Characterization of Squid Crystallin Genes

COMPARISON WITH MAMMALIAN GLUTATHIONE S-TRANSFERASE GENES*

Stanislav I. Tomarev, Rina D. Zinovieva, and Joram Piatigorsky

From the Laboratory of Molecular and Developmental Biology, National Eye Institute, National Institutes of Health, Bethesda, Maryland 20892

(Received for publication, October 30, 1991)

Previous experiments have indicated that the crystallins of the squid lens (S-crystallins) are evolutionarily related to glutathione S-transferases (GST) (EC 2.5.1.18). Here we confirm by peptide sequencing that the crystallins of the lens of the squid Ommastrephes sloani pacificus comprise a family of GST-like proteins. Squid lens extracts showed 400 times less GST activity than those of liver using 1-chloro-2,4-dinitrobenzene as a substrate, suggesting that the abundant GST-like crystallins lack enzymatic activity. Four different cDNAs (pSL20-1, pSL18, pSL11, and pSL4) showed 20-25% similarity in homologous regions with mammalian GST polypeptides. pSL20-1, pSL18, and pSL4 each encode an S-crystallin with a unique internal peptide that is unrelated to mammalian GSTs or any other sequence in GenBank. The S-crystallin family is encoded in a minimum of 9-10 genes, and the exon-intron structures of at least two of these (SL20-1 and SL11) are similar to those of the mammalian GST genes. The SL20-1 gene has six exons, with the its unique internal peptide encoded precisely in exon 4; the SL11 gene lacks a unique internal peptide and has five exons. Experiments using bacterial chloramphenicol acetyltransferase as a reporter gene showed that at least 84 and 111 base pairs of 5'-flanking sequence are needed for function of the SL20-1 and SL11 promoters, respectively, in a transfected rabbit lens epithelial cell line (N/N1003A). Within these regions each has a putative TATA box and an upstream AP-1 site overlapping with antioxidant responsive-like elements, which are regulatory elements in the rat GST Ya and quinone reductase genes responsive to oxidative stress.

Crystallins are a surprisingly diverse group of proteins comprising 80-90% of the soluble protein of the transparent eye lens (1-5). All vertebrates have a- , b- , and g-crystallins, with the proportional amounts of each depending upon the species. There are also a number of crystallins which are present only in certain taxonomic groups or selected species of vertebrates. These taxon-specific crystallins are particularly fascinating in that they have been recruited from metabolic enzymes and are thus known as enzyme-crystallins (6, 7). In some cases enzyme-crystallins retain their catalytic activity when expressed in the lens as crystallins, while in other cases they appear metabolically inert in the lens.

The complex eyes of some invertebrates also contain transparent cellular lenses (8, 9). Cephalopods (10) and cubomedusian jellyfish (11) are two striking examples of convergent evolution leading to eyes and lenses with resemblances to those of vertebrates. In contrast to the extensive investigations which have been performed on the vertebrate crystallins, very few studies have been conducted on invertebrate crystallins. The major cephalopod and cubomedusian jellyfish crystallins differ from each other and from vertebrate crystallins (11-15). The octopus lens, however, has one less abundant crystallin (called g-crystallin) (16), which is related if not identical to aldehyde dehydrogenase (17), an enzyme also recruited as a lens crystallin in the mammalian elephant shrew (18). In addition to being directly relevant to understanding the basis of lens transparency, further knowledge of invertebrate lens crystallins may provide insights into possible similarities and differences during the convergent evolution of the eye lens of vertebrates and invertebrates, possibly leading to a greater understanding of the evolution of new organs in general.

Previous studies have shown that the squid lens contains one major family of soluble crystallins (called S-crystallins) with polypeptide molecular masses of 27-36 kDa (15, 19) which are clearly related to the family of glutathione S-transferases (GSTs) (20, 21). Here we show by peptide analysis and by cDNA and genomic DNA sequencing that squid S-crystallins comprise a family of GST-like polypeptides encoded by a minimum 9-10 S-crystallin genes. We demonstrate that four analyzed S-crystallin genes are expressed predominantly, if not exclusively, in the lens and most probably do not code for an active squid GST. We characterize the exon-intron structure of two S-crystallin genes (SL20-1 and SL11), compare them with the exon-intron structures of vertebrate GST genes, and identify promoters for both squid genes which function in a transfected rabbit lens epithelial cell line.

MATERIALS AND METHODS

Obtaining Squid—Ommastrephes sloani pacificus was collected at the Union of Soviet Socialist Republics Pacific shore (Vladivostok) and Loligo opalescens was collected at the Hopkins Marine Station (Pacific Grove, CA). The lenses and specified organs were removed from the squids, stored in liquid nitrogen, and brought to the National Institutes of Health, Bethesda, MD on dry ice, where the present experiments were conducted.

Preparation of Crystallins—Lenses were homogenized in 5 volumes of 10 mM Tris-HCl, pH 7.5, and 1 mM EDTA, stirred for 30 min, and centrifuged for 15 min in an Eppendorf 5415 centrifuge at maximum speed, all at 4 °C. The supernatant fraction was used for SDS-

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

The abbreviations used are: GST, glutathione S-transferase; PCR, polymerase chain reaction; SDS, sodium dodecyl sulfate.

* Present address: N. K. Koltzov Institute of Developmental Biology, USSR Academy of Sciences, Moscow, USSR.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBankTM/EMBL Data Bank with accession number(s) M74315-M74327.
polycrylamide gel electrophoresis in a 12.5% running gel. After electrophoresis the proteins were either stained with Coomassie Blue or electroblotted onto a nitrocellulose filter and stained with Ponceau. S. Partial amino acid sequences were obtained from specific bands on the filter, as shown below.

**Partial Sequencing of Lens Crystallins**—Proteins present in the electrophoretically derived bands on the filters were eluted and digested with trypsin. The resultant tryptic peptides were separated by high performance liquid chromatography and amino acid sequencing which were performed as a service by Dr. W. Lane at Microchemistry Facility, Harvard University and the Cambridge Prochem (Cambridge, MA) (see Ref. 11).

**GST Assay**—Lenses and digestive glands were homogenized in 10 volumes of 0.1 M potassium phosphate, pH 7.0, and centrifuged for 15 min in an Eppendorf 5415 centrifuge at maximum speed at 4 °C. The supernatant fractions were assayed for GST activity with 1-chloro-2,4-dinitrobenzene as a substrate, as described (22, 23).

**Isolation of Nucleic Acids**—High molecular weight DNA was isolated from squid testis by standard procedures (24). Phage and plasmid DNAs were isolated using Qiagen columns (QIAGEN Inc., Switzerland, CA). For transfections, plasmid DNAs were purified by filtration the proteins were either stained with Coomassie Blue or electroblotted onto a nitrocellulose filter and stained with Ponceau. S. Partial amino acid sequences were obtained from specific bands on the filter, as shown below.

**Fig. 1. Comparison of amino acid sequences of the squid crystallin encoded by plSL20-1, plSL11, pSL18, and pSL4 cDNAs with the rat GST Ya (40) and Yp (41) subunits. Invariant residues (60) are boxed. Dashes show the gaps which were introduced to maximize similarity. Positions of introns in the SL20-1, SL11, rat Ya (61), and rat Yp (62) genes are shown by arrows. Numbering of the crystallin encoded in pSL4 assumes that its N-terminus to the maximum extent to which the two species of squid were size-fractionated on 1.5% agarose, 2.2 M formaldehyde gels and transferred to Nitran nylon filters as recommended by the manufacturers (Schleicher & Schuell), cross-linked by ultraviolet irradiation using Stratalinker (Strategene, La Jolla, CA), and hybridized to the indicated cDNA inserts. EcoRI-HindIII restriction fragments of cDNA inserts were gel-purified using GeneClean (BIO 101, La Jolla, CA) and labeled with [γ-32P]dCTP (3000 Ci/mmol, Amersham Corp.) using the random primer method (Bethesda Research Laboratories kit) to a specific activity near 10^10 cpm/μg DNA. Hybridization occurred overnight at 42 °C in 40% formamide, 6 × SSC (standard saline citrate (0.15 M NaCl, 0.015 M sodium citrate)), 5 × Denhardt’s solution, 10% dextran sulfate, 5 μl EDTA, 10 mM sodium phosphate, pH 6.8, 0.1% SDS, 100 μg/ml fractionated denatured salmon sperm DNA, and 1–2 × 10^5 cpm/ml of cDNA probe. Filters were washed twice in 2 × SSC and 0.1% SDS at room temperature for 30 min, twice in 1 × SSC and 0.1% SDS at 45 °C for 40 min, and once in 0.1 × SSC and 1% SDS at 45 °C for 40 min.

**TABLE 1**

| GST Ya | GST Yp |
|--------|--------|
| SL20-1 | 55     |
| SL11   | 55     |
| SL18   | 55     |
| GS Yp  | 56     |

**TABLE 2**

| GST Ya | GST Yp |
|--------|--------|
| SL20-1 | 131    |
| SL11   | 114    |
| SL18   | 117    |
| GS Yp  | 117    |

**TABLE 3**

| GST Ya | GST Yp |
|--------|--------|
| SL20-1 | 111    |
| SL11   | 117    |
| SL18   | 117    |
| GS Yp  | 117    |

**TABLE 4**

| GST Ya | GST Yp |
|--------|--------|
| SL20-1 | 111    |
| SL11   | 117    |
| SL18   | 117    |
| GS Yp  | 117    |

**TABLE 5**

| GST Ya | GST Yp |
|--------|--------|
| SL20-1 | 111    |
| SL11   | 117    |
| SL18   | 117    |
| GS Yp  | 117    |

**TABLE 6**

| GST Ya | GST Yp |
|--------|--------|
| SL20-1 | 111    |
| SL11   | 117    |
| SL18   | 117    |
| GS Yp  | 117    |

**TABLE 7**

| GST Ya | GST Yp |
|--------|--------|
| SL20-1 | 111    |
| SL11   | 117    |
| SL18   | 117    |
| GS Yp  | 117    |

**TABLE 8**

| GST Ya | GST Yp |
|--------|--------|
| SL20-1 | 111    |
| SL11   | 117    |
| SL18   | 117    |
| GS Yp  | 117    |

**TABLE 9**

| GST Ya | GST Yp |
|--------|--------|
| SL20-1 | 111    |
| SL11   | 117    |
| SL18   | 117    |
| GS Yp  | 117    |

**TABLE 10**

| GST Ya | GST Yp |
|--------|--------|
| SL20-1 | 111    |
| SL11   | 117    |
| SL18   | 117    |
| GS Yp  | 117    |

**TABLE 11**

| GST Ya | GST Yp |
|--------|--------|
| SL20-1 | 111    |
| SL11   | 117    |
| SL18   | 117    |
| GS Yp  | 117    |
Fig. 2. Coomassie Blue-stained SDS-polyacrylamide gel electrophoresis of squid water-soluble lens proteins. Two differentiation of the M, (lane 1) and 50 kDa (lane 2)) of protein were used. A-D (cut) designate the regions of the gel which were used for protein sequencing. Position of the marker proteins are shown on the left.

In some cases where the background remained high, the filters were autoradiographed after additional washes in 0.1 X SSC and 0.1% SDS at 60 °C for 40 min.

Southern Blot Analysis—Restricted genomic or recombinant phage DNAs were fractionated on 0.8% agarose gels in 0.04 M Tris-acetate, pH 7.2, 2 mM EDTA and transferred to nitrocellulose or Nitrin nylon filters as recommended by the manufacturers (Schleicher & Schuell). Labeling of EcoRI-HindIII fragments of cDNA inserts and the conditions of hybridization and washing were identical to those used for the Northern blot analysis, except that all washes in 1 X SSC and 0.1 X SSC were conducted at 60 °C for 40 min. Oligonucleotide probes were radioactively labeled with [γ-32P]ATP (7000 Ci/mmol, INC, Irvine, CA) using T4 polynucleotide kinase (Bethesda Research Laboratories). Hybridization was for 2-3 h or overnight at 50 °C in 6 X SSC, 0.1% SDS. Filters were washed four times for 5 min each time in 6 X SSC, 0.01 M sodium phosphate, pH 6.8, 1 mM EDTA, pH 8.0, and 0.5% SDS. Filters were washed four times for 5 min each time in 6 X SSC and 0.5% SDS at room temperature followed by washing for 1 min at 50 °C in the same solution.

Construction and Screening of the Squid Genomic Library—Squid genomic DNA was partially digested with Mbol and approximately 8-11-kilobase pair fragements were isolated by sucrose gradient centrifugation and cloned in the EMBL-3 bacteriophage vector (27). 2 X 109 independent plaques were obtained. The library was constructed and amplified as a service by Clonetech (Palo Alto, CA). Using 24 polyonucleotide kinase (Bethesda Research Laboratories). Hybridization was for 2-3 h or overnight at 50 °C in 6 X SSC, 0.1% SDS. Filters were washed four times for 5 min each time in 6 X SSC and 0.5% SDS at room temperature followed by washing for 1 min at 50 °C in the same solution.

Construction and Screening of the Squid Genomic Library—Squid genomic DNA was partially digested with Mbol and approximately 8-11-kilobase pair fragements were isolated by sucrose gradient centrifugation and cloned in the EMBL-3 bacteriophage vector (27). 2 X 109 independent plaques were obtained. The library was constructed and amplified as a service by Clonetech (Palo Alto, CA). Using 24 polyonucleotide kinase (Bethesda Research Laboratories). Hybridization was for 2-3 h or overnight at 50 °C in 6 X SSC, 0.1% SDS. Filters were washed four times for 5 min each time in 6 X SSC and 0.5% SDS at room temperature followed by washing for 1 min at 50 °C in the same solution.

Construction and Screening of the Squid Genomic Library—Squid genomic DNA was partially digested with Mbol and approximately 8-11-kilobase pair fragements were isolated by sucrose gradient centrifugation and cloned in the EMBL-3 bacteriophage vector (27). 2 X 109 independent plaques were obtained. The library was constructed and amplified as a service by Clonetech (Palo Alto, CA). Using 24 polyonucleotide kinase (Bethesda Research Laboratories). Hybridization was for 2-3 h or overnight at 50 °C in 6 X SSC, 0.1% SDS. Filters were washed four times for 5 min each time in 6 X SSC and 0.5% SDS at room temperature followed by washing for 1 min at 50 °C in the same solution.

DISCUSSION

RESULTS AND DISCUSSION

Multiple cDNAs Encoding Squid Crystallins—Previously, several cDNAs with GST-like sequences derived from at least three genes were isolated from a squid lens cDNA library (20, 39). Here (Fig. 1) we compare the sequence of the deduced proteins of another newly isolated squid lens cDNA, pSL4, with those of pSL11, pSL20-1 (originally called pSL20) and pSL18. Fig. 1 also compares these four deduced squid crystallins with the rat Ya (40) and Yp (41) GST subunits, which they most resemble.

The deduced polypeptide molecular masses encoded in these cDNAs are approximately 24 kDa for pSL11, 26 kDa for pSL20-1, 36 kDa for pSL18, and 34 kDa for pSL4. pSL4 is incomplete since it lacks an initiating codon near its 5' end and stays in frame from the beginning. If one assumes that the full-length NH2-terminal sequence of the undecoupled protein partially encoded by pSL4 is the same as that in the other crystallins of this family, the molecular mass of the complete polypeptide would be about 45 kDa. The alignment in Fig. 1, as noted earlier (20), considerable conservation of the NH2 and COOH termini of the squid crystallins with each other and with the rat GST subunits (except that the Ya and Yp polypeptides are 16 and 8 residues longer at their COOH termini than the squid polypeptides, respectively). Although the pSL11 polypeptide can be aligned with the rat Ya and Yp polypeptides along their entire sequences, the pSL20-1, pSL18, and pSL4 polypeptides have unique inserts in their central regions (shaded in Fig. 1). Insert sizes for the different crystallins are 17 residues for pSL20-1, 103 residues for pSL18, and 193 residues for pSL4. These inserts are not homologous to the rat GST polypeptides or to any other sequence in GenBank. The insert of pSL4 S-crystallin contained four complete and two incomplete repeats of 10 amino acids with consensus RGGGGYXXVG/S. The NH2-terminal part of this repeat, RGD, present four times in the insert, was found previously in the variety of cell adhesives proteins like fibronectin (42) and is believed to be a part of...
Peptide Evidence for a Family of GST-like Squid Lens Crystallins—In order to substantiate directly the presence of a family of related crystallin polypeptides with differing sizes in the squid lens and to attempt identifying the polypeptides encoded in the cDNAs described above, the soluble proteins of the squid lens were subjected to SDS-polyacrylamide gel electrophoresis, electroblotted onto nitrocellulose, eluted, digested with trypsin, and some chromatographically isolated peptides sequenced. The electrophoretic patterns of two concentrations of squid crystallins show a group of major bands with a molecular mass of about 25 kDa and several minor polypeptides with higher molecular masses (Fig. 2). Three to four tryptic peptides from each of four sizes of electroblotted proteins (cuts A-D in Fig. 2) were sequenced and compared with amino acid sequences deduced from the cDNAs described above encoding proteins with molecular masses consistent with those obtained from the polyacrylamide gel. Only two of the tryptic peptides matched perfectly with a region encoded in the cDNAs; these were NT67(2) and NT21 which were both encoded in pSL20-1. The other peptide sequences had a maximum of 50–87% identity with the deduced sequences found in the cDNAs, with the variations generally involving conservative amino acid changes (not shown). These results suggest the existence of a family of GST-like squid crystallins comprising more than the four members presented in Fig. 1.

Since the major 24–27-kDa squid crystallins as well as less abundant higher molecular mass crystallins (about 32–34 kDa) are all related to mammalian GST and to each other, we conclude that the squid lens consists essentially of one class of GST-like polypeptides. Our recent studies on octopus lens crystallins gave similar but not identical results: besides the major GST-like S-crystallins there is another less abundant aldehyde dehydrogenase-like crystallin (Ω-crystallin) in octopus lens (17).

Enzyme Activity Tests for Squid Crystallins—In view of the relationship of the squid S-crystallins with GST, extracts of the squid lens were assayed for enzymatic activity and compared with comparable extracts from the squid digestive gland (functional analog of vertebrate liver) and from mammalian lenses using 1-chloro-2,4-dinitrobenzene as a substrate. Recently, GST with very high specific activity using 1-chloro-2,4-dinitrobenzene was isolated and partially characterized from the digestive gland of the squid Loligo vulgaris (44). Our results showed that the squid lens extract had 400 times less GST activity than the digestive gland extract (0.26 ± 0.02 versus 106.3 ± 1.0 OD/min/mg protein, respectively).
This is similar to previous results obtained with the octopus (17). It is also noteworthy that the squid crystallins do not bind to a glutathione-agarose column, as does mammalian and squid GST (44). These data are consistent with a specialized structural rather than enzymatic role in the lens for cephalopod S-crystallins, although we cannot exclude the possibility that S-crystallins have some enzyme activity when tested with another of the possible substrates for GST (see Ref. 45).

Tissue Distribution of Squid Crystallin mRNAs—We next examined by Northern blotting the tissue distribution of RNAs which hybridize to the four cloned cDNAs (Fig. 3). Total RNA was used from lens, digestive gland, testis, and muscles of the homologous squid, O. pacificus, and from lens, gills, heart, ovary, giant axons, and brain of the heterologous squid, L. opalescens. The Northern blots showed intense hybridization of the cDNAs to both the homologous and heterologous lens RNAs. Although not evident from Fig. 3, the strongest hybridization was with pSL20-1, consistent with it being one of the most prevalent crystallin mRNAs; in addition, this RNA appears highly conserved between the two species of squid being examined. The pSL11, pSL18, and pSL4 cDNAs also cross-hybridized with the heterologous lens RNAs; however, in each case the patterns were different in O. pacificus and L. opalescens (see Fig. 3).

The squid crystallin cDNAs hybridized either more strongly (pSL11 and pSL4) or exclusively (pSL20-1 and pSL18) to the RNA from the lens than from the other tissues tested. Multiple hybridization bands obtained with pSL4 cDNA may be due to repetitive sequence in its central part (see previous section). It is unlikely that any RNAs were seriously degraded because the control octopus α-tubulin RNA hybridized as a single band of approximately 2 kilobases (Fig. 3). Further experiments are necessary in order to determine whether the hybridizations of pSL4 and pSL11 cDNAs in non-lens tissues represent expression of these genes.

Squid Crystallin Genes—In order to begin characterizing the squid crystallin genes, Southern blots of genomic DNA digested with EcoRI, HindIII, or BamHI were hybridized with EcoRI-HindIII fragments isolated from pSL11, pSL20-1, pSL18, or pSL4 (Fig. 4). These probes lacked 243 bp (SL18), 70 bp (SL18), and 160 bp (SL20-1) of the 5' ends of their cDNAs. The hybridization patterns obtained confirmed that each cDNA recognized specifically a different set of genomic sequences and suggested that some of them (especially pSL20-1) may be hybridizing to more than one gene or pseudogene, although we could not rule out allelic polymorphisms or the presence of restriction sites within introns dividing genes into pieces from this data alone.

We next screened 5 × 10⁶ plaques of an EMBL-3 genomic library for clones hybridizing to each cDNA. All four cDNAs hybridized either more strongly (pSL11 and pSL4) or exclusively (pSL20-1 and pSL18) to the RNA from the lens than from the other tissues tested. Multiple hybridization bands obtained with pSL4 cDNA may be due to repetitive sequence in its central part (see previous section). It is unlikely that any RNAs were seriously degraded because the control octopus α-tubulin RNA hybridized as a single band of approximately 2 kilobases (Fig. 3). Further experiments are necessary in order to determine whether the hybridizations of pSL4 and pSL11 cDNAs in non-lens tissues represent expression of these genes.

Squid Crystallin Genes—In order to begin characterizing the squid crystallin genes, Southern blots of genomic DNA digested with EcoRI, HindIII, or BamHI were hybridized with EcoRI-HindIII fragments isolated from pSL11, pSL20-1, pSL18, or pSL4 (Fig. 4). These probes lacked 243 bp (SL18), 70 bp (SL18), and 160 bp (SL20-1) of the 5' ends of their cDNAs. The hybridization patterns obtained confirmed that each cDNA recognized specifically a different set of genomic sequences and suggested that some of them (especially pSL20-1) may be hybridizing to more than one gene or pseudogene, although we could not rule out allelic polymorphisms or the presence of restriction sites within introns dividing genes into pieces from this data alone.

We next screened 5 × 10⁶ plaques of an EMBL-3 genomic library for clones hybridizing to each cDNA. All four cDNAs hybridized either more strongly (pSL11 and pSL4) or exclusively (pSL20-1 and pSL18) to the RNA from the lens than from the other tissues tested. Multiple hybridization bands obtained with pSL4 cDNA may be due to repetitive sequence in its central part (see previous section). It is unlikely that any RNAs were seriously degraded because the control octopus α-tubulin RNA hybridized as a single band of approximately 2 kilobases (Fig. 3). Further experiments are necessary in order to determine whether the hybridizations of pSL4 and pSL11 cDNAs in non-lens tissues represent expression of these genes.

Squid Crystallin Genes—In order to begin characterizing the squid crystallin genes, Southern blots of genomic DNA digested with EcoRI, HindIII, or BamHI were hybridized with EcoRI-HindIII fragments isolated from pSL11, pSL20-1, pSL18, or pSL4 (Fig. 4). These probes lacked 243 bp (SL18), 70 bp (SL18), and 160 bp (SL20-1) of the 5' ends of their cDNAs. The hybridization patterns obtained confirmed that each cDNA recognized specifically a different set of genomic sequences and suggested that some of them (especially pSL20-1) may be hybridizing to more than one gene or pseudogene, although we could not rule out allelic polymorphisms or the presence of restriction sites within introns dividing genes into pieces from this data alone.

We next screened 5 × 10⁶ plaques of an EMBL-3 genomic library for clones hybridizing to each cDNA. All four cDNAs hybridized either more strongly (pSL11 and pSL4) or exclusively (pSL20-1 and pSL18) to the RNA from the lens than from the other tissues tested. Multiple hybridization bands obtained with pSL4 cDNA may be due to repetitive sequence in its central part (see previous section). It is unlikely that any RNAs were seriously degraded because the control octopus α-tubulin RNA hybridized as a single band of approximately 2 kilobases (Fig. 3). Further experiments are necessary in order to determine whether the hybridizations of pSL4 and pSL11 cDNAs in non-lens tissues represent expression of these genes.
Fig. 6. Polymorphic positions in the SL20-1 and SL11 crystallin genes. Numbering corresponds to the positions in the cDNA sequences, with position 1 being the mRNA cap site. Polymorphic positions in the coding region are shown in the context of their codons.

The genomic library was next screened with a mixture of the EcoRI-HindIII probes derived from the pSL20-1 cDNA, which is neither present in the pSL4, pSL11, and pSL18 cDNAs. After hybridization of the same filters used for screening with pSL20-1 probe, two clones hybridizing to the pSL11 probe and two hybridizing to the pSL4 probe were obtained. Although the two SL11 and the two SL4 genomic sequences each gave different patterns of BamHI or BglII bands hybridizing to their respective probes (data not shown), further analysis by hybridization with oligonucleotides corresponding to parts of the cDNAs and sequencing indicated that each pair comprised overlapping fragments of a single gene.

We conclude that S-crystallins in the squid are encoded by a gene family consisting of at least 9–10 members. pSL20-1 cDNA belongs to a subfamily of genes which are more similar to each other than to the pSL4, pSL11, and pSL18 cDNAs. The SL20 subfamily appears to comprise the major S-crystallins in the lens.

Characterization of the SL20-1 and SL11 Genes—The BglII restriction fragments from the SL20-1 and SL11 genes were cloned into the BamHI site of the BlueScript plasmid and sequenced. The lengths of the SL20-1 and SL11 genes are approximately 13.5 and 17 kilobase pairs, respectively. The exon-intron junctions were determined by comparing the gene and cDNA sequences (Fig. 5). It is interesting to note that all intron sequences near the exons (with the exception of intron 2 in SL11) contained more than 50% T residues. Since other squid genomic sequences have not been characterized, it is not known at present whether this is a common feature of squid introns or if it is a specific characteristic of this family of squid genes. The size of each intron was estimated by a combination of direct sequencing, restriction mapping, and PCR. The SL20-1 gene contains six exons, while the SL11 gene has five exons. The central peptide in the deduced protein of the pSL20-1 cDNA, which is neither present in the pSL11 cDNA nor has a homologue in the rat GST subunits, is encoded precisely in exon 4 of the SL20-1 gene. Moreover, partial sequencing showed that a 14-amino acid central peptide in SL20-4, which is not represented in the rat GST proteins, is also encoded in a separate exon (data not shown).

The exon-intron boundaries of the SL20-1 and SL11 genes coincided identically with respect to the homologous amino acid sequences that they encode (Fig. 1). Indeed, they were both very similar to those present in the rat GST genes, with greatest similarity occurring with the Yp gene. The position of intron 1 in squid S-crystallin genes coincides exactly with the position of intron 2 in rat GST Ya gene, while the position of all other introns (with the exception of intron 4 in SL20-1 S-crystallin gene) are identical in squid S-crystallin genes and rat GST Yp gene.

The additional exon 4 in the squid SL20-1 cDNA belongs to a subfamily of genes which are more similar to each other than to the pSL4, pSL11, and pSL18 cDNAs. After hybridization of the same filters used for screening with pSL20-1 probe, two clones hybridizing to the pSL11 probe and two hybridizing to the pSL4 probe were obtained. Although the two SL11 and the two SL4 genomic sequences each gave different patterns of BamHI or BglII bands hybridizing to their respective probes (data not shown), further analysis by hybridization with oligonucleotides corresponding to parts of the cDNAs and sequencing indicated that each pair comprised overlapping fragments of a single gene.

We conclude that S-crystallins in the squid are encoded by a gene family consisting of at least 9–10 members. pSL20-1 cDNA belongs to a subfamily of genes which are more similar to each other than to the pSL4, pSL11, and pSL18 cDNAs. The SL20 subfamily appears to comprise the major S-crystallins in the lens.

Characterization of the SL20-1 and SL11 Genes—The BglII restriction fragments from the SL20-1 and SL11 genes were cloned into the BamHI site of the BlueScript plasmid and sequenced. The lengths of the SL20-1 and SL11 genes are approximately 13.5 and 17 kilobase pairs, respectively. The exon-intron junctions were determined by comparing the gene and cDNA sequences (Fig. 5). It is interesting to note that all intron sequences near the exons (with the exception of intron 2 in SL11) contained more than 50% T residues. Since other squid genomic sequences have not been characterized, it is not known at present whether this is a common feature of squid introns or if it is a specific characteristic of this family of squid genes. The size of each intron was estimated by a combination of direct sequencing, restriction mapping, and PCR. The SL20-1 gene contains six exons, while the SL11 gene has five exons. The central peptide in the deduced protein of the pSL20-1 cDNA, which is neither present in the pSL11 cDNA nor has a homologue in the rat GST subunits, is encoded precisely in exon 4 of the SL20-1 gene. Moreover, partial sequencing showed that a 14-amino acid central peptide in SL20-4, which is not represented in the rat GST proteins, is also encoded in a separate exon (data not shown).

The exon-intron boundaries of the SL20-1 and SL11 genes coincided identically with respect to the homologous amino acid sequences that they encode (Fig. 1). Indeed, they were both very similar to those present in the rat GST genes, with greatest similarity occurring with the Yp gene. The position of intron 1 in squid S-crystallin genes coincides exactly with the position of intron 2 in rat GST Ya gene, while the position of all other introns (with the exception of intron 4 in SL20-1 S-crystallin gene) are identical in squid S-crystallin genes and rat GST Yp gene. The additional exon 4 in the squid SL20-1
Fig. 8. Nucleotide sequences of the 5'-flanking regions of SL20-1 (A) and SL11 (B) squid crystallin genes. TATA-box is boxed, and the AP-1 site, triplet for initiator methionine and repetitive sequences (CAT), and (TA), are underlined. Promoter deletion constructions are shown by the arrows with angles. Positions of major transcription initiation sites correspond to position 1, and other transcription initiation sites are indicated by vertical arrows.

Fig. 9. Relative CAT activity in extracts from transfected rabbit lens epithelial cells N/N1003A. The results of three experiments are presented with their standard errors.
tion that they represent multiple transcription initiation sites of their respective genes.

The 5'-flanking sequences of the SL20-1 and SL11 genes show considerable differences although both are rich in T (Fig. 8, A and B). Both genes have a putative TATA sequence (boxed A and B), as do all vertebrate crystallin genes, at an appropriate distance upstream of the postulated first transcription initiation site (arrows in A and B). Interestingly, a 13-bp sequence (5'-GACTGCATAGATA-3') immediately downstream of the TATA box in SL20-1 gene has only 1 bp mismatch with a sequence at positions -35/-23 in the mouse αB-crystallin gene (36). Both squid genes have an AP-1 consensus sequence (underlined in A and B) 15 bp (SL20-1) and 45 bp (SL11) upstream of the designated TATA box. The promoter sequence of another squid S-crystallin gene, SL20-3, also contains an AP-1 consensus sequence 15 bp upstream of TATA box (data not shown). Many crystallin genes of vertebrates also contain an AP-1 site in their 5'-flanking sequence (46), as do functional 5' regulatory sequences of the human γ (47) and rat Yp γ (48) of GST. Mutagenesis of an AP-1-like site in the chicken βB1-crystallin promoter greatly reduces its activity in transfected embryonic chicken lens epithelial cells and its ability to bind putative transcriptional factor in chicken lens nuclear extracts (49); moreover, deletion or site-directed mutagenesis of these sites in the rat (48, 50) or human (47) GST genes also reduced their function.

The AP-1 site overlaps with a sequence similar to a regulatory cis-element called the antioxidant responsive element (GTGACNNNGC) in the SL20-1, SL20-3, and SL11 genes and, in addition, with a sequence similar to the xenobiotic responsive element (TNGCGTG) in the SL11 gene. Both the antioxidant responsive element and xenobiotic responsive element are involved in the induction of rat GST Ya and quinone reductase gene expression (51-54). An antioxidant responsive element-like element is also present in the 5'-flanking region of the guinea pig z-crystallin gene, which has quinone reductase activity (55). Since oxidative insult is a cause of cataract in vertebrate lenses (56), use of a cis-element associated with the SL20-1 gene (Fig. 8) and -80 (46), 15 bp (SL20-1) and -60 (SL11) abolished promoter function despite the continued presence of the putative TATA boxes. Interestingly, the consensus AP-1 sites are contained in the region of the promoters required for function. Reversing the orientation of the largest promoter of each gene resulted in loss of activity (Fig. 9). All SL20-1 and SL11 promoter constructions functioned 15-20 and 3-5 times better, respectively, in the lens cells than the kidney cells (data not shown). Additional transfection experiments indicated that CAT constructs driven by promoters of the mouse major histocompatibility (H2K), chicken α-actin (58), and chicken β-globin (59) genes also functioned considerably better in rabbit lens N/N1003A cells than in rabbit kidney RK13 cells (data not shown). Thus, while our transfection experiments delimited a functional promoter in the squid genes capable of working in vertebrate lens cells, they do not establish lens preference.

Acknowledgments—We thank Dr. Douglas Lee for help with GST assays, Drs. Christina Sax, John Klement, and Robert Dubin for advice on transfection, Drs. Martin Breitman, Robert Kim, Bruce Paterson, and Mark Reitman for providing promoter constructions of different genes, Joan McDermott, Drs. P. Vasantha Rao, J. Samuel Piatigorsky, Dr. Joseph Horwitz, Pedro Gonzalez, Ignacio Rodriguez, and Terete Borras for providing us with their unpublished results, and Dawn Chiocchirichi for secretarial assistance.

REFERENCES

1. Bloemendal, H. (1982) CRC Crit. Rev. Biochem. 12, 1-38
2. de Jong, W., and Hendriks, W. (1986) J. Mol. Biol. 24, 121-129
3. Lubsen, N. H., Aarts, H. J. M., and Schoenmakers, J. G. J. (1988) Prog. Biophys. Mol. Biol. 51, 47-76
4. Piatti, J., and Piatigorsky, J. (1989) Annu. Rev. Biochem. 57, 479-504
5. de Jong, W. J. H., and Hendriks, W., Mulders, J. M., and Bloemendal, H. (1989) Trends Biochem. Sci. 14, 365-368
6. Piatigorsky, J., and Wistow, G. (1989) Cell 57, 197-199
7. Alper, M. A. (ed) (1982) Photochemistry and Vision in Invertebrates, NATO Advanced Science Institute, Series A, Vol. 74, pp. 1-858, Plenum Press, New York
8. Cronin, T. W. (1996) An. Zool. 26, 403-415
9. Parkard, A. (1972) J. Biol. Chem. 247, 241-307
10. Piatigorsky, J., Horwitz, J., Kuwabara, T., and Cutress, C. E. (1989) J. Comp. Physiol. A 164, 577-587
11. Brem, A. K. (1978) J. Embryol. Exp. Morph. 46, 11-118
12. Halbert, S. P., and Manski, W. (1985) Progr. Allergy 7, 107-186
13. Manski, W., and Halbert, S. P. (1965) in Proteids of the Biological Fluids (Peeters, H, ed) Proceedings of the Twelfth Colloquium, Bruges, 1964, pp. 117-134, Elsevier, Amsterdam
14. Siezen, R. J., and Shaw, D. C. (1982) Biochim. Biophys Acta 704, 394-399
15. Chiou, S. H. (1988) FEBS Lett. 241, 261-264
16. Tomarev, S. I., Zinovieva, R. D., and Piatigorsky, J. (1991) J. Mol. Biol. 266, 2426-24231
17. Wistow, G., and Kim, H. (1991) J. Mol. Biol. 32, 262-269
18. Chou, S.-H. (1984) J. Biochem. (Tokyo) 95, 75-82
19. Tomarev, S. I., and Zinovieva, R. D. (1988) Nature 336, 86-88
20. Tomarev, G., and Piatigorsky, J. (1987) Science 236, 1554-1556
21. Habig, W. B., Pabst, M. J., and Jakoby, W. B. (1974) J. Biol. Chem. 249, 730-739
22. Ashwasti, Y. C., Saneto, R. P., and Srivastava, S. K. (1980) Exp. Eye Res. 30, 29-38
23. Srinivas, P., and Sashidharan, M. (1976) Nucleic Acids Res. 3, 2303-2308
24. Birnboim, H. C., and Doly, J. (1979) Nucleic Acids Res. 7, 1513-1523
25. Chromczynski, P., and Sacchi, N. (1987) Anal. Biochem. 162, 156-159
26. M. L. Breitman, unpublished results.

P. Gonzalez, L. Rodriguez, and T. Borras, T., unpublished observations.
J. McDermott and J. Piatigorsky, unpublished observations.
Squid Crystallin Genes

27. Frischauf, A.-M., Lebrach, H., Poustka, A., and Murray, N. (1983) *J. Mol. Biol.* 170, 827-842
28. Maxam, A. M., and Gilbert, W. (1980) *Methods Enzymol.* 65, 499–560
29. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U. S. A.* 74, 5463-5467
30. Kanehisa, M. (1988) *IDEAS-88: User Manual*, Frederick Cancer Research Facility, Frederick, MD

Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. R., and Erlich, H. A. (1988) *Science* 239, 487-491
32. Lok, S., Stevens, W., Breitman, M. L., and Tsui, L. C. (1989) *Nucleic Acids Res.* 17, 3563-3582
33. Reddan, J. R., Chepelinsky, A. B., Dziedzic, D. C., Piatigorsky, J., and Goldenberg, E. M. (1986) *Differentiation* 33, 168-174
34. Beale, A. J., Christofinis, G. C., and Furminger, I. G. S. (1963) *Lancet* 2, 640-641
35. Herbomel, P., Bourachot, B., and Yaniv, M. (1984) *Cell* 39, 653-662
36. Dubin, R. A., Wawrousek, E. F., and Piatigorsky, J. (1988) *Biochim. Biophys. Acta* 888, 215-225
37. Neumann, J. R., Morency, C. A., and Russian, K. O. (1987) *Mol. Cell. Biol.* 7, 1463-1469
38. Piatigorsky, J., and Zelenka, P. (1991) *Adv. Deu. Biochem.* 1, 211–256
39. Morrow, C. S., Goldsmith, M. E., and Cowan, K. H. (1990) *Proc. Natl. Acad. Sci. U. S. A.* 87, 3826-3830
40. Roth, H. J., Das, O. C., and Piatigorsky, J. (1991) *Mol. Cell. Biol.* 11, 1468-1499
41. Okuda, A., Imagawa, M., Sakai, M., and Muramatsu, M. (1989) *J. Biol. Chem.* 264, 16919-16926
42. Harris, J., Coles, B., Meyer, D. J., and Ketterer, B. (1991) *Comp. Biochem. Physiol.* 98B, 511–515
43. Mannervik, B., and Danielson, U. H. (1988) *CRC Crit. Rev. Biochem.* 23, 283-337
44. Piattigorsky, J., and Zelenka, P. (1991) *Adv. Dev. Biochem.* 1, 211–256
45. Morrow, C. S., Goldsmith, M. E., and Cowan, K. H. (1990) *Proc. Natl. Acad. Sci. U. S. A.* 87, 3826-3830
46. Roth, H. J., Das, O. C., and Piatigorsky, J. (1991) *Mol. Cell. Biol.* 11, 1468-1499
47. Okuda, A., Imagawa, M., Sakai, M., and Muramatsu, M. (1990) *EMBO J.* 9, 1131-1135
48. Favreau, L. V., and Pickett, C. B. (1991) *J. Biol. Chem.* 266, 4556-4561
49. Rushmore, T. H., and Pickett, C. B. (1990) *J. Biol. Chem.* 265, 14648-14653
50. Rushmore, T. H., King, R. G., Paulson, K. E., and Pickett, C. B. (1990) *Proc. Natl. Acad. Sci. U. S. A.* 87, 3826-3830
51. Rushmore, T. H., Morton, M. R., and Pickett, C. B. (1991) *J. Biol. Chem.* 266, 11632-11639
52. Rao, P. V., Krishna, C. M., and Zigler, J. S. (1992) *J. Biol. Chem.* 267, 96-102
53. Spector, A. (1991) in *Oxidative Stress: Oxidants and Antioxidants* (Sies, H., ed) pp. 829–558 Academic Press, New York
54. Kim, R. Y., Lietman, T., Piatigorsky, J., and Wistow, G. J. (1991) *Gene* (Amst.) 103, 193-200
55. Quitschke, W. W., DePonti-Zilli, L., Lin, Z. Y., and Paterson, B. M. (1989) *Mol. Cell. Biol.* 9, 3218-3230
56. Hesse, J. E., Nickol, J. M., Lieber, M. R., and Felsenfeld, G. (1986) *Proc. Natl. Acad. Sci. U. S. A.* 83, 4312-4316
57. Rhoeds, D. M., Zarliengo, R. P., and Tu, C.-P. D. (1987) *Biochim. Biophys. Res. Commun.* 145, 474-481
58. Telakowski-Hopkins, C. A., Rothkopf, G. S., and Pickett, C. B. (1986) *Proc. Natl. Acad. Sci. U. S. A.* 83, 9393-9397
59. Okuda, A., Sakai, M., and Muramatsu, M. (1987) *J. Biol. Chem.* 262, 3858-3863
Characterization of squid crystallin genes. Comparison with mammalian glutathione S-transferase genes.
S I Tomarev, R D Zinovieva and J Piatigorsky

*J. Biol. Chem.* 1992, 267:8604-8612.

Access the most updated version of this article at [http://www.jbc.org/content/267/12/8604](http://www.jbc.org/content/267/12/8604)

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 0 references, 0 of which can be accessed free at [http://www.jbc.org/content/267/12/8604.full.html#ref-list-1](http://www.jbc.org/content/267/12/8604.full.html#ref-list-1)