eye proteins may not strictly qualify for this classification. Little is known about these proteins. Some at least are

probably conserved among widely divergent species since immunochemical methods suggest that the compound eyes of diverse arthropods contain related proteins.[140] A 52 kDa calcium-binding glycoprotein of the extracellular corneal lens of *Drosophila* compound eye has been partially characterized and named drosocrystallin.[141] It is not clear what the superfamily relationships of this protein might be or whether it is related to the common arthropod protein. Finally, there has been some analysis of the acellular lens of the mollusk *Aplysia californica* which contains protein subunits of two size ranges, about 60 kDa and 80 kDa.[142] However, since these subunits have identical N-terminal protein sequences they are likely to be products of the same gene, perhaps derived by post-translational modification. The limited sequence available does not reveal the relatedness, if any, of these proteins.

RECRUITMENT THROUGH MODIFIED GENE EXPRESSION

In the broadest sense of recruitment all that is necessary is for a protein to achieve very high concentrations in the lens. This could be accomplished by either transcriptional or post-transcriptional events. The latter are purely hypothetical but might in principle include tissue-specific enhancement of protein stability, or the stability or translational efficiency of mRNA. In fact no such mechanisms have yet been observed in the lens. Instead all the examples of recruitment hinge on tissue-specific increases in gene transcription (see Fig. 2.3).

TATA-Box Promoters

In some cases it is possible to compare homologous genes from species in which recruitment has and has not occurred. One intriguing observation which arises from this is that most of the enzyme crystallin genes, like those which encode the ubiquitous α-, β- and γ-crystallins, make use of TATA boxes to define the starting point for transcription while their non-recruited homologues lack TATA boxes and instead have GC-rich promoters of the kind associated with "housekeeping" genes (Fig. 4.8).

For example, the α-enolase/τ-crystallin gene of the duck has a TATA box and a single predominant transcription start site.[52] In contrast the promoter for human α-enolase, the only other α-enolase gene to have been sequenced, has a GC-rich promoter and multiple transcription start sites.[61] Similarly in chickens and ducks the genes for both the non-enzymatic δ1-crystallin and the enzymatically active ASL/δ2-crystallin have TATA boxes[14,15,29,143] while the genes for human and rat ASL again have "housekeeping" promoters.[144,145] In the case of ζ-crystallin in both guinea pig and llama, which have been recruited through acquisition of an alternative lens promoter, the upstream "enzyme promoter" is a GC-rich type while the downstream lens promoter again has a TATA box.[97,146,147]

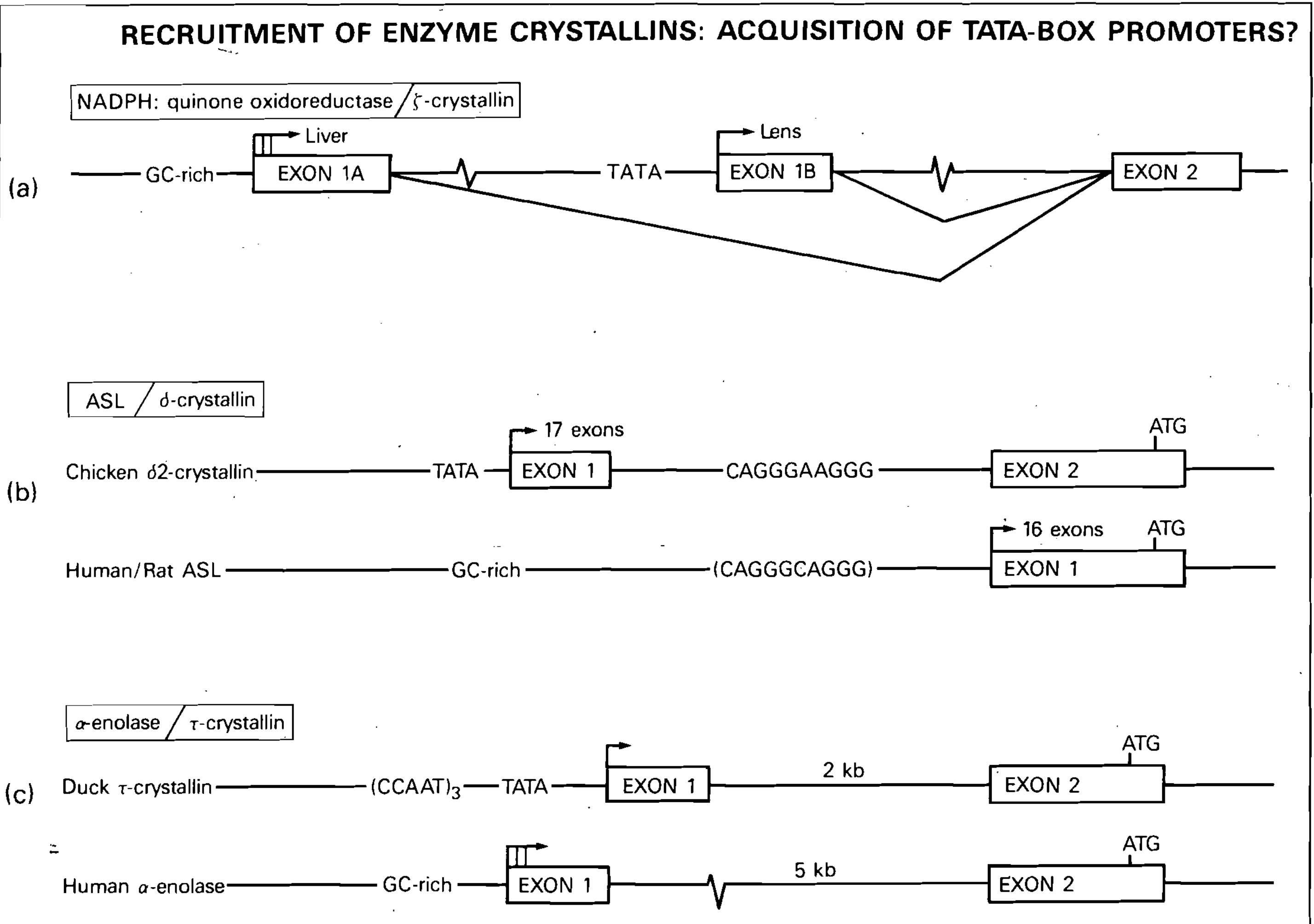


Fig. 4.8. Gene recruitment of enzyme crystallins is often associated with the presence of TATA box promoters. Taken from Harrington L et al, Dev Biol 1991; 148:508-516.

This suggests that at least one major route to recruitment requires the presence of a TATA box. This could reflect a preferred interaction between a lens-specificity factor and some component of the TFIID complex which is associated with binding to a TATA box but not to initiation complexes which form on non-TATA box promoters. However, there are always exceptions to any generalization about crystallins and it appears that duck LDHB/ε-crystallin lacks a TATA box altogether.[5]

New Promoters

In the case of the enzyme crystallins it is clear that genes expressed in many tissues have undergone sequence modification resulting in lens-specific overexpression. The most dramatic example of this is the acquisition of a second, lens-specific promoter by the NADPH:quinone oxidoreductase gene which gave rise to ζ-crystallin.[97,146,147] As described above, the 5' UTRs of ζ-crystallin mRNA in guinea pig lens and liver are different. They arise from alternative first exons spliced to a common second exon. Sequence analysis shows that the lens first exon lies downstream of that used in liver. Thus it appears to have been inserted into what would have been the first intron of the enzyme gene (Fig. 4.8). The guinea pig promoter has been defined by functional analysis.[134,137] It was noted that the complete promoter was neatly flanked by 9 bp direct repeats, one of which was upstream of all the functional elements while the other was in the first exon. When the ζ-crystallin promoter region was compared with the first intron of the homologous gene from mouse in which it has not been recruited as a crystallin there was no sequence similarity within the promoter itself.[146] However just downstream of the promoter region there was some conservation of sequence between species. This similarity begins close to the position of the direct repeat in the first exon of guinea pig ζ-crystallin.[147] This is at least consistent with the possibility that the lens promoter derives from a sequence which was inserted by transposition into the enzyme gene intron in an ancestor of guinea pigs.

ζ-Crystallin is also expressed in the lenses of camelids, including llama. Hystricomorph rodents like guinea pig and camelids like llama are sufficiently distant in evolutionary terms that the recruitment of ζ-crystallin in both must either represent an ancient ancestral feature of most mammals or else independent recruitment in two lineages. Yet gene sequencing shows that llama ζ-crystallin has also been recruited by insertion of an alternative promoter and first exon into the first intron of the same gene.[97] The insertion is close to the same position in guinea pig but not identical. Furthermore sequence alignments show that there is little conservation of promoter sequences but some conservation of both the alternative first exons. At present it is difficult to unravel the histories of the two genes. Their similarities are striking but so are their differences and many questions are yet to be answered.

Was ζ-crystallin ancestral to many mammals but lost in all but two lineages? Was the recruitment independent but directed in a similar way because of some feature of the enzyme gene which made it particularly prone to accepting transposons in its first intron? Both guinea pigs and camelids are of South American origin. Were they both subjected to a similar environmental pressure which led to the recruitment of a particular gene as a crystallin, or is it possible that they actually share a closer ancestry than expected?

Whatever the full story of ζ-crystallin recruitment turns out to be, it is clear that possession of two promoters has certain advantages for an enzyme crystallin gene. The two functions of enzyme and crystallin are separated allowing each promoter to specialize for its role without setting up an adaptive conflict at the level of gene expression. In spite of this, however, most other crystallin genes seem to make use of a single promoter for both modes of expression. Thus duck δ1-, δ2- and ε-crystallins all use the same transcription start site in lens and non-lens expression.[5,148]

However it is not out of the question that the δ-crystallins at one time made use of alternative promoters. When bird δ-crystallin genes are compared with their human homologues the most striking difference is in the number of exons.[66] The δ-crystallins have 17 while the mammalian ASL genes have only 16.[14,15,143-145] The difference lies in the presence of a 5' UTR exon in the bird genes (Fig. 4.8). It is conceivable that the recruitment of ASL occurred when the enzyme gene in an ancestral reptile gained a TATA-box containing lens promoter 5' to its housekeeping promoter. Both promoters could have co existed for a time. Eventually however the older, downstream promoter was lost. Intriguingly, a sequence proposed as a functional element for ASL and related genes in mammals can be found in the first intron of δ-crystallin genes in chicken and duck.[66,144] Could this be a sequence required for non-lens expression of these genes, a leftover from the original promoter which now works in cooperation with the newer upstream promoter?

Lens Specificity

Whatever the mechanism for promoter modification in crystallin gene recruitment, the result is that the gene acquires regulatory elements which respond to the peculiar transcriptional environments of the lens with high expression. In spite of the technical problems inherent in studying a tissue in which much of the gene expression occurs in terminally differentiated cells which will not grow in culture, the last few years have seen a great increase in our understanding of the detailed molecular mechanisms of crystallin gene expression. In particular there are exciting indications that lens expression depends upon tissue-restricted DNA binding proteins with roles in tissue determination, such as Pax-6 and SOX-2.[133-136,149] These results are described in the next section.

References

1. Stapel SO, Zweers A, Dodemont HJ, Kan JH, de Jong WW. ε-crystallin, a novel avian and reptilian eye lens protein. Eur J Biochem 1985; 147:129-36.
2. Wistow G, Anderson A, Piatigorsky J. Evidence for neutral and selective processes in the recruitment of enzyme-crystallins in avian lenses. Proc Natl Acad Sci USA 1990; 87:6277-80.
3. 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.
4. Hendriks W, Mulders JWM, Bibby MA, Slingsby C, Bloemendal H, de Jong WW. Duck lens ε-crystallin and lactate dehydrogenase B4 are identical: Single-copy gene product with two distinct functions. Proc Natl Acad Sci USA 1988; 85:7114-8.
5. Kraft HJ, Hendriks W, de Jong WW, Lubsen NH, Schoenmakers JGG. Duck lactate dehydrogenase B/ε-crystallin gene: lens recruitment of a GC-promoter. J Mol Biol 1993; 229:849-59.
6. Sibley CG, Ahlquist JE, Monroe BL. A classification of the living birds of the world based on DNA-DNA hybridization studies. Auk 1988; 105:409-23.
7. Chen DM, Collins JS, Goldsmith TH. The ultraviolet receptor of bird retinas. Science 1984; 225:337-40.
8. Walls GL. The Vertebrate Eye and Its Adaptive Radiation. facsimile of 1942 edition. New York, NY: Hafner, 1967.
9. Harding JJ, Crabbe MJC. The lens: Development, proteins, metabolism and cataract. In: Davson H, ed. The Eye. v. 1B. New York: Academic Press, 1984:207-492.
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. Piatigorsky J. δ-Crystallin and their nucleic acids. Mol Cell Biochem 1984; 59:33-56.
12. Yu NT, East EJ, Chang RC, Kuck JF. Raman spectra of bird and reptile lens proteins. Exp Eye Res 1977; 24:321-34.
13. Simpson A, Bateman O, Driessen H et al. The structure of avian lens δ-crystallin reveals a new fold for a superfamily of oligomeric enzymes. Nature Struct Biol 1994; 1:724-34.
14. Nickerson JM, Wawrousek EF, Hawkins JW et al. The complete sequence of the chicken δ1-crystallin gene and its flanking region. J Biol Chem 1985; 260:9100-5.
15. Nickerson JM, Wawrousek EF, Borras T et al. Sequence of the chicken δ2 crystallin gene and its intergenic spacer. Extreme homology with the δ1 crystallin gene. J Biol Chem 1986; 261:552-7.
16. Hawkins JW, Nickerson JM, Sullivan MA, Piatigorsky J. The chicken δ-crystallin gene family. Two genes of similar structure in close chromosomal approximation. J Biol Chem 1984; 259:9821-5.

17. Yasuda K, Okada TS. Structure and expression of chicken crystallin genes. Oxford Survey on Eukaryotic genes 1986; 3:183-209.
18. Parker DS, Wawrousek EF, Piatigorsky J. Expression of the δ-crystallin genes in the embryonic chicken lens. Dev Biol 1988; 126:375-81.
19. Beacham IR, Schweitzer BW, Warrick HM, Carbon J. The nucleotide sequence of the yeast ARG4 gene. Gene 1984; 29:271-9.
20. Wistow G, Piatigorsky J. Recruitment of enzymes as lens structural proteins. Science 1987; 236:1554-6.
21. Yeh LS, Elzanowski A, Hunt LT, Barker WC. Homology of δ-crystallin and argininosuccinate lyase. Comp Biochem Physiol [B] 1988; 89:433-7.
22. Doolittle RF. Lens proteins. More molecular opportunism. Nature 1988; 336:18.
23. Wistow G. Lens crystallins: a model system for gene recruitment. In: Zimmer EA, White TJ, Cann RL, Wilson AC, eds. Molecular Evolution. Producing the Biochemical Data. v. 224. San Diego, CA: Academic Press, 1993; 563-75. (Methods in Enzymology).
24. Rao GN, Cotlier E. Urea cycle enzymes in retina, ciliary body-iris, lens and senile cataracts. Exp Eye Res 1984; 39:483-95.
25. Hecker M, Sessa WC, Harris HJ, Anggard EE, Vane JR. The metabolism of L-arginine and its significance for the biosynthesis of endothelium-derived relaxing factor: cultured endothelial cells recycle L-citrulline to L-arginine. Proc Natl Acad Sci USA 1990; 87:8612-6.
26. Piatigorsky J, O'Brien WE, Norman BL et al. Gene sharing by δ-crystallin and argininosuccinate lyase. Proc Natl Acad Sci USA 1988; 85:3479-83.
27. Williams LA, Piatigorsky J. Heterogeneity of δ-crystallins of the embryonic mallard lens. Correlation between subunit compositions and isoelectric points. Biochemistry 1979; 18:1438-42.
28. Wistow G, Piatigorsky J. Gene conversion and splice-site slippage in the argininosuccinate lyases/δ-crystallins of the duck lens: members of an enzyme superfamily. Gene 1990; 96:263-70.
29. Li X, Wistow GJ, Piatigorsky J. Linkage and expression of the argininosuccinate lyase/δ-crystallin genes of the duck: insertion of a CR1 element in the intragenic spacer. Biochim Biophys Acta 1995; 1261:25-34.
30. Barbosa P, Wistow GJ, Cialkowski M, Piatigorsky J, O'Brien WE. Expression of duck lens δ-crystallin cDNAs in yeast and bacterial hosts: δ2-crystallin is an active argininosuccinate lyase. J Biol Chem 1991; 266:22319-22.
31. Kondoh H, Araki I, Yasuda K, Matsubasa T, Mori M. Expression of the chicken 'δ2-crystallin' gene in mouse cells: evidence for encoding of argininosuccinate lyase. Gene 1991; 99:267-71.
32. Chiou SH, Lee HJ, Chu H, Lai TA, Chang GG. Screening and kinetic analysis of δ-crystallins with endogenous argininosuccinate lyase activity in the lenses of vertebrates. Biochem Int 1991; 25:705-13.
33. Lin CW, Chiou SH. Sequence analysis of pigeon δ-crystallin gene and its deduced primary structure. Comparison of avian δ-crystallins with and without endogenous argininosuccinate lyase activity. FEBS Lett 1992; 311:276-80.

34. Chiou SH, Lo CH, Chang CY, Itoh T, Kaji H, Samejima T. Ostrich crystallins. Structural characterization of δ-crystallin with enzymic activity. Biochem J 1991; 273:295-300.
35. Li X, Beebe DC. Transcriptional control of δ-crystallin gene expression in the chicken embryo lens: demonstration by a new method for measuring mRNA metabolism. Mol Cell Biol 1993; 13:3282-90.
36. Li X, Zelenka PS, Piatigorsky J. Differential expression of the two δ-crystallin genes in lens and non-lens tissues: shift favoring δ2 expression from embryonic to adult chickens. Dev Dyn 1993; 196:114-23.
37. Cuthbertson RA, Tomarev SI, Piatigorsky J. Taxon-specific recruitment of enzymes as major soluble proteins in the corneal epithelium of three mammals, chicken, and squid. Proc Natl Acad Sci USA 1992; 89:4004-8.
38. Barbosa P, Cialkowski M, O'Brien WE. Analysis of naturally occuring and site-directed mutations in the argininosuccinate lyase gene. J Biol Chem 1991; 266:5286-90.
39. Abu-Abed M, Turner MA, Atkinson J, Dole K, Howell PL. Crystallization and preliminary X-ray analysis of a mutant duck δ-II crystallin. J Mol Biol 1994; 243:944-6.
40. Gilbert W, Marchionni M, McKnight G. On the antiquity of introns. Cell 1986; 46:151-3.
41. Dorit RL, Schoenbach L, Gilbert W. How big is the universe of exons? Science 1990; 250:1377-82.
42. Palmer JD, Logsdon J JR. The recent origins of introns. Curr Opin Genet Dev 1991; 1:470-.
43. D'Souza SE, Ginsberg MH, Plow EF. Arginyl-glycyl-aspartic acid (RGD): a cell adhesion motif. Trends Biochem Sci 1991; 16:246-50.
44. Carper D, Nishimura C, Shinohara T et al. Aldose reductase and ρ-crystallin belong to the same protein superfamily as aldehyde reductase. FEBS Lett 1987; 220:209-13.
45. Carper DA, Wistow G, Nishimura C et al. A superfamily of NADPH-dependent reductases in eukaryotes and prokaryotes. Exp Eye Res 1989; 49:377-88.
46. Watanabe K, Fujii Y, Nakayama K et al. Structural similarity of bovine lung prostaglandin F synthase to lens ϵ-crystallin [renamed ρ-crystallin] of the European common frog. Proc Natl Acad Sci USA 1988; 85:11-5.
47. Bohren KM, Bullock B, Wermuth B, Gabbay KH. The aldo-keto reductase superfamily. cDNAs and deduced amino acid sequences of human aldehyde and aldose reductases. J Biol Chem 1989; 264:9547-51.
48. Farber GK, Petsko GA. The evolution of α/β barrel enzymes. Trends Biochem Sci 1990; 15:228-34.
49. Stapel SO, de Jong WW. Lamprey 48-kDa lens protein represents a novel class of crystallins. FEBS Lett 1983; 162:305-9.
50. Williams LA, Ding L, Horwitz J, Piatigorsky J. τ-Cystallin from the turtle lens: Purification and partial characterization. Exptl Eye Res 1985; 40:741-9.
51. Wistow GJ, Lietman T, Williams LA et al. τ-crystallin/α-enolase: One

gene encodes both an enzyme and a lens structural protein. J Cell Biol 1988; 107:2729-36.
52. Kim RY, Lietman T, Piatigorsky J, Wistow GJ. Structure and expression of the duck α-enolase/τ-crystallin encoding gene. Gene 1991; 103:193-200.
53. Kim RY, Wistow GJ. Expression of the duck α-enolase/τ-crystallin gene in transgenic mice. FASEB J 1993; 7:464-9.
54. Zieske JD, Bukusoglu G, Yankauckas MA, Wasson ME, Kettmann HT. α-Enolase is restricted to basal cells of stratified squamous epithelium. Dev Biol 1992; 151:18-26.
55. Johnstone SA, Waisman DM, Rattner JB. Enolase is present at the centrosome of HeLa cells. Exp Cell Res 1992; 202:458-63.
56. Knull HR, Walsh JL. Association of glycolytic enzymes with the cytoskeleton. Curr Top Cell Regul 1992; 33:15-30.
57. Masters C. Microenvironmental factors and the binding of glycolytic enzymes to contractile filaments. Int J Biochem 1992; 24:405-10.
58. 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.
59. Rudner G, Katar M, Maisel H. Enolase in the avian and turtle lens. Curr Eye Res 1990; 9:139-50.
60. Wistow G, Kim H. Lens protein expression in mammals: taxon-specificity and the recruitment of crystallins. J Mol Evol 1991; 32:262-9.
61. Giallongo A, Oliva D, Cali L, Barba G, Barbieri G, Feo S. Structure of the human gene for a-enolase. Eur J Biochem 1990; 190:567-73.
62. Lebioda L, Stec B. Crystal structure of enolase indicates that enolase and pyruvate kinase evolved from a common ancestor. Nature 1988; 333:683-6.
63. Lebioda L, Stec B, Brewer JM. The structure of yeast enolase at 2.25A resolution. J Biol Chem 1989; 264:3685-93.
64. Straus D, Gilbert W. Genetic engineering in the precambrian: structure of the chicken triosephosphate isomerase gene. Mol Cell Biol 1985; 5:3497-506.
65. Wold F. Enolase. In: Boyer PD, ed. The Enzymes. v. 5. New York: Academic Press, 1971:499-538.
66. Wistow G. Lens crystallins: gene recruitment and evolutionary dynamism. Trends Biochem Sci 1993; 18:301-6.
67. Walls GL. The reptilian retina. I. A new concept of visual cell evolution. Amer J Ophthal 1934; 17:892-915.
68. Röll B. Crystallins in lenses of gekkonid lizards (Reptilia, Gekkonidae). Herpetol J 1995:in press.
69. Jimenez-Asensio J, Gonzalez P, Zigler JS, Garland DL. Glyceraldehyde 3-phosphate dehydrogenase is an enzyme crystallin of diurnal geckos of the genus Phelsuma. Biochem Biophys Res Commun 1995; 209:796-802.
70. Rao CM, Zigler JS Jr. Levels of reduced pyridine nucleotides and lens photodamage. Photochem Photobiol 1992; 56:523-8.

71. Carroll RL. Vertebrate Paleontology and Evolution. New York: Freeman, 1988.
72. Norman D. The Illustrated Encyclopedia of Dinosaurs. New York: Crescent Books, 1985.
73. Novacek MJ. Morphology, paleontology, and the higher clades of mammals. Chap. 11. In: Genoways HH, ed. Current Mammalogy. v. 2. New York: Plenum Publishing Corporation, 1990:507-43.
74. de Jong WW, Leunissen JAM, Wistow GJ. Eye lens crystallins and the phylogeny of placental orders: Evidence for a macroscelid-paenungulate clade? In: Szalay FS, Novacek MJ, McKenna MC, eds. Mammal Phylogeny: Placentals. New York, NY: Springer Verlag, 1993:5-12.
75. Nowak RM. Walker's Mammals of the World. 5th ed. v. 1. Baltimore and London: The Johns Hopkins University Press, 1991.
76. Graham C, Hodin J, Wistow G. Gene recruitment of cytoplasmic (retin)aldehdye dehydrogenase as η-crystallin in elephant shrew eye lens. 1995:in preparation.
77. Hempel J, Nicholas H, Lindahl R. Aldehyde dehydrogenases: widespread structural and functional diversity within a shared framework. Protein Sci 1993; 2:1890-900.
78. Weiner H, Wang TT, Farres J. Molecular biological studies on liver mitochondrial aldehyde dehydrogenase. Alcohol Alcohol Suppl 1991; 1:91-5.
79. McCaffery P, Lee MO, Wagner MA, Sladek NE, Drager UC. Asymmetrical retinoic acid synthesis in the dorsoventral axis of the retina. Development 1992; 115:371-82.
80. Champer RJ, Huang J, Asson-Batres MA, Saari JC. Cloning and sequencing of a cDNA encoding bovine retinal aldehyde dehydrogenase. Invest Ophthalmol Vis Sci 1994; 35:1465.
81. Gudas LJ. Retinoids and vertebrate development. J Biol Chem 1994; 269:15399-402.
82. Chambon P. The retinoid signaling pathway: molecular and genetic analyses. Semin Cell Biol 1994; 5:115-25.
83. McCaffrery P, Posch KC, Napoli JL, Gudas L, Drager UC. Changing patterns of the retinoic acid system in the developing retina. Dev Biol 1993; 158:390-9.
84. Tini M, Otulakowski G, Breitman ML, Tsui LC, Giguere V. An everted repeat mediates retinoic acid induction of the γF-crystallin gene: evidence of a direct role for retinoids in lens development. Genes Dev 1993; 7:295-307.
85. Tini M, Tsui LC, Giguere V. Heterodimeric interaction of the retinoic acid and thyroid hormone receptors in transcriptional regulation on the γF-crystallin everted retinoic acid response element. Mol Endocrinol 1994; 8:1494-506.
86. Balkan W, Klintworth GK, Bock CB, Linney E. Transgenic mice expressing a constitutively active retinoic acid receptor in the lens exhibit ocular defects. Dev Biol 1992; 151:622-5.
87. Perez-Castro AV, Tran VT, Nguyen-Huu MC. Defective lens fiber dif-

ferentiation and pancreatic tumorigenesis caused by ectopic expression of the cellular retinoic acid-binding protein I. Development 1993; 119:363-75.

88. Huang QL, Russell P, Stone S, Zigler JS. ζ-Crystallin, a novel lens protein from the guinea pig. Curr Eye Res 1987; 6:725-32.
89. Rodriguez IR, Gonzalez P, Zigler JS Jr, Borras T. A guinea pig hereditary cataract contains a splice-site deletion in a crystallin gene. Biochim Biophys Acta 1992; 1180:44-52.
90. Rodokanaki A, Holmes RK, Borras T. ζ-Crystallin, a novel protein from the guinea pig lens is related to alcohol dehydrogenases. Gene 1989; 78:215-24.
91. Borras T, Persson B, Jornvall H. Eye lens ζ-crystallin relationships to the family of long-chain alcohol/polyol dehydrogenases. Biochemistry 1989; 28:6133-9.
92. 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.
93. Lee DC, Gonzalez P, Rao PV, Zigler JS Jr, Wistow GJ. Carbonyl-metabolizing enzymes and their relatives recruited as structural proteins in the eye lens. In: Weiner H, ed. Enzymology and Molecular Biology of Carbonyl Metabolism. v. 4. New York: Plenum Press, 1993:159-68. (Advances in Experimental Medicine and Biology; 284).
94. Rao PV, Krishna CM, Zigler JS. Identification and characterization of the enzymatic activity of ζ-crystallin from guinea pig lens. A novel NADPH:quinone oxidoreductase. J Biol Chem 1992; 267:96-102.
95. Garland D, Rao PV, Del Corso A, Mura U, Zigler JS Jr. ζ-Crystallin is a major protein in the lens of *Camelus dromedarius*. Arch Biochem Biophys 1991; 285:134-6.
96. de Jong WW, Lubsen NH, Kraft HJ. Molecular Evolution of the Eye Lens. In: Chader G, Osbourne N, eds. Progress in Retinal and Eye Research. v. 13. : Elsevier Science Ltd, 1994:391-442.
97. Gonzalez P, Rao PV, Nunez SB, Zigler JS. Evidence for independent recruitemnt of ζ-crystallin/quinone reductase (CRYZ) as a crystallin in camelids and hystricomorph rodents. Mol Biol Evol 1995:in press.
98. Mulders JW, Hendriks W, Blankesteijn WM, Bloemendal H, de Jong WW. λ-Crystallin, a major rabbit lens protein, is related to hydroxyacyl-coenzyme A dehydrogenases. J Biol Chem 1988; 263:15462-6.
99. Hendriks W, Sanders J, de Leij L, Ramaekers F, Bloemendal H, de Jong WW. Monoclonal antibodies reveal evolutionary conservation of alternative splicing of the αA-crystallin primary transcript. Eur J Biochem 1988; 174:133-7.
100. Kim RY, Gasser R, Wistow GJ. μ-crystallin is a mammalian homologue of *Agrobacterium* ornithine cyclodeaminase and is expressed in human retina. Proc Natl Acad Sci USA 1992; 89:9292-6.
101. Schröder J. Plant hormones in plant-microbe interactions. In: Kosuge T, Nester EW, eds. Plant-microbe Interactions. Molecular and Genetical Perspectives. v. 2. New York: MacMillan, 1986:40-63.

102. Schindler U, Sans N, Schröder J. Ornithine cyclodeaminase from octopine Ti plasmid Ach5: identification, DNA sequence, enzyme properties, and comparison with gene and enzyme from nopaline Ti plasmid C58. J Bacteriol 1989; 171:847-54.
103. Sans N, Schindler U, Schröder J. Ornithine cyclodeaminase from Ti plasmid C58: DNA sequence, enzyme properties and regulation of activity by arginine. Eur J Biochem 1988; 173:123-30.
104. Wistow G, Segovia L. μ-crystallin: an unexpected marker for photoreceptors. Invest Ophthalmol Vis Sci 1994; 35:2073.
105. Jahn D, Verkamp E, Söll D. Glutamyl-transfer RNA: a precursor of heme and chlorophyll biosynthesis. Trends Biochem Sci 1992; 17:215-8.
106. Barnstable CJ. Glutamate and GABA in retinal circuitry. Curr Opinion Neurobiol 1993; 3:520-5.
107. Inana G, Hotta Y, Zintz C et al. Molecular basis of ornithine aminotransferase defect in gyrate atrophy. Prog Clin Biol Res 1991; 362:191-219.
108. Kanno C, Ishiguro S, Shiono T, Kikuchi M, Tamai M. Decrease of opsin content in the developing rat photoreceptor cells by systemic administration of L-glutamate. Cell Struct Funct 1991; 16:399-403.
109. Chen H, Phillips HA, Callen DF, Kim RY, Wistow GJ, Antonarakis SE. Localization of the human gene for μ-crystallin to chromosome 16p. Genomics 1992; 14:1115-6.
110. Yokoyama Y, Narahara K, Tsuji K, Ninomiya S, Seino Y. Autosomal dominant congenital cataract and microphthalmia associated with a familial t(2; 16) translocation. Hum Genet 1992; 90:177-8.
111. Ali MA, ed. Photoreception and Vision in Invertebrates. New York: Plenum Press, 1984 (NATO ASI Series A: Life Sciences; 74).
112. Land MF, Fernald RD. The evolution of eyes. Annu Rev Neurosci 1992; 15:1-29.
113. 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.
114. Tomarev SI, Zinovieva RD. Squid major lens polypeptides are homologous to glutathione S-transferases subunits. Nature 1988; 336:86-8.
115. Tomarev SI, Chung S, Zinovieva RD, Piatigorsky J. Structure and in vitro expression of S-crystallins and glutathione S-transferases of the squid. Invest Ophthalmol Vis Sci 1994; 35:1997.
116. Tomarev SI, Zinovieva RD, Piatigorsky J. Primary structure and lens-specific expression of genes for an intermediate filament protein and a β-tubulin in cephalopods. Biochim Biophys Acta 1993; 1216:245-54.
117. Tomarev SI, Zinovieva RD, Piatigorsky J. Characterization of squid crystallin genes. Comparison with mammalian glutathione S-transferase genes. J Biol Chem 1992; 267:8604-12.
118. Chiou SH. A novel crystallin from octopus lens. FEBS Lett 1988; 241:261-4.
119. Tomarev SI, Zinovieva RD, Piatigorsky J. Crystallins of the octopus lens. Recruitment from detoxification enzymes. J Biol Chem 1991; 266: 24226-31.

120. Zinovieva RD, Tomarev SI, Piatigorsky J. Aldehyde dehydrogenase-derived Ω-crystallins of squid and octopus. Specialization for lens expression. J Biol Chem 1993; 268:11449-55.
121. Montgomery MK, McFall-Ngai MJ. The muscle-derived lens of a squid bioluminescent organ is biochemically convergent with the ocular lens. Evidence for recruitment of aldehyde dehydrogenase as a predominant structural protein. J Biol Chem 1992; 267:20999-1003.
122. Davson H. The Lens. In: Physiology of the Eye. 5th ed. London: Macmillan Press, 1990:139-201.
123. Alexander RJ, Silverman B, Henley WL. Isolation and characterization of BCP54, the major soluble protein of bovine cornea. Exp Eye Res 1981; 32:205-16.
124. Silverman B, Alexander RJ, Henley WL. Tissue and species specificity of BCP 54, the major soluble protein of bovine cornea. Exp Eye Res 1981; 33:19-29.
125. Abedinia M, Pain T, Algar EM, Holmes RS. Bovine corneal aldehyde dehydrogenase: the major soluble corneal protein with a possible dual protective role for the eye. Exp Eye Res 1990; 51:419-26.
126. Cooper DL, Baptist EW, Enghild JJ, Isola NR, Klintworth GK. Bovine corneal protein 54K (BCP54) is a homologue of the tumor-associated (class 3) rat aldehydrogenase (RATALD) gene. Gene 1991; 98:201-7.
127. Lindahl R, Hempel J. Aldehyde dehydrogenases: what can be learned from a baker's dozen sequences? Adv Exp Med Biol 1991; 284:1-8.
128. Rabaey M, Segers J. Changes in the polypeptide composition of the bovine corneal epithelium during development. In: Trevor-Roper PD, ed. Vith Congress of the European Society of Ophthalmology. London: Academic Press, 1981:41-4.
129. Benedek GB. Theory of transparency of the eye. Appl Optics 1971; 10:459-73.
130. Delaye M, Tardieu A. Short-range order of crystallin proteins accounts for eye lens transparency. Nature 1983; 302:415-7.
131. Walther C, Gruss P. Pax-6, a murine paired box gene, is expressed in the developing CNS. Development 1991; 113:1435-49.
132. Li HS, Yang JM, Jacobson RD, Pasko D, Sundin O. Pax-6 is first expressed in a region of ectoderm anterior to the early neural plate: implications for stepwise determination of the lens. Dev Biol 1994; 162:181-94.
133. Cvekl A, Sax CM, Bresnick EH, Piatigorsky J. A complex array of positive and negative elements regulates the chicken αA-crystallin gene - involvement of Pax-6, USF, CREB and/or CREM and AP-1 proteins. Mol Cell Biol 1994; 14:7363-76.
134. Richardson J, Cvekl A, Wistow G. Pax-6 is essential for lens specific expression of ζ-crystallin. Proc Natl Acad Sci USA 1995; 92:4676-80.
135. Cvekl A, Kashanchi F, Sax CM, Brady JN, Piatigorsky J. Transcriptional regulation of the mouse αA-crystallin gene: Activation dependent on a cyclic AMP responsive element (DE1/CRE) and a Pax-6 binding site. Mol Cell Biol 1995; 15:653-60.

136. Cvekl A, Sax CM, Li X, McDermott JB, Piatigorsky J. Pax-6 and lens-specific transcription of the chicken δ1-crystallin gene. Proc Natl Acad Sci USA 1995; 92:4681-5.
137. Weber C. Lens of the hydromedusan *Cladonema* studied by SDS gel electrophoresis and immunofluorescent technique. J Exp Zool 1981; 217:15-21.
138. Piatigorsky J, Horwitz J, Kuwabara T, Cutress CE. The cellular eye lens and crystallins of cubomedusan jellyfish. J Comp Physiol [A] 1989; 164:577-87.
139. Piatigorsky J, Horwitz J, Norman BL. J1-crystallins of the cubomedusan jellyfish lens constitute a novel family encoded in at least three intronless genes. J Biol Chem 1993; 268:11894-901.
140. Edwards JS, Meyer MR. Conservation of antigen 3G6: a crystalline cone constituent in the compound eye of arthropods. J Neurobiol 1990; 21:441-52.
141. Komori N, Usukura J, Matsumoto H. Drosocrystallin, a major 52 kDa glycoprotein of the *Drosophila melanogaster* corneal lens. Purification, biochemical characterization and subcellular localization. J Cell Sci 1992; 102:191-201.
142. Cox RL, Glick DL, Strumwasser F. Isolation and protein sequence identification of *Aplysia californica* lens crystallins. Biol Bull (Woods Hole) 1991; 181:333-5.
143. Ohno M, Sakamoto H, Yasuda K, Okada TS, Shimura Y. Nucleotide sequence of a chicken δ-crystallin gene. Nucleic Acids Res 1985; 13:1593-606.
144. Abramson RD, Barbosa P, Kalumuk K, O'Brien WE. Characterization of the human argininiosuccinate lyase gene and analysis of exon skipping. Genomics 1991; 10:126-32.
145. Matsubasa T, Takiguchi M, Amaya Y, Matsuda I, Mori M. Structure of the rat argininosuccinate lyase gene: close similarity chicken δ-crystallin genes. Proc Natl Acad Sci USA 1989; 86:592-6.
146. Gonzalez P, Hernandez-Calzadilla C, Rao PV, Rodriguez IR, Zigler JS Jr, Borras T. Comparative analysis of the ζ-crystallin/quinone reductase gene in guinea pig and mouse. Mol Biol Evol 1994; 11:305-15.
147. Lee DC, Gonzalez P, Wistow G. ζ-Crystallin: A lens-specific promoter and the gene recruitment of an enzyme as a crystallin. J Mol Biol 1994; 236:669-78.
148. Hodin J, Wistow G. 5'-RACE PCR of mRNA for three taxon-specific crystallins: For each gene one promoter controls both lens and non-lens expression. Biochem Biophys Res Commun 1993; 190:391-6.
149. Kamachi Y, Sockanathan S, Liu Q, Breitman M, Lovell-Badge R, Kondoh H. Involvement of SOX proteins in lens-specific activation of crystallin genes. EMBO J 1995:in press.
150. Abad-Zapatero C, Griffith JP, Sussman JL, Rossmann MG. Refined crystal structure of dogfish M4 apo-lactate dehydrogenase. J Mol Biol 1987; 198:445-67.

151. Harrington L, Klintworth GK, Secor TE, Breitman ML. Developmental analysis of ocular morphogenesis in αA-crystallin/diphtheria toxin transgenic mice undergoing ablation of the lens. Dev Biol 1991; 148:508-16.

CHAPTER 5

Crystallin Gene Expression: The Pax-6 Connection

In spite of their diversity, taxon-specific and ubiquitous crystallin genes of mammals and birds are able to direct specific expression across species in lens cells of other mammals, birds and even amphibians.[1-6] This shows that vertebrates have conserved the ancestral gene cascades required for the development and differentiation of the eye and the lens and that consequently the transcriptional machineries of the lens are held in common among distantly related vertebrate species. Therefore when genes are recruited as crystallins they experience common transcriptional environments no matter in what vertebrate species the recruitment occurs.

However this does not necessarily mean that all crystallin genes use exactly the same mechanisms of expression. The lens maintains several different cell types throughout life and the expression profiles of crystallin genes vary with developmental stage and with the state of differentiation of the lens (see Figs. 1.3 and 1.4). Thus the six γ-crystallin genes γA-F are expressed preferentially in embryonic lens fibers[7-11] while γs-crystallin is expressed later in secondary fibers of the mature lens[12] and αA-crystallin is expressed in both epithelium and fibers throughout life.[13,14] Crystallins such as αA and ζ-crystallin[15] which are expressed in both epithelial cells and in fibers may also experience a boost in gene expression or protein synthesis during fiber cell differentiation to respond to the increased demand for protein in rapidly expanding cell volumes.

Different crystallin genes may therefore respond to different transcriptional environments and one gene may itself experience differences in these environments as it is expressed in different parts of the lens. Furthermore some genes, like αB-crystallin, contain elements for lens and non-lens expression in the same promoter region. This has created complexity and diversity in the expression mechanisms of crystallin genes in spite of all they have in common. For this reason it may be easier to discern some fundamentals of lens-specific expression in a

recently recruited taxon-specific crystallin, like guinea pig ζ-crystallin, than in more ancient ubiquitous crystallins like αA- and αB-crystallin. Nevertheless, recent results have begun to indicate a surprising degree of consensus for a widespread role for Pax-6 as a transcription factor involved in crystallin gene expression.[15-18] Other pattern-forming gene products with expression in lens, such as SOX-2,[19] are also likely to be involved in fine-tuning tissue-specificity and in conferring differential expression patterns within the lens.

PAX-6, EYE DEVELOPMENT AND THE EXPRESSION OF CRYSTALLINS

The differentiation of early embryos into complex, specialized tissues depends upon families of pattern forming "master genes," exemplified by those of the *Hox* families which encode proteins capable of sequence specific DNA-binding through structural motifs known as homeodomains (HD).[20] The homeodomain was first recognized as a conserved DNA-binding motif encoded by homeotic genes of *Drosophila*, such as *antennapedia* and *bithorax*, and was subsequently identified in many other families of DNA-binding proteins.[21-24] Genes for homeodomain-containing proteins are found in conserved clusters throughout the metazoa.[22-24] They are expressed very early in development and have essential roles in establishing segmentation patterns in embryogenesis.[22-24] This is often achieved by establishing overlapping regions of expression of different genes whose products exert positive and negative effects on the expression of their own and other pattern-forming genes as well as a variety of target genes.

Pax genes encode a similar family of proteins which are characterized by another DNA-binding motif, the paired-domain (PD), often, but not always, in conjunction with a homeodomain[25] (Fig. 5.1). The paired-domain too was first identified in *Drosophila*, in the paired gene[25,26] from which its name derives. *Pax* genes are also expressed early in development and play important roles in organogenesis. *Pax-6* is expressed from very early stages in eye and CNS in mouse, chicken and zebrafish[27-30] and its expression in the earliest precursors of chicken lens cells make it a candidate for one of the essential molecular determinants of lens competence.[29] Mutants in *Pax-6* have severe eye defects in mouse (*Small eye*)[31] and humans (aniridia and Peter's anomaly).[32-34] Remarkably it has now been found that homologues of Pax-6 play similar roles in invertebrates.[35] The *Drosophila* gene eyeless has been cloned and shows over 90% identity in predicted amino acid sequence to vertebrate Pax-6 proteins in its PD and HD regions. Mutants in eyeless fail to develop eyes. *Pax-6* homologues have also been detected in cephalopods, flat worms[35] and even in *C.elegans*[36,37] (A. Chisholm, personal communication).

Most dramatically of all, ectopic expression of *Drosophila* or mouse Pax-6 in various parts of *Drosophila* results in the induction of complete

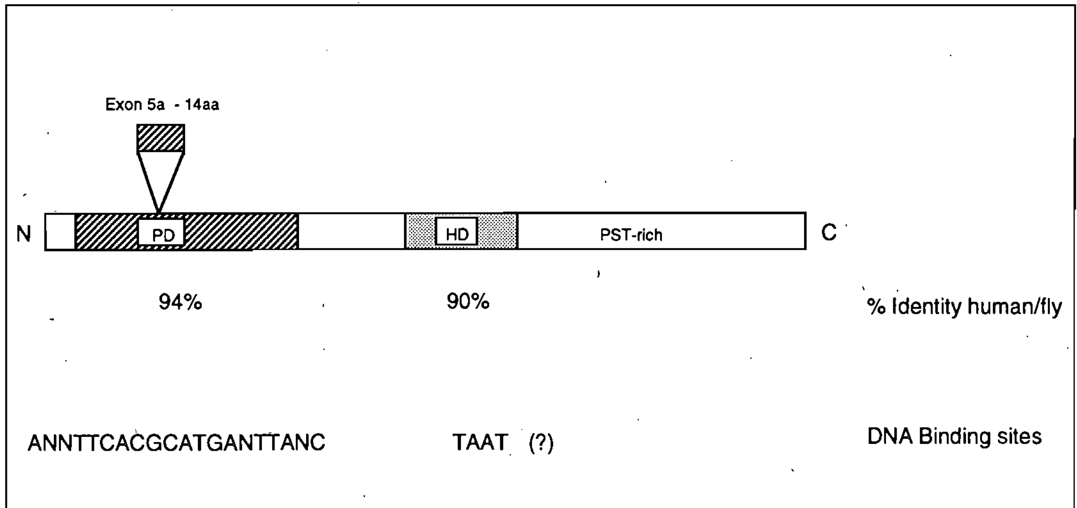


Fig. 5.1. The domain structure of Pax-6. PD: paired domain; HD: homeodomain; PST: the proline, serine, threonine-rich C-terminal region which may have a role in transactivation. Preferred consensus binding sites for PD and HD are shown.

compound eyes.[38] These observations suggest that *Pax-6* has an ancient ancestral role in the establishment of eyes and that all eyes may share a common ancestry in a simple light sensitive organ determined at least in part by *Pax-6* expression (see Fig. 1.1). Even as eyes became more sophisticated and divergent in many lineages the essential role of *Pax-6* was conserved and played a role in the evolutionary elaboration of new structures such as the cellular lens and ciliary body in vertebrates.

An in vitro consensus binding site for the PD of Pax-6 has been derived[39] (Fig. 5.1). This 20 bp sequence is long by the standards of many other transcription factors and reflects the fact that the PD has a bipartite structure. Independent binding of the two sub domains is possible as has been shown by the x-ray structure analysis of the PD itself bound to DNA.[40] The DNA sequence used in this analysis is CGTCACGGTTGA but since it only binds the N-terminal part of the PD it presumably does not represent a full binding site. Most binding studies of PDs from various Pax proteins have used an even shorter sequence, GTTCC.[41] Furthermore, alternative splicing of *Pax-6* transcripts gives rise to a variant which has an insertion in the N-terminal subdomain of the PD (Fig. 5.1) and as a result binds a different consensus sequence from the unspliced form.[42]

Like other HD-containing proteins, members of the Pax family seem to be transcription factors. In vitro experiments have shown that Pax-6 can act as a transcription activator.[43,44] Recently, functional Pax-6 binding sites have been detected in the promoter of the mouse gene

for the neural cell adhesion molecule L1[45] although the expression of this gene is not restricted to Pax-6 expressing cells. These sites conform quite well to the in vitro consensus.[39] However the most proximal site contains a TAAT sequence which also allows binding of Hoxa-1, a HD protein.[45]

Given its essential high-level pattern-forming role in early embryogenesis, Pax-6 was at first sight an improbable candidate for a transcription factor involved in expression of crystallin genes. Crystallins are probably at one of the end points of the molecular cascade in lens development and must maintain expression throughout life. However it is now clear that in at least some cases Pax-6 itself does act as transcription factor in crystallin gene expression.[15-18]

Our present understanding of the transcriptional control of crystallin gene expression is reviewed in the following sections, starting with the taxon-specific crystallins which have undergone more recent recruitment and ending with the more ancient ubiquitous crystallins which seem to have elaborated more complex control mechanisms.

TAXON-SPECFIC CRYSTALLINS

ζ-Crystallin: Pax-6 and the Recruitment of an Enzyme Crystallin

ζ-Crystallin is one of those taxon-specific enzyme crystallins which was first observed as a crystallin and only later proved to be an enzyme, in this case a novel NADPH:quinone oxidoreductase.[46-48] Sequence analysis revealed that this gene uses two separate promoters for lens and non-lens expression in guinea pig tissues[49,50] (Fig. 5.2). The lens promoter is located in what would otherwise be the first intron of the enzyme gene. This intron maps to the untranslated region of the mRNA and transcripts from both promoters splice to the same second exon which contains the initiator methionine codon. Thus the same protein is produced in both cases.

Since this gene makes use of widely separated alternative promoters for lens and non-lens expression, the lens-promoter does not need to accommodate additional binding sites for other functions. Although there is no evidence that other genes have made similar use of alternative promoters, ζ-crystallin illustrates important features of gene recruitment and lens-specific expression which may be generally applicable.

At first sight the lens-specificity of ζ-crystallin could have been achieved by several possible mechanisms. For example, the gene could have acquired a lens-specific enhancer somewhere in the gene which might have activated a TATA sequence already present in the first intron. In this scenario there might be no functional elements in the sequences upstream of the lens-specific alternative first exon. Alternatively the guinea pig could have experienced species-specific modification

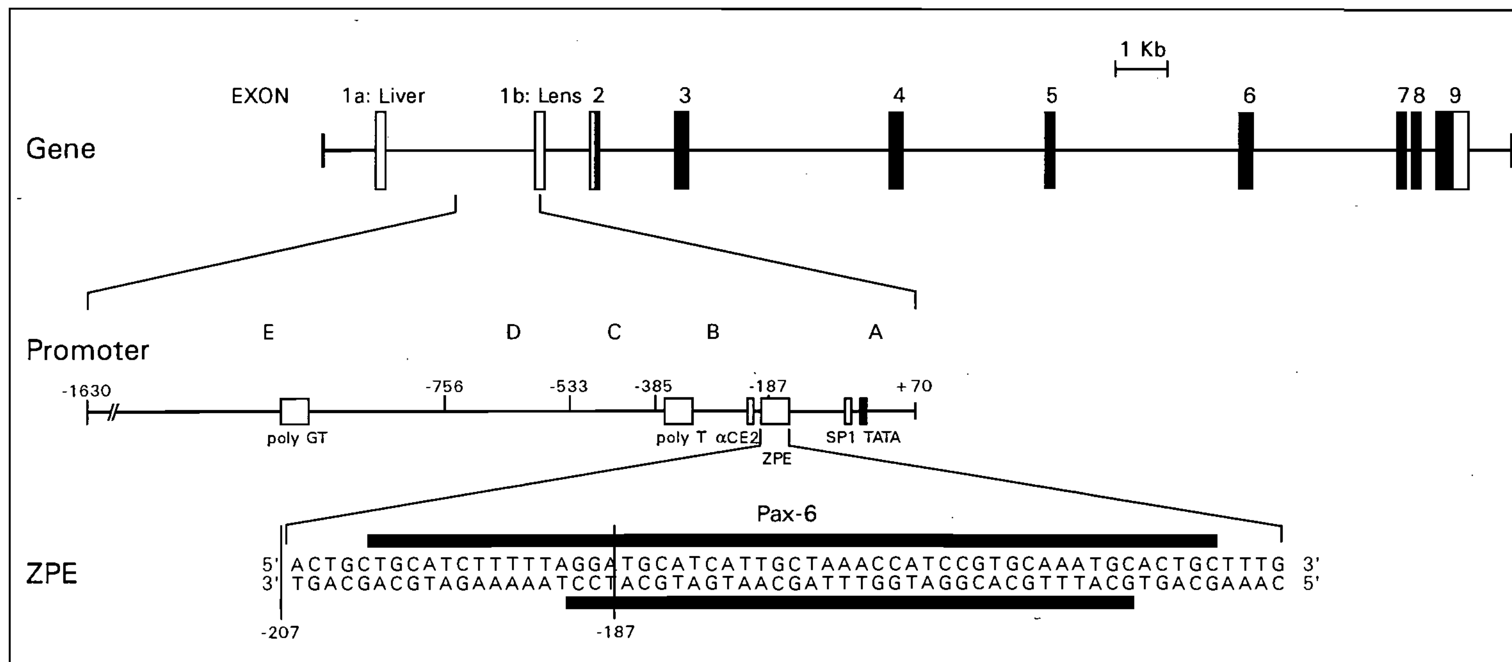


Fig. 5.2. The gene and lens promoter for guinea pig ζ-crystallin. Top: Gene structure. Exons are boxed and coding sequence is shaded. Alternative first exons are marked. Middle: Layout of the lens promoter. αCE2 refers to a sequence identical to one identified as functionally important in the chicken αA-crystallin gene (see Fig. 5.9), Bottom: The ZPE/Pax-6 binding site. This figure is updated from Lee DC, Gonzalez P, Wistow G, J Mol Biol 1994; 236:669-781.

of the complement of transcription factors expressed in its lens. This might have activated a cryptic promoter in the first intron. In such a case the promoter would function only in guinea pig lens and would not exhibit lens-specific expression in other species such as mice.

Fortunately, the mechanism of recruitment proved to be the simplest and most accessible possibility. The recruitment of this enzyme crystallin occurred through acquisition of a lens-specific alternative promoter which does not require host-specific factors[51] (Fig. 5.2). The lens promoter is neatly flanked by 9 bp direct repeats and when the guinea pig promoter is compared to the intron sequence of the homologous but unrecruited gene in mouse no similarity in sequence is observed upstream of the direct repeat in the guinea pig first exon while there is limited sequence similarity downstream of this point. This raises the possibility that the lens promoter may have been inserted into the gene by a transposon-mediated event.[51]

The strong tissue preference of the lens promoter is apparent in both transient transfections of cells in culture and in transgenic mice.[51] While proximal regions of the promoter (-385/+70) have some activity in the brain of transgenic mice this is abolished by the addition of more distal regions (Fig. 5.3). The minimal active lens promoter is differentially footprinted by extracts from lens and non-lens cells. In lens cell extracts a single 50 bp element, the ζ protected element or ZPE, is protected from DNase I digestion (Fig. 5.2). In fibroblast extracts the ZPE is incompletely protected and is flanked by two additional protected elements, the upstream and downstream boxes (UB and DB). This suggests that in lens a tissue-specific (or preferred) factor binds to the promoter at the ZPE to form a transcriptionally active complex while in non-expressing cells, some competing factor(s) occupies the ZPE and the suppression of promoter activity is completed by additional binding of other factors at the UB and DB sites (Fig. 5.4).

The ZPE is a Pax-6 Site

In electrophoretic mobility shift assay (EMSA) the ZPE forms two specific complexes. Complex I is found in extracts of non-lens tissues such as liver and lung. This shows that the competing factor which binds the ZPE in non-expressing cells does not require the cooperation of the UB and DB factors in order to bind. The protein composition of complex I is not yet known. However, the ZPE contains consensus binding sites for several families of general transcription factors[51] including the C/EBP family, octamer family and HLH family.[52]

Complex II is formed in lens extracts while in extracts of lens-derived cells and brain both complexes I and II are present.[15] At this stage it is not known whether both complexes can exist in one cell type or whether they are mutually exclusive. Brain consists of multiple cell types and even the cultured lens-derived cells may contain lens-like and non-lens like populations.

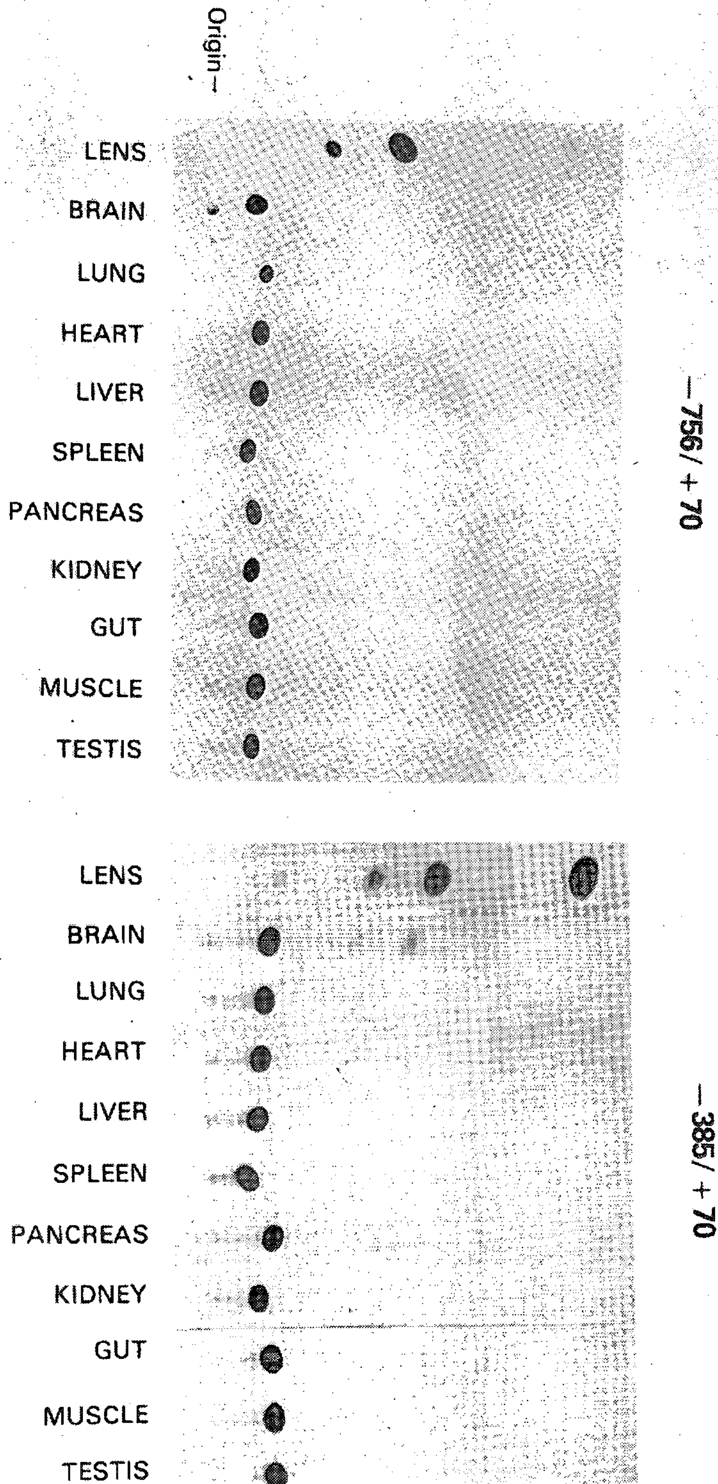


Fig. 5.3. Thin layer chromatography analysis of CAT reporter gene activity in tissues of transgenic mice bearing ζ-crystallin promoter constructs. -756/+70.ZP/CAT (top) is a "full-length" wild type promoter with lens-specificity.[51] -385/+70.ZP/CAT (bottom) exhibits some expression in brain. Reprinted with permission from Lee DC, Gonzalez P, Wistow G, J Mol Biol 1994; 236:669-78

By competitive EMSA a region of the ZPE, designated ZE-1, was defined as essential for formation of complex II. The ZE-1 site represents the core of a consensus Pax-6 binding site in the ZPE.[15,39] Antisera to Pax-6 abolish complex II without affecting complex I. Recombinant human Pax-6 gives an identical footprint on the ζ-crystallin promoter to that formed by mouse lens extract (in preparation). Mutation of the ZPE which abolishes Pax-6 binding in vitro also abolishes promoter activity in vivo[15] (Fig. 5.5). The identification of Pax-6 with complex II is consistent with its expression in both brain and eye. Western blot analysis shows that Pax-6 protein is present in lens and brain and in extracts of αTN4-1 and N/N1003A cells. RT-PCR is also able to detect mRNA for Pax-6 in these tissues and cells.[15] Thus Pax-6 is essential for expression of the ζ-crystallin lens promoter, however other factors which footprint poorly or which rely on protein-protein interactions rather than DNA-binding could also be involved in gene activation (Fig. 5.4).

Pax-6 Expression in Mature Lens

Pax-6 is expressed in mature lens appropriately for a continuing role in tissue-specific gene expression.[15] Pax-6 protein is present in adult guinea pig lens. By immunohistochemical staining Pax-6 is detected most prominently in lens epithelial cells where the nuclei make up a large fraction of cell volume but it is also detectable in the nuclei of the elongating fiber cells in the equatorial region. Immunochemistry for ζ-crystallin in the same system shows that the crystallin is present in the cytoplasm of the epithelial cells and in fiber cells.[15] Its most intense staining is in the equatorial fibers. Thus Pax-6 is present in cells which express ζ-crystallin and the maximum expression of Pax-6 seems to occur prior to the maximum expression of ζ-crystallin protein during lens cell differentiation.

Pax-6 mRNA is present in adult mouse lens and brain but exhibits an interesting tissue-specific pattern of alternative splicing. In mouse brain both of the alternatively spliced mRNAs which correspond to the alternative PD forms of Pax-6 are detected at essentially equal abundance. In contrast, in adult mouse lens and in lens-derived cultured cells the mRNA corresponding to the form of Pax-6 able to bind the ZPE greatly predominates. Thus, while the expression of the *Pax-6* gene is not tissue-specific, lens-specific differences in the abundance of alternatively spliced forms of *Pax-6* mRNA may contribute to tissue discrimination in binding activity.

Pax-6 Binding Sites in Several Genes

In addition to its central importance in the lens-specific expression and gene recruitment of guinea pig ζ-crystallin[15] Pax-6 has been detected in complexes of factors binding to functional elements of the promoters for chicken[16] and mouse αA-crystallins[17] and to the enhancer

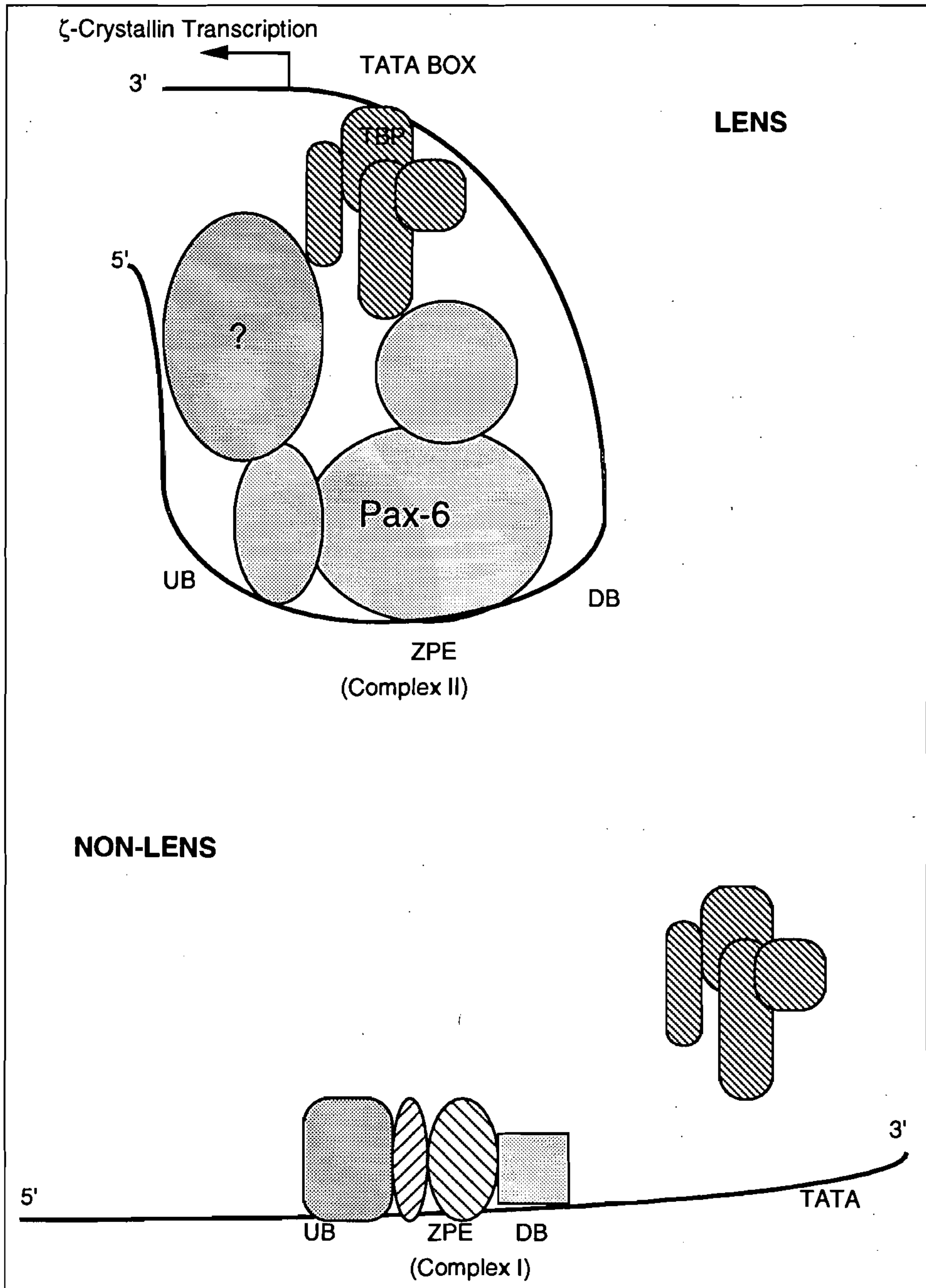


Fig. 5.4. A cartoon to illustrate some features of the ζ-crystallin lens promoter. Top: In lens Pax-6 binds the ZPE and activates the promoter, perhaps in concert with other factors not yet identified which may or may not contact DNA. Bottom: In other tissues a different complex occupies the ZPE together with other factors at the flanking upstream box (UB) and downstream box (DB) and the promoter is inactive.

MUT-756/+70

Origin →

LENS BRAIN LUNG HEART LIVER SPLEEN PANCREAS KIDNEY GUT MUSCLE TESTIS

Fig. 5.5. Thin layer chromatography analysis of CAT reporter gene activity in tissues of transgenic mice bearing the -756MUT/+70.ZP/CAT construct which contains specific mutations in the ZPE which abolish Pax-6 binding and promoter activity.

	t	t																		
A	N	N	T	T	C	A	C	G	C	W	T	S	A	N	T	K	M	N	Y	:E
A	T	T	T	G	C	A	C	G	G	A	T	G	G	T	T	T	A	G	C	:ζ
T	T	A	T	T	C	A	C	T	A	A	T	G	G	C	T	G	C	A	C	:L1
A	A	T	C	C	C	A	C	T	A	A	T	G	C	C	T	T	C	A	T	:α
A	T	T	T	N	C	A	C	K	N	A	T	G	N	N	T	T	M	N	Y	:con

Fig. 5.6. Alignment of Pax-6 recognition sequences.
E: The in vitro binding sequence of Epstein et al.[39] Positions of equal preferences are shown by standard nomenclature as follows: W (A or T); S (C or G); K (G or T); M (A or C); Y (C or T). Positions 2 and 3 have a slight preference for T (shown as lower case above the line) which matches other sequences as shown.
ζ: The site from ζ-crystallin in complement.
L1: The proximal site from mouse L1.[45]
α: The -61/-40 site from chicken αA-crystallin.[16]
con: a consensus derived from these sequences.

region of chicken δ1-crystallin,[18] as discussed below. Figure 5.6 shows the consensus Pax-6 PD binding sequence which was determined in vitro[39] compared with the ζ-crystallin ZPE, the most proximal Pax-6 site of mouse neural adhesion protein L1[45] and the chicken αA-crystallin -61/-40 site, which is the closest matching site to the in vitro consensus found so far in other crystallin genes. If positions with no clear preference are omitted from the in vitro consensus, there is 13/16 identity with the ζ-crystallin Pax-6 site (15/18 if positions 2 and 3 of the consensus which have slight preferences for T are included). Interestingly, the L1 and αA-crystallin sites in this alignment share a striking feature absent from both ζ-crystallin and from the in vitro consensus. They contain a consensus HD binding site (TAAT) which in the L1 gene has been shown to bind another HD protein, Hoxa-1.[45]

It seems likely that there is a range of higher and lower affinity sites for Pax-6 binding. The close match of the ζ-crystallin site to the in vitro consensus suggests that it may be a relatively high affinity site and this may explain the very strong footprinting of this site seen in protection analyses using lens cell extracts.[51] "Indeed, in contrast to the partial binding of a PD revealed by the recent x-ray structure,[40] it seems likely that the 50 bp ZPE of ζ-crystallin binds the PD, HD and possibly other C-terminal regions of a single Pax-6 molecule in a contiguous site (in preparation)."

Lens-specificity Through Acquisition of a Pax-6 Binding Site

In the case of ζ-crystallin, which is the result of relatively recent evolutionary events, a simple model of the process of gene recruitment can be envisaged. The initial event may have been the acquisition of a binding site for Pax-6 in an intron of an enzyme gene. Since Pax-6 expression is not limited to lens this could also have conferred expression in other tissues such as iris, retina and brain. However a lens-preferred pattern of expression could have occurred through selective binding of one form of Pax-6 resulting from tissue-specific alternative splicing in lens. Subsequently the expression of the recruited gene could have been further fine-tuned by the addition of other *cis*-elements to the recruited promoter which eliminated expression in other Pax-6 containing tissues. Indeed, the ζ-crystallin promoter does contain upstream sequences which suppress expression in transgenic mouse brain (see ref. 51, unpublished).

It is well known that changes in the expression of pattern-forming genes can produce significantly altered developmental programs. In the same way, acquisition of binding sites for master gene factors like Pax proteins could radically alter the protein composition of a tissue, such as the lens, in one evolutionary step.

PAX-6 AND CATARACT

Defects in Pax-6 expression can have serious effects in the eye. Mice homozygous for small eye have no eye or orbit at all and there is severe facial malformation.[31] Heterozygotes have microphthalmia. Heterozygote humans with aniridia lack a properly formed iris but do not have microphthalmia.[32,33,53] This species difference is interesting in itself. It shows that even such a fundamental gene as *Pax-6* may not function identically in species as closely related as two mammals.

Although the lens is usually unaffected at birth, as patients with aniridia age they also develop cataract, apparently with considerable heterogeneity.[53] This suggests that a single gene dosage of *Pax-6* is sufficient for normal lens development but not for development of the iris in humans. However, the progression of cataract also suggests that a single functional copy of *Pax-6* is not sufficient for the maintenance of a healthy transparent lens. Although analyses of mouse and guinea pig lens show that *Pax-6* expression continues in mature mammalian lens[15] it is also clear from studies in chicken embryos that levels of *Pax-6* mRNA decline during embryogenesis.[29] In neither birds nor mammals is there a complete picture of the level of expression of this gene throughout life. However the following hypothesis is not inconsistent with what is presently known.

Let us suppose that *Pax-6* expression is maximal during embryogenesis during rapid organogenesis and tissue differentiation but that later its expression declines as it adopts a maintenance role. Let us also suppose that in any dependent tissue there is a minimal threshold level of Pax-6 protein required for normal expression of some important target genes. In lens, this level is exceeded even by a single gene dose of *Pax-6* for some time after birth. However as gene expression declines, the single gene cannot maintain the maintenance threshold level. As it falls below this level gene expression is disrupted in lens epithelia and in newly differentiated fibers and opacities form.

Clearly this hypothetical model also has implications for the normal lens with two functioning *Pax-6* genes. Just as in aniridia, the level of *Pax-6* expression would decline. The approach to the critical threshold would be postponed because of the higher gene dosage but eventually, at a later age, the threshold would still be reached. This could be a cause of some cases of senile cataract. Premature senile cataract could occur when some minor difference in gene expression or in mRNA or protein stability led to a more rapid fall in Pax-6 levels in lens.

OTHER TAXON-SPECIFIC CRYSTALLINS

δ-CRYSTALLINS: RECRUITMENT THROUGH A LENS-SPECIFIC ENHANCER

As a result of gene duplication and specialization there are two genes for δ-crystallins in birds. One encodes ASL/δ2-crystallin and in

many birds, including chicken, this gene is not expressed in lens at crystallin-like levels.[5,54-56] The expression of ASL/δ2-crystallin has not been extensively studied. Instead, most work has concentrated on the lens-specialized δ1-crystallin gene which may contribute 90% of the soluble protein in the embryonic chick lens.[57] Initial studies on chicken δ1-crystallin identified a lens-preferred promoter and an upstream negative element.[58,59] However, the key to the high expression of this gene has proved to be an enhancer located in the third intron of the δ1-crystallin gene[60] (Fig. 5.7).

Surprisingly the δ1-enhancer is also present in the much less active ASL/δ2-crystallin gene of the chicken and combinations of promoters and enhancers from both genes were all found to be preferentially expressed in lens.[61] Even in the duck, the δ-enhancer is highly conserved in the third intron of the ASL/δ2-crystallin gene.[62] Promoter sequences are also highly conserved between chicken and duck δ2-crystallins. Indeed, the only major difference between the sequences separating δ1 and δ2 genes is the insertion of a CR1-type repetitive element in the duck locus.[62] Given this high degree of similarity, what then determines the differences in lens expression seen among different δ-crystallins in different species? The answer is not known although a search is underway for a possible silencer element responsible for suppressing expression of chicken δ2-crystallin in lens.

The δ1-enhancer: SOX and Pax

The δ1-crystallin enhancer contains at least two overlapping binding sites for factors designated δEF1 and δEF2[63-65] (Fig. 5.7). A protein capable of binding to the δEF1 site has been cloned and turns out to be a general factor which may have a role in suppressing nonlens expression.[65] In particular, δEF1 is able to compete with bHLH proteins for binding at a class of E-boxes.[66] One such E-box binding protein is USF which is able to bind the δEF1 site in lens extracts.[18]

Multimers of the δEF2 element can act as a lens-specific enhancer for a heterologous promoter.[65] The lens-preference of the enhancer apparently depends upon binding of multi-component complexes.[65] One binding activity, designated δEF2a, was found to be highly enriched in chicken lens cells and has been cloned revealing identity with SOX-2.[19] The SOX family, whose name derives from SRY-box, is another group involved in pattern-formation and organogenesis during embryonic development.[67-69] Instead of a HD or PD, these proteins use a DNA-binding domain first identified in high mobility group (HMG) proteins. Like other pattern-forming gene products they exhibit restricted and overlapping patterns of expression which contribute to formation of tissue-specificity. Overexpression of SOX-2 increased δ1-crystallin enhancer activity in lens cells but not in fibroblasts.[19] This suggests that lens specificity results from a combination of lens-preferred factors. Indeed, two Pax-6 sites have also been identified in the δ1-crystallin enhancer region.[18] These sites do not correspond to the functionally

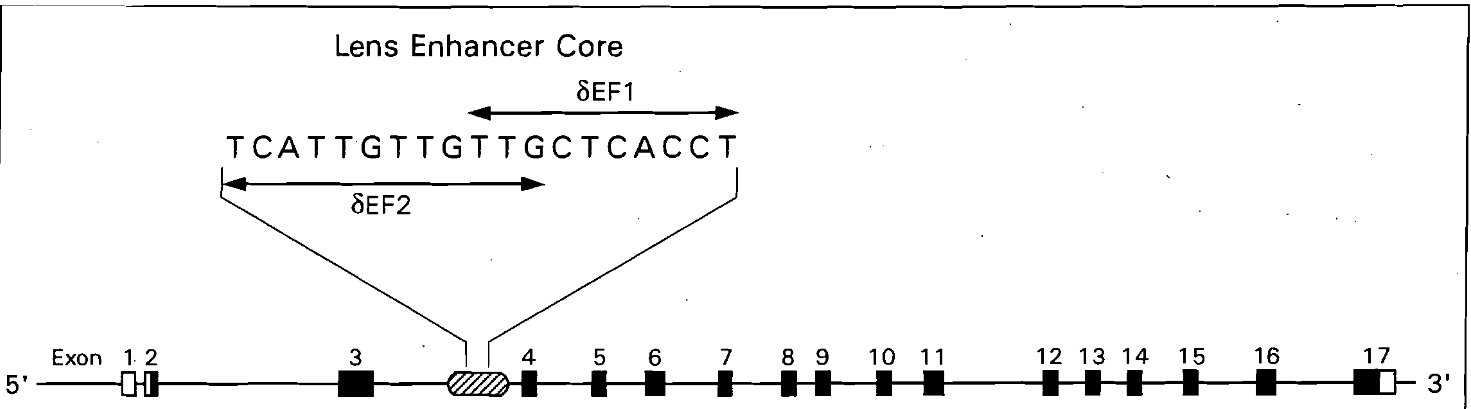


Fig. 5.7. The chicken δ1-crystallin gene and the position of the enhancer. Taken from Wistow G, Richardson J, Jaworski C, Graham C, Sharon-Friling R, Segovia L. In: Tombs MP, ed. Biotechnology and Genetic Engineering Reviews. v. 12. Andover, Hants: Intercept Ltd, 1994; 1-38.

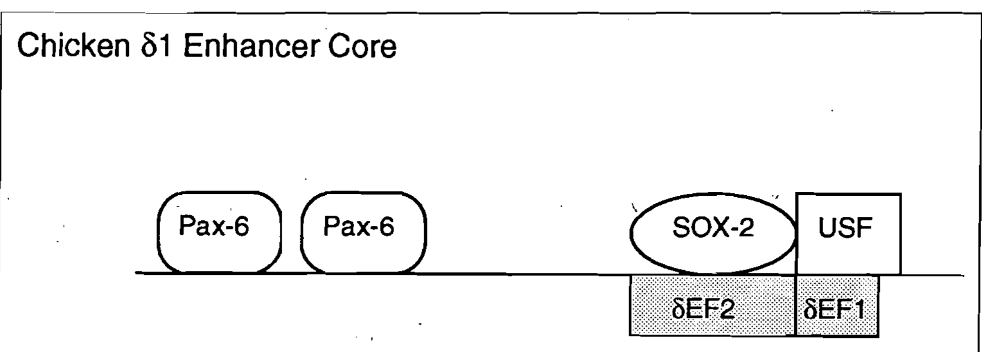


Fig. 5.8. A cartoon showing binding sites identified for the δ1-crystallin enhancer in lens.

defined δEF2 elements but they lie in sequences which are conserved in both δ1 and δ2-crystallin genes in chicken and duck.[18] It thus seems that the δ1-crystallin enhancer may operate through complex binding of different tissue-restricted, pattern-forming gene products involving both Pax-6 and SOX-2 (Fig. 5.8).

α-Enolase/τ-Crystallin

As discussed above, α-enolase/τ-crystallin is not a typical crystallin but it has some interesting features. In the duck lens it is highly abundant in epithelial cells but contributes only a small amount in overall crystallin content to adult lens.[70,71] The duck gene for α-enolase/τ-crystallin has been cloned and some promoter analysis carried out.[70] Although the endogenous gene shows preferential expression in embryonic duck lens epithelia, no lens-preferred expression of the recombinant gene promoter was observed in either transfected cultured cells or in transgenic mice.[70,72] In spite of these unpromising characteristics, this gene does serve to illustrate the way in which different regions of the lens express different complements of proteins. In fact its gene expression appears to be subject to differentiation-specific control and part of this may be due to the myc family of proto-oncogenes.

There is considerable circumstantial evidence associating high levels of expression of α-enolase and c-myc in various cell types.[73] This also applies to the lens in which the ratio of c-myc to N-myc mRNA drops as lens cells differentiate from epithelial cells where α-enolase is abundant to fiber cells in which it is at lower levels.[74] Furthermore, although there is otherwise little sequence similarity between them, the gene promoters of human and duck α-enolase[70,75] both contain an E-box, a potential myc family binding site,[76,77] at equivalent positions about 600 bp upstream of the transcription start site. C-myc can bind to this E-box in the duck gene and co-expression of c-myc induces expression of the duck α-enolase/τ-crystallin gene promoter.[78] This induction is abolished by mutation of the E-box. Clearly, control of α-enolase expression is not the main purpose of c-myc. Instead c-myc and its relatives are part of the essential processes of tissue differentiation. As a possible target of these factors the expression of α-enolase is subservient to that higher function. Its expression in lens may therefore be at least partly a side effect of other processes of differentiation, a lesson that may also be applicable to other crystallins.

Lactate Dehydrogenase-B/ε-Crystallin

The gene for duck LDHB/ε-crystallin has been cloned and its expression examined in chicken lens and heart cells.[79] Identical start sites for transcription were found in both tissues,[79,80] although a second start site 28 bp upstream was also found in heart. High expression of the promoter in both cell types required sequences from the first intron although since it is unable to enhance a heterologous promoter the intron does not contain an enhancer as classically defined.[79] Although

the gene is expressed at higher levels in lens than heart in vivo, no difference in promoter activity was observed in cultured cells. This suggests that the cultured chicken lens cells do not appropriately mimic the characteristics of intact lens required for ε-crystallin overexpression.

When levels of endogenous LDHB mRNA were measured in various tissues of chicken and duck it was found that expression in duck exceeded that in chicken in three tissues, lens, retina and pancreas.[81] This is intriguing since these tissues are all sites of Pax-6 expression in birds.[29,82]

UBIQUITOUS CRYSTALLINS

Since their initial recruitment to the lens occurred in a common ancestor of all vertebrates, the genes for the ubiquitous crystallins have had much longer than those of the enzyme crystallins to complicate and elaborate their transcriptional machinery. This increased sophistication may explain the apparent complexity of their promoters compared to that of a recently recruited gene such as ζ-crystallin.

αA-CRYSTALLINS: CONSERVATION AND COMPLEXITY IN UBIQUITOUS CRYSTALLINS

Of the two α-crystallin genes expressed in all vertebrate lenses αA-crystallin is the most tissue-specialized. Presumably a single sHSP/α-crystallin gene was recruited to the lens in a distant common ancestor of all vertebrates. This original single gene would probably have produced both lens and non-lens transcripts from one promoter with a mixture of functional elements. In this respect it would have resembled the modern αB-crystallin. At some point shortly after this initial recruitment there was a gene duplication and specialization of one of the pair of genes for lens. This gene became αA-crystallin. Over time the evolving αA-crystallin gene would have lost its non-lens expression for which there was no selective advantage and would have progressively fine-tuned its expression in lens.

In many species αA-crystallin is the single major provider of protein to the lens.[83] As such it may have been under special pressure to modify its expression during development and differentiation to suit lenses as different as those of mice and chickens. Thus, the promoter of this ancient gene might very well have acquired an unusual degree of complexity with various functional elements overlaying each other to modulate expression in different lineages. Indeed, this is the kind of picture which has emerged from a large number of studies of the gene expression of αA-crystallin.

αA-crystallin gene promoters have been cloned from mouse,[84] hamster,[85] mole rat,[86] chicken[87] and human.[84] Low levels of αA-crystallin have been detected in spleen and thymus in rat,[88] however the expression of this gene is otherwise highly lens-preferred. Both mammalian and chicken αA-crystallin genes share this high expression in the lens

and the chicken αA promoter is expressed in the lens of transgenic mice.[4] This suggests that all vertebrate αA promoters should share conserved functional promoter elements required for lens expression. These elements should be apparent in the approach known as "phylogenetic footprinting"[89] comparing promoter sequences for conserved regions. Alignments of promoter sequences for mouse, human and chicken uncovered four elements clustered close to the transcription start site which were significantly conserved in both sequence and position in all species[84,90] (Fig. 5.9). These elements correspond roughly to the minimal promoter regions of mouse and chicken αA-crystallin genes required for function in cell culture: -111 bp for mouse[91] and -162 bp for chicken.[4,92]

On a larger scale, all three αA-crystallin promoters also share the presence of species-specific repetitive elements at approximately 1 kb upstream of the coding regions.[84] A variety of other short sequence motifs are also common to all three genes but their position and number are not conserved. Whether these are the result of motif shuffling or whether they have no functional significance is not yet known. However, there is evidence that the four major conserved elements are indeed important for gene expression in lens.

In vivo and in vitro footprinting of the mouse αA-crystallin promoter have shown similar protected regions in αTN4-1 cells, mouse lens nuclear extract and in nuclear extracts from L929 fibroblasts.[93] In spite of this similarity, electrophoretic mobility shift assays (EMSA) suggested that different proteins were bound to the same sites in both lens and non-lens environments.

Binding Sites in the Mouse αA-crystallin Promoter

Several elements in the mouse αA-crystallin promoter have been defined by a variety of binding and functional studies. The first of these were the distal (-111/-88) and proximal (-88/-60) elements.[91] Both of these elements were found to be necessary for expression in transient transfection of PLEs. Since then several other studies have led to a more complex picture. The current view of this promoter defines several additional elements (Fig. 5.9).

The most 5' of these, DE1 (-111/-97), is part of the original distal element and contains one of the phylogenetically conserved blocks of sequence. It also corresponds to part of an enhancer which was defined in parallel studies of the closely related hamster αA-crystallin gene. In heterologous promoter experiments, the hamster αA-crystallin enhancer (-180/-85 in that gene) was able to activate the minimal promoter of mouse γF-crystallin in transfections in chicken PLEs.[91,94] DE1 has sequence similarity to a cAMP responsive element (CRE) and recent work has shown that this site can bind general CREB factors thereby activating transcription[17] (Fig. 5.10).

Overlapping the TATA box is the TATA/PE1 (-35/-19) element, while the 5' end of the transcribed region of the gene contains the

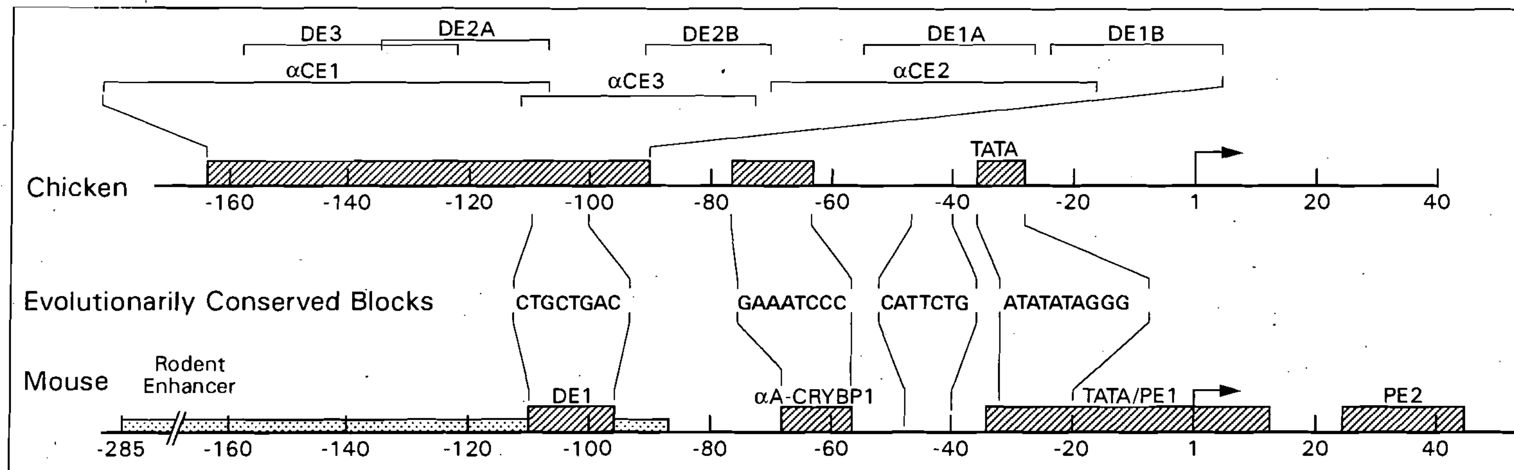


Fig. 5.9. The promoter regions of chicken and mouse αA-crystallin genes. Shaded regions were defined by protection and mutational assays. Sequences evolutionarily conserved among αA-crystallin promoters are also shown. Adapted from Wistow G, Richardson J, Jaworski C, Graham C, Sharon-Friling R, Segovia L. Crystallins, In: Tombs MP, ed. Biotechnology and Genetic Engineering Reviews. v. 12. Andover, Hants: Intercept Ltd, 1994; 1-38.

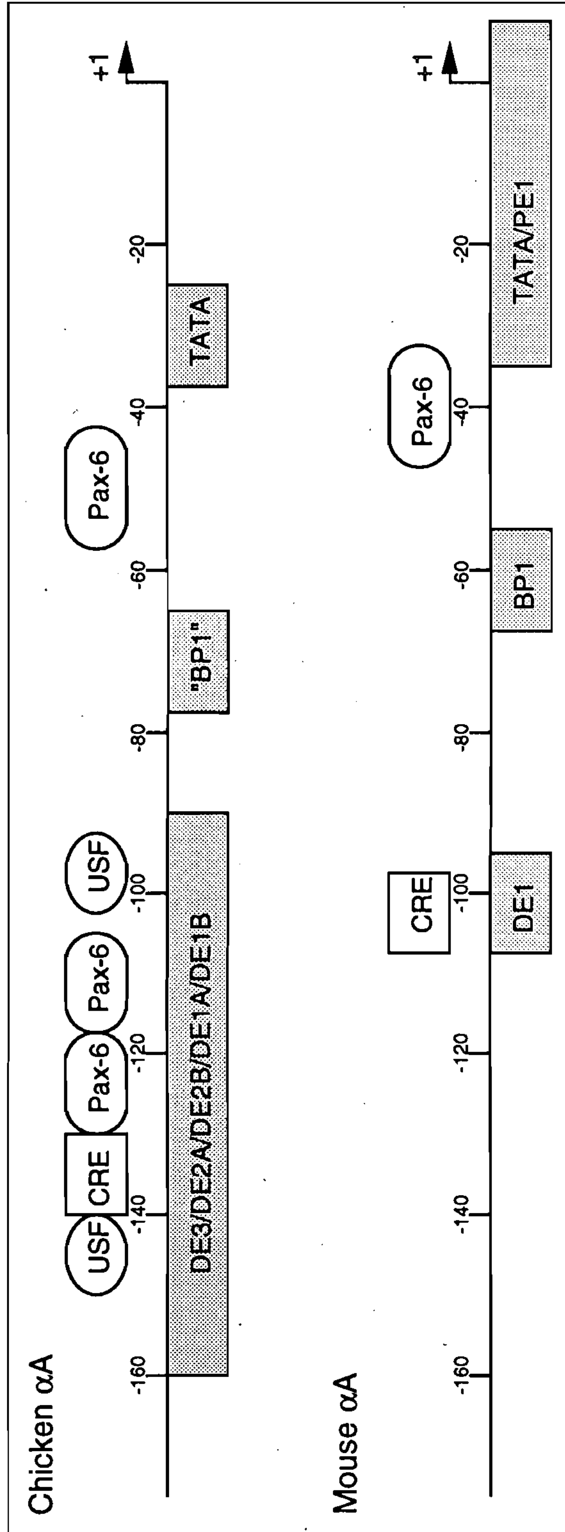


Fig. 5.10. Factor binding sites identified with lens-specific expression of αA-crystallin promoters. BP1 represents the αA-CRYBP1 binding site in mouse. A similar site in chicken is shown as "BP1."

element PE-2 (+24/+43).[90] The PE1 and PE2 sequences are conserved among mammals but not in chicken.[84,90] The PE1 region has been further defined into the TATA box (-31/-26) and the PE1B sequence (-25/-12).[95] Deletions within either element eliminated promoter activity in transient transfections of lens-derived cells but did not eliminate activity in transgenic mice.[95] DNase I footprinting and EMSA analyses showed similar patterns of protein binding in the region in both lens and fibroblast extracts although there was no evidence that TATA-binding protein (TBP) was present in any complex.[95]

αA-CRYBP1

The originally defined proximal element contains an evolutionarily conserved sequence designated motif D[84] which is similar to binding sites of the NFkB family.[52] A factor which binds this site was cloned from αTN4-1 cells and named αA-CRYBP1.[96] This protein, which appears to be the mouse homologue of a human transcription factor called PRDII-BF or MBP-1 which is involved in expression of immune response genes,[97,98] binds the sequence GGGAAATCCC at positions -66/-57 in the mouse promoter. Mutation in the αA-CRYBP1 region of the mouse gene reduces but does not eliminate promoter activity in PLEs and in transfected αTN4-1 cells.[96,99] However the significance of this factor for lens expression in vivo is not clear. αA-CRYBP1 is essentially ubiquitous in its tissue distribution.[96] The enhancer activity of the αA-CRYBP1 site is quite low in transfection studies and is apparent only in αTN4-1 cells in which the αA-CRYBP1 factor is abundant, although multimers of the site can activate expression of a heterologous promoter in a variety of other cells.[100] Furthermore, although αA-CRYBP1 binds a site in the mouse gene promoter it does not bind the equivalent conserved sites in the chicken, mole rat and human promoters which differ from the mouse sequence by only one base.[84,86,100] It thus seems possible that a different factor binds motif D in vivo for normal lens expression and that αA-CRYBP1 has a more general role.

Pax-6

Recently a binding site for Pax-6 has been identified just downstream of the motif D/αA-CRYBP1 site at position -49/-33 in the mouse αA-crystallin promoter[17] (Fig. 5.10). Part, but not all, of this region is well conserved among species and as described below there is evidence that the equivalent region of the chicken αA-crystallin promoter is also able to bind Pax-6.[16] The Pax-6 binding site is flanked by the αA-CRYBP1 and TATA/PE1 elements but mutational and binding analyses of this region had not previously identified this sequence itself as a discrete element in the mouse αA-crystallin promoter.[90]

Cooperativity and Redundancy

One theme which has emerged from studies of the mouse αA-crystallin promoter is that of cooperative binding at multiple elements and functional redundancy in vivo. For example, although both proximal and distal elements were found to be essential for expression in PLEs[91] the -88/+44 promoter construct which lacks the distal element showed lens-specific expression in transgenic mice.[101] This observation has been refined to show that the DE1 and αA-CRYBP1 sites are functionally redundant in transgenic mouse experiments.[101] The reason for the discrepancy between PLE and transgenic mouse studies is not known.

Binding Sites in the Chicken αA-crystallin Promoter

The promoter of the chicken αA-crystallin gene contains evolutionarily conserved sequences which corresponds to the DE1 element of the mouse αA-crystallin promoter (Fig. 5.9). Different laboratories have defined these sequence elements either as DE1A and DE1B[90,102] or as αCE2.[103] Surprisingly, in spite of the conservation of sequence, the mouse DE1 sequence failed to compete with chicken DE1A/B for binding of chicken nuclear extracts in EMSA[102] suggesting that mouse and chicken genes may bind different factors or that there is species specificity in binding of common factors. Upstream of the DE1A,B/αCE2 elements is another set of elements known as DE3 and DE2A[90,102] or as αCE1.[103] The αCE2 element requires the presence of αCE1 for activity[103] suggesting a cooperative interaction. A binding activity specific to the αCE1 element has been identified.[104,105]

Pax-6 and General Factors

Recent work has examined the binding of known factors whose consensus binding sites resemble sites in the chicken αA-crystallin promoter.[16] This has resulted in a picture of a complex array of general and tissue-restricted factors showing differential binding to the functional elements of the chicken αA-crystallin promoter in lens and nonlens extracts (Fig. 5.10). In lens extracts the general factor USF (upstream factor) is able to bind a non-canonical E-box in the DE2A element while cAMP-response element (CRE) binding factors bind immediately downstream. In fibroblast extracts the same sites bind USF and AP1. In both lens and fibroblasts USF complexed with an unknown protein binds to the DE1B site. These studies also showed that another element DE2B together with DE1A and a previously unidentified downstream site at -57/-41 were able to bind Pax-6. Of these, the -57/-41 site shows the closest match to the in vitro binding site for the paired domain (PD) of Pax-6[39] and corresponds to the single Pax-6 binding site identified in the mouse αA-crystallin promoter[17] (Fig. 5.10). Surprisingly, as in the mouse gene, this consensus Pax-6 site was not previously identified as a discrete element in functional or binding studies of the chicken gene.[90]

Both mouse and chicken αA-crystallin promoters bind complex arrays of general factors which may all be important for function. In spite of their similarities in sequence and in patterns of expression there seem to be considerable differences in the detailed molecular mechanisms of the two promoters. These differences probably result from their relative antiquity. Over time these genes have added complexity and redundancy to their transcriptional machinery to fine tune their patterns of expression in different species.

αB-CRYSTALLIN

If αA-crystallin is essentially specific to lens in its expression, the same cannot be said for αB-crystallin.[106-109] This gene seems to retain more of the non-lens expression of the sHSP family ancestor of this family, just as in protein sequence αB-crystallin is closer than αA-crystallin to sHSPs, p20 and the mysterious *C.elegans* sequence as discussed in chapter 3. In many adult tissues of rodents and man αB-crystallin is expressed constitutively.[90,110] It may also be induced by a variety of stresses, such as heat and osmotic shock in cultured mammalian cells and in various disease states.[90,110] In birds too, αB-crystallin mRNA is present in non-lens tissues in hatched ducks, although non-lens expression was undetectable by Northern blot in embryonic tissues.[81,111] However in contrast to similar cultured mammalian cells (NIH 3T3 cells) duck embryonic fibroblasts have no constitutive expression of αB-crystallin mRNA, nor is there any inducibility by either heat or osmotic stress.[112] It has been suggested that these phenomena may be linked and that some or all of the constitutive expression of αB-crystallin mRNA in mammalian cells in culture is due to the stresses of culture.[112]

The combination of several modes of expression contributes to a complex promoter structure in mammalian αB-crystallin genes. The maintenance of clusters of elements for different functions may also be responsible for bestowing much higher conservation of sequence in αB-crystallin promoters of mammals than is found in αA-crystallin genes of the same species. However this high degree of conservation does not extend to the promoter of the duck gene for αB-crystallin, and phylogenetic footprinting of homologous mammalian and avian aB-crystallin genes reveals only a few discrete islands of similarity[112] (Fig. 5.11). This probably reflects the taxon-specific differences in expression patterns of this gene. It seems that in birds αB-crystallin expression has specialized further than in mammals and what may have been the ancestral condition of stress-inducibility has been lost.

Multiple Transcripts

In mouse lens, heart, skeletal muscle and kidney there is a single major transcription start site downstream of a TATA box in the αB-crystallin gene promoter.[108] In addition, heart and skeletal muscle make use of a minor start site between 40 to 50 bp upstream of the major

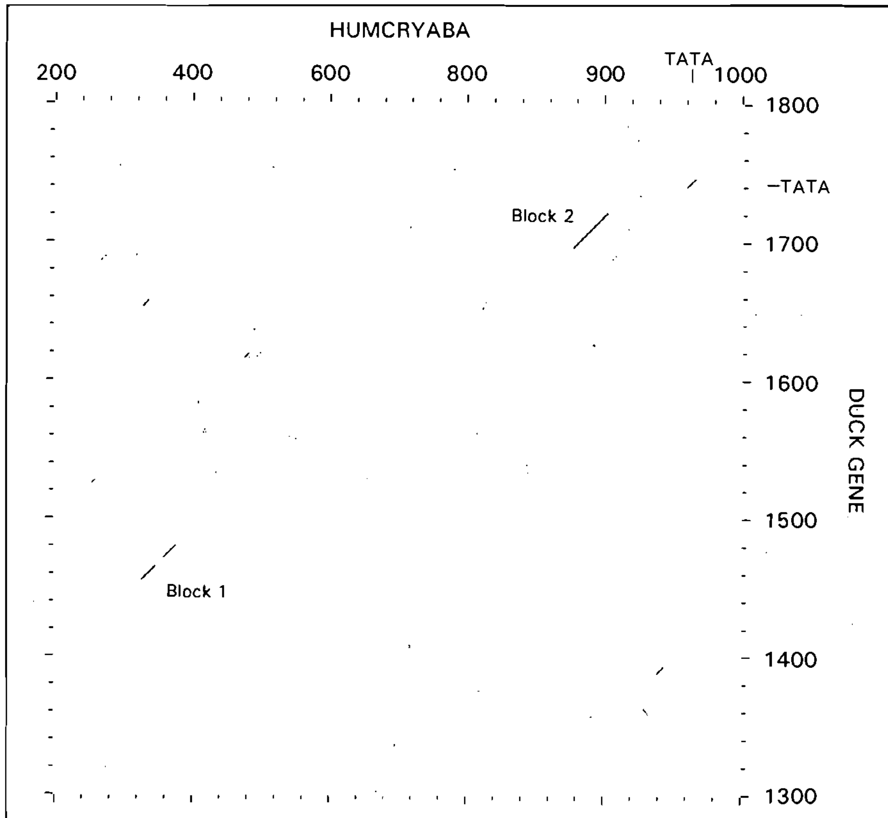


Fig. 5.11. Comparison of promoter regions of duck and human (GenBank entry HUMCRYABA) αB-crystallin genes. Comparison with the equivalent region of the mouse gene gives a very similar result.[112] Block 1, Block 2 and TATA mark the discrete regions of sequence similarity between bird and mammal genes.

site.[108] However, there is also evidence for much longer transcripts in some mammalian tissues which appear to start as much as 474 bp upstream of the major start site.[113] These long transcripts are apparently the predominant form, albeit at low levels, in lung, brain and spleen.[108,113,114] However it has recently been suggested that the far upstream transcripts in mammals are not polyadenylated mRNAs.[115] Furthermore, in duck the major mRNA for αB-crystallin in lung is the same size as in other tissues.[111] Birds do have a long form of αB-crystallin mRNA but this has only been observed in lens, perhaps because of the greater abundance of αB-crystallin in that tissue.[111,116] This longer mRNA arises from use of a downstream alternative site for polyadenylation at the 3' end of the gene.[111,112,117] The major upstream site lies in sequences capable of forming a hairpin loop in RNA transcripts and it has been suggested that this may occasionally hide the site during processing thereby promoting use of the downstream sequence.[111]

Promoter Elements

Although there has been some functional analysis of the human αB-crystallin gene[118] most work has focused on the mouse gene. In transgenic mice the -661/+44 region of the mouse gene promoter was able to direct reporter gene expression in lens and skeletal muscle even when present as a single copy.[119] For detectable expression in heart and other tissues multiple copies of the transgene were required. Transient transfection of deletion mutants showed that -426/-257 contains an enhancer necessary for expression in myotubes while sequences downstream of -115 were essential for lens expression.[119] Binding studies using DNase I protection defined four sites in the enhancer. αBE-1 (-407/-397), αBE-2 (-360/-327) and αBE-3 (-317/-306) were protected by extracts of both myotubes and TN4-1 cells while the fourth region, MRF (-300/-288) which contains a consensus E-box sequence was protected only in myotube extract[120] (Fig. 5.12). In muscle, the enhancer responded to activation by MyoD and myogenin binding through the E-box. In contrast to previous results showing that -115/+44 was essential for lens expression, DNase I protection using αTN4-1 extract, mutagenesis and transient transfection experiments defined -147/-118 as the lens-specific region (LSR).[121]

Most of these sequences are also well conserved in the human gene but in spite of a high degree of general conservation the human gene lacks the E-box of the mouse MRF region (Fig. 5.12). To gain a wider perspective on which elements are well conserved, the mammalian promoters were compared to a more distantly related homologue, the duck αB-crystallin gene[112] (Fig. 5.12). Only the αBE-2 site, which contains a consensus GATA-factor binding sequence[52] was conserved as part of Block 1. However, in contrast to the human gene, the duck gene has an identical E-box in the same general region as the mouse MRF although flanking sequences are not conserved. Other defined elements are much less well conserved. However there is a strikingly well conserved (27/33 identical) block of sequence, Block 2, just upstream of the TATA box in all three genes which lies within the -115/+44 construct originally identified as essential for lens expression. This seems to be a good candidate for an element important in lens preference. Whether Block 2 binds Pax-6 or other lens-preferred factors remains to be seen, although it does not closely resemble known high-affinity Pax-6 sites.

Since αB-crystallin is induced by heat-shock in NIH 3T3 cells it was not surprising that a perfect consensus heat-shock response element (HSRE) (alternative triplet repeats of NGAAN and its complement) is present at positions -53 to -39 in both mouse and human genes[122] (Fig. 5.12). However these sequences and other putative HSREs are absent from the duck gene promoter and accordingly the duck αB-crystallin gene lacks the stress responses exhibited by its mammalian homologues.[112]

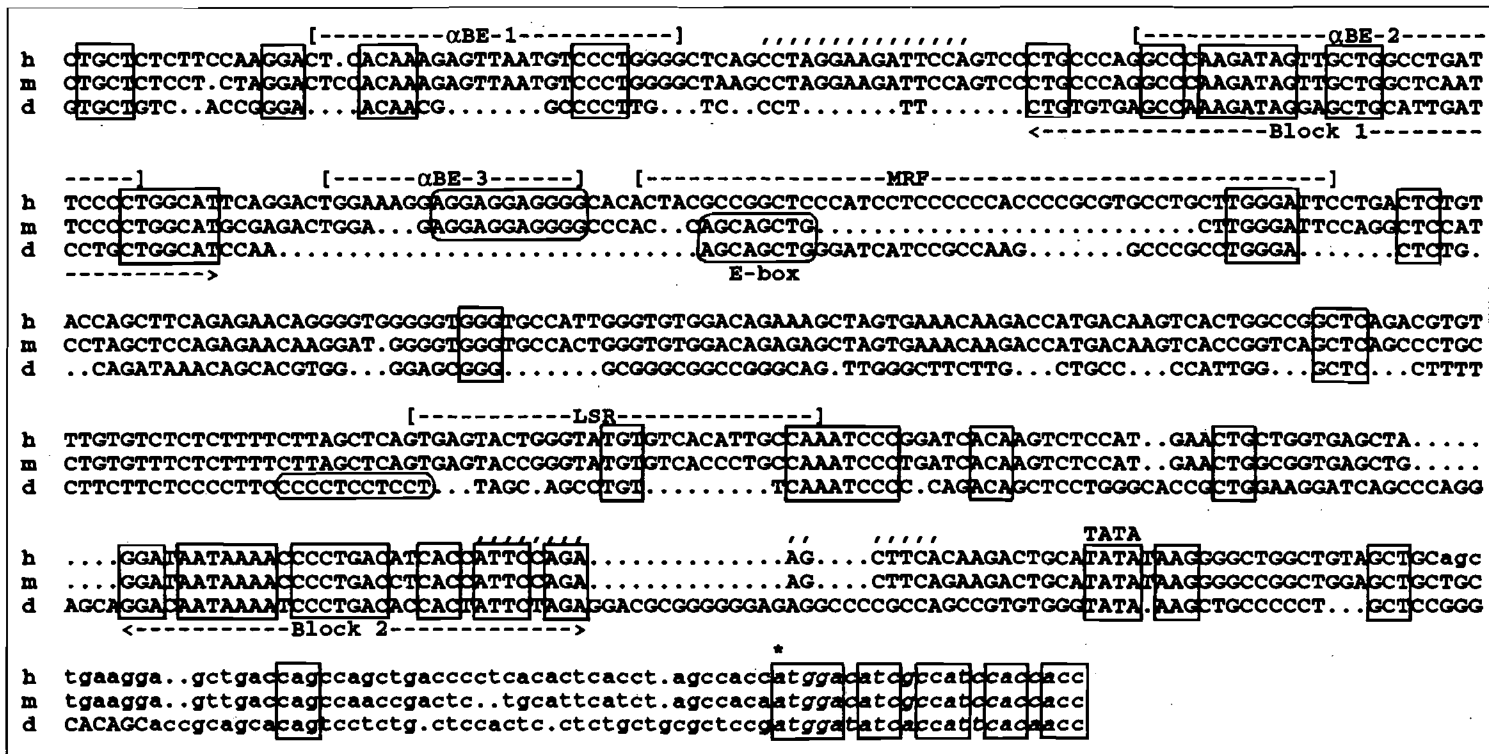


Fig. 5.12. Sequence alignment of promoter regions of human, mouse and duck αB-crystallin genes. αBE-1, αBE-2, αBE-3, MRF, E-box and LSR show functionally defined regions from the mouse promoter. Block 1 and Block 2 show regions of conservation from dot matrix analysis. Putative heat shock response elements are shown by tick marks above the sequences. Transcribed sequences are shown in lower case and protein coding sequence in italics. An asterisk indicates the position of the initiator methionine. Square edged boxes show at least three consecutive bases identical in all three sequences. The E-box which is conserved between mouse and duck and the αBE-3 motif which is present in complement downstream in the duck promoter are shown in round-edged boxes. Reprinted with permission from Wistow G, Graham C. Biochim Biophys Acta 1995: in press.

β-Crystallins

The expression of this multigene family has not been studied with the same intensity as that of other ubiquitous crystallins. β-Crystallins are highly lens-preferred, although it has now been found that detectable levels of various β-crystallin polypeptides are present in a number of non-lens tissues of various developmental origins, including retina, brain, liver and kidney.[123] This means that the promoters for these genes may, like that of αB-crystallin, contain element for expression in different tissues. Members of the family also show differential expression during development and differentiation in the lens adding another level of complexity to studies of their transcriptional regulation.[12,124-126]

Although studies of β-crystallin gene expression are at a relatively early stage, there has been some promoter analysis for two genes from chicken. Several deletion fragments of the chicken βB1-crystallin gene from -434/+30 to -126/+30 were able to drive expression of the CAT reporter gene more efficiently in primary explants of embryonic chicken PLEs than in muscle fibroblasts or HeLa cells.[127] Four functional elements, PL-1, PL-2, OL-1 and OL-2, defined by transfection studies, footprinting and EMSA, were located between positions -126 and -53 of the promoter. Both -2448/+30 and -434/+30 promoter constructs were lens-specific in transgenic mice[128] providing another example of promoter recognition across species.

Two β-crystallin polypeptides, βA3- and βA1-crystallins, are encoded by a single gene in both mammals and birds.[129,130] A promoter construct containing the fragment -382/+22 of the chicken gene promotes expression of the CAT reporter gene in chicken PLEs but not in dermal fibroblasts.[130] Deletion of sequences between -382 and -143 greatly reduces promoter activity. Other than a putative AP-1 site in chicken βA3/A1-crystallin,[130] no transcription factors have yet been identified for β-crystallin genes.

γ-Crystallins

In contrast to other ubiquitous crystallins, the embryonic γ-crystallins of mammals show no evidence of non-lens expression. Everything about their function and evolutionary history suggests that they have been the most highly specialized lens proteins from a very early stage in vertebrate evolution. Although the same thing has been assumed and proved wrong about one crystallin after another, at this point in our understanding γ-crystallins are lens-specific. Furthermore unlike α-crystallins, for example, γ-crystallins are expressed in only one cell type, differentiated fiber cells. The only variability in their expression is a pattern of developmental regulation.[7-11,131] In view of this, it is not surprising that γ-crystallin gene promoters are generally well conserved both among family members and between species.[132] All have a block of similar sequence extending about 90 bp upstream of the transcription start site.

One challenge in studying these genes is that their natural cellular background, the differentiated fiber cell, is not amenable to transient transfection or other cell culture methods. Instead, researchers have been obliged to make use of other systems. The most apparently heterologous of these is the chicken primary lens epithelial explant (PLE) system. Since γ-crystallins are fiber cell specific and since embryonic γ-crystallins are not even present in the chicken this would appear to be a very different environment for these genes. Nevertheless important data have been garnered from these experiments and confirmed in transgenic mouse studies. Presumably some population of cells derived from chicken PLEs acquires fiber cell-like character in culture and this transcriptional environment is evolutionarily conserved with fiber cells of mice.

In mammals, there are six tightly clustered γ-crystallin genes with standardized names γA-γF.[133-136] These are the classic γ-crystallins which are an important part of the embryonic lens and which, as described above, are absent from bird lenses. In rodents all six γ-crystallins are induced as elongating fiber cells form in the embryonic lens. After birth their expression decreases differentially until only γB-crystallin transcripts are detectable in the adult rat lens.[11,137]

The expression of the mouse γF-crystallin gene (formerly designated γ2-crystallin) has been examined in detail in chicken PLEs.[138] In this system the γF-crystallin gene was shown to have a lens-specific promoter consisting of two upstream enhancer-like elements and a proximal promoter[94] (Fig. 5.13). The enhancer elements were successfully substituted by similar regions of the hamster αA-crystallin gene. When these results were extended to studies in transgenic mice it was found that at early stages in lens development either the enhancers or the proximal promoter could direct gene expression, while later in development cooperation between these elements was required for expression in fiber cells.[139]

The proximal promoter of the γF gene contains a 23 bp element, γF-1, which when multimerized can direct reporter gene expression to lens fiber cells and to hindbrain.[139,140] A factor capable of binding the γF-1 site was cloned from chicken brain and named γFBP.[141] γFBP is a zinc finger protein expressed in the sclerotome during early somitogenesis. Its transcripts undergo alternative splicing and one variant form expressed in lens with developmental regulation acts as a transcriptional repressor. It was suggested that γFBP has a role in regulation of γF-crystallin expression and in sclerotome differentiation.[141]

It was also noticed that the mouse γF-crystallin gene enhancer region contained an "everted repeat" which resembled retinoic acid response elements (RARE) (Fig. 5.13). This was investigated and it was found that expression of the γF-crystallin promoter in PLEs was enhanced by retinoic acid, a potent inducer of cell differentiation. The γF-crystallin element was designated as a novel type of RARE and named

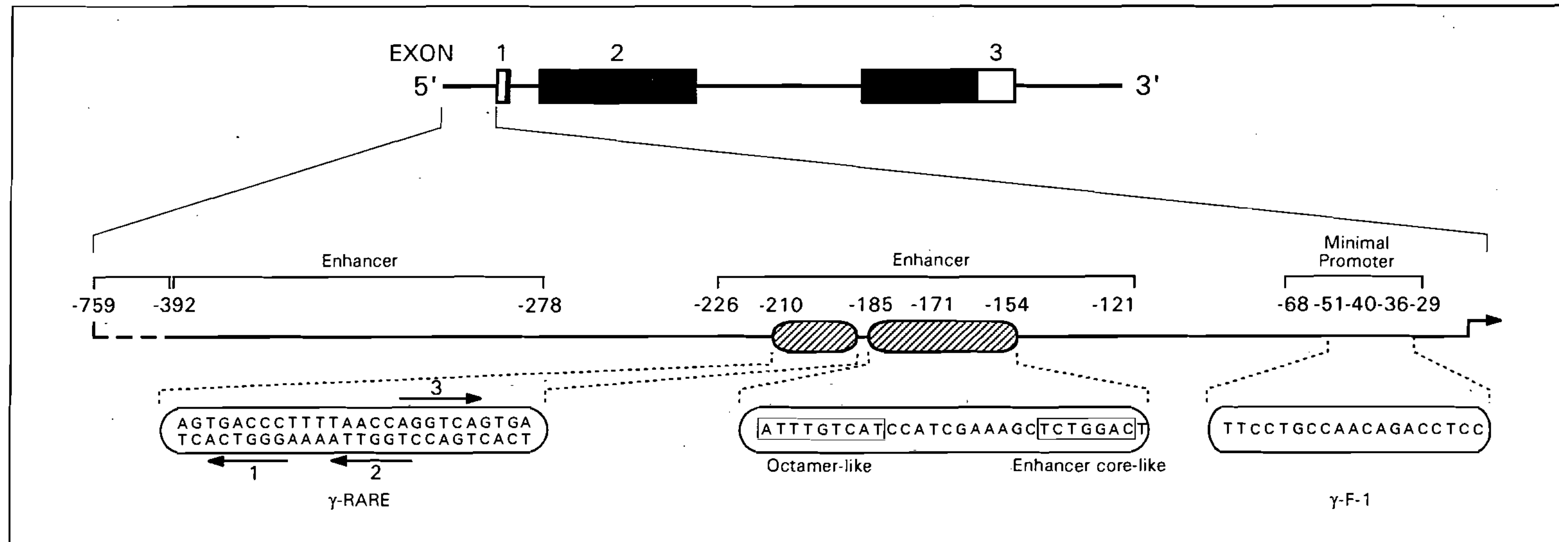


Fig. 5.13. Promoter region of the mouse γF-crystallin gene.

γF-RARE.[142] This element has subsequently been shown to bind heterodimers of the thyroid hormone T3 receptor with either RXR or RAR retinoic acid receptors.[143] Retinoic acid binding proteins seem to have an important role in lens cell differentiation since expression of cellular retinoic acid-binding protein I (CRAB-I)[144] or retinoic acid receptor,[145] under control of the mouse αA-crystallin promoter disrupts fiber cell differentiation in transgenic mice. The CRAB-I transgenic mice also exhibited tumorigenesis in pancreas[144] which is of interest since the important lens transcription factor Pax-6 is also expressed in pancreas.[82]

The other γ-crystallin gene which has been studied intensively is rat γD-crystallin. The promoter was studied by transient transfection into primary mouse lens epithelial cells or into another system, chicken neural retina cells, which can be induced to transdifferentiate into lens-like cells in culture.[146] Functional analyses showed that the γD-crystallin promoter contained at least three important elements but with different requirements for these elements in the two cultured cell systems.[146] The presence of a non-lens silencer was also suggested.[147] Corresponding to the mouse γF-1 site, a 12 nucleotide sequence in the rat γD gene immediately upstream of the TATA box was found to bind a factor found in rat lens but not in retina or brain. Non-lens factors were found to bind to the putative silencing region.[147] The rat γD- and mouse γF-crystallin promoters were also used to express the CAT reporter gene in transgenic *X.laevis* tadpoles, another impressive illustration of the conservation of lens recognition through evolution.[6]

SOX Proteins and γ-Crystallins

Very recently γF-crystallin has joined the group of crystallins whose expression seems to involve transcriptional activation by the products of pattern forming genes. Just as SOX-2 was found to be able to increase lens-specific activity of the δ1-crystallin enhancer, it was also shown to be essential for lens-specific expression of γF-crystallin.[19] This is an important observation since γ-crystallins are strictly specific to lens fiber cells while Pax-6 is expressed in lens epithelial cells.[15] While genes which are activated in the epithelia may be under the control of Pax-6 itself, fiber-specific genes might be expected to be regulated by other factors which could be downstream of Pax-6 in the eye cascade and in lens differentiation.

TRANSGENICS

Crystallin gene promoters have been used to direct expression of foreign genes to the transgenic mouse lens. As discussed above, some of these experiments were principally concerned with in vivo functional analysis of the promoters themselves, using the bacterial genes chloramphenicol acetyltransferase (CAT) or β-galactosidase (lacZ) as reporters. However crystallin promoters have also been used to express

a variety of oncogenes, viral proteins and toxins in lens to investigate processes of differentiation and development. In many cases this directed expression has made it possible to seriously disrupt the lens without compromising viability.

The most widely used promoter for these experiments has been the extensively characterized and strongly lens-preferred mouse αA-crystallin gene promoter. The -366/+45 fragment of this promoter has been used to transform lens cells with SV40 large T antigen[148] and polyoma virus large T antigen.[149] While SV40 large T was oncogenic in lens polyoma large T was not. The dbl oncogene[150] and human papilloma virus type 16 E6 and E7 oncogenes[151,152] have been used to disrupt lens differentiation. A fusion of retinoic acid receptor-α and lacZ,[145] cellular retinoic acid-binding protein I,[144] human or murine urokinase-type plasminogen activator (uPA)[153] and γ-interferon[154] have all been targeted to mouse lens also resulting in failure of fiber cell differentiation. Human immunodeficiency virus TAT protein has been expressed by the mouse αA-crystallin promoter and has been able to transactivate expression from the HIV long terminal repeat in double transgenics.[155]

The toxins diphtheria toxin A (DT-A)[156] and ricin A[157] under the control of the mouse αA-crystallin promoter have been used to ablate transgenic mouse lens cells resulting in microphthalmia. DT-A expression can completely eliminate lens cells while transgenic mice expressing ricin apparently retain some lens cells. The more complete ablation with DT-A may result from the ability of the toxin to be released from expressing cells and to kill neighboring cells. The FGF (fibroblast growth factor) family members FGF-3/Int-2[158] and aFGF[159] have also been expressed in lens. Lens development was again disrupted. In particular a secreted form of aFGF was able to induce differentiation in lens epithelial cells although a non-secreted form had no effect.[159] These results, together with those comparing DT-A and ricin expression[156,157] suggest that in transgenic animals the -366/+45 mouse αA-crystallin promoter may not be expressed at high levels in epithelial cells.

The -347/+43 fragment of the hamster αA-crystallin promoter has also been used to express CAT in transgenic mouse lens and to ablate cells with DT-A.[94,160] The DT-A lens ablation results are similar to those in mouse, with microphthalmia, loss of lens and several surrounding lens tissues.

The -759/+45 fragment of the mouse γF-crystallin promoter has been used to direct expression of CAT, lacZ, DT-A and SV40 T-antigen to the fiber cells of transgenic mice.[9,94,139,161] A large fragment of the chicken δ1-crystallin promoter (-2200/+51) directed expression of the xanthine-guanine phosphoribosyl transferase gene primarily to the lens of chimeric transgenic mice.[162]

The promoter for a mammalian enzyme crystallin, guinea pig ζ-crystallin, also shows clear lens preference in transgenic mice.[51] Since the endogenous ζ-crystallin gene is expressed at high levels in lens

epithelia[15] this promoter offers the potential to target higher expression of transgenes to epithelial cells than has sometimes been possible with αA-crystallin constructs.

A powerful new technique in transgenic research involves the use of sequence specific prokaryotic DNA recombinases, such as Cre.[163] When directed to a specific tissue the Cre recombinase can be used to splice other transgenic recombinant sequences to activate or inactivate another transgene carrying the recombinase recognition sequence. Already this system has been used to produce targeted activation of the oncogenic SV40 T antigen in lens.[164,165] In the future, homologous recombination could be used to replace a gene with an engineered copy containing recombinase sites which could then be specifically spliced out in the lens by tissue-specific expression of Cre.

POST-TRANSCRIPTIONAL CONTROL

Most studies have concentrated on transcriptional mechanisms for tissue-specific gene expression of crystallins and these have been quite successful. However they may not tell the whole story. For example, although crystallin promoters confer lens-specific expression in transgenic mice, the level of expression of reporter genes does not seem to be comparable to that of the crystallin itself. While this could be due to the absence from the recombinant constructs of enhancers or other positive elements present in the complete gene, it is also possible that post-transcriptional events are also important in high level crystallin expression. These could include enhanced crystallin mRNA stability, specific mechanisms for processing or translation of crystallin mRNA or enhanced crystallin protein stability.

Evidence for such mechanisms comes from observations of a marked discordance between levels of mRNA and protein for αA- and αB-crystallins. In rat lens it was found that the overall level of mRNA for αB was higher than that for αA, the reverse of the relative abundances at the protein level.[12] In the same experiments mRNA for βB2-crystallin was also found to be relatively overrepresented. When rat lens epithelial cells were induced to differentiate in culture using bFGF it was found that crystallin genes were induced at specific time points in the order α, β, γ.[166] This corresponds well to the pattern of crystallin expression in the lens, with γ-crystallins being the most fiber-specific and α-crystallins the least. The mRNAs were stable for several days in culture then all disappeared. It was suggested that bFGF caused a pulse of gene expression for each class of gene. Crystallin mRNAs were then stable until removed by a differentiation specific mechanism.[166] Since crystallin mRNA in the lens is apparently very stable there would be no need for continual gene expression to provide for continuing protein synthesis.

The possibility of a burst of protein synthesis of a taxon-specific crystallin has been suggested by immunochemical localization of ζ-crystallin in adult guinea pig lens.[15] The crystallin was detected in

cytosol in epithelial cells and in cortical fibers, but staining was markedly more intense in the newly elongating fiber cells of the equatorial region. Whether by transcriptional or post-transcriptional mechanisms it makes sense to see an increase in crystallin synthesis as fiber cells elongate and the demand for protein in a single cell increases enormously. Indeed such boosted protein synthesis could contribute to cell elongation in several ways. Increased protein concentrations could contribute to an osmotic pressure gradient increasing the inflow of water and increasing cell volume. Certain crystallins could also participate in elaboration and organization of cytoskeleton for the elongating cell.

SUMMARY

The high expression of crystallins in the lens is mainly the result of tissue-specific transcriptional activation. Specificity results from the interplay of lens-preferred factors including Pax-6 and the products of other pattern-forming genes including SOX-2, together with general factors. The recruitment of crystallin genes occurs through the acquisition of binding sites for these factors.

REFERENCES

1. Kondoh H, Yasuda K, Okada TS. Tissue-specific expression of a cloned chick δ-crystallin gene in mouse cells. Nature 1983; 301:440-2.
2. Lok S, Breitman ML, Chepelinsky AB, Piatigorsky J, Gold RJM, Tsui L-C. Lens-specific promoter activity of a mouse γ-crystallin gene. Mol Cell Biol 1985; 5:2221-30.
3. Kondoh H, Katoh K, Takahashi Y et al. Specific expression of the chicken δ-crystallin gene in the lens and the pyramidal neurons of the piriform cortex in transgenic mice. Dev Biol 1987; 120:177-85.
4. Klement JF, Wawrousek EF, Piatigorsky J. Tissue-specific expression of the chicken αA-crystallin gene in cultured lens epithelia and transgenic mice. J Biol Chem 1989; 264:19837-44.
5. Piatigorsky J. Lens crystallins. Innovation associated with changes in gene regulation. J Biol Chem 1992; 267:4277-80.
6. Brakenhoff RH, Ruuls RC, Jacobs EH, Schoenmakers JG, Lubsen NH. Transgenic *Xenopus laevis* tadpoles: a transient in vivo model system for the manipulation of lens function and lens development. Nucleic Acids Res 1991; 19:1279-84.
7. Van Leen RW, Breuer ML, Lubsen NH, Schoenmakers JGG. Developmental expression of crystallin genes: in situ hybridization reveals a differential localization of specific mRNAs. Develop Biol 1987; 123:338-45.
8. Murer-Orlando M, Paterson RC, Lok S, Tsui LC, Breitman ML. Differential regulation of γ-crystallin genes during mouse lens development. Dev Biol 1987; 119:260-7.
9. Breitman ML, Bryce DM, Giddens E et al. Analysis of lens cell fate and eye morphogenesis in transgenic mice ablated for cells of the lens lineage. Development 1989; 106:457-63.
10. 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.
11. Goring DR, Breitman ML, Tsui LC. Temporal regulation of six crystallin transcripts during mouse lens development. Exp Eye Res 1992; 54:785-95.
12. Aarts HJM, Lubsen NH, Schoenmakers JGG. Crystallin gene expression during rat lens development. Eur J Biochem 1989; 183:31-6.
13. Zwaan J. The appearance of α-crystallin in relation to cell cycle phase in the embryonic mouse lens. Dev Biol 1983; 96:173-81.
14. Oguni M, Setogawa T, Hashimoto R, Tanaka O, Shinohara H, Kato K. Ontogeny of α-crystallin subunits in the lens of human and rat embryos. Cell Tissue Res 1994; 276:151-4.
15. Richardson J, Cvekl A, Wistow G. Pax-6 is essential for lens specific expression of ζ-crystallin. Proc Natl Acad Sci USA 1995; 92:4676-80.
16. Cvekl A, Sax CM, Bresnick EH, Piatigorsky J. A complex array of positive and negative elements regulates the chicken αA-crystallin gene - involvement of Pax-6, USF, CREB and/or CREM and AP-1 proteins. Mol Cell Biol 1994; 14:7363-76.
17. Cvekl A, Kashanchi F, Sax CM, Brady JN, Piatigorsky J. Transcriptional regulation of the mouse αA-crystallin gene: Activation dependent on a cyclic AMP responsive element (DE1/CRE) and a Pax-6 binding site. Mol Cell Biol 1995; 15:653-60.
18. Cvekl A, Sax CM, Li X, McDermott JB, Piatigorsky J. Pax-6 and lens-specific transcription of the chicken δ1-crystallin gene. Proc Natl Acad Sci USA 1995; 92:4681-5.
19. Kamachi Y, Sockanathan S, Liu Q, Breitman M, Lovell-Badge R, Kondoh H. Involvement of SOX proteins in lens-specific activation of crystallin genes. EMBO J 1995:in press.
20. Krumlauf R. Hox genes in vertebrate development. Cell 1994; 78:191-201.
21. McGinnis W, Krumlauf R. Homeobox genes and axial patterning. Cell 1992; 68:283-302.
22. McGinnis W, Garber RL, Wirz J, Kuroiwa A, Gehring WJ. A homologous protein-coding sequence in *Drosophila* homeotic genes and its conservation in other metazoans. Cell 1984; 37:403-8.
23. Manak JR, Scott MP. A class act: conservation of homeodomain protein functions. Development 1994; 1994(Supplement):61-71.
24. Gehring WJ, Affolter M, Burglin T. Homeodomain proteins. Annu Rev Biochem 1994; 63:487-526.
25. Hill RE, Hanson IM. Molecular genetics of the Pax gene family. Curr Opinion Cell Biol 1992; 4:967-72.
26. Noll M. Evolution and role of Pax genes. Curr Opin Genet Dev 1993; 3:595-605.
27. Walther C, Gruss P. Pax-6, a murine paired box gene, is expressed in the developing CNS. Development 1991; 113:1435-49.
28. Gruss P, Walther C. Pax in development. Cell 1992; 69:719-22.
29. Li HS, Yang JM, Jacobson RD, Pasko D, Sundin O. Pax-6 is first expressed in a region of ectoderm anterior to the early neural plate: implications for stepwise determination of the lens. Dev Biol 1994; 162:181-94.

30. Krauss S, Johansen T, Korzh V, Moens U, Ericson JU, Fjose A. Zebrafish pax[zf-a]: a paired box-containing gene expressed in the neural tube. EMBO J 1991; 10:3609-19.
31. Hill RE, Favor J, Hogan BL et al. Mouse *small eye* results from mutations in a paired-like homeobox-containing gene. Nature 1991; 354:522-5.
32. Ton CC, Hirvonen H, Miwa H et al. Positional cloning and characterization of a paired box- and homeobox-containing gene from the aniridia region. Cell 1991; 67:1059-74.
33. Glaser T, Walton DS, Maas RL. Genomic structure, evolutionary conservation and aniridia mutations in the human PAX6 gene. Nat Genet 1992; 2:232-9.
34. Hanson IM, Fletcher JM, Jordan T et al. Mutations at the PAX6 locus are found in heterogeneous anterior segment malformations including Peters' anomaly. Nat Genet 1994; 6:168-73.
35. Quiring R, Walldorf U, Kloter U, Gehring WJ. Homology of the *eyeless* gene of *Drosophila* to the *Small eye* gene in mice and Aniridia in humans. Science 1994; 265:785-9.
36. Zhang Y, Emmons SW. mab-18 contains a homeo domain similar to vertebrate pax and *Drosophila* paired box containing genes. Worm Breeder's Gazette 1993; 13:54.
37. Chisholm A, Horvitz B. The worm without a face: an update on vab-3. Worm Breeder's Gazette 1993; 13:48.
38. Halder G, Callaerts P, Gehring WJ. Induction of ectopic eyes by targeted expression of the *eyeless* gene in *Drosophila*. Science 1995; 267:1788-92.
39. Epstein J, Cai J, Glaser T, Jepeal L, Maas R. Identification of a Pax paired domain recognition sequence and evidence for DNA-dependent conformational changes. J Biol Chem 1994; 269:8355-61.
40. Xu W, Rould MA, Jun S, Desplan C, Pabo CO. Crystal structure of a paired domain-DNA complex at 2.5 A resolution reveals structural basis for Pax developmental mutations. Cell 1995; 80:639-50.
41. Chalepakis G, Wijnholds J, Gruss P. Pax-3-DNA interaction: flexibility in the DNA binding and induction of DNA conformational changes by paired domains. Nucleic Acids Res 1994; 22:3131-7.
42. Epstein JA, Glaser T, Cai J, Jepeal L, Walton DS, Maas RL. Two independent and interactive DNA-binding subdomains of the Pax6 paired domain are regulated by alternative splicing. Genes Dev 1994; 8:2022-35.
43. Plaza S, Dozier C, Saule S. Quail Pax-6 (Pax-QNR) encodes a transcription factor able to bind and trans-activate its own promoter. Cell Growth Differ 1993; 4:1041-50.
44. Glaser T, Jepeal L, Edwards JG, Young SR, Favor J, Maas RL. Pax6 gene dosage effect in a family with congenital cataracts, aniridia, anophthalmia and central nervous system defects. Nat Genet 1994; 7:463-71.
45. Chalepakis G, Wijnholds J, Giese P, Schachner M, Gruss P. Characterization of Pax-6 and Hoxa-1 binding to the promoter region of the neural cell adhesion molecule L1. DNA Cell Biol 1994; 13:891-900.
46. Huang QL, Russell P, Stone S, Zigler JS. ζ-Crystallin, a novel lens protein from the guinea pig. Curr Eye Res 1987; 6:725-32.

47. Garland D, Rao PV, Del Corso A, Mura U, Zigler JS Jr. ζ-Crystallin is a major protein in the lens of *Camelus dromedarius*. Arch Biochem Biophys 1991; 285:134-6.
48. Lee DC, Gonzalez P, Rao PV, Zigler JS Jr, Wistow GJ. Carbonyl-metabolizing enzymes and their relatives recruited as structural proteins in the eye lens. In: Weiner H, ed. Enzymology and Molecular Biology of Carbonyl Metabolism. v. 4. New York: Plenum Press, 1993:159-68. (Advances in Experimental Medicine and Biology; 284).
49. Rodokanaki A, Holmes RK, Borras T. ζ-Crystallin, a novel protein from the guinea pig lens is related to alcohol dehydrogenases. Gene 1989; 78:215-24.
50. Gonzalez P, Hernandez-Calzadilla C, Rao PV, Rodriguez IR, Zigler JS Jr, Borras T. Comparative analysis of the ζ-crystallin/quinone reductase gene in guinea pig and mouse. Mol Biol Evol 1994; 11:305-15.
51. Lee DC, Gonzalez P, Wistow G. ζ-Crystallin: A lens-specific promoter and the gene recruitment of an enzyme as a crystallin. J Mol Biol 1994; 236:669-78.
52. Faisst S, Meyer S. Compilation of vertebrate-encoded transcription factors. Nucleic Acids Res 1992; 20:3-26.
53. Nelson LB, Spaeth GL, Nowinski TS, Margo CE, Jackson L. Aniridia. A review. Surv Ophthalmol 1984; 28:621-42.
54. Parker DS, Wawrousek EF, Piatigorsky J. Expression of the δ-crystallin genes in the embryonic chicken lens. Dev Biol 1988; 126:375-81.
55. Kondoh H, Araki I, Yasuda K, Matsubasa T, Mori M. Expression of the chicken 'δ2-crystallin' gene in mouse cells: evidence for encoding of argininosuccinate lyase. Gene 1991; 99:267-71.
56. Wistow G, Piatigorsky J. Lens crystallins: evolution and expression of proteins for a highly specialized tissue. Ann Rev Biochem 1988; 57:479-504.
57. Piatigorsky J. δ-Crystallin and their nucleic acids. Mol Cell Biochem 1984; 59:33-56.
58. Borras T, Nickerson JM, Chepelinsky AB, Piatigorsky J. Structural and functional evidence for differential promoter activities of the two linked δ-crystallin genes in the chicken. EMBO J 1985; 4:445-52.
59. Borras T, Peterson CA, Piatigorsky J. Evidence for positive and negative regulation in the promoter of the chicken δ1-crystallin gene. Dev Biol 1988; 127:209-19.
60. Hayashi S, Goto K, Okada TS, Kondoh H. Lens-specific enhancer in the third intron regulates expression of the chicken δ1-crystallin gene. Genes Dev 1987; 1:818-28.
61. Thomas G, Zelenka PS, Cuthbertson RA, Norman BL, Piatigorsky J. Differential expression of the two δ-crystallin/argininosuccinate lyase genes in lens, heart and brain of chicken embryos. New Biologist 1990; 2:903-14.
62. Li X, Wistow GJ, Piatigorsky J. Linkage and expression of the argininosuccinate lyase/δ-crystallin genes of the duck: insertion of a CR1 element in the intragenic spacer. Biochim Biophys Acta 1995; 1261:25-34.
63. Goto K, Okada TS, Kondoh H. Functional cooperation of lens-specific

and nonspecific elements in the δ1-crystallin enhancer. Mol Cell Biol 1990; 10:958-64.
64. Funahashi JI, Kamachi Y, Goto K, Kondoh H. Identification of nuclear factor δEF1 and its binding site essential for lens-specific activity of the δ1-crystallin enhancer. Nucleic Acids Res 1991; 19:3543-7.
65. Kamachi Y, Kondoh H. Overlapping positive and negative regulatory elements determine lens-specific activity of the δ1-crystallin enhancer. Mol Cell Biol 1993; 13:5206-15.
66. Sekido R, Murai K, Funahashi J et al. The δ-crystallin enhancer-binding protein delta EF1 is a repressor of E2-box-mediated gene activation. Mol Cell Biol 1994; 14:5692-700.
67. Rex M, Uwanogho D, Cartwright E, Pearl G, Sharpe PT, Scotting PJ. Sox gene expression during neuronal development. Biochem Soc Trans 1994; 22:252S.
68. Vriz S, Lovell-Badge R. The zebrafish Zf-Sox 19 protein: a novel member of the Sox family which reveals highly conserved motifs outside of the DNA-binding domain. Gene 1995; 153:275-6.
69. Wright EM, Snopek B, Koopman P. Seven new members of the Sox gene family expressed during mouse development. Nucleic Acids Res 1993; 21:744.
70. Kim RY, Lietman T, Piatigorsky J, Wistow GJ. Structure and expression of the duck α-enolase/τ-crystallin encoding gene. Gene 1991; 103:193-200.
71. Rudner G, Katar M, Maisel H. Enolase in the avian and turtle lens. Curr Eye Res 1990; 9:139-50.
72. Kim RY, Wistow GJ. Expression of the duck α-enolase/t-crystallin gene in transgenic mice. FASEB J 1993; 7:464-9.
73. Giallongo A, Feo S, Moore R, Croce CM, Showe LC. Molecular Cloning and Nucleotide Sequence of a Full-length cDNA for Human α-Enolase. Proc Natl Acad Sci USA 1986; 83:6741-5.
74. Harris LL, Talian JC, Zelenka PS. Contrasting patterns of c-myc and N-myc expression in proliferating, quiescent and differentiating cells of the embryonic chicken lens. Development 1992; 115:813-20.
75. Giallongo A, Oliva D, Cali L, Barba G, Barbieri G, Feo S. Structure of the human gene for α-enolase. Eur J Biochem 1990; 190:567-73.
76. Blackwell TK, Kretzner L, Blackwood EM, Eisenman RN, Weintraub H. Sequence-specific DNA binding by the c-Myc protein. Science 1990; 250:1149-51.
77. Halazonetis TD, Kandil AN. Determination of the c-MYC DNA-binding site. Proc Natl Acad Sci USA 1991; 88:6162-6.
78. Warwar RE, Kim RY, Wistow GJ, Zelenka PS. The τ-crystallin/α-enolase gene: a candidate for regulation by c-myc. Invest Ophthalmol Vis Sci 1992; 33:794.
79. Kraft HJ, Hendriks W, de Jong WW, Lubsen NH, Schoenmakers JGG. Duck lactate dehydrogenase B/ε-crystallin gene: lens recruitment of a GC-promoter. J Mol Biol 1993; 229:849-59.
80. Hodin J, Wistow G. 5'-RACE PCR of mRNA for three taxon-specific crystallins: For each gene one promoter controls both lens and non-lens

expression. Biochem Biophys Res Commun 1993; 190:391-6.

81. Kraft HJ, Voorter CE, Wintjes L, Lubsen NH, Schoenmakers JG. The developmental expression of taxon-specific crystallins in the duck lens. Exp Eye Res 1994; 58:389-95.
82. Turque N, Plaza S, Radvanyi F, Carriere C, Saule S. Pax-QNR/Pax-6, a paired box- and homeobox-containing gene expressed in neurons, is also expressed in pancreatic endocrine cells. Mol Endocrinol 1994; 8:929-38.
83. 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.
84. Jaworski CJ, Chepelinsky AB, Piatigorsky J. The αA-crystallin gene: conserved features of the 5'-flanking regions in human, mouse, and chicken. J Mol Evol 1991; 33:495-505.
85. van den Heuvel R, Hendriks W, Quax W, Bloemendal H. Complete structure of the hamster αA-crystallin gene. Reflection of an evolutionary history by means of exon shuffling. J Mol Biol 1985; 185:273-84.
86. Hendriks W, Leunissen J, Nevo E, Bloemendal H, de Jong WW. The lens protein αA-crystallin of the blind mole rat, *Spalax ehrenbergi*: Evolutionary change and functional constraints. Proc Natl Acad Sci USA 1987; 84:5320-4.
87. Thompson MA, Hawkins JW, Piatigorsky J. Complete nucleotide sequence of the chicken αA-crystallin gene and its 5' flanking region. Gene 1987; 56:173-84.
88. Kato K, Shinohara H, Kurobe N, Goto S, Inaguma Y, Ohshima K. Immunoreactive αA crystallin in rat non-lenticular tissues detected with a sensitive immunoassay method. Biochim Biophys Acta 1991; 1080:173-80.
89. Gumucio DL, Shelton DA, Bailey WJ, Slightom JL, Goodman M. Phylogenetic footprinting reveals unexpected complexity in trans factor binding upstream from the ε-globin gene. Proc Natl Acad Sci USA 1993; 90:6018-22.
90. Sax C, Piatigorsky J. Expression of the α-crystallin/small heat shock protein/molecular chaperone genes in the lens and other tissues. In: Advances in Enzymology and Related Areas in Molecular Biology. v. 69. New York, NY: John Wiley & Sons Inc, 1994; 155-201.
91. Chepelinsky AB, Sommer B, Piatigorsky J. Interaction between two different regulatory elements activates the murine αA-crystallin gene promoter in explanted lens epithelia. Mol Cell Biol 1987; 7:1807-14.
92. Yasuda K, Kitamura M, Okazaki K, Takeuchi M. Mechanism involved in regulation α-crystallin gene expression. In: Piatigorsky J, Shinohara T, and Zelenka PS, eds. UCLA Symp. on the Molecular Biology of the Eye Lens: Genes, Vision, and Ocular Disease; Ed. v. Vol. 8. New York.: Alan R. Liss, Inc., 1988; 205-13.
93. Kantorow M, Cvekl A, Sax C, Piatigorsky J. Protein-DNA interactions of the mouse αA-crystallin control regions. J Mol Biol 1993; 230:425-35.
94. Yu CC-K, Tsui L-C, Breitman ML. Homologous and heterologous enhancers modulate spatial expression but not cell-type specificity of the murine γF-crystallin promoter. Development 1990; 110:131-9.

95. Sax CM, Cvekl A, Kantorow M et al. Lens-specific activity of mouse αA-crystallin promoter in the absence of a TATA box: functional and protein binding analysis of the mouse αA-crystallin PE1 region. Nucleic Acids Res 1995; 23:442-51.
96. Nakamura T, Donovan DM, Hamada K et al. Regulation of the mouse αA-crystallin gene: isolation of a cDNA encoding a protein that binds to a *cis* sequence motif shared with the major histocompatibility complex class I gene and other genes. Mol Cell Biol 1990; 10:3700-8.
97. Fan C-M, Maniatis T. A DNA-binding protein containing two widely separated zinc finger motifs that recognize the same DNA sequence. Genes Dev 1990; 4:29-42.
98. Baldwin AS Jr, LeClair KP, Singh H, Sharp PA. A large protein containing zinc finger domains binds to related sequence elements in the enhancers of the class I major histocompatibility complex and kappa immunoglobulin genes. Mol Cell Biol 1990; 10:1406-14.
99. Donovan DM, Sax CM, Klement JF, Li X, Chepelinsky AB, Piatigorsky J. Conservation of mouse αA-crystallin promoter activity in chicken lens epithelial cells. J Mol Evol 1992; 35:337-45.
100. Sax CM, Klement JF, Piatigorsky J. Species-specific lens activation of the thymidine kinase promoter by a single copy of the mouse αA-CRYBP1 site and loss of tissue specificity by multimerization. Mol Cell Biol 1990; 10:6813-6.
101. Sax CM, Ilagan JG, Piatigorsky J. Functional redundancy of the DE-1 and αA-CRYBP1 regulatory sites of the mouse αA-crystallin promoter. Nucleic Acids Res 1993; 21:2633-40.
102. Klement JF, Cvekl A, Piatigorsky J. Functional elements DE2A, DE2B, and DE1A and the TATA box are required for activity of the chicken αA-crystallin gene in transfected lens epithelial cells. J Biol Chem 1993; 268:6777-84.
103. Matsuo I, Yasuda K. The cooperative interaction between two motifs of an enhancer element of the chicken αA-crystallin gene, αCE1 and αCE2, confers lens-specific expression. Nucleic Acids Res 1992; 20:3701-12.
104. Matsuo I, Kitamura M, Okazaki K, Yasuda K. Binding of a factor to an enhancer element responsible for the tissue-specific expression of the chicken αA-crystallin gene. Development 1991; 113:539-50.
105. Matsuo I, Takeuchi M, Yasuda K. Identification of the contact sites of a factor that interacts with motif I (αCE1) of the chicken αA-crystallin lens-specific enhancer. Biochem Biophys Res Comm 1992; 184:24-30.
106. Duguid JR, Rohwer RG, Seed B. Isolation of cDNAs of scrapie-modulated RNAs by subtractive hybridization of a cDNA library. Proc Natl Acad Sci USA 1988; 85:5738-42.
107. Iwaki T, Kume-Iwaki A, Liem RK, Goldman JE. αB-crystallin is expressed in non-lenticular tissues and accumulates in Alexander's disease brain. Cell 1989; 57:71-8.
108. Dubin RA, Wawrousek EF, Piatigorsky J. Expression of the murine αB-crystallin gene is not restricted to the lens. Mol Cell Biol 1989; 9:1083-91.

109. Iwaki T, Kume-Iwaki A, Goldman JE. Cellular distribution of αB-crystallin in non-lenticular tissues. J Histochem Cytochem 1990; 38:31-9.
110. de Jong WW, Leunissen JAM, Voorter CEM. Evolution of the α-crystallin/small heat-shock protein family. Mol Biol Evol 1993; 10:103-26.
111. Lee DC, Kim RY, Wistow GJ. An avian αB-crystallin. Non lens expression and sequence similarities with both small (hsp27) and large (HSP70) heat shock proteins. J Mol Biol 1993; 232:1221-6.
112. Wistow G, Graham C. The duck gene for αB-crystallin shows evolutionary conservation of discrete promoter elements but lacks heat and osmotic stress response. Biochim Biophys Acta 1995:in press.
113. Frederikse PH, Dubin RA, Haynes JI, Piatigorsky J. Structure and alternate tissue-preferred transcription initiation of the mouse αB-crystallin/small heat-shock protein gene. Nucleic Acids Res 1994; 22:5686-94.
114. Iwaki A, Iwaki T, Goldman JE, Liem RK. Multiple mRNAs of rat brain α-crystallin B chain result from alternative transcriptional initiation. J Biol Chem 1990; 265:22197-203.
115. Srinivasan AN, Bhat SP. Complete structure and expression of the rat αB-crystallin gene. DNA Cell Biol 1994; 13:651-61.
116. Dodemont H, Groenen M, Jansen L, Schoenmakers J, Bloemendal H. Comparison of the crystallin mRNA populations from rat, calf and duck lens. Evidence for a longer αA2-mRNA and two distinct aB2 in the birds. Biochim Biophys Acta 1985; 824:284-94.
117. Sawada K, Agata K, Eguchi G. Crystallin gene expression in the process of lentoidogenesis in cultures of chicken lens epithelial cells. Exp Eye Res 1992; 55:879-87.
118. Dubin RA, Ally AH, Chung S, Piatigorsky J. Human αB-crystallin gene and preferential promoter function in lens. Genomics 1990; 7:594-601.
119. Dubin RA, Gopal-Srivastava R, Wawrousek EF, Piatigorsky J. Expression of the murine αB-crystallin gene in lens and skeletal muscle: identification of a muscle-preferred enhancer. Mol Cell Biol 1991; 11:4340-9.
120. Gopal-Srivastava R, Piatigorsky J. The murine αB-crystallin/small heat shock protein enhancer: identification of αBE-1, αBE-2, αBE-3, and MRF control elements. Mol Cell Biol 1993; 13:7144-52.
121. Gopal-Srivastava R, Piatigorsky J. Identification of a lens-specific regulatory region (LSR) of the murine αB-crystallin gene. Nucleic Acids Res 1994; 22:1281-6.
122. Klemenz R, Frohli E, Steiger RH, Schafer R, Aoyama A. αB-crystallin is a small heat shock protein. Proc Natl Acad Sci USA 1991; 88:3652-6.
123. Head MW, Peter A, Clayton RM. Evidence for the extralenticular expression of members of the β-crystallin gene family in the chick and a comparison with δ-crystallin during differentiation and transdifferentiation. Differentiation 1991; 48:147-56.
124. Carper D. Deficiency of functional messenger RNA for a developmentally regula β-crystallin polypeptide in a hereditary cataract. Science 1982; 217:463-4.
125. Hejtmancik JF, Beebe DC, Ostrer H, Piatigorsky J. d- and β-crystallin mRNA levels in the embryonic and posthatched chicken lens: temporal and spatial changes during development. Develop Biol 1985; 109:72-81.

126. Voorter CE, de Haard-Hoekman WA, Hermans MM, Bloemendal H, de Jong WW. Differential synthesis of crystallins in the developing rat eye lens. Exp Eye Res 1990; 50:429-37.
127. Roth HJ, Das GC, Piatigorsky J. Chicken βB1-crystallin gene expression: presence of conserved functional polyomavirus enhancer-like and octamer binding-like promoter elements found in non-lens genes. Mol Cell Biol 1991; 11:1488-99.
128. Duncan MK, Roth HJ, Thompson M, Kantorow M, Piatigorsky J. Chicken βB1 crystallin: gene sequence and evidence for functional conservation of promoter activity between chicken and mouse. Biochim Biophys Acta 1995; 1261:68-76.
129. Hogg D, Tsui L-C, Gorin M, Breitman ML. Characterization of the human β-crystallin gene Hu βA3/A1 reveals ancestral relationships among the βγ-crystallin superfamily. J Biol Chem 1986; 261:12420-7.
130. McDermott JB, Peterson CA, Piatigorsky J. Structure and lens expression of the gene encoding chicken βA3/A1-crystallin. Gene 1992; 117:193-200.
131. 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.
132. 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.
133. Lok S, Tsui L-C, Shinohara T, Piatigorsky J, Gold R, Breitman M. Analysis of the mouse γ-crystallin gene family: Assignment of multiple cDNAs to discrete genomic sequences and characterization of representative gene. Nucleic Acids Res 1984; 12:4517-29.
134. Moormann RJ, den Dunnen JT, Heuyerjans J et al. Characterization of the rat γ-crystallin gene family and its expression in the eye lens. J Mol Biol 1985; 182:419-30.
135. Breitman ML, Lok S, Wistow G et al. γ-Crystallin family of the mouse lens: structural and evolution relationships. Proc Natl Acad Sci USA 1984; 81:7762-6.
136. Graw J, Liebstein A, Pietrowski D, Schmitt-John T, Werner T. Genomic sequences of murine γB- and γC-crystallin-encoding genes: promoter analysis and complete evolutionary pattern of mouse, rat and human γ-crystallins. Gene 1993; 136:145-56.
137. van Leen RW, van Roozendaal KEP, Lubsen NH, Schoenmakers JGG. Differential expression of crystallin genes during development of the rat eye lens. Dev Biol 1987; 120:457-64.
138. Lok S, Stevens W, Breitman ML, Tsui LC. Multiple regulatory elements of the murine γ2-crystallin promoter. Nucleic Acids Res 1989; 17:3563-82.
139. Goring DR, Bryce DM, Tsui LC, Breitman ML, Liu Q. Developmental regulation and cell type-specific expression of the murine γF-crystallin gene is mediated through a lens-specific element containing the γF-1 binding site. Dev Dyn 1993; 196:143-52.

140. Liu QR, Tini M, Tsui LC, Breitman ML. Interaction of a lens cell transcription factor with the proximal domain of the mouse γF-crystallin promoter. Mol Cell Biol 1991; 11:1531-7.
141. Liu Q, Shalaby F, Puri MC, Tang S, Breitman ML. Novel zinc finger proteins that interact with the mouse γF-crystallin promoter and are expressed in the sclerotome during early somitogenesis. Dev Biol 1994; 165:165-77.
142. Tini M, Otulakowski G, Breitman ML, Tsui LC, Giguere V. An everted repeat mediates retinoic acid induction of the γF-crystallin gene: evidence of a direct role for retinoids in lens development. Genes Dev 1993; 7:295-307.
143. Tini M, Tsui LC, Giguere V. Heterodimeric interaction of the retinoic acid and thyroid hormone receptors in transcriptional regulation on the γF-crystallin everted retinoic acid response element. Mol Endocrinol 1994; 8:1494-506.
144. Perez-Castro AV, Tran VT, Nguyen-Huu MC. Defective lens fiber differentiation and pancreatic tumorigenesis caused by ectopic expression of the cellular retinoic acid-binding protein I. Development 1993; 119:363-75.
145. Balkan W, Klintworth GK, Bock CB, Linney E. Transgenic mice expressing a constitutively active retinoic acid receptor in the lens exhibit ocular defects. Dev Biol 1992; 151:622-5.
146. Peek R, van der Logt P, Lubsen NH, Schoenmakers JGG. Tissue- and species-specific promoter elements of rat γ-crystallin genes. Nucleic Acids Res 1990; 18:1189-97.
147. Peek R, Kraft HJ, Klok EJ, Lubsen NH, Schoenmakers JGG. Activation and repression sequences determine the lens-specific expression of the rat γD-crystallin gene. Nucleic Acids Res 1992; 20:4865-71.
148. Mahon KA, Chepelinsky AB, Khillan JS, Overbeek PA, Piatigorsky J, Westphal H. Oncogenesis of the lens in transgenic mice. Science 1987; 235:1622-8.
149. Griep AE, Kuwabara T, Lee EJ, Westphal H. Perturbed development of the mouse lens by polyomavirus large T antigen does not lead to tumor formation. Genes Dev 1989; 3:1075-85.
150. Eva A, Graziani G, Zannini M, Merin LM, Khillan JS, Overbeek PA. Dominant dysplasia of the lens in transgenic mice expressing the *dbl* oncogene. New Biol 1991; 3:158-68.
151. Griep AE, Herber R, Jeon S, Lohse JK, Dubielzig RR, Lambert PF. Tumorigenicity by human papillomavirus type 16 E6 and E7 in transgenic mice correlates with alterations in epithelial cell growth and differentiation. J Virol 1993; 67:1373-84.
152. Pan H, Griep AE. Altered cell cycle regulation in the lens of HPV-16 E6 or E7 transgenic mice: implications for tumor suppressor gene function in development. Genes Dev 1994; 8:1285-99.
153. Miskin R, Axelrod JH, Griep AE et al. Human and murine urokinase cDNAs linked to the murine αA-crystallin promoter exhibit lens and non-lens expression in transgenic mice. Eur J Biochem 1990; 190:31-8.

154. Egwuagu CE, Sztein J, Chan CC et al. Ectopic expression of γ-interferon in the eyes of transgenic mice induces ocular pathology and MHC class II gene expression. Invest Ophthalmol Vis Sci 1994; 35:332-41.
155. Khillan JS, Deen KC, Yu SH, Sweet RW, Rosenberg M, Westphal H. Gene transactivation mediated by the TAT gene of human immunodeficiency virus in transgenic mice. Nucleic Acids Res 1988; 16:1423-30.
156. Kaur S, Key B, Stock J, McNeish JD, Akeson R, Potter SS. Targeted ablation of α-crystallin-synthesizing cells produces lens-deficient eyes in transgenic mice. Development 1989; 105:613-9.
157. Landel CP, Zhao J, Bok D, Evans GA. Lens-specific expression of recombinant ricin induces developmental defects in the eyes of transgenic mice. Genes Dev 1988; 2:1168-78.
158. Chepelinsky AB, Robinson ML, Overbeek PA et al. FGF-3 ectopic expression induces differentiation of central lens epithelia and appearance of secretory epithelia in the eyes of transgenic mice. Invest Ophthalmol Vis Sci 1994; 35:1998.
159. Robinson ML, Overbeek PA, Verran DJ et al. Extracellular FGF-1 acts as a lens differentiation factor in transgenic mice. Development 1995; 121:505-14.
160. Harrington L, Klintworth GK, Secor TE, Breitman ML. Developmental analysis of ocular morphogenesis in αA-crystallin/diphtheria toxin transgenic mice undergoing ablation of the lens. Dev Biol 1991; 148:508-16.
161. Bryce DM, Liu Q, Khoo W, Tsui LC, Breitman ML. Progressive and regressive fate of lens tumors correlates with subtle differences in transgene expression in γF-crystallin-SV40 T antigen transgenic mice. Oncogene 1993; 8:1611-20.
162. Takahashi Y, Hanaoka K, Hayasaka M et al. Embryonic stem cell-mediated transfer and correct regulation of the chicken δ-crystallin gene in developing mouse embryos. Development 1988; 102:259-69.
163. Barinaga M. Knockout mice: round two. Science 1994; 265:26-8.
164. Pichel JG, Lakso M, Westphal H. Timing of SV40 oncogene activation by site-specific recombination determines subsequent tumor progression during murine lens development. Oncogene 1993; 8:3333-42.
165. Lakso M, Sauer B, Mosinger B Jr et al. Targeted oncogene activation by site-specific recombination in transgenic mice. Proc Natl Acad Sci USA 1992; 89:6232-6.
166. Peek R, McAvoy JW, Lubsen NH, Schoenmakers JG. Rise and fall of crystallin gene messenger levels during fibroblast growth factor induced terminal differentiation of lens cells. Dev Biol 1992; 152:152-60.
167. Wistow G, Richardson J, Jaworski C, Graham C, Sharon-Friling R, Segovia L. Crystallins: The overexpression of functional enzymes and stress proteins in the eye lens. In: Tombs MP, ed. Biotechnology and Genetic Engineering Reviews. v. 12. Andover, Hants: Intercept Ltd, 1994; 1-38.

CHAPTER 6

A Brief History of Lens and Crystallin Recruitment

Equipped with photosensitive eyespots whose development was under the control of a gene cascade originating with *Pax-6,* simple metazoans began to thrive. Even a primitive level of vision conferred great advantages. Improvements in optics and in processing the sensory data which was gathered conferred even greater advantages. In the ancestors of vertebrates part of the enhancement of the eye occurred through the evolution of a cellular lens. A single change in the differentiation program of a layer of cells covering the retina caused swelling and cell elongation to form a simple concentrating lens. Associated with the stress of this event, certain gene families were induced in the elongating cells. These included genes involved with synthesis and protection of cytoskeleton and with other aspects of osmotic stress. In the vertebrate lineage, these genes included a small heat shock protein homologous to αB-crystallin and at least one member of the βγ-crystallin superfamily, perhaps homologous to βB2-crystallin. These were highly stable proteins able to interact closely with other components of the cytomatrix.

The stress responses in these elongating cells would initially have been no different from those inducible by similar stresses in other cells and the induced genes would not yet have become crystallins. However the function of the new lens would have been greatly improved by increasing its protein concentration and refractive index. The stress proteins which were already necessary for the primitive lens would have been among the easiest targets for further induction. This could have occurred in a tissue-preferred way if these genes acquired binding sites for transcription factors expressed in the eye as part of the "eye cascade." Several different evolutionary experiments must have occurred and most would have failed. For example, high expression of stress genes outside the lens in the retina might have had a deleterious effect on vision while some combinations of eye-related transcription factor binding sites might have caused overexpression in parts of the central

nervous system and elsewhere. However selection of successful experiments produced a fine-tuned, lens-specific system.

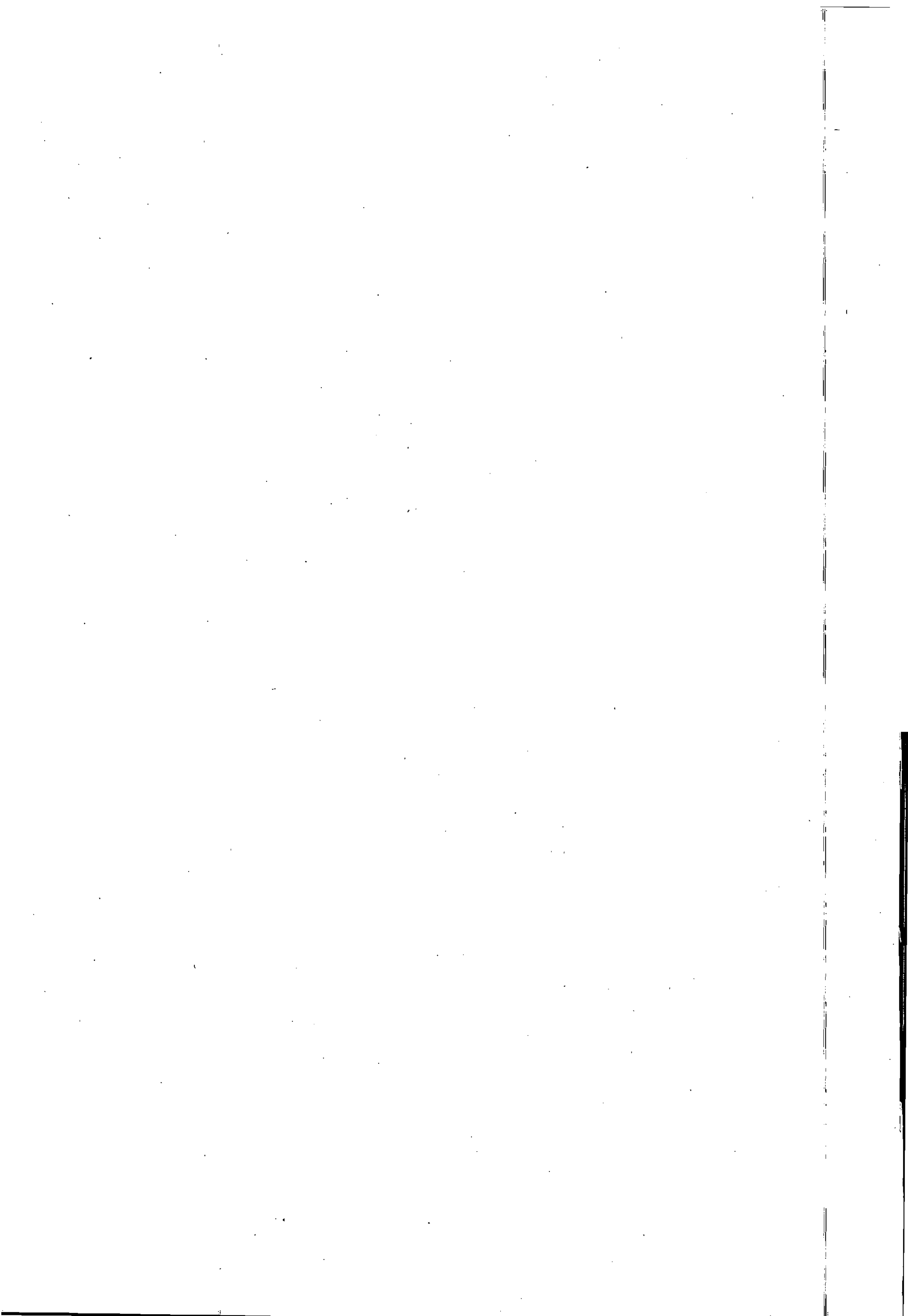
These modifications in gene expression would have involved not only the proto-crystallins but also the transcription factors themselves. Expression patterns of members of the eye cascade must also have been modified to produce a new tissue, the lens. Indeed, it appears that *Pax-6* itself was co-opted for this new program. While continuing its ancient role in defining eye as a whole it acquired an additional role in specifying lens.

Once the path for the development of the lens was established it continued to be refined, setting up differential expression patterns within the tissue and throughout its development. Terminally differentiated fiber cells evolved to provide the core of the lens. These cells acquired a different complement of transcription factors to allow for different patterns of gene expression within the lens. The original complement of recruited proteins began as dual function stress proteins and crystallins. With time and with new demands they experienced adaptive conflicts which were resolved by gene duplication and specialization. A lens-specialized αA-crystallin evolved to play a particular role in association with lens cytoskeleton. The β-crystallins multiplied allowing for more complex patterns of protein expression during development. A very highly specialized protein family specific to the fiber cells also arose, perhaps by duplication from the ancestral βγ-crystallin of the earliest lens. The members of this family were the γ-crystallins which acquired specializations in structure allowing the lens to achieve even higher protein concentrations and higher refractive index.

At this stage the lens had become the sophisticated optical device of the aquatic vertebrate eye. However its evolution was not over. As the vertebrate eye moved from an interface with water to one with air the lens and other parts of the eye adapted to the new environment. Lens-hardening γ-crystallins were replaced or diluted by newly recruited crystallins. This time the recruitment process, occurring in an established tissue, made use of a different class of protein. There was no benefit to be gained from further introductions of stress proteins. Instead metabolic enzymes were recruited. Just as before, the genes for the new crystallins acquired binding sites for transcription factors involved in lens specification and their expression in lens was enhanced while their non-lens expression continued. Unlike the crystallins, these proteins were not specifically adapted for a high protein concentration environment and they contributed to a general softening and hydration of the lens. However it is likely that they were still required to be able to form stable interaction with the cytoskeleton which formed the essential scaffolding of the lens cells. Indeed, even outside the lens many metabolic enzymes are found anchored to cytoskeleton.

While the primary impetus to the recruitment of enzyme crystallins was modification of the optical properties of the lens, other factors were also involved. Some reptiles adopted a habit in which UV glare became a problem. The δ-crystallin in their soft lenses did not provide any opportunity for resolving this problem. However a second enzyme crystallin recruitment proved beneficial. When LDHB was recruited as ε-crystallin, archosaur lenses acquired a mechanism for increased UV filtering. Similar patterns occurred in other lineages. In mammals there was a complex history involving lens-softening, a reversion to hard spherical lenses and then a series of re-modifications with reductions in γ-crystallin expression and various enzyme crystallin recruitments.

Although the lens is a highly specialized tissue in a highly specialized organ, the interplay of development and tissue specification with gene expression, molecular evolution and the environment which has shaped it must have occurred over and over again in different living systems during the history of life on this planet.



INDEX

Page numbers in italics denote figures (f) or tables (t).

A

α-crystallin, 37
  αA-CRYBP1, 142
  and small heat shock proteins, 39-40, 41, 45, *46f*, 47
  gene structure, 41, *43f*
  insert exon, 42, *43f*
  ontogeny, 38
  promoters, 138-139, *140f*, *141f*, 143
  protein structure, 44-45
  structure, 38
αB-crystallin, 144-145, *145f*
  promoters, 146, *147f*
ALDH (aldehyde dehydrogenase) superfamily, 107-108

B

β-crystallin, 47-48, 59, 148
  ancestral, 65
  conformation, *55f*, *56f*
  dimers, 55
  expression patterns, 50-51
  sequence alignment, *54f*
  structure, 52-53
  taxon-specific differences, 48-49
βg-crystallin superfamily, 59-61, *61f*, *62f*
  evolution, 64
  introns, 62-64

C

Crystallins, 1, 25-26, *84f*. *See also* specific crystallins.
  gecko, 97
  gene promoters, 151-153
  gene recruitment, 26, 28, 30, *33f*, 112
  mechanisms of expression, 123-126
  others, 108-109
  recruitment, 165-167
  refraction, 12
  sequence modifications, *87f*
  stress responses, 13, *14t*
  taxon-specific, 12-13, *13t*, *85f*, *100f*, 126, 134
  transcriptional mechanisms, 153
  transparency, 12
  ubiquitous, 12, *13t*

D

δ-crystallin, 89-92
  enhancer, 134-135, 137, *136f*
  His89/Gln89, 92-93
  intron-slippage, 93
  neutrality, 93

E

ε-crystallin, 83, 85, 137
  role, 88-89
  sequence modifications, 85-88, *86f*,
Eye, 15
  evolution of, *3f*, 4

G

γ-crystallin, 26, 47, 49-50, 57-59, 65-68, 148-151, *150f*
  conformation, *55f*
  expression patterns, 50-51
  monomers, 53
  sequence alignment, *54f*, *66f*
  structure, *51f*, 52-53

H

Human eye
  development, *8f*
  structure, *5f*, 4-6

L

λ-crystallin, 103
Lens
  differentiation, *9f*, 10
  evolution of, 11, 17-18, 27, 99, *100f*
  growth factors, 10
  recruitment, 165-167
  regeneration, 16
  structure and development, 6-7, 9
  transdifferentiation, 16

M

μ-crystallin, 104-106
Molecular chaperones, 40

N

η-crystallin, 101-102

P

Pax-6, 14, 17, 130, *132f*, 133-134, 142-144
Pax-6, 14-17, 124-126, *125f*, 130, 134
Protein evolution, *29f*, *30f*, 31-32

R

ρ-crystallin, 94
Refractive index, 7, 12
Retinal, 1

S

S-crystallin, 106
SOX proteins, 151

T

τ-crystallin, 94-96, *95f*, 137
    structure, 96-97, *98f*
TATA box, 109, *110f*, 111

Z

ζ-crystallin, 103, 111-112, 126, *127f*, 128, *129f*, *131f*
    ZPE (ζ protected element), 128-129, *132f*