Provided for non-commercial research and education use. Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

http://www.elsevier.com/copyright



Available online at www.sciencedirect.com





Comparative Biochemistry and Physiology, Part D 3 (2008) 155-171

Molecular diversity and genomic organisation of the α , β and γ eye lens crystallins from the Antarctic toothfish *Dissostichus mawsoni*

Andor J. Kiss*, C.-H. Christina Cheng

Department of Animal Biology, 515 Morrill Hall, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA

Received 4 December 2007; received in revised form 3 February 2008; accepted 7 February 2008 Available online 19 February 2008

Abstract

The eye lens of the Antarctic toothfish living in the -2 °C Southern Ocean is cold-stable. To investigate the molecular basis of this cold stability, we isolated, cloned and sequenced 22 full length crystallin cDNAs. We found two α crystallins ($\alpha A, \alpha B$), six β crystallins ($\beta A1, \beta A2, \beta A4, \beta B1, \beta B2, \beta B3$) and 14 γ crystallins ($\gamma N, \gamma S1, \gamma S2, \gamma M1, \gamma M3, \gamma M4, \gamma M5, \gamma M7, \gamma M8a, \gamma M8b, \gamma M8c, \gamma M8d, \gamma M8e, and <math>\gamma M9$). Alignments of α, β and γ with other known crystallin sequences indicate that toothfish α and β crystallins are relatively conserved orthologues of their vertebrate counterparts, but the toothfish and other fish γM crystalling genes screened from a toothfish BAC library indicated α crystalling genes occurred in a single genomic region of ~266 kbp, β crystalling genes in ~273 kbp, while the γ crystallin gene family occurred in two separate regions of ~180 and ~296 kbp. In phylogenetic analysis, the γM isoforms of the ectothermic toothfish displayed a diversity not seen with endothermic mammalian γ crystallins. Similar to other fishes, several toothfish γ crystallins are methionine-rich (γM isoforms) which may have predisposed the toothfish lens to biochemically attenuate γ crystallin hydrophobicity allowing for cold adaptation. In addition to high methionine content, conservation of $\alpha\beta$ crystallins could have allowed for greater evolutionary plasticity resulting in increased polydispersity of γ crystallins contributing to the cold-stability of the Antarctic toothfish lens.

© 2008 Elsevier Inc. All rights reserved.

Keywords: Lens; Crystallins; Antarctic toothfish; Alpha; Beta; Gamma; cDNA; Genome organisation; Bacterial artificial chromosome; Tetraodon

1. Introduction

The Antarctic toothfish *Dissostichus mawsoni* is a large predatory fish that is endemic to the waters of the Southern Ocean. It is a member of the teleost suborder Notothenioidei which includes about 45% of the fish species and accounts for over 95% of the fish biomass in the sub-zero (-2 °C) continental shelf waters of Antarctica. The Antarctic notothenioid fishes represent an important and diverse taxonomic group of cold adapted ectothermic vertebrates (Eastman, 1993; Parker et al., 2002; di Prisco et al., 2007). Many of the fishes which

E-mail address: kissaj@muohio.edu (A.J. Kiss).

inhabit the perennially sub-zero seawater have undergone a myriad of physiological and biochemical adaptations (Detrich, 1991; Behan-Martin et al., 1993; Eastman, 1993; Cheng and Chen, 1999), most notable of which was the evolution of a blood borne antifreeze glycoprotein which prevents these fish from freezing (DeVries, 1988; Cheng and Chen, 1999). Adaptations in the visual system of these fishes have been recently described, specifically within the retinal organisation of several Antarctic species (Pointer et al., 2005) and in the eye lens of the Antarctic toothfish *D. mawsoni* (Kiss et al., 2004).

The eye lenses of endothermic vertebrates, as well as ectothermic tropical fishes display a phenomenon known as cold cataract (Delaye et al., 1982; Banh and Sivak, 2004; Kiss et al., 2004) which does not occur within the lenses of the Antarctic toothfish at their normal body temperature of -2 °C (Kiss et al., 2004). The phenomenon of cold-cataracts has been used to

^{*} Corresponding author. Current Address: Department of Zoology, Laboratory for Ecophysiological Cryobiology, Miami University, Oxford, Ohio 45056, USA. Tel.: +1 513 529 3195; fax: +1 513 529 6900.

¹⁷⁴⁴⁻¹¹⁷X/\$ - see front matter @ 2008 Elsevier Inc. All rights reserved. doi:10.1016/j.cbd.2008.02.002

model cataracts of non-thermal aetiology as well as other related protein aggregation diseases (sickle cell anaemia, Alzheimer's dementia) in humans (Benedek, 1997). The common denominator in these seemingly unrelated pathologies is a protein instability caused by changes in the micro-environment of the particular proteins, which in the case of cataracts are the lens crystallins. Vertebrate eye lenses are composed of fibre cells arranged in an onion-like structure that are packed full of α , β and γ crystallins (Bloemendal et al., 2004). It is one or more of the γ crystallin proteins that are responsible for cold-cataracts in mammals (Broide et al., 1991). Cold-cataracts in endothermic mammals were well-documented with current opinion that coldcataracts can be attributed to a liquid-liquid phase separation of cold-sensitive crystallin proteins (Clark and Benedek, 1980; Delaye et al., 1982; Siezen and Benedek, 1985; Liu et al., 1996; Banh and Sivak, 2004). Because the Antarctic toothfish lens does not display a cold-cataract, it presents us with the inverse model system for understanding not only lens crystallin thermal stability, but possibly protein stability in general.

Previous work has suggested that a reduced proportion of hydrophobic residues in many proteins could be an adaptation to prevent cold denaturation at temperatures close to or below 0 °C (Privalov, 1990). Although this proposal is very attractive and has been convincing in limited *in vitro* biophysical studies, it remains to be conclusively demonstrated as a general strategy for cold adaptation of structural (non-enzymatic) proteins in cold stenothermal organisms such as the Antarctic toothfish.

Our previous results (Kiss et al., 2004) suggested that biochemical differences between γ crystallins of the toothfish, tropical fishes and endothermic mammals were the principal factors contributing to the cold stability of the toothfish lens. If biochemical differences of the toothfish γ crystallins are in fact responsible for the increased stability, these differences might be encoded at the level of the primary structure. Moreover, as the eye lens is incredibly dense in crystallins (close to 1000 mg mL^{-1} in fishes (Kroger et al., 1994)) contributions to cold stability undoubtedly arise from interactions between all three $(\alpha\beta\gamma)$ crystallin molecules. To address the hypothesis that biochemical properties encoded by the toothfish crystallins impart low temperature stability, we isolated, cloned and sequenced the toothfish lens crystallin cDNAs. To frame the evolution of the γ toothfish crystallins, we have analysed them in conjunction with the available sequences from the tropical zebrafish and a number of previously unidentified sequences from the temperate Tetraodon nigroviridis (spotted green pufferfish).

2. Materials and methods

2.1. cDNA library construction

Lenses of the Antarctic toothfish, *D. mawsoni* (Norman) were collected from live specimens caught in McMurdo Sound, Antarctica, and stored at -80 °C. Total lens RNA was isolated using Ultraspec RNA isolation reagent (Biotecx, TX), and the poly(A)⁺ fraction was isolated from the total RNA using oligo (dT)-cellulose (Collaborative Research, MA) following standard protocols (Sambrook and Russell, 2001). A cDNA library was

constructed from poly(A)⁺ RNA using the Universal Riboclone cDNA Synthesis System (Promega, WI) per manufacturer instructions except for the following modifications. In place of the *Eco*RI adaptors, *Bam*HI/*Xmn*I non-palindromic adaptor formed with complimentary oligonucleotides 5'-d(GATCCGAAGG-GGTTCG)-3' and 5'-d(pCGAACCCCTTCG)-3' (New England Biolabs, MA) was ligated to the blunt-ended doubled stranded cDNA, and then partially filled-in with dGTP and dATP. The cDNAs were then ligated to the phagemid pBK-CMV (Stratagene, CA) previously digested with *Xho*I and partially filled-in with dTTP and dCTP, transformed into XL1-BLUE Supercompetent cells as per manufacturer's protocol (Stratagene, CA).

2.2. Library screening for crystallin cDNA clones

2.2.1. αB and αB crystallins

Recombinant clones from the lens cDNA library were screened to identify α , β , and γ crystallin cDNA clones by a combination of PCR-amplification with crystallin primers and Southern analyses of restriction digested plasmid DNA. For α crystallin, clones of the αA isoform were inadvertently identified by PCRscreening of 96 clones using a primer pair that targeted the full length aB coding sequence - alpha5 (5'-ATGGAAATTTCTATC-CAGCATCCCTGG-3') and alpha3 (5'-TCCACAGATGATAG-GGATGCTGCG-3') based on reported zebrafish (Danio rerio) αB crystallin cDNA sequence (Posner et al., 1999). AlphaB crystallin was isolated using primers DRAB1 (5'-CCTGGATGT-GAAGCACTTCT-3') and DRAB1-2 (5'-CAACACGCCGT-CAGAGGATA-3') specific to sequence sites within the conserved α crystallin domain of zebrafish αB cDNA. An expected amplicon of ~250 bp was obtained from clones and subsequently sequenced to contain full length αB crystallin cDNA clone.

2.2.2. β and γ crystallins

Since β and γ crystallins are members of the same superfamily and they share sequence homology in their core structural regions (Slingsby and Clout, 1999), we were able to identify putative β/γ crystallin cDNA clones by low stringency hybridisation of a Southern blot of SstI & XhoI digested clones from the library using a ³²P-labeled 407 bp partial γ crystallin cDNA from D. mawsoni. This 407 nt partial y crystallin cDNA was obtained by RT-PCR amplification using previously reported primers CPG5 (5'-GAGGACAGGAACTTCC-AGGG-TCGC-3') and CPG3 (5'-GCCTCTGTAGTGGGGGCTGCTCA-TAC-3') designed to the Asian (or common) carp Cyprinus carpio yM2 crystallin (Chang et al., 1988); this partial cDNA was later found to correspond to nucleotide 46-453 of the toothfish yM8e. Initially, 50 randomly selected clones were restriction digested and screened in this low stringency manner. Of the 50 clones, 36 were identified by positive hybridisation which when sequenced gave five full length (3γ and 2β) and 11 partial unique β/γ crystallin cDNAs. Tentative identification of the sequences as β/γ crystallin isoforms came from nucleotide comparisons at GenBank/NCBI (BLASTN algorithm). These 11 partial β/γ cDNAs were used as templates to design isoform specific oligonucleotide primers ($T_{\rm m} \ge 65$ °C) to obtain full length cDNA sequences by 5' and 3' RACE.

Most of the known vertebrate β crystallin cDNA isoform sequences are conserved homologues, thus partial cDNA sequences that were not initially identified in the low stringency Southern blot were obtained by designing isoform specific primers based on an alignment of known mammalian and chicken β isoform sequences we obtained from GenBank/NCBI. Three toothfish β crystallins (β A2, β A4, β B3) were obtained by PCR screening of 288 clones in three 96-well plates using the following pairs of primers: β A2F 5'-CCCTCGCTACGAGGCCTGGAGCGGAAA-3', β A2R 5'-CGGATGGACTGCACCTGGTTGGTGTGAGCC-3'; β A4F 5'-CTGGAAGATCGTGGTGTGGGGATGA-3', β A4R 5'-CGTGCGTGTCGCTGTCCAG-3'; β B3F 5'-GCCCCCCTCCA-GATCGACAGCC-3', and β B3R 5'-GCGGCGCACCGACTG-CATCA-3'.

Sequences of several partial γ crystallins cDNA (γ S2, γ M7, γ M9) were used to design primers to obtain the full length crystallin cDNA sequences by PCR of clones from the cDNA library. For the γ S2 isoform we designed sense and antisense primers (γ S2F 5'-TGATAGCATCAAGTCATGCCGCTCTAT-CCAAAA-3', γ S2R 5'-AAATTCAGGGTTCTGCAGCAGGG-GG-3'). For the γ M7 and γ M9 we used isoform specific primers Dm γ M7-R 5'-TTCGTGCATCTGACCGACATGTCGGACGC-3' and γ M9-R 5'-TTAGCAGATTTTGCAAACATTAGTTGC-3' paired with a vector (pBK-CMV) specific primer T7AS 5'-CAGTGAATTGTAATACGACTCACTATAGGGC-3'. A total of 384 clones in four 96-well microtitre plates were screened by this PCR method to obtain the full length cDNAs of these 3 isoforms.

The remainder of the γ cDNA clones reported herein were identified by Southern Blot screening of recombinant clones with the 407 nt γ M8e probe. Recombinant crystallin clones were sequenced in both directions using ABI Prism BigDye 3 Terminator reagent (Applied Biosystems, CA, USA) and the reactions analysed on a ABI3730*xl* sequencer (Applied Biosystems) at the W. M. Keck Center for Comparative and Functional Genomics (University of Illinois, Urbana-Champaign). Screening approximately 800 clones by Southern blot and PCR methods yielded two α , six β and fourteen γ crystallin cDNA sequences.

2.3. Genomic Southern Blots of D. mawsoni DNA

Approximately 20 µg of genomic DNA from *D. mawsoni* liver was digested overnight with 50 units of *Eco*RI. Digests were resolve on a 1% agarose gel, vacuum blotted to Hybond-N nylon membrane (Amersham Biosciences), UV-crosslinked and hybridised with ³²P-labeled probes (see figure legends) in PerfectHyb (Sigma–Aldrich, St Louis, MO, USA). The blot was washed to 55°C in 0.1XSSC/0.5%SDS, and autoradiographed on Kodak Biomax X-ray film.

2.4. Screening and analyses of BAC clones containing crystallin genes

A large-insert DNA bacterial artificial chromosome (BAC) library was constructed for *D. mawsoni* using red blood cells based on previously published methods (Miyake and Amemiya,

2004). A total of 67,584 recombinant clones, equivalent to about 6X genome (2C) coverage were robotically picked and archived in 384-well master culture plates, as well as printed on nylon hybridisation membranes as a macro-array for screening target genes. The α crystallin gene family was screened with full length $^{32}\text{P-labelled cDNAs}$ for αA and $\alpha B,$ and the β crystallin gene family with full length BA2 and BB2. Gamma crystallin gene regions were screened with the same 32 P-labelled γ M8e probe used in the initial identification (at higher stringency) of the 14 γ crystallin isoforms. Putative positive clones were then re-streaked from archived glycerol stocks to single colonies. Recombinant BAC plasmids were prepared from liquid cultures of single colonies, digested with NotI and electrophoresed on a pulsed field system (CHEF Mapper XA System, BioRad). Low range PFG (pulsed field gel) markers (NEB) were used to estimate the size of the BAC clone inserts. To confirm the clones were positive for crystallin genes, the NotI digested BAC DNA was vacuum blotted from the gel onto Hybond-Nylon membrane, and the membrane was hybridised with the same α,β or γ probes (as appropriate) used in the initial BAC library macro-array screening. The verified positive clones were then digested with HindIII, and separated on a 1% agarose gel (40 cm) along with DNA Marker II for Genomic DNA Analysis (Fermentas, MD), stained with SYBR Green I (Molecular Probes/Invitrogen) and photographed using Kodak 1D Image Analysis System (Eastman Kodak Company, New Haven, CT, USA). The BAC clone fingerprint (HindIII restriction fragment mobilities) images were edited with IMAGE v3.10 (www.sanger.ac.uk/Software/Image/), and imported into FPCv8.2 (Fingerprint Contig; www.agcol. arizona.edu/software/fpc/) for construction of clone order and overlap (Nelson and Soderlund, 2005).

Table 1

Dissostichus mawsoni eye lens crystallin cDNA sequences and their GenBank accession numbers

Sequence name	cDNA length (bp)	Amino acid length (predicted)	GenBank accession
αA	1012	176	DQ143965
αB	1452	164	DQ147910
β A 1	1301	196	DQ147911
β A 2	850	196	DQ143966
βA4	712	196	DQ143967
β B 1	777	230	DQ143968
β B2	864	208	DQ143669
β B 3	1015	253	DQ143670
γN	752	183	EU016230
γS1	737	176	DQ143671
γS2	891	175	DQ143672
γM1	620	178	DQ143673
γM3	603	176	DQ143675
γM4	613	180	DQ143676
γM5	765	176	DQ143678
γM7	645	175	DQ143679
γM8a	656	182	DQ268581
γM8b	623	184	DQ143681
γM8c	677	183	DQ143682
γM8d	619	184	DQ143683
γM8e	628	177	DQ143674
γM9	672	178	DQ143677

2.5. Nomenclature

Nomenclature of the α and β crystallins from *D. mawsoni* was based on homology with vertebrate crystallins. Gamma crystallin designations for toothfish were based on the previous nomenclature " γ M" of zebrafish and carp as much as possible (Chang et al., 1988; Wistow et al., 2005). Further designations were based on BLASTN (GenBank/NCBI), our phylogenetic reconstruction, identity matrices and manual inspection of aligned amino acid sequences.

2.6. Bioinformatics

DNA sequences were edited for quality and to remove vector and other non-cDNA portions. Amino acid translations of the crystallin cDNAs were aligned using ClustalW in BioEdit v7.0.5.2 (Hall, 1999), and manually checked for alignment quality. Amino acid identity matrices (IDENTIFY, BLOSUM62, GONNET) were generated using BioEdit v7.0.5.2.

For phylogenetic analyses, γ crystallins from *D. mawsoni* and other ectothermic species (see Table 1 Suppl. Tables 1 and 2 for GenBank accession nos.) were first aligned by the translated amino acid sequences. Nucleotide sequences were aligned based on this amino acid alignment using CodonAlign2.0 (Hall, 2004). All phylogenetic reconstructions were performed using DNA sequence alignments unless indicated otherwise. Tree construction was done by MrBayes3.1.1 using the evolutionary model GTR+I+ Γ (Waddell and Stell, 1997) as selected by Mr. Modeltest v2.0 (Nylander, 2004). Four runs using four chains and 4,000,000 generations were performed. All other parameters in MrBayes 3.1.1 were left at default. Stationarity was assessed using the "sump burnin=2500" command and examining the



Fig. 1. Alignment of the *in silico* translation of cDNAs of the Antarctic toothfish (Dm) αA (A) and αB (B) crystallins with zebrafish (Dr), Asian catfish (Cf, Cb), cow (Bt) and human (Hs). (A) The identity between the αA homologues is very high (black background) in the typically more divergent N-terminal domain, as well as the conserved α crystallin domain. (B) Although the α crystallin domain of the αB sequences is also highly conserved, the more divergent N-terminal domain and the C-terminal extension are more variable than in the αA crystallin. The domain region designations are based on previous reports (Caspers et al., 1995; Narberhaus 2002). Grey regions indicate \geq 50% similarity between sequences.

plot of the generation versus the log likelihood values. A 50% majority rule consensus tree was generated using the command "sumpt burnin=2500".

3. Results

3.1. a Crystallins

The *in silico* translation of the α A and α B lens crystallin cDNA sequences from the Antarctic toothfish *D. mawsoni* are shown in Fig. 1 (GenBank Accession nos. in Table 1). Toothfish sequences were aligned with zebrafish *D. rerio* (Dr) (Posner et al., 1999; Runkle et al., 2002), catfish *Clarias fuscus* (Cf) and catfish *Clarias batrachus* (Cb) (Yu et al., 2004), with cow *Bos taurus* (Bt) and human *Homo sapiens* (Hs) sequences (GenBank accession nos: NP_776714, NP_000385, NP_776715, AAP36581). Designations of the N-terminal region/domain, α crystallin domain and C-terminal extensions are based on previously published reports (de Jong et al., 1993; Caspers et al., 1995; Narberhaus, 2002).

The α A and α B alignments (Fig. 1) highlight the conservation of the α crystallin domain between these five fish and mammalian taxa. The amino acid identities between the α A sequences range from a low of 68.7% between toothfish and human, to a high of 94.2% between cow and human (Table 2). Among the fish species analysed, α B sequences are more divergent from each other than are the α A sequences (Table 2). The greatest variation is in the N-terminal region of the α B crystallin (Fig. 1B). In contrast, the N-terminal region of the α A crystallin is highly conserved between fish and mammals, thus accounting for the greater sequence similarity of α A between these phylogenetically distant species (Fig. 1A).

Southern blot of *Eco*RI digested *D. mawsoni* genomic DNA probed with ³²P-labeled full length *D. mawsoni* cDNAs for α A and α B resulted in hybridisation to three distinct bands (3, 1, 2) in decreasing intensity (Fig. 2). The presence of three bands is consistent with the presence of three α isoforms in zebrafish and catfish (Yu et al., 2004; Smith et al., 2006). It is tempting to suggest that the intensity of the three bands in the Southern blot is representative of the gene copy number, isoform type and

Table 2 Amino acid identities (percent identical) based upon the alignments of αA and αB crystallins presented in Fig. 1

Sequence name	Dm αA	Dr aA	Cf αA	Bt αA	Hs αA
Dm αA	100	77.2	72.7	71.0	68.7
Dr αA		100	81.5	73.2	72.1
CfαA			100	71.0	70.4
Bt αA				100	94.2
Hs αA					100
Sequence name	$Dm \ \alpha B$	$Dr \; \alpha B$	$Cf\alpha B$	Bt αB	Hs αB
Dm αB	100	47.0	51.4	48.2	47.7
Dr αB		100	56.5	60.5	59.4
Cb αB			100	62.5	61.9
Bt αB				100	97.1
Hs αB					100



Fig. 2. Genomic Southern blot of *D. mawsoni Eco*RI digested DNA. Digested DNA (20 μ g) was probed with full length ³²P-labelled cDNA probes from α A and α B from cloned from *D. mawsoni*. The number arrows to the left of the blot indicate the three bands that hybridised to the probes. It is known that there are at least two (α A and α B) genes in toothfish (this study) and zebrafish (Dahlman et al., 2005), in addition we isolated a partial fragment of a third α B₂ (see Results section).

possibly expression levels. However, attempts to isolate a complete second αB cDNA sequence were unsuccessful. A partial second αB sequence was isolated giving an *in silico* translation tentatively designated αB_b "TKDGVVEITGKHEDRKDEHG-FVSRSFTRKYTLPSNTDVEKVNSSL", corresponding to position 86 to 130 of the full length *D. mawsoni* αB sequence (Fig. 1) which spans the conserved ' α crystallin' domain.

3.2. B Crystallins

The β crystallin protein sequences from the toothfish are presented in Fig. 3 and aligned with the zebrafish, cow and human sequences. The alignment highlights the highly conserved four domains (coloured underscores) which form the Greek-key fold motifs. Within the β crystallins, the β A3 and β A1 differ by their initiation codons; the translational start of the β A3 transcript is upstream of the β A1 (McDermott et al., 1997; Bloemendal et al., 2004). The six β crystallin isoforms isolated from toothfish are clear homologues of the β crystallin family. They are unambiguously identified when contrasted with each other (Table 3) and compared to mammalian orthologues (Fig. 3). The toothfish β A1, β A2, β A4 sequences show a high degree of sequence identity (100% identity – black background) or sequence similarity (\geq 50% similarity – grey background) to both the zebrafish and mammalian (cow and human) β crystallins (Fig. 3). The β B1, β B2, and β B3 toothfish sequences again show conservation within their domains (coloured underscores) to zebrafish, cow and human (Fig. 3). The proline– arginine rich repeat, designated a 'PAPA-arm' (Bloemendal et al., 1984; Hejtmancik et al., 1986; Coop et al., 1998) that is found N-terminal to the first domain of the β B1 mammalian crystallin, is similar to the PAPA domain found in the Cterminal domain of the toothfish β B3 crystallin. An analogous PNPN domain is also found in the C-terminal domain of zebrafish β B3 (Fig. 3).

	10	20	30	40	50	60	70	80	90	100	110	120
Dm BA1				QINPMP		LCP W	KITVYDOEYFOGRR	MEETACCONIN	EC-CMENIRS	LKVDC	CAWVGYEHSSE	CCOOFVLE
Dr_BA1			IQIA	LTNPMP		LGP W	KLTVYDQENFQGKR	MEETAACQNIN	iec-gmdnvrs	LKVEC	GAWAGYEHSSFO	CGQQFILE
Bt_βA3 (βA1)]	METQTVQQELE	SLPTTKMA	QTNPMPGS		VGP W	KITIYDQENFQGKR	MEFTSSCPNVS	BR-NFDNVRS	LKVBC	GAWVGYEHTSFO	CGQQFVLE
Нѕ_βАЗ (βА1)]	METQAEQQELE	TLPTTK	QTNPTPGS		LGP	KITIYDQENFQGKR	MERTSSCPNVS	DR-SFDNVRS	LKVDS	CAWIGYEHTSFO	CGQQFILE
Dm BA2			MK	I-TQQMEQ		MCQ F	KITVWEEENFQGKR	CEEMLECONIN	ER-GFNKIRS	IK <mark>VE</mark> N	GPWVGYEYPEF	QGQQFILE
Dr_BA2			idiN	QREQMEQ		QCQ W	RITVWEEENFQGKR	CEFRLECPNII	DR-DFQKIRS	IKVDN	GPWVGYEYPEY	QGQQFILE
Bt_BA2			iMS	-SAPAQG		PAP A	SLTLWDEEDFQGRR	CRLLSDCANIG	BRGGLRRVRS	VKVDN	CAWVAFEYPDF	QGQQFILE
Hs_BA2				-SAPAPG		PAP A	SLTLWDEEDFQGRR	CRLLSDCANVO	ERGGLPRVRS	VKVDN	CVWVAFEYPDF(QGQQFILE
Dm BA4				HHCTKF		SCH W	KIIVFDEECFQGRR	HEFTSECCNVN	EF-GFETVRS	MRVES	CAWVGYEHASY	QGQQFVLE
Dr_BA4				HHCTKF		S <mark>G</mark> H W	KIIVYDEECFQGR <mark>H</mark>	heftseconvi	ief-gfesvrs	LRVES	Gawvgyehasy(QGHQFVLE
Bt_BA4		MSGMFSGS	ISETSG	LQCTKS		AGH W	KIVVWDEEGFQGRR	HEFTAECPSVI	EL-GFETVRS	LKVLS	CAWVGEDHAGE	QGQQYVLE
HS_DA4				LQCTKS		AGPW	KMVVWDEDGFQGRR	HDITADCPSVI	DL-GFDTVRS	LKVLS	CANVGEDHAGE	QCQQY1115
Dm_βB1	MSQTAKSASSQ	- GTDAKDKGAP	- PAATSKA	TKTGEPG		MGS F	RMMMFDQENFQGKM	IEVQNECMNVC	DR-GMDRVRS	IIVEC	GPFVAFEQTNFI	RGEMFILE
Dr_βB1	MSQTAKSATNQ	- GTDAKEKGAP	APAATSKA	SKTGDPGF		MGN Y	KIFLFDQENFQGRM	MEVQNECMNVC	ER-GMDRVRS	IIVEC	GPFVAEEQTNE!	RGEMFILD
Bt_BB1	MSQPA-AKASATAAVN	PGPDGKGKAGP	PPGPAPGS	GPAPAPAPAP	AQPAPAAKAE	LPPGS	KLVVFEQENFQGRR	VIDIOSCIECLNLC	DR-GFIDRVRS	IIVTS	GPWVAI=DQSNI5I	RGEMFVLE
HS_PB1	MSQAARASASATVAVN	FGFDIKGKGAP	FAGISFSF	GIILAPIIVP	TISAKAAE	IIPPGN I		NUT SGIC SILLA	DR-GFDRVRS	TTOPAN	SPWVAPEQSNE!	KG/SMETHS
Dm_BB2		MAADH	-QNPASKQ	QQPGTSA		F	KLVIYEQENFQGR	HELTGPCNNVQ	DEA-GVEKVGS	ILVLC	GPWVGYE <mark>QANC</mark> I	KGEQYVYE
Dr_BB2		MATDH	-QNPATKC	KQPVASA		F	KLVIYEQENFQGR	HELTGPCNNLQ	DEA-GVEKVGS	VLVQC	GPWVGFEQPGC	KGEQYVFE
Bt_BB2		MASDH	-QTQAGKF	QPLNP			KIIIFEQENFQGHS	HOLNGPOPNLE	UPT - GVIPKAGS	VLVQA	GPWVGYPQANCE GPWVGYPQANCE	KGEQFVFE
no_por		- Indon	21 grioni	20 Ditt			ATTT DO MILOOM				STATE PLATE	
Dm_BB3		MSEQQSAP	EQQAAGKS	QGGAGAS		צ	KVFLFEFENFQG CK	aefsaeckdvr	19K-GLEKVGS	VIVES	GPWAGYDRHGF?	IGEQFILE
Dr_BB3		MSDQQGVP	DQQVAGKS	QGGAGAI		¥	KVSLYEFENFRCKK	LELSAECKDLI	EK-NLEKVGS	IIVES	GPWVGFEQKGFI	LGEQFVLE
вс_рвз		MAEQHSTP	EQAAAGKS	HGGLGGS		¥	KVILYELENFOGKR	CRUSAROPSLI	DS-LLOKVGS	TOVES	GPWLAFDRRAFT GPWLAFDSRAFT	RGEONVLE
			- 2			1						
	130	140	150	160	170	180	190	200	210	22	0 230	240
-										.		
Dr BA1	RGDYPRFEATSGSNSTR	IERIMSERPI	CSANHKE	SKITTVEREN	INSGROFFIC-	DDYPSLO	AMGWANNEVQSMQI		PGIRGIQIIM	PCDURG	GEVICIREFESI	HAOTFOVO
Bt_βA3 (βA1)	RGEYPRWDAWSGSNAYH	IERLMSFRPI	CSANHKE	SKITIFEKEN	FIGROWEIC-	DDYPSLO	AMGWPNNEVGSMKI	QC GAWVCYQY	PGYRGYQYIL	PCDHHG	GDYKHWREWGSI	HAQTSQIC
Hs_βA3 (βA1)	rgeyprwdawsgsnayh	IERLMSFRPI	CSANHKE	SKMTIFEKEN	FIGRQWEIS-	DDYPSLQ	amgw <mark>fnnevgs</mark> mki	QS GAWVCYQY	PGYRGYQYIL	PCDHHG	GDYKHWREWGSI	HAQTSQIQ
Dm Ba2	KONYPRVIEAWSONSSVP	TRUM	KSANHSD	SEVENDORD	POGREEPMC -	TRAVAL	AMONCSKEWDSTKW		PGYRGYOV	PHRENC		OMHTNOVO
Dr BA2	KGDYPCYOAWSGNSSYR	TEHMLSFRPI	KCANHSD	SKITMYECED	MMGRKFEMC -	DDYPSLM	AMGWCSKEVPSIKV	NS GAWVGYOF	PGYRGYQYIF	ERDRRO	GEYRKYYEFCT(OAHSNOIC
Bt_BA2	KGDYPRW <mark>SAWSG</mark> SAGHH	SDQLLSFRPV	lc <mark>anh</mark> sd	SRVTLFEGEN	FQGCKFELN-	DDYPSLF	SMGWASKDVGSLKV	SS GAWVAYQY	PGYRGYQYVL	ERDHHS	GEFRNYSEFGT	QAHTGQLQ
Hs_BA2	KGDYPRWSAWSGSSSHN	SNQLLSFRPV	LCANHND	SRVTLFEGDN	FQGCKFDLV-	DDYPSLF	SMGWASKDVGSLKV	SS GAWVAYQY	PGYRGYQY	ERDRHS	GPFCTYGPLCT	QAHTGQLQ
Dm BA4	RGEYPOCDAFGGSNAVH	IERMTSFRPI	ACANHRIP	CRMUITEREN	DLGRKGDLS-	DDYPSLC	AMVWCNNEVGSLKT	OS GARVOYOY	PGYRGYOY	CORHC	GEFKHERDEGS	HCOTPOIO
Dr_BA4	RGEYPQCDSFGGSNAYH	IERMTSFRPI	SCANHRE	CRMTIYEREN	YLGRKGELS -	DDYPSL	AMGWCNNEVGSLRV	QS GAFVCYQF	PGYRGYQYIM	ECDRHC	GEYKQFREFGSI	HSQTPQIQ
Bt_BA4	RGEYPSWDAWSGNTSYP	ARRITSERDU	ACANHRD	SRUTTEROEN	INT ODVORT O	DDVDGT	AMONDONEWOSEHW		DOVDOROWIT		GDYKHFREWCSI	HAQTFQVQ
Hs_BA4				51121 41 6 X 61	ILTOXICOTTS -	DUIPSLO		HSGAWVCSOF	PGIRGPQIVI	ECDHHS		
	RGEYPSWDAWGGNTAYP	AERLTSERPA	AC <mark>ANH</mark> RD	SRLTIFEQEN	FLGKKGBLS-	DDIPSLO	AMGWEGNEVGSFHV	HS GAWVCSOF HS GAWVCSOF	PGYRGFQYVL	econns econns	GDYKHFREWGS1	HAPTFQVQ
Dm BB1	rgeypswdawggntayp kgeyprwdtwsnsyr	aerltsfrpa	AC <mark>ANH</mark> RD	SRLTIFEQEN	FLGKKGELS - FKGNKMEIQE	DDIPSLO	AMGWEGNEVGSFHV AHGFCDR - VGSVRV	HS GAWVCSOF HS GAWVCSOF PG GSWVGYQY	PGYRGFQYVL	BCDHHS BCDHHS	GDYKHFREWGS1 GDYKHYNDFS	HAPTFOVO - AYQPOMO
Dm_βB1 Dr_βB1	RGEYPSWDAWGGNTAYP KGEYPRWDTWSNSYR KGEYPRWDTWSNSYR	AERLTSERPA SDCLMSLRPV SDCLMSLRPF	ACANIIRD RMDSL-E RMDPM-E	SRLTIFEQEN HKICLYELSI HKICLFELSI	FRGNKMETQE FRGNKMETQE	DDVPSLQ DDVPTLW DDVPTLW	AMGWEGNEVGSFHV AHGFCDR - VGSVRV AHGFCDR - VGSVRV	HS GAWVCSOF HS GAWVCSOF PG GSWVGYQY PG GAWVGYQY	PGYRGFQYVL PGYRGFQYVL PGYRGYQYLF PGYRGYQYLF	BCDHHS BCDHHS BC	gdykhfræggsi gdykhyndfs- gdyrhynefc	HAPTFQVQ -AYQPQMQ -AFQPQIQ
Dm_βB1 Dr_βB1 Bt_βB1	RGEYPSWDAWGCNTAYP KGEYPRWDTWS NSYR KGEYPRWDTWS NSYR KGEYPRWDTWS SSYR	AERLTSFRPA SDCLMSLRPV SDCLMSLRPF SDRLMSFRPI	ac <mark>anh</mark> rd RMDSL-E RMDPM-E KMDAQ-E	SRLTIFEQEN HKICLYELSI HKICLFELSI HKLCLFEGAN	IFLEKKGELS FRENKMEIQE FRENKMEIQE FRENTMEIQE	DDYPSLQ DDYPSLQ DDVPTLW DDVPTLW DDVPSLW	AMGWEGNEVGSFHV AHGFCDR - VGSVRV AHGFCDR - VGSVRV IVYGFCDR - VGSVRV	HS GAWVCSON HS GAWVCSON PG GSWVGYQY SS GTWVGYQY	PGYRGFQYVL PGYRGYQYLF PGYRGYQYLF PGYRGYQYLL	BCDHHS BCDHHS BC BC BP	gdykhfrewgsi gdykhyndfs - gdyrhynefc - gdfrhwnewg -	HAPTFQVQ -AYQPQMQ -AFQPQIQ -AFQPQMQ
Dm_βB1 Dr_βB1 Bt_βB1 Hs_βB1	RGEYPSWDAWGGNTAYP KGEYPRWDTWSNSYR KGEYPRWDTWSNSYR KGEYPRWDTWSSSYR KGEYPRWNTWSSSYR	ABRUTSERPA SDCIMSLRPV SDCIMSLRPF SDRIMSERPI SDRIMSERPI	ACANHIRD RMDSL-E RMDPM-E KMDAQ-E KMDAQ-E	SRLTIFEQEN HKICLYELSD HKICLFELSD HKLCLFEGAN HKISLFEGAN	FILCKROBLS FILCKROBLS FIRCNROBLOB FIRCNTOBLOB FIRCNTIBLOG	DDIPSIC DDYPSIC DDVPTLW DDVPTLW DDVPTLW DDVPSLW	AMGWEGNEVGSFHV AHGFCDR - VGSVRV AHGFCDR - VGSVRV IVYGFCDR - VGSVRV IVYGFCDR - VGSVRV	HS GAWVCSOF HS GAWVCSOF PG GSWVGYQY PG GAWVGYQY SS GTWVGYQY SS GTWVGYQY	PGYRGFQYVL PGYRGFQYVL PGYRGYQYLF PGYRGYQYLF PGYRGYQYLL PGYRGYQYLL	9CDHHS 9CDHHS 9C 9C 9P 9P	GDYKH FRDWCS GDYKH YNDFS GDYRH YNDFC GDFRHWNDWC GDFRHWNDWC	HAPTFOVO -AYQPOMQ -AFQPOIQ -AFQPOMQ -AFQPOMQ
Dm_βB1 Dr_βB1 Bt_βB1 Hs_βB1 Dm_βB2	RGEYPSWDAWGGNTAYP KGEYPRWDTWS - NSYR KGEYPRWDTWS - NSYR KGEYPRWDTWS - SSYR KGEYPRWTWS - SSYR KGEYPRWTWS - SSYR	ABRUTSFRPA SDCIMSLRPV SDCIMSLRPF SDRIMSFRPI SDRIMSFRPI SDRILAFRPI	ACANHIRD RMDSL-E RMDPM-E KMDAQ-E KMDAQ-E KMDAQ-E	SRLTIFEQEN HKICLYELSD HKICLPELSD HKLCLPEGAN HKISLPEGAN HKIVLYENPS	FILCKROBLS FROMMETOR FROMMETOR FROMMETOR FROMTISTOR FROMTISTOR	DDIFSIG DDIFSIG DDVPTLW DDVPTLW DDVPSLW DDAPSLW DDAPSFH	AMGWEGNEVGSFHV AHGFCDR - VGSVRV AHGFCDR - VGSVRV IVYGFCDR - VGSVRV IVYGFSDR - VGSVRV IAHGYQEK - VSSVRV	HS GAWVCSOL HS GAWVCSOL PG GSWVGYQY PG GAWVGYQY SS GTWVGYQY SS GTWVGYQY QS GTWVGYQY	PGYRGFQYVL PGYRGFQYVL PGYRGYQYLF PGYRGYQYLF PGYRGYQYLL PGYRGYQYLF	90000000000000000000000000000000000000	GDYKH FREWGS GDYKHYNDFC GDYRHYNDFC GDFRHWNEWG GDFRHWNEWG SEYKDSSEFG	HAPTFOVO - AYQPOMQ - AFQPOTQ - AFQPOMQ - AFQPOMQ - AEIFQIQ
Dm_βB1 Dr_βB1 Bt_βB1 Hs_βB1 Dm_βB2 Dr_βB2	RGEYPSWDAWGGNTAYP KGEYPRWDTWS - NSYR KGEYPRWDTWS - NSYR KGEYPRWDTWS - SSYR KGEYPRWDTWS - SSYR KGEYPRWDSWT - NSRR KGEYPRWDSWT - NSRR	ABRUTSFRPA SDCIMSURPV SDCIMSURPF SDRIMSFRPI SDRIMSFRPI SDTILAFRPI SDCIVAFRPI	ACANIERD RMDSL-E RMDPM-B KMDAQ-B KMDAQ-B KVDSQ-D KVDSQ-D	SRLTIFEQEN HKICLYELSD HKICLFEGAN HKISLFEGAN HKIVLYENPS HKIVLYENPS	FILGKKGELS FRONT DE FRONT DE F	DDTPSIC DDTPSIC DDVPTIW DDVPTIW DDVPSIW DDAPSIW DDVPSFH	AMGWEGNEVGSFEV AHGFCDR - VGSVEV AHGFCDR - VGSVEV IVYGFDR - VGSVEV VHGYDE - VGSVEV AHGYDEK - VSSVEV	HS GAWVCSOF HS GAWVCSOF PG GSWVGYQY PG GAWVGYQY SS GTWVGYQY SS GTWVGYQY QS GTWVGYQY QS GTWVGYQY	PGYRGFQYVL PGYRGFQYVL PGYRGYQYLF PGYRGYQYLL PGYRGYQYLL PGYRGYQYLF PGYRGFQYLF	ВСОННS ВСОННS ВС ВР ВР ВР БК БК	gdykh prewgsi gdyrh yndfs - gdyrh yndfc - gdfrh wnewg - gdfrh wnewg - gdfrh wnewg - gefr bsefg - gefr bsefg - gefr bsefg -	HAPTFOVO -AYOPOMO -AFOPOIO -AFOPOMO -AFOPOMO -AEIPOIO -AEIPOIO
Dm_βB1 Dr_βB1 Bt_βB1 Hs_βB1 Dm_βB2 Dr_βB2 Bt_βB2 Hz_βD2	RGEYPSWDAWGGNTAYP KGEYPRWDTWS - NSYR KGEYPRWDTWS - NSYR KGEYPRWDTWS - SYR KGEYPRWDTWS - SYR KGEYPRWDSWT - NSRR KGEYPRWDSWT - NSRR KGEYPRWDSWT - SSR	AERUTSFRPA SDCUMSLRPV SDCUMSLRPF SDRUMSFRPI SDRUMSFRPI SDTILAFRPI SDCIVAFRPI TDSUSSIRPI	ACANHRD RMDSL-B RMDPM-B KMDAQ-B KMDAQ-B KMDAQ-B KVDSQ-D KVDSQ-D KVDSQ-B	SREATPEORN HKICTYBLSE HKICTPBLSE HKICTPBLSE HKICTPBAN HKISTFBAN HKIVTYBNPS HKITYBNPS HKITYBNPS	FRIGREGES FRIGREGES FREGREGES FREGREGES FREGREGES FREGREGES FREGRES FR	DDYPSIC DDYPSIC DDYPSIC DDYPSIC DDYPSIC DDYPSIC DDYPSIC DDYPSIC DDYPSIC	AMGWEGNEVGSFEV AHEGPCDR - VGSVRU AHEGPCDR - VGSVRU VVGPCDR - VGSVRU VVGPSDR - VGSVRU AHGYQEK - VSSVRU AHGYHEK - VSSVRU AHGYDEK - VSSVRU	HS GAWVCSOF HS GAWVCSOF PG GAWVCSOF SS GTWVCYQY SS GTWVCYQY SS GTWVCYQY SS GTWVCYQY SS GTWVCYQY SS GTWVCYQY	PGYRGFQYVL PGYRGFQYVL PGYRGYQYLF PGYRGYQYLL PGYRGYQYLL PGYRGYQYLF PGYRGFQYLF PGYRGFQYLF	9CDHHS 9CDHHS 9C 9P 9P 9P 9K 9K 9K	GDYKEPREWGSI GDYKEYNDFS- GDFREWNEWG- GDFREWNEWG- GDFREWNEWG- GEFRESEFG- GEFRESEFG- GEFRESEFG- GDYREGGFG-	HAPTFOVO -AYOPONO -AFOPOIO -AFOPONO -AFOPONO -AEIPOIO -AEIPOIO -APOPOVO
Dm_BB1 Dr_BB1 Bt_BB1 Hs_BB1 Dm_BB2 Dr_BB2 Bt_BB2 Hs_BB2	RGEYPSWDAWGGNTAYP KGEYPRWDTWS - NSYR KGEYPRWDTWS - SSYR KGEYPRWDTWS - SSYR KGEYPRWDSWT - NSRR KGEYPRWDSWT - NSRR KGEYPRWDSWT - SSRR KGEYPRWDSWT - SSRR	ABELTSFRPA SOCIASIRPY SORIASIRPP SORIASFRPI SORIASFRPI SOCIAFRPI SOCIAFRPI TDSISSIRPI	ACANHRD RMDSL-S RMDPM-S KMDAQ-S KMDAQ-S KMDAQ-S KVDSQ-S KVDSQ-S KVDSQ-S KVDSQ-S	SRLTIFEQEN HKICLFELSI HKICLFEGAN HKISLFEGAN HKIVLYENPS HKIVLYENPS HKITLYENPN HKITLYENPN HKITLYENPN	HIGH KOBIS FILG KROBIS FILG KROBIS FIRG NIMELOF FIRG NIMELOF FIRG NIMELOF FIRG KROBIT FIRG KROBIT FIRG KROBIT	DDTPSIC DDTPSIC DDTPTIK DDTPTIK DDTPSIK DDTPSIK DDTPSFH DDTPSFH	Angene vos ve Angeron - vos ve Vygeron - vos ve Vygeron - vos ve Vygeron - vos ve Vygeron - vos ve Angygek - vss ve Angygek - vss ve Angygek - vss ve Angygek - vss ve	HE GAWUCSO HS GAWUCSO PG GSWUGYOY SG GTWUGYOY SS GTWUGYOY SS GTWUGYOY SS GTWUGYOY SS GTWUGYOY SS GTWUGYOY SS GTWUGYOY	PGYRGFQYVL PGYRGFQYVL PGYRGYQYLL PGYRGYQYLL PGYRGYQYLL PGYRGYQYLF PGYRGYQYLF PGYRGFQYLF PGYRGLQYLL PGYRGLQYLL	SCDHHS SCDHHS SC BP BP SK SK SK	GDYKEPREWGS GDYKEYNDFC - GDYREYNDFC - GDFREWNEWG - GDFREWNEWG - GDFREWNEWG - GEFRECSEFG - GDFRESSEFG - GDYRDSGDFG - GDYRDSGDFG -	HAPTFOVO -AYOPONO -AFOPOIO -AFOPONO -AFOPONO -ABIPOIO -AALPOIO -APOPOVO -APOPOVO -APOPOVO
Dm_βB1 Dr_βB1 Bt_βB1 Hs_βB1 Dm_βB2 Dr_βB2 Bt_βB2 Hs_βB2 Hs_βB2 Dm_βB3	RGEYPSWDAWGGNTAYP KGEYPRWDTWS - NSYR KGEYPRWDTWS - SSYR KGEYPRWDTWS - SSYR KGEYPRWDSWT - SSYR KGEYPRWDSWT - NSRR KGEYPRWDSWT - SSRR KGEYPRWDSWT - SSRR KGEYPRWDTWT - NSON	ARRITSFRPA SDCLASURPY SDCLASURPP SDRLASFRPI SDRLASFRPI SDRLASFRPI SDRLASFRPI SDCLAAFRPI SDCLAAFRPI TDSLSSURPI SYSLESURPI	ACANHRD RMDSL-S RMDPM-B KMDAQ-B KMDAQ-B KVDSQ-D KVDSQ-D KVDSQ-C KVDSQ-S KVDSQ-S KVDSQ-S	SRLTIFEQEN HKICIFELSI HKICIFELSI HKICIFEGAN HKISIFEGAN HKIVIFENPS HKIVIFENPS HKITIFENPS HKIIFENPS	HIGHROEDS FELGREGELS FREMEMETOP FREMEMETOP FREMEMETOP FREME FREME FTGREMETIF FTGREMETIF FTGREMETUT	DDYPSIC DDYPSIC DDYPSIC DDYPSIC DDYPSIC DDYPSIC DDYPSIC	AARGWEGNEVGSFEU AARGFCDR - VGSVRU AARGFCDR - VGSVRU VYYGFCDR - VGSVRU VYYGFCDR - VGSVRU VARGYQEK - VSSVRU AARGYQEK - VSSVRU AARGYQEK - VSSVRU AARGYQEK - VSSVRU AARGYQEK - VSSVRU	HE CANVCSOL HE CANVCSOL PC CENVCSOL PC CENVCSOL SE CTWVCSOL SE CTWVCSOL SE CTWVCSOL SE CTWVCSOL SE CTWVCSOL SE CTWVCSOL SE CTWVCSOL	PGTRGFQYVL PGTRGFQYVL PFGTRGYQYLF PFGTRGYQYLF PFGTRGYQYLF PFGTRGFQYLF PFGTRGFQYLF PFGTRGFQYLF PFGTRGFQYLF	COHHS CC PC EP EP PK PK BK BK	GDYKHERENGS GDYKHYNESC- GDYRHYNESC- GDFRHMNENG- GDFRHMNENG- GEYKDSSEFG- GFRECSEFG- GDYKDSSEFG- GDYKDSSEFG- GDYKDSSEFG- GDYKMNNWE-	HAPTFOVO -AYOPONO -AFOPONO -AFOPONO -AFOPONO -AFOPONO -ALIPOIO -APHPOVO -APHPOVO -APHPOVO -APHPONO
Dm_βB1 Dr_βB1 Bt_βB1 Hs_βB1 Dm_βB2 Dr_βB2 Bt_βB2 Hs_βB2 Hs_βB2 Dm_βB3 Dr_βB3 Dr_βB3 Dr_βB3	RGEYPSWDAWGGNTAYP KGEYPRWDTWS - NSYR KGEYPRWDTWS - SSYR KGEYPRWDTWS - SSYR KGEYPRWDSWT - SSYR KGEYPRWDSWT - NSRR KGEYPRWDSWT - SSRR KGEYPRWDSWT - SSRR KGEYPRWDTWT - NSON KGEYPRWDTWT - NSON	ARRITSFRPA SDCLASURPY SDCLASURPY SDRLASFRPI SDRLASFRPI SDRLASFRPI SDRLASFRPI SDCLAAFRPI SDCLAAFRPI SDCLAAFRPI STSLSURPI STSLFSURPI STSLFSURPI	ACANHRD RMDSL-B RMDPM-B KMDAQ-B KMDAQ-B KMDAQ-B KMDSQ-B KVDSQ-B KVDSQ-B KVDSQ-B KVDSQ-B KVDSQ-B	SRITIFEOR HKICLYBLSI HKICLFEGAN HKISLFEGAN HKISLFEGAN HKIVLYBNPS HKITLYBNPS HKITLYBNPS HKITLYBNPS HKILLFENPG	HIGK COLO FILGER COLO FILGER COLO FILGER COLO FILGER COLO FILGER COLO FILGER FILE FILGER FILE FILE FILE FILE FILE FILE FILE FILE	DDVPSIA DDVPSIA DDVPATA DDVPATA DDVPSIA DDVPSIA DDVPSIA DDVPSPH DDVPSPH DDVPSPH	AARGWEGNEVOSFEV AARGFCDR - VOSVRU AARGFCDR - VOSVRU VYYGFCDR - VOSVRU VYYGFCDR - VOSVRU VARGYDEK - VSSVRU AARGYDEK - VSSVRU AARGYDEK - VSSVRU IGEGFDDR - VASVRA VYRGFDDR - VASVRA	HE CANVCSOL HE CANVCSOL PC SAVVCSOL SE GAWVGSOL SE CTWVGSOL SE CTWVGSOL SE CTWVGSOL SE CTWVGSOL SE CTWVGSOL SE CTWVGSOL SE CTWVGSOL LN CTWVGSOL	PGTRGTQYVL PGTRGTQYUF PGTRGTQYLF PGTRGTQYLF PGTRGTQYLF PGTRGTQYLF PGTRGTQYLF PGTRGTQYLF PGTRGTQYLF PGTRGTQTFF	ECDHHS ECDHHS EC EP EP EK EK EK EH	GDYKHPRENGS GDYKHYNESC 	HAPTFOVO -AFOPONO -AFOPONO -AFOPONO -AFOPONO -AFOPONO -AALPONO -APAPONO -APAPONO -APAPONO -APAPONO -APAPONO
Dm_βB1 Dr_βB1 Bt_βB1 Hs_βB1 Dm_βB2 Dr_βB2 Bt_βB2 Hs_βB2 Dm_βB3 Dr_βB3 Bt_βB3 Bt_βB3 Hs_βB3	RGEYPSWDAWGGNTAYP KGEYPRWDTWS - NSYR KGEYPRWDTWS - SSYR KGEYPRWDTWS - SSYR KGEYPRWDSWT - NSRR KGEYPRWDSWT - NSRR KGEYPRWDSWT - SSRR KGEYPRWDSWT - SSRR KGEYPRWDTWT - NSON KGEYPRWDTWT - NSON KGEYPRWDTWT - NSON	ARRITSPRPA SDCLASTRPA SDCLASTRPA SDCLASTRPA SDRLASTRPA SDRLASTRPA SDRLASTRPA SDRLASTRPA SDRLASTRPA SDRLASTRPA STATESTRPA STATESTRPA SDSLASTRPA	ACANHRD RM SL-S RM PM-B KNDAQ-B KNDAQ-B KNDAQ-B KNDSQ-B KNDSQ-B KNDSQ-B KNDSQ-B KNDSQ-B KNDSQ-B HD GP-D ND GP-D	SRITIFEOR HKICLFEOR HKICLFEGAN HKISLFEGAN HKISLFEGAN HKITLYRNP HKITLYRNP HKILFENPG HKLHLFENPG HKLHLFENPG	HIGK GBLS - FRIGK GBLS - FRIGK MEIOF FRIGHT BIO FRIGHT BIO FRIGHT BIO FRIGK BEIT FRIGK MEIT FRIGK MEIT FRIGK MEIT FRIGK MEIT FRIGK MEIT FRIGK MEIT	DUPPSIC DUPPSIC DUPPSIC DUPPSIC DUPSIC DUPSIC DUPSIC DUPSIC DUPSIC DUPSIC	AARGWEGNEVOSFEU AARGFCDR - VOSVRU AARGFCDR - VOSVRU VYYGFCDR - VOSVRU VYGFCDR - VOSVRU VARGYDEK - VSSVRU AARGYDEK - VSSVRU AARGYDEK - VSSVRU AARGYDEK - VSSVRU AARGYDE - VASVRA AARGPDR - VASVRA AARGPDR - VASVRA	He CAWUCSOL HE CAWUCSOL PG CSWUCYCO SE CTWUCYCO SE CTWUCYCO SE CTWUCYCO SE CTWUCYCO SE CTWUCYCO SE CTWUCYCO SE CTWUCYCO SE CTWUCYCO LN CTWUCYM IN CTWUCYM IN CTWUCYM	PGTRGTQYVL PGTRGTQYUL PGTRGTQYLL PGTRGTQYLL PGTRGTQYLL PGTRGTQYLL PGTRGTQYLL PGTRGTQYLL PGTRGTQYLL PGTRGTQYLL PGTRGTQYLL PGTRGTQYLL PGTRGTQYLL PGTRGTQYLL PGTRGTQYLL	ECDHHS ECDHHS EC EP EP ER EK EK EK ER ER ER ER ER ER	GDYKHPRENGS GDYKHYNDFS- GDYRHYNEFC- GDPRHYNEFC- GDPRHYNEFC- GDPRHYNEFC- GDYKDSSEFG- GDYKDSSEFG- GDYKDSSEFG- SDYKHYNEFC- SDYKHYNEFC- SEYRHYNEFC- SEYRHYNEFC-	HAPTFOVO -AYOPONO -AFOPONO -AFOPONO -AFOPONO -AFOPONO -AFOPONO -APOPONO -APOPONO -APOPONO -APOPONO -ATOPONO -ANOPONO -ANOPONO

Fig. 3. Alignment of the *in silico* translation from cDNAs of the Antarctic toothfish (Dm) β crystallins against zebrafish (Dr), cow (Bt) and human (Hs). Immediately apparent is the extent of the identity of residues across the phylogenetically diverse taxa. Within the four conserved domains regions (alternating blue and green bars) there is high levels of identity across all of the β crystallins (black background) as well as substantial similarity (\geq 50% grey background). It is likely that only the β A1 is present for the toothfish (Dm_ β A1) as an upstream start codon (making an ' β A3' isoform) was not found in the cloned cDNA. The N-terminal regions of the toothfish β A1, β A2, β A4, β B2 and β B3 crystallins are highly conserved across all four taxa. Toothfish β B1 crystallin shows a high degree of conservation with zebrafish of the N-terminal linker element. Note the presence of a large poly-(PA) element in the C-terminal region of the toothfish β B3 crystallin, similar to the poly-(PN) element in the zebrafish β B3 crystallin.

A.J. Kiss, C.-H.C. Cheng / Comparative Biochemistry and Physiology, Part D 3 (2008) 155-171



Fig. 3 (continued).

3.3. y Crystallins

A total of 14 distinct γ crystallin cDNAs from the lens of the Antarctic toothfish (Dm) were isolated, cloned and sequenced (Table 1). In silico translations of the γ cDNAs and their alignment with three γ crystallins from the Asian (common) carp C. carpio (Cc) (Chiou et al., 1986) and 16 γ crystallins from zebrafish D. rerio (Dr) (Wistow et al., 2005), 14 'unamed protein products' (which we propose are in fact γ ; see below and Fig. 4) crystallins from spotted green pufferfish T. nigroviridis, as well as several other γ crystallins from lipshark, blind Mexican cavefish, African clawed frog, Japanese firebelly newt, iguana and sea squirt are shown (Fig. 4). There is a high degree of sequence identity among the various toothfish isoforms (40.7% to 92.3%; Table 4) as well as high conservation between ectothermic species (black background in Fig. 4) and similarities $(\geq 50\%$ grey background Fig. 4). Of the fourteen Antarctic γ crystallin sequences, two (γ S1 and γ S2), are likely paralogues of the γS isoform and one was clearly an orthologue of a γN isoform (Pan et al., 1997; Bloemendal et al., 2004; Wistow et al., 2005). Based on alignment (Fig. 4), percent identities (Table 4) and phylogenetic analysis (Fig. 5), the remaining eleven

Table 3

Amino acid identities toothfish β crystallins (percent identical) based upon pairwise alignment against each other

Sequence name	β A 1	βΑ2	β A 4	β B 1	β B 2	β B3
βA1	100.0	59.1	67.8	32.0	36.2	26.6
β A 2		100.0	54.0	35.8	38.5	28.8
βA4			100.0	32.0	36.6	28.8
β B 1				100.0	48.2	40.2
β B 2					100.0	42.6
β B 3						100.0

Antarctic γ crystallin cDNA sequences were designated (γ M) originally proposed for methionine rich γ crystallins from carp (Chang et al., 1988), but now adopted for other fishes as well as other aquatic animals such as some frogs (Lu et al., 1996). In naming the toothfish γ crystallins we used the orthologous name as assigned previously for carp and zebrafish (γ M1, γ M3, γ M4, γ M5, γ M7) where possible. In cases where there were not orthologues of toothfish γ crystallins, we continued to use the same γ M crystallin nomenclature, but used the next available " γ M#".

Five related γ crystallins from the toothfish, four of which all share a conserved nine amino acid N-terminal extension "MS (N/T)T(G/D)M(N/S)M(R/-)" and did not have clear homologues in zebrafish or carp, were assigned as $\gamma M8a \sim e$ (Figs. 4 and 5). These four $\gamma M8$ toothfish crystallins ($\gamma M8a \sim d$) consistently formed a distinct clade in phylogenetic trees, regardless of the method of tree construction (Fig. 5). In addition to the toothfish specific $\gamma M8$ isoforms there are clearly identified $\gamma M1$ crystallin with the three residue conserved "MMF" insertion between positions 102 and 104, which is also present in yM1 of zebrafish, carp, pufferfish and lipshark (Ccx_y1) (Fig. 4). No yMX sequence was identified from our toothfish lens cDNA library. However, given the resolution of our phylogram, we place the previously reported zebrafish yMX isoform (Wistow et al., 2005) very close to the γ S group indicating that this might be a γS 'caught in the act' of evolving into a more γM -like isoform (Fig. 5).

A phylogram of ectothermic and largely obligate aquatic γ crystallins is presented (Fig. 5) as a Bayesian 50% majority rule consensus tree. The evolutionary model used was the Generalised Time Reversible plus Invariant plus Gamma (GTR+I+ Γ) model of substitution as chosen by MrModelTest (Nylander 2004). We used the coding regions of the nucleotide sequences

of the cDNAs (*sans* 5'-UTRs, 3'-UTRs and stop codons) aligning the nucleotide sequences based on the amino acid alignments (Suppl. Fig. 1). The outgroup choice for this tree was urochordate sea squirt *Ciona intestinalis* because it represents an $\beta\gamma$ sequence from a basal chordate and an aquatic marine

ectotherm. (Franck et al., 2004; Shimeld et al., 2005; Riyahi and Shimeld, 2007).

In the construction of the phylogram we included the toothfish cDNAs cloned and sequences from this study as well as the recently reported zebrafish γ crystallins (Wistow et al., 2005).

	1	0 20	30	40	50	60	70	8
a i a i								
			G-KITEVEL	DVEFGGRRLELEIS		-YLSRVGSTRVES	GCEMVYER	NSYMON
Dr YM1			G-KVIFYEI	ORNFQGRNYECMSI	DCSDISS-	YLGRVGSIRVES	GCFIVYER	NGFMGN
Ccx_71		R	G-KIIFYEI	DRNFQGRSYECMSI	DCSDITT-	YMSRCQSCRVES	GCFMVYER	PNFMGM
Tn_YM1		· R	G-KIIFYE	RNFQGRNYECMSI	DCSDMSY-	-YLNRCQSCRVES	GCFMVYER	PNFMGN
$Tn_{\gamma M3}$		· R	G-KIIFYEI	KNFQGRSYECMSI	DCSDMTS-	-YLNRCSSCRVES	GCFMVYDR	PNFMGN
$Tn_{\gamma M11}$ $Tn_{\gamma M13}$			G-KIIFYEL	PNFOGRSYETSSI	DCSELTS-	- YLSRCOSCRVES	GCEMVYDR	PNFMGN
Tn YM10			G-KIIFYEI	ORNFQGRSYETSSI	DCSDMTS-	YLSRCQSCRVES	GCVMVYER	PNFMGN
$Tn_{\gamma M12}$		MI	G-KIIFYE	RNFQGRSYECMSI	DCSDMSS-	YLSRCQSCKVES	GCFIVYDR	PNFMGN
Tn_YM9		·	G-KIIFYEI	ORNFQGRSYE <mark>TS</mark> SI	DCADMSS-	YLSRCHSCRVES	GCFMVYDR	TNYMGN
Dr_YM2b		MHG	KVIFYEI	DRNFQGRSYECMSI	DCADMSS-	-YLSRCHSCRIES	GCFMMYDR	PNYMGN
Dr VM2a		MHG	KVIFYEL	DRNFOGRSYECMSI		- YLSRCHSCRIES	GCEMMYDR	PNYMGN
Cc YM2			M-KVTFYEI	DRNFQGRSYDCMSI	DCADFSS-	YMSRCHSCRVHS	GCWMMYDO	PNYMGN
Am_y1			IFYEI	ORNEMGRSWECSGI	DCADMSS-	YMNRCHSFRVMS	GCWMFYDQ	риумсн
Cf_YM2-1		· R	G-KIIFYE	RNFTGRSWECSGI	DCPNMSS-	YLNRCYSCRVES	GFWMVYDR	PNFMGN
Cf_γM2-2		·	G-KVIFYEI	DRNFMGRSYECSSI	DCSDMSS-	YLSRCHSCKVER	GCWMVYDR	PNFMGN
CE_YS2			G-KVIFYEL	DRNFMGRSYECSSI		YLSRCHSCKVGR	GCWMVYDR	PNFMGN
$Tn \gamma M2$			S-KIIFYE	RNFOGRSYECKSI	DCPDMSS-	YLSRCHSCRVER	GCFVVYDR	TNYMGN
Dm YM8c		MSNTGMNMH	G-KITFYE	KNFQGRSYECMNI	DCSDMSS-	SLSRSQSCRVES	GCFMVYDR	NNYMGN
Dm_YM8d		MSNTGMNM	G-KITFYE	KNFQGRSYELMNI	DCSDMTS-	NLSRCQSCRVES	GCEMVYDR	SNYMGN
Dm_YM8b		MSNTDMNM	G-KIIFYE	KNFQGRSYECMNI	DCSDISS-	-YLSRCQSVRVES	GCFMVYDR	NNYMGN
Dm_YM8a		MSTTDMSM-	G-KIIFYE	RNFQGRSHECMSI		YLSRSQSCRVES	GCFMVYDR	NNYMGN
Dm_YM1 Dm_YM8e		MT	G-KINFIE	RNFOGRSYETSSI	DCSDMSS-	YLSRCUSCRVES	GCEMVYDH	NNYMGN
Dm YM3		MI	I G-KIIFYEI	KNFQGRSYETSSI	DCADMAS-	HLSRCHSCKVES	GYFMVYDR	TNYIGN
Dm_YM9		MT	G-KIIFYEI	DRNFQGSQYETSSI	DCPELTS-	YLTRCHSCRVES	GNFMVCDR	PNYMGN
Dr_YM3		· 1	G-KIIFYEI	ORNFQGRSYEC STI	DCADMST-	YLSRCNSCRVES	GCFVVYDR	PNFMGN
Cc_YM3		·	G-KIIFYEI	DRNFQGRSYECSSI	DCSDMST-	-YLSRCHSCRVES	GCFVVYDV	PNYMGN
Dr_YM4 Dm VM4		MD	G-RITFYER	KNFOGRSYETSNI	DOPELSQ-	YLSRCSSCRVEN	GLEMVYEK	PNFMGN
Tn YM5	MGIKVRVWLGF	VPEQSGKELDTR	G-RIIFYEI	KNFQGRSYETSSI	DCAELTS-	YLSRCNSCRVES	GCFMVYER	PNYMGH
Dr_YM5			G-KIIFYEI	ORNFQGRNFESSGI	DCPELTA-	YLSRCCSCRVES	GCFMVYEH	SNFIGH
$Dm_{\gamma M5}$		MS	G-KIVFFE	RNFQGRSYECMSI	DCSEITS-	HLGRCSSCRVES	GTFMVYDQ	PNFTGQ
Dr_YM6		MAI	G-KIIFEEI	DRNFQGRHHECSSI	DSADLHP-	-YFTQCNSIRVES	GCEMVYEH	PNYMGQ
$Dm_{\gamma M7}$			G-KIIFYEL	DRNFQGRSHECSSI	DSADLHS- DCADLHS-	-YFNRCNSIRVES	GCEMUYER	TNEMCE
Tn YM14			G-KIIFYEI	ORNFQGRSYECSS	ECSDLHS-	HFSRCNSIKVDS	GDWVVYEK	PNYMGY
Dm_YS1		MS	S-KITFYED	ORNFQGRSHECDTI	DCPDMHP-	HFSRCNSIKVES	GCWVLYEK	PNYTGY
Tn_YS3		MS	S-KIIFYED	ORNFQGRSYEC <mark>ATI</mark>	DCPDMHP	HFSRCNSIKVES	GCWVLYEK	TNYTGY
Dr_YMX			S-KITFYE	KNFQGRHYDCTGI	DCADMQS-	HFNRCNSIRVDS	CSWVAYEK	PNFSGY
$Tn_{\gamma MI5}$		MAYNHSAAT	G-KIIFYE	RNFQGRHWECNII	DOMDTFR-	-HFNCCNSIRVSG	CHWVAMSK	PFYMGY
Cp Y2			G-KIIFYEI	KNFLGRSYECSTE	ECADLTS-	YFSRCNSIRVES	GNWILYEH	PNFRGH
x1_yb		k	G-KIFFYEI	KNFQGRSYECNSI	DCSDLSS-	YFNRCNSIRVES	GNWILYEQ	PSYRGH
x1_γ3		· N	G-KIFFYEI	KNFQGRSYECNSI	DCSDLSS-	YFNRCNSIRVES	GNWILYEQ	PSYRGH
x1_γ1		·	1 G-KIFFYEI	DRNFQGRSYECSSE	ECSDLSS	YFNRCNSIKVDN	GNWILYEQ	PSYRGH
x1_γ4 x1_γ5			G-KINFYER G-KINFYER	RNFOGRCYECSSI	ECSDLSS-	-YFNRCNSTRVDG	GNWILYEO	PSYRCH
x1 γ2			G-KIIFYEI	ORNFQGRSYECNS	ECPDMSS-	-NFRRCNSIRVES	GDWILYEH	PNYRGH
Dr_YSa			G-RIIFFEI	KNFQGRRHECDSI	DCSDFHT-	YLNRCNSIRVES	GAWVVYER	PNFIGY
Dr_YSb			G-KIVFFEI	KNFQGRRYEC DSI	DCSDFHT-	YLSRCNSIRVES	GAWVVYER	PNYTGN
Ii_YS		MSK1	GNRITFYE	KNFQGRRYESDRI	DCLDFHT - ·	DLSRCNSIRVEG	GAWVVYER	PNFAGN
Tn YS2			G-KINFFEL	KNFOGRSYECATI	DTPDLRT-	FFSRCNSIKVOS	GCWLLYFH	PNYTGN
Dr YSc		MKLAKNMDR	G-KIVFYEI	ORNFQGRSFECSLI	DCPELSS-	HFTRCNSIRVEN	GAWVLYER	PNYMGF
Tn_YS1		MDV	G-KIVFYEI	KNFQGHSFECSSI	DCPELIT-	-HFRRCNSVQVES	GAWVLYER	PNYLGY
Dr_ySd		MA1	E-SIVFYEI	DRNFQGRSYECKGI	DTSDLHS-	-FF <mark>SRC</mark> NSARVKG	GFWVLYER	PNYMGY
Dm_YN		MSQYS	G-KIVFYE	KCHTGRKLEICSI		GFMNRVNSVRVES	GAYVCHDH	PDFKGQ
Dr YN2		MSQYS	G-KICFYE	RCFTGRCLIPVYC	DCDNFODR	GFMNRVNSIRVES	CAWICHDH	PDFKGQ
Tn_YN		MSQYS	G-KITFYE	KCFTGRKLEVRG	ECDNFQDW	GFM-KVNSIRVES	GAWICYDH	PDFKGO
II_YN		MSQYS	G-KIVFYE	KCETCKKLEISGE	PCDNFQER	GFTNRVNSICVQS	GAWVGFSH	ADFRGQ

Fig. 4. Amino acid alignment of the 14 distinct γ crystallins from the Antarctic toothfish (Dm) with 51 other γ crystallin sequences from ectothermic animals. The four major Greek–Key fold motifs are indicated in alternating blue/green bars at the bottom of the figure. Boundaries between the domains are in red (). The numbering (+1) of the crystallins begins where the first glycine (G) residue appears in the majority of sequences. The toothfish appears to have unique class of γ M8 crystallins which are absent in the other species. Identity (100%) between each species' γ crystallins is indicated by a black background, whereas similarity (\geq 50%) is indicated by grey background. Species are designated by a two (or three in the case of lipshark, i.e. *Chiloscyllium colax*=Ccx_ γ 1) letter abbreviation composed of *Genus species* name and one of the five different isoform classes (γ +#, M, S, N, MX).

	90	100	110	120	130	140	150	160
	·····				· · · · · · · · · · ·	. <u></u> .	[]	
ci_γ1	SYKLTPGKYPNPGSW-				GG	NDDE	LSSVKQQ	
Cc_YM1	QFFLRRGEYHDMQRM-	-MSMGMMFD	TIRSCRMI	PPYRGS	YRMRIYERDTFGG	QMHEVMDDC	ONIMERYRMSI)-WQSC
Dr_YM1	QFFLRRGEYHDIQRM-	-MSMGMMFD	TIRSCRMI	PSYRGS	FRMRIYERDNFGG	QMYELMDDCI	SIQERYRMSI)-CQSC
Ccx_71	QFFLRRGEYHDMQRM-		SIRSCR	YPYR-A	FRMRIYEREYFGG	QMSELMGDCI	DSIMDRFRMSI)-CMYC
In_YM1	QYFLRRGEYNDMORL -	-FSMGMMFD	SIRSCRII	PHHRGQ	FRMRIYERENFSG	QMHELMDDCI	DNIQDRYRMSI)-CLSS
In_YM3	QYFLRRGEYSDY	-ISFGMS-D	SIRSCRUI	PQi‡iRGT	YRIRIYERENFQG	QSQELMDDC	DNIQDRYRMSI)-CQSC
In_YM11	QYFMKRGEYADYMSM-	-MGLTGG	-IRSCRMI	PMitiRGQ	FRMRIYERENFSG	QMNEL TDDC	DSIQDRYRMSI)-CLSS
In_YMI3	QYFLRRGEYSDYMSM-	-MCMSAG	-IRSCRMI	PMIRCQ	FRMRIYERENFSG	QMNELTONC	DNIQDRYRMSI	J-CLSS
In_YMIO	QIRMRRGENADIMSM-	-MGMRDC	TRSCRMI	PMRRGQ	FRMRIYERENFGG		SN LODRYRMSI	
Th VM9	OFFWERGEVEDVO-H		TRSCRMI		FRMRIYERDINFSG			
Dr VM2b	OYFERKGDYADYMSM-	-FCMND	GTRSCRUT	PMHRGS	YRTKIYERENEGG.	OMVELMDDC	ONTMORYRMSE	
Dr VM2c	OYFFRKGDYADYMSM-	-FCMND	CIRSCRMI	PMIURGS	YRLKIYERENFGG	OMYELMDDC	ONIMORYRMSE	I-COSC
Dr YM2a	OYFFRRGDYADYMSM-	-FGMND	CIRSCRMI	PMHRGT	FRMRIYERENFGG	OMHEMMDDCI	ONIMORYRMSI	I-COSC
Cc YM2	QYFFRRGEYADYMSM-	-FGMSN	CIRSCRMI	PMHRGS	YRMRIYERENFMG	QMYEMADDCI	OSIMORYRMPH	I-CQSC
Am_Y1	QYCFRRGEYSDYMSM-	-W <mark>G</mark> SSS	WVRSCRMI	PRYSGH	YRMRMYERDNYMG-	QMMEMNDDC	onfmnrynwsi	IGCQSC
Cf_γM2-1	QYFLKKGEYNDYIGT-	-WGMNG	WIRSCRWI	PMHISGP	HKMRLYPRENFMG	QMMEMSDDC	DSFMDRYNRS H	IGCMSW
Cf_γM2-2	QYFIKRGEYPDYMSM-	-WGWGNN	CIRSCRMI	PMYRGS	YTTTMYRRENSMA	QMMDVIEDC	DSIMDRYHWSI	GCHSC
Cf_γS2	QFFIRRGEYPDYMSA-	-WGWGNN	CIRSCRMI	PMYRGA	YRMKIYERENFLG	QMMEISDDC	DSIMDRYRWSO	GCHSC
cf_γs1	QYFMRRGEYADYMSM-	-WGWGNN	CIRSCRMI	PMYKGS	YRMKVYERENFNG	SRQMDVMDDCI	OSFMORYHWS	ISFMSC
rn_γM2	QYFLRRGDYADYMSM-	-MCMSD	CIRSCRMI	PMHRGS	YRMRIYERENFGG	QMSELMEDCI	ONVMERYRMS	I-CMSC
Dm_YM8c	QYFMKRGEYSDYMSM-	-MGMRD	CIKSCRMI	PMHRGQ	FRMKIYEKENFDG	QSHDMMEDCI	ONIMORYRMNE	-CQSC
Dm_YM8d	QCFMKRGEYSDYMSM-	-MGMRD	CIKSCRMI	PMitiRGQ	YRMKIYEKENFGG		DNIMDRYRMNE	
Dm_YM8D	QYFMRRGEYSDYMSM-	-MGMRE		PMHIRGQ	FRMKIYERENFGG		IDERYRMSI	
Dm_yMoa	OFFMRRGEYSDIMSM-		TRACCAMI	PM HRGQ	FRMRIYERENFGG			
Dm VM8e	OVEWRKGEVSDYOR	Mems	CTKSCRMT	PM HRCO	FRMRIHEKENFGG		ONMORYHMNE	
Dm VM3	OYEMKRGEYSDYORM-	-MGFGD	CIRSCRMI	PMHKGS	YKLRIYERENFGG	OMNEVNEDCI	ONIODRYHMSI	
Dm YM9	OYFMKRGEYSDCMSM-	-MGMSD	CIRSCRUI	PMHRGS	YKMRIYEKENFGG	OMNEMNEDCI	ONIOERYSMSI	-CMSC
Dr YM3	OFFMRRGEYADYMRMO	MSD	GIRSCRVV	PQYRGP	YRMRIYERDNFGG	QMYEL TDDC	OSFMORYRISI	-cosc
Сс_үМЗ	QFFMRRGEYADYMRMG	MSD	GIRSCRMV	PQYRGP	YRMRIYERENFGG	QMYDLTDDC	OSFVDRYRMSI	-CQSC
Dr_YM4	QFFLRRGEYADCKRM	LSD	SIRSCRTI	PQHRGA	FRMRIYEKENFGG	MSYELTEDCI	ESTADRFRLSE	-MRSC
Dm_YM4	QMLVRRGEYPDNQRL-	-MGMT-TSD	CIRSSRMI	PMHKGP	FRMKIYEKENFGG	QMHELMDDC	ONMODRFRMNE	-CQSC
Γn_γM5	QMLVRRGEYPDNORL-	-MGMS-MSD	CIRSCRMI	PMHRGP	FRMRIYEKENFGG-	QMNELMDDCI	ONIQDRYRISI	-CQSA
Dr_YM5	QMLVRRGDYHDNKRI-	-MGMS-TSD	CIKSCKMI	PMHKGT	FRMRIFEKENFVG	QKYELMDDC	SIQERFYMSI)-CQSC
Dm_jM5	QYLLTRGEYPEYQNT -	-IGFNE	CIOSCRMV	PEHKGP	FKMRIYERANFEG	QMQELTDDC	DSIQDQYQMSI)-MQSC
Dr_YM6	QFFLRRGDYSDCORM-	-IGFSN	SIRSCRUI	PMYNGN	YKMRLYDQADMGG	-QMIEVTEDCI	PNIMORFHTSI	D-IHSC
Dm_YM7	OVEL DRCEVDDVMDT	- IGLN D	CVRSCRMI	PQBRGS	YKMRLYERSDMSG	OMHEMLDDC	NIQURLSMSI	-FNSC
$T_{\rm D} \gamma M14$	OVEL REGEVEDYORW-		CVRSCRMI	DM SKVA	HEMMLYERDERCC	RVMDLMEDVI	OSLVEHENST	
Dm VS1	OYWLTRGEYPDYORW-	-MGFND	TRSCRIFT	SY-TSECP	YRMRIYERPNFOG	OMMERSED	SVOENFCSH	D-TYSC
In VS3	OYVLTPGEYPDYOCW-	-MGYND	TICSCRTE	SY-TSECP	YRIRIYCRPNFOG	OMMEFSDDC	SVOLOFHSH	2-IYSC
Dr YMX	OYMLFKGEYPDFOHW-	-AGFND	CIRSCRVV	PAYTGN	YRMKIFERADFGG	OAMELNEDCI	DLRORFHNG	-ISSA
In YM15	QYILGPGEYPDYHSW-	-MGFNN	CVRSCOML	SPYRGS	YKMRIYNRPDLMG	NMMEFSDDC		-IYSC
Cp_y1	QYYLRRGEYPDFQHW-	-MGFND	SIRSCRLT	PQHRGS	YRIRVCERDNFGG	QMMEFSEDC	PHVYEQFRYN	-IHSC
Cp_γ2	QYYLRRGEYPDFQHW-	-MGFND	SIRSCRLT	PQHRGS	YRIRVYERDNFGG	QMMEFSEDC	PHVYEQFRYN	D-IHSC
кі_үв	QYYLWKGEYPDFQRW-	-MGFND	SIRSCRMS	PYHQGQ	YKMRIYERGDFQG	QNMEFFEDC	PNTYDRFRFR	D-IHSC
x1_γ3	QYYLWKGEYPDFQRW-	-MGFND	SIRSCRMS	PYVSHQGQ	YKMRISERGDFQG	QNMEFFEDC	PNTYDRFRFR	-IHSC
x1_γ1	QYYLWKGEYPDFQRW-	-MGFND	SIRSCRMS	PY∺QGQ	YKMRIYERGDFQG	QMMEFFDDC	PNTYDRFRFH)-IHSC
x1_γ4	QYYLWQGEYPDFQRW-	-MGFND	NIRSCRFI	PQHNGQ	YKMRIYEKGDYQG	QMMEFFDDC	PNTYDRFRFH	D-IHSC
x1_γ5	QYYLWKGEYPDFORW-	-MGFNE	YIRSCRFI	PQitiHGQ	YKMRIYEKGEFQG	-QMMBFSDDC	PNTYDRFNFR	D-IHSC
x1_γ2	QYYLRRGEYPDFQQW-	-MGFND	CIKSCRLS	PQHQGS	FRMKIYEREDFRG	-QMMEFTEDC	PNVYERFNFQ)-IHSC
Dr_ysa	DIVLIRGENPUNCKW-		CLOSCENT	HF-VSGSE	HKIQLYDKGDGAG		SVVERFRIRE	
ti vs	MYVITHCEVPEYORW-	MGLN	RI.SSCRIT	OL-SCCAR	YOTOFYEKCDECC.	OMVESTEDC	SVOLAFALKI	
Dm VS2	OYILSSGEVPDHOOW	-MGEND	STKSCRST	ON-VYCKS	WKIRFYPNKDPEC	OAAECVEDC	SVYETEKEOF	-VHSS
In VS2	OYILSHGEVPDHOOW-	-MGFND	SIKSCRAT	KN-VYGKS	WKIRFYDKODEGG	OTAFCVVDC	SVYETIKLE	- FHSC
Dr YSc	OYILTRGEYPDYORW-	-MGYND	TIRSCRMV	RN-HTG-S	FRIRLYPRPDPOG	OTMPSSEDWI	SLYDRFRORE	-VHSC
In YS1	QYVLTOGEYPDYOHW-	-MGYND	SVRSCRLT	RN-TSS-V	FKIRVYERPDFSG	OMLESTEDLI	RDLADYWHRH	-VHSA
Dr YSd	QYILGPGEYPDYOHW-	-IGFND	CVRSCRLV	RH-VIG-D	LKLKLFERPNFDG	QTWEVTEST	SIQERFLCR	- INSC
Dm_γN	QYILEHGEYPEFORWN	AHND	HMGSCKPI	RMHGEH	YRMELFEGENFSG-	OCVELCEDC	FLQARGLTKS	SCVNSV
Dr_YN1	QYMLEKGEYPDFORWN	IGHND	HMGSCKPI	KMHGEQ	YRMELFEGQNFTG-	QCVELCDDC	FLQSTGFSK	ICLNSI
Dr_YN2	QYILERGEYPEFQRWN	SHND	HMGSCRPI	RMHGEQ	YRMELFEGCNYTG	QCMELCDDC	PFLQSRGFNT	ICVNSV
In_YN	QYILEHGEYPEFQRWN	ISHND	HMGSCKPI	RMHGEH	YRIELFDACNFSG	QCVEICDDT	FLQSRGLSKN	ICINSV
ΙΊ γΝ	OYILEHGEY PSFYRWN	ICHND	RMGSCKPV	GMHGEH	YRIELYEGKYESG	RSOFTKDC	FLSRQGWAKN	IMINAI

A.J. Kiss, C.-H.C. Cheng / Comparative Biochemistry and Physiology, Part D 3 (2008) 155-171

Fig. 4 (continued).

In addition, we included a number of other fishes and ectotherms, specifically; *C. fuscus*, whitespotted clarias; *Chiloscyllium colax*, lipshark; *Astyanax mexicanus*, blind Mexican cavefish; *C. carpio*, Asian (common) carp, *Xenopus laevis*, African clawed frog; *Cynops pyrrhogaster*, Japanese firebelly newt and *Iguana iguana*, iguana (Suppl. Table 2). Furthermore, there are 15 previously 'unidentified proteins' from *T*. *nigroviridis* (spotted green pufferfish) used in this analysis which were deposited into the GenBank/NCBI database by Genoscope (France) and the Whitehead Institute (Massachusetts Institute of Technology) during the course of the *Tetraodon* genome project (www.genoscope.cns.fr/externe/English/ Projets/Projet_C/organisme_C.html). We have identified these 15 sequences as γ crystallins and thus propose isoforms names

		170	180	190 	200	210	22	20
Ci_γ1		<u> </u> -				Í		
Cc_YM1	HVM-D	GHWLFYEQP	HYRGRMWYFRP	GEYRSFRDMG-Y	78 NMRFMS	SMRRI	IDMC	
Dr_YM1	HVM-D	GHWLMFEQP	HYRGRMIYFRP	GEYRSFRDMG-F CEYRSFRDMG-N	SNMRFIS	MRRI		
Tn VM1		GHWLUYEOP	HERGRMMYLRP	GEYRSFRDMG-F GEYRSFRDOG-F	SGMRFM	SMRRI	IDIC	
Γn γM3	NVL-D	GHWLLYEQP	HFRGRMMYLRP	GEYRSIRDVG-H	SGMRLS	SIRRI	MDS	
rn_γM11	QVM-D	GHWLCTSSP			SGAG			
In_YM13	QVM-D	GHWLLYEQP	HFRGRMIYLRP	GEYRSFRDLG-V	78 GMRFM	SMRRI	MDMC	
In_γM10		GHWLLYEQP	QYRGSMIYLRP(GEVRNERDLG-V	/SSSRFMS	SMRRI I		
In_γMI2 Γn γM9		GHWLUYEOP		GEYRSFRDLG-F GEYRSFRDLG-V	79 GMREMS 7N GMRVMS	SMRRI	MDPCS	
Dr YM2b	⊻ HVM-D	GHWLMYEQP	OYRGRMMYLRP	GEYRSFREN	GGTRFM	MRRI	IDSMY	
Dr_YM2c	HVM-D	GHWLMYEQP	QYRGRMMYLRP(GEYRSFREN	IGGT <mark>R</mark> FM	SMRRI	IDSMY	
Dr_YM2a	HVM-E	GHWLVYEQP	QYRGRMMYMRP(GEYRSPREN	IGGMRFL	SMRRI	NDSFY	
Cc_γM2	HVM-D	GHWLMYEQP	HYRGRMWYFRP	GEYRSESNN	IGGMRFM	SMRRI I	MDSWY	
Am_γ⊥ ີf vM2-1	PRD-G	CHWLMYEEP	NYMGRMWYFGPO	GHYRNISSNYG Covrsyrhmmgn	SCMRFOS			
$Cf \gamma M2 - 2$	HVM-D	GHWLMYEOP	HYRGRMWYFRP	GEYRSFRETMGN	18 GMRFMI	RMRRI	MDSWY	
cf_γs2	HVT-G	GHWLMYEQP	HYRGRMWYFRP	GEYRSFRDF	GNTNFM	MRRI	MA	
Cf_γs1	NVM-D	AHWLMYEHP	HYRGRMWYFRP	GEYRNFRD Y	(G GMRFM	SMRRM	VA	
Γn_γM2	NVM-E	GHWLMYEQP	HYRGRMMYVRP	GEYRNFMS-TIC	SNMRVIS	SMRRI	rDSCQ	
Dm_YM8c	NVM-D NVM-D	GHWLMYEQP	QFRGKMMYMRP(OFRCKMMYMRP)	CEWKNEMDMC - N	ISCORFMS	MRRI		
Dm YM8b	NVM-D	GHWLMYEOP		GEYRSFREMG-M	SGURFMS		IDSCN	
Dm YM8a	NVM-D	GHWLMYEQA	QFRGKMMYMRP	GEYRNFREMG-M	ISGQRFM	SMRRI	rDMC	
Dm_YM1	NVM-E	GHWLMYEQP	QFRGKMMYMRP	GEYRNFKDMG-M	18 GQRFM	MRRI	rdMC	
Dm_γM8e	NVM-D	GHWLMYEQP	QYKGKMMYLRP	GEYKSFRDMG-N	18 GQRFM	SMKRI	MDSCY	
Dm_γM3	NVM-D	GHWLMYEQP	QFRGKMMYLRP	GEYKSERDMG-Y	DAMRFS	SIRRI	IDSC	
Dr VM3	NVM-D HVM-D	GHWLMYEOP	HYRGRMTYFRP	GEYRSERDMG-Y	15 GORFMS	SVRRT	VDLC	
Cc YM3	HVM-D	GHWLMYEQP	HYRGRIVYFRP	GEYRSFRDMG - Y	SNVKFS	SVRRI 1	MDLC	
Dr_YM4	HVM-D	GHWLLYEQP	HYRSRMLYLRP	GEYRSFRDMG-1	ISFRFS	SLRRI	MEPCN	
Dm_YM4	NVM-E	GHWLMFEQP	NYRGKMMYLRP	GEYRNLRETG-M	SNMTK-FS	SMRRI	MDSC	
In_γM5	HVM-D	GHWLLYEQP	HFRGRMMYLRP	GEYRSMRDMG-M	GPMDMRIG	SIRRI I	MDSC	
Dr_YM5	NVT-H	GHWLLYERP	NFEGRMINIRPO	GEMSTESDING-I CRIVENT.KEMN_S	GSLKIAS	SVRRI	MESC	
Dr YM6		GHWLLIYEOP	NYRGRMEYLGP	GEYRKYSDWCGN	APRIG		ISFN	
Dm_γM7	NVM-E	GHWLLYDQP	NYKGRAYYMRP	GEYRRFSDWGGV	/SPRVG	SIRRI	SDLN	
Dr_YM7	HVM-D	GHWLVYEQP	NYTGRQFYLRP	GEYRSYNDWGGV	7 T SRMG S	SIRRI	rDL	
In_YM14	NVI-G	GHWI FYEHP	HYRGRQYLMGP	GQYRRFNEWGSI	SPRVG	SIRHI	VC	
Dm_YS1	NVM-E	GYWTLYEQP	NYRGROYFMRP	GEYRKFSDWGAI	CATTG	SFRRI	ref	
Dr VMX	NVM-D	GYWTLHEHP	NYTGROFFLER	GEVERYTEWCSC	SPTIG		TDPK	
rn γM15	NIL-E	GFWIFYEHP	NYKGROYFLRP	GEYRACGDWGCH	INPMVG	FRRI	RTLM	
Cp_γ1	NEQ-D	GHWVFYEEP	NYRGR <mark>QY</mark> YLRP(GEYRRYSDWGAS	SSPKVG	SFRRV	RDLY	
Cp_γ2	NVQ-D	GHWVFYEEP	NYRGR <mark>QY</mark> YLRP(GEYRRYSDWGAS	SSPKVG	SFRRV	RDLY	
xl_γb	NVF-D	GNWMFYEEP	NYRGR <mark>QY</mark> YLRP	GEYRRYSDWGAS	SSARIG	SFRRV	HHMF	
κ⊥_γ3 κ1_γ1		GNWMFYEEP	NYRGROYYLRP	GEYRRYSDWGAS	SSARIG	SFRRV .	HHIF	
x1 $\gamma 4$	NVF-D	GHWMFYEEP	NYRGROYYLRP	GEYRKYSDWGAS	SSPRIG	SFRRV	YHKFKSTQTF	TNTFVLVO
x1_γ5	NVS-D	GHWMFYEEP	NYRGRQYYLRH	GEYRRFSDWGAS	SARIG	SFRRV	HHMF	
x1_γ2	QVL-D	GYWMFYEEP	NYRGR <mark>QY</mark> YLRP	GEYRRYTDWGA1	INPRIG	SFRHV	YHR	
Dr_YSa	KVL-D	GIWIFYEHP	NYRGRQYLLQK	GEYRKPVDWGAV	/CPTVQ	SFKRL	re	
Dr_YSb ti vs	KWL-D	GYWIFYEHP	NYRGHQYFLEK	ENYRKPVDWEAI Gevisk Dmewica z	CPSVQ	SFRRF	re	
Dm YS2	VVM-D	GAWVLYEOP	NYCGHOYFLER	GEYNNY TDWGAT	SPAVG	SFRMI	rkf	
Γn_γS2	VVM-D	GAWVLYEQP	NYHGHQYFLER	GEYHNYTDWGAT	SPAAG	SFRMI	rDF	
Dr_ysc	NVL-D	GAWIFFEHP	NYRGRQYLLEK	GEYRCFTDWNAM	HPTVG	SIRRI	QDF	
Fn_γS1	QVL-D	GAWVFHELS	NFHGRQYLLEK	GKYRRFTEWAAM	IN PKVG	FRCA	V	
Dr_ysd	KWH-E	APCVFFDHA	NYRGRQYFLEK	GEYRRHTEWCAN RNVSTHTEWOAR	HPTVG		1"1"D	
Dr YN1	KVYGD	GAMAMYDEP	NYRGRMYIVER	CNYCADTEWOSE	SNPSIQ	SIRRV	VNYF	
Dr_YN2	RVFGD	GAWVMYEEP	NFRGRMYIVER	CNYCGHNEWQAQ	NPHIQ	SIRRI	VNYF	
Γn_γN	KVFGD	GAWVMYEEP	NFRGRMYVVER	GDYCSHNEWQAQ	QNPNIQ	SIRRV	VNYF	
ii_γN	KWYGD	GAWVLYEEP	NYRGQMYVVER	GDYSSCNEWQAS	SNANIQ	IRRV	INYF	

Fig. 4 (continued).

based on the γM nomenclature used in this study in Suppl. Table 1.

The γ crystallins can be divided into three major isoform groups (γN , γS , and γM which includes amphibian γ) largely based on the phylogram (Fig. 5). The γN isoforms form a distinct clade (green, Fig. 5) indicating their ancestral position and high conservation across species. Most of the γ S isoforms (blue, Fig. 5) are basal to the γ M isoforms (purple, Fig. 5) with exception of the γ S1 and γ S2 from *C. fuscus* whitespotted clarias. The γ crystallins included from the amphibians (*X. laevis* and *C. pyrrhogaster*) form a clade nested within the γ M crystallins of the fishes (red, Fig. 5). From the Bayesian tree

165

Ammo acid identi	mino acu identities (percent identical) based upon pan wise anginitent of toburnsh y crystallins													
Sequence name	$\gamma M1$	γM3	γM4	$\gamma M5$	$\gamma M7$	γM8a	γM8b	γM8c	γM8d	γM8e	γM9	$\gamma S1$	γ S2	γN
γM1	100.0	69.8	69.6	56.9	58.4	76.8	73.2	76.3	75.9	76.7	69.6	51.1	42.7	49.6
γM3		100.0	64.4	61.9	61.3	69.9	69.0	70.4	72.2	75.2	76.4	53.6	45.7	50.2
γM4			100.0	61.1	55.0	63.4	65.2	65.5	66.8	71.2	66.2	49.1	40.8	53.2
γM5				100.0	53.9	55.1	57.6	55.1	54.3	55.0	55.6	51.4	42.3	51.9
γM7					100.0	55.7	55.4	53.0	52.1	58.4	60.1	58.1	50.8	53.9
γM8a						100.0	87.5	86.3	83.6	73.3	72.8	47.8	40.7	49.9
γM8b							100.0	85.3	83.6	73.9	71.7	48.1	41.0	50.1
γM8c								100.0	92.3	73.9	69.5	47.8	41.3	49.1
γM8d									100.0	76.0	72.8	47.0	41.0	48.9
γM8e										100.0	75.2	48.0	42.4	50.5
γM9											100.0	48.6	41.8	52.7
γS1												100.0	59.6	51.9
yS2													100.0	49.0
γN														100.0

(Fig. 5) the phylogenetic distance (expected changes per site) between the γ crystallin cDNA coding sequences can be inferred from the branch lengths in the tree (Fig. 5), thus it appears that the amphibian γ crystallins are more closely related to the γ S isoforms than to some of the more derived γ M fish crystallins. The remaining γ M crystallins show clustering based on their relatedness to the previously named carp and zebrafish sequences.

A Southern blot of *Eco*RI digested toothfish genomic DNA probed with a conserved centre section (407 nt of γ M8e) of a toothfish γ crystallin revealing 17 bands (Fig. 6). The number of distinct bands in the Southern blot is within the magnitude of number of γ cDNAs isolated and sequenced.

3.4. Genomic organisation

Table 4

A BAC library macro-array was screened with ³²P-labeled toothfish α (α A and α B), β (β A2 and β B2) and γ (γ M8e) crystallin probes. Seventeen BAC a clones were initially identified from the macro-array and further re-analysed by HindIII digest Southern blot (Suppl. Fig. 2). Thirteen of these clones (a1, a3~a11, a13~a15, a 17) were found to be overlapping and form one contig 46 consensus band CB units long (Fig. 7A). Each CB unit was estimated to be ~5800 bp long based on the mean restriction fragment sizes in the HindIII DNA fingerprint analysis (Suppl. Figs. 2, 3 and 4), therefore 46 CB units would cover a region of 266.8 kbp (Castellarin et al., 2006). Clones a12, a16 did not overlap with the single contig. Clone a2 was also not in the contig and was not analysed further as it had only two bands at ~8000 and 5700 bp in the HindIII DNA fingerprint and very weak Southern Blot hybridisation (Suppl. Fig. 2). Southern blot analysis of the other 16 α clones indicated strong hybridisation for all the clones except weak hybridisation for clone a16. Hybridisation patterns were similar between the clones, except for clones a5 and a12 which also had different HindIII DNA fingerprint patterns (Suppl. Fig. 2). Only clone a12 showed both a different fingerprint/hybridisation pattern than that of the other α BAC clones.

Six β crystallin clones identified in the initial macro-array screen had extremely similar *Hin*dIII fingerprints and almost identical hybridisation patterns by Southern blot analysis (Suppl.

Fig. 3). These data were confirmed in FPC analysis which placed all the β BAC clones within the same 47 CB (272.6 kbp) contig (Fig. 7B).

There were twenty seven positive clones found for γ crystallins in the BAC macro-array screen. Based on our FPC analysis the γ crystallin BAC clones divide up into three groups. The first group forms a contig of 31 CB units (179.8 kbp) and includes clones g1, g3-5, and g12 (Fig. 7C and Suppl. Fig. 4) and have similar Southern Blot hybridisation (Suppl. Fig. 4). The second group forms a separate contig of 51 CB units (295.8 kbp) and includes clones g6-11, g13-20 and g22 (Fig. 7D and Suppl. Fig. 4). All of these clones within this second contig group also have a similar Southern Blot hybridisation, which differs from the first γ contig group. Two clones that did not contig with either of the two previous groups (g2 and g21, Suppl. Fig. 4) also shared similar Southern Blot hybridisation patterns to each other. These two clones failed to contig even after re-analysis with adjusted Tolerance and Cut-Off parameters in the FPC analysis program (Soderlund et al., 2000).

4. Discussion

The Antarctic toothfish *D. mawsoni* lives in the perennially sub-zero seawater (-2° C) of the Southern Ocean at the lower thermal limit of marine vertebrate ectotherms. At this subzero temperature, the toothfish has a completely transparent lens composed of α , β and γ lens crystallins that are similar to other mammalian vertebrates (Kiss et al., 2004). To investigate the possible molecular basis of the cold stable toothfish lens, we have obtained the crystallin sequences by cDNA cloning and sequencing.

There were two complete α crystallin (α A and α B) cDNAs isolated from toothfish lens. Toothfish α A crystallin show high sequence identity to α A from both fishes and mammals in contrast to α B sequences which are more divergent (Table 2). The homology of the toothfish α A amino acid sequence with other vertebrates suggests that α A crystallin is under greater functional constraint than the α B isoform, therefore maybe acting as the primary sHSP within the eye lens as has been previously hypothesized (Bova et al., 1997; Rajaraman et al.,





Fig. 6. Genomic Southern blot of *D. mawsoni Eco*RI digested DNA. DNA was probed with 407 nucleotide ³²P-labeled PCR amplicon from the coding region of γ M8e (*D. mawsoni*). Labeled arrows to the left of the blot indicate bands that hybridised to the toothfish DNA. There were 17 positive bands.

2001; Dahlman et al., 2005). The converse, that α B likely has a diversity of non-lenticular roles (Bennardini et al., 1992; Piatigorsky 1998; Dahlman et al., 2005) is also supported by our toothfish α B crystallin sequence data. Interestingly, we had some difficultly obtaining α B cDNA from the toothfish lens, which may reflect the low expression levels of the α B message (mRNA). Proteomic analysis of α A and α B crystallins from toothfish as well as detailed chaperone-like assays of recombinant toothfish α A and α B crystallins is in progress and will be published elsewhere.

Beta crystallin protein sequences from toothfish lens are aligned with zebrafish, cow and human (Fig. 3). They display a high degree of conservation both for each β isoform and between each species analysed. We did not isolate a BA3 isoform, which in mammals has a longer N-terminal extension that normally 'incorporates' the BA1 isoform by means of an earlier, upstream translational start codon (Bloemendal et al., 2004). There are three basic β crystallin isoforms (β B1, β B2, β B3) in toothfish, as in mammals. The 'A or 'B' designation after the ' β ' referring to either an acidic (βA #) or basic (βB #) class of the β crystallin isoform (Bloemendal et al., 2004). The N-terminal PAPA-arm of β B1 (Fig. 3) of cow and human lenses has been shown to be strongly associated with the membrane (aqueous insoluble) portion of the lens (Bloemendal et al., 1984; Hejtmancik et al., 1986; Coop et al., 1998; Bateman et al., 2001). We did observe a significant proportion of β crystallins associate with the membrane (insoluble) component of the toothfish lens during isolation of the lens proteins (Kiss, 2005). Although the N-terminal extension in the toothfish BB1 does not have a PAPA-arm, there is a long PAPA-arm in the Cterminal of the toothfish β B3 crystallin. Interestingly, this toothfish C-terminal PAPA-arm appears to have a counterpart C-terminal 'PNPN-arm' in zebrafish BB3 crystallin (Fig. 3). Although the hydrophobic PAPA-arm of bovine BB1 is thought to insert itself into the lipid membrane, the substitution of asparagine (N) for alanine (A) in the zebrafish PNPN-arm would make the PNPN-arm very polar and thus its function could be quite different than what is hypothesized for the mammalian and possibly toothfish PAPA-arms.

Recent structural studies of recombinant mammalian β crystallins (Bateman et al., 2001; Bateman et al., 2003; Van Montfort et al., 2003) have suggested that linker regions as well as the extensions (N- and C-terminal) are involved in oligomerisation of the β crystallins. While it is clear that within recombinant experiments this is the case, it remains unclear what are the native function(s) of N-terminal and C-terminal extensions in an intact lens.

The 14 distinct γ crystallin isoforms found in the Antarctic toothfish lens fall into three isoform groups. We found a single γN isoform, which based on gene structure is believed to be an evolutionary bridge between the β and γ isoforms (Wistow et al., 2005). Two more cDNAs belonged to the γS (formerly β_S (Björk, 1961)) group and the remaining eleven were γM isoforms (Chang et al., 1988; Wistow et al., 2005) so named originally for their high methionine content. The very high methionine content in the γM crystallins of the toothfish and other fishes is a curious adaptation (Chang et al., 1988), and so far remains unexplained. Unlike other non-polar residues, methionine has special properties which may in fact contribute significantly to the stability of fish lenses, both in terms of their

Fig. 5. Phylogram of ectothermic γ crystallins from the Antarctic toothfish, zebrafish, pufferfish and select other fishes, amphibians and iguana. Tree was constructed from the aligned nucleotides of the coding region from the cDNAs (Suppl. Fig. 1). The 50% majority-rule consensus tree was generated by the program Mrbayes, implementing a GTR+I+ Γ evolutionary model. Each major isoform class in the tree is colour coded (γ N=green, γ S= blue, amphibian γ = red, γ M= purple). Only the γ N isoform is monophyletic. The cross-species conserved γ S isoforms are polyphyletic in this analysis. The amphibian γ crystallins (red) form a clade nested basally within the γ M isoform group, close to the γ S isoform class indicating that they may be intermediate between the ancestral γ S and the fish γ M isoforms. Numbers at branch points are Bayesian consensus values (similar to bootstrap values) and the scale bar indicates 0.2 expected changes per site. Images of species used are included (but are not to relative scale).



Fig. 7. Genomic organisation of the α , β and γ lens crystallin genes from the Antarctic toothfish as assayed by BAC library analysis. (A) The single contig of the α crystallin gene region found to be 46 CB units (266.2 kbp) long. All but three clones (a7 and a12) were part of the contig. (B) Single contig region of β crystallin genes inclusive of all positive BAC clones was 47 CB units (272.6 kbp) long. (C) Two contigs of 31 CB units (contig group 1; 179.8 kbp) and 51 CB units (contig group 2; 295.8 kbp) encompassed all but g2 and g21 γ crystallin clones. Designations following the clone name indicate that: '*' clone was buried, '=' same bands as a parent clone, '~' approximately same bands as a parent clone.

high density and in terms of their ability to readily adapt to a range of temperatures. Methionine has both a structural plasticity as well as the ability to have its hydrophobicity altered by reversible oxidation of its thioether (Gellman, 1991). In the context of the extremely dense fish lens where protein concentrations are upwards of 1000 mg mL⁻¹ (Kroger et al., 1994), the multiple isoforms of γM crystallins would undoubtedly contribute to the protein stability by increasing the crystallin polydispersity, thus preventing catastrophic crystallization at such high concentrations. Furthermore, γM crystallins would impart 'flexibility', or sponginess to the yM crystallin surface which would further discourage crystal lattice formation. Additionally, the abundance of methionines may have biochemically predisposed the vM crystallins to cold adaptation by allowing the fish lens to reversibly oxidize their thioethers via enzymatic means (Marchetti et al., 2005; Sagher et al., 2006) thereby affording a mechanism to attenuate the hydrophobicity of these γ crystallins. Consequently, a primary adaptation of toothfish lenses by increasing the protein stability at high density by generating polydisperse γM crystallin isoforms, could have leant itself to a secondary adaptation of cold stability.

Phylogenetic analysis of the γ crystallin isoforms was done using a Bayesian approach employing a GTR+I+ Γ model of molecular evolution (Waddell and Stell, 1997). This model of evolution is particularly well-suited for ancient protein coding sequences as it acknowledges that (i) there has been reversion, (ii) that there are invariant sites and (iii) that not all sites are under the same selective pressures. Furthermore, using the nucleotides that code for the amino acids aligned as codon units may facilitate greater resolution depth in the tree (Simmons et al., 2002). Phylogenetic reconstruction of the toothfish γ

crystallins in comparison with other ectothermic species shows that the γN isoforms form a monophyletic clade, whereas conserved yS crystallins are polyphyletic. The placement of the recently discovered γN class of crystallins ancestrally to the γS clade suggests that they pre-date the evolution of γS , as has been suggested from comparative analysis of their gene structure in D. rerio (Wistow et al., 2005). Based on phylogenetic analysis, there is a range of γS isoforms with a decreasing gradient of similarity from γN to γM . The two C. fuscus γ S isoforms (Cf_ γ S1 and Cf_ γ S2, Figs. 4 and 5) could very well be misnamed as they seem to be well-situated with the γ M group (specifically γ M2) in both our alignment and Bayesian phylogenetic analysis. The multiplicity of the toothfish γM crystallin isoforms are emphasized in the phylogram, many of which are distinct from the zebrafish and pufferfish (Fig. 5). In our previous study (Kiss et al., 2004), we proposed that the γ crystallin component of the toothfish lens as the most likely candidate for the transparency of the toothfish lens at -2° C. However, while there are several yM crystallins isoforms unique to the toothfish, comparative hydrophobicity plots did not suggest major differences from other fish crystallins. It is possible that a few select amino acid changes, in addition to post-translational modifications of the γ crystallins could have profound implications on the cold stability of the lens. Regions between the domains of the γ crystallins termed 'linker' regions, and N-terminal extensions (such as the long methionine rich ones in the γM class) may also have significant influence on cold stability by affecting the solution dynamics (Wu et al., 2005). To fully address these issues, extensive biochemical analysis and proteomics experiments would likely be informative.

An outline of the genomic organisation of the α , β and γ crystallin genes was obtained by screening a BAC library and analyses of the DNA fingerprints by FPC. Using this approach we found that α and β crystallin separately formed a single α or β contig each whereas the γ genes formed two contigs (Fig. 7). The two γ contigs differ both in *Hin*dIII fingerprint pattern as well as Southern blot hybridisation patterns. The second γ contig appears to have a Southern blot hybridisation pattern similar to that of the β contig Southern blot (Suppl. Fig. 3). This similarity is not unexpected as both β and γ crystalling share conserved domains and are members of the same gene superfamily (Bloemendal et al., 2004; Wistow et al., 2005). The probe used to hybridize to the γ gene regions did have variable hybridisation strength to some BAC clones, of which most are represented in the second contig. Consequently, these results suggest that some β and γ genes are in spatial proximity, or alternatively that the second contig of γ BAC clones could very well be β crystallins. Two γ clones (g2 and g21) do not form a contiguous group with either of the γ gene contigs even though their Southern blot hybridisation pattern is similar to the β and second γ gene contig groups. Within mouse, rat and humans, the γ crystallin genes are clustered as $\gamma A \sim F$ with $\gamma S (\beta_S)$ a short distance apart, interspersed with highly repetitive sequences (Willard et al., 1985; den Dunnen et al., 1987; Skow et al., 1988; den Dunnen et al., 1989). Based on annotation in NEIBank, zebrafish vS isoforms are present on two different chromosomes (16 and 9) (Wistow, 2002). Without more information at the DNA sequence level, it is difficult to state with certainty how the $\beta\gamma$ gene organisation in the Antarctic toothfish is arranged. Current and proposed BAC projects argue persuasively for the sequencing of the Antarctic toothfish which will enable questions regarding γ crystallin evolution to be addressed more fully (NRC, 2003; Clark et al., 2004; Cheng et al., 2007).

Comparison of toothfish $\alpha\beta\gamma$ crystallin sequences with other vertebrates indicates that α and β isoforms are well-conserved both in number and isoform type. Our current work illustrates the homology of α and β toothfish crystallin sequences to other phylogenetically distant vertebrate species (Bloemendal et al., 1984; Hejtmancik et al., 1986; Behrens et al., 1998; Chen et al., 2001; Runkle et al., 2002) reiterating the likely conservation of function of these two types of crystallins. Conversely, toothfish γ crystallins exist as multiple polydisperse isoforms, some of which appear to be paralogues and some unique. Given this nature of the toothfish γ crystallins, it is attractive to suggest that they are more evolutionarily plastic without the functional constraints of the $\alpha\beta$ isoforms. Thus, the combination of an increased methionine content allowing for attenuation of protein hydrophobicity, along with increased polydispersity of γM isoforms in the dense eye lens may have predisposed the toothfish lens for biochemical adaptations to the extreme cold.

Acknowledgments

The authors wish to thank Prof. Mason Posner (Dept. Biology, Ashland University, Ashland, Ohio, USA) for suggestions on redesigning of the toothfish αB isoform specific primers (DRAB-1 and DRAB1-2), as well as helpful discussions during the preparation of this manuscript. In addition the authors thank the anonymous reviewers for their helpful comments and critiques. This work was supported by NSF grant OPP 02-31006 to C-H.C.C. and A.L.D.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.cbd.2008.02.002.

References

- Banh, A., Sivak, J.G., 2004. Laser scanning analysis of cold cataract in young and old bovine lenses. Mol. Vis. 10, 144–147.
- Bateman, O.A., Lubsen, N.H., Slingsby, C., 2001. Association behaviour of human betaB1-crystallin and its truncated forms. Exp. Eye Res. 73, 321–331.
- Bateman, O.A., Sarra, R., van Genesen, S.T., Kappe, G., Lubsen, N.H., Slingsby, C., 2003. The stability of human acidic beta-crystallin oligomers and hetero-oligomers. Exp. Eye Res. 77, 409–422.
- Behan-Martin, M.K., Jones, G.R., Bowler, K., Cossins, A.R., 1993. A near perfect temperature adaptation of bilayer order in vertebrate brain membranes. Biochim. Biophys. Acta 1151, 216–222.
- Behrens, M., Wilkens, H., Schmale, H., 1998. Cloning of the alphaA-crystallin genes of a blind cave form and the epigean form of *Astyanax fasciatus*: a comparative analysis of structure, expression and evolutionary conservation. Gene 216, 319–326.
- Benedek, G.B., 1997. Cataract as a protein condensation disease: the Proctor Lecture. Invest. Ophthalmol. Vis. Sci. 38, 1911–1921.

Author's personal copy

A.J. Kiss, C.-H.C. Cheng / Comparative Biochemistry and Physiology, Part D 3 (2008) 155-171

- Bennardini, F., Wrzosek, A., Chiesi, M., 1992. Alpha B-crystallin in cardiac tissue. Association with actin and desmin filaments. Circ. Res. 71, 288–294.
- Björk, I., 1961. Studies on g-crystallin from calf lens: I. Isolation by gel filtration. Exp. Eye Res. 1, 145–154.
- Bloemendal, H., Berbers, G.A., De Jong, W.W., Ramaekers, F.C., Vermorken, A.J., Dunia, I., Benedetti, E.L., 1984. Interaction of crystallins with the cytoskeletal-plasma membrane complex of the bovine lens. Ciba Found. Symp. 106, 177–190.
- Bloemendal, H., De Jong, W., Jaenicke, R., Lubsen, N.H., Slingsby, C., Tardieu, A., 2004. Ageing and vision: structure, stability and function of lens crystallins. Prog. Biophys. Mol. Biol. 86, 407–485.
- Bova, M.P., Ding, L.L., Horwitz, J., Fung, B.K., 1997. Subunit exchange of alphaA-crystallin. J Biol Chem 272, 29511–29517.
- Broide, M.L., Berland, C.R., Pande, J., Ogun, O.O., Benedek, G.B., 1991. Binary-liquid phase separation of lens protein solutions. Proc. Natl. Acad. Sci. U. S. A. 88, 5660–5664.
- Caspers, G.J., Leunissen, J.A., de Jong, W.W., 1995. The expanding small heatshock protein family, and structure predictions of the conserved "alphacrystallin domain". J. Mol. Evol. 40, 238–248.
- Castellarin, S.D., Di Gaspero, G., Marconi, R., Nonis, A., Peterlunger, E., Paillard, S., Adam-Blondon, A.F., Testolin, R., 2006. Colour variation in red grapevines (*Vitis vinifera* L.): genomic organisation, expression of flavonoid 3'-hydroxylase, flavonoid 3',5'-hydroxylase genes and related metabolite profiling of red cyanidin-/blue delphinidin-based anthocyanins in berry skin. BMC Genomics 7, 12.
- Chang, T., Jiang, Y.J., Chiou, S.H., Chang, W.C., 1988. Carp gamma-crystallins with high methionine content: cloning and sequencing of the complementary DNA. Biochim. Biophys. Acta 951, 226–229.
- Chen, J.-Y., Chang, B.-E., Chen, Y.-H., Lin, C.J.-F., Wu, J.-L., Kuo, C.-M., 2001. Molecular cloning, developmental expression, and hormonal regulation of zebrafish (*Danio rerio*) [beta] crystallin B1, a member of the superfamily of [beta] crystallin proteins. Biochem. Biophys. Res. Commun. 285, 105–110.
- Cheng, C.H., Chen, L., 1999. Evolution of an antifreeze glycoprotein. Nature 401, 443–444.
- Cheng, C.-H.C., Nicodemus, J., Silic, S., Ghigliotti, L., Pisano, E., 2007. Genomic analysis of the evolution of antifreeze glycoprotein genes in antarctic notothenioid fish. Joint Genomics Institute User Meeting, vol. 2. US Department of Energy, Mariott Hotel, Walnut Creek, California.
- Chiou, S.H., Chang, T., Chang, W.C., Kuo, J., Lo, T.B., 1986. Characterization of lens crystallins and their mRNA from the carp lenses. Biochim. Biophys. Acta 871, 324–328.
- Clark, J.I., Benedek, G.B., 1980. The effects of glycols, aldehydes, and acrylamide on phase separation and opacification in the calf lens. Invest. Ophthalmol. Vis. Sci. 19, 771–776.
- Clark, M.S., Clarke, A., Cockell, C.S., Convey, P., Detrich III, H.W., Fraser, K.P.P., Johnston, I.A., Methe, B.A., Murray, A.E., Peck, L.S., Römisch, K., Rogers, A.D., 2004. Antarctic genomics. Compar. Funct. Genom. 5, 230–238.
- Coop, A., Goode, D., Sumner, I., Crabbe, M.J., 1998. Effects of controlled mutations on the N- and C-terminal extensions of chick lens beta B1 crystallin. Graefe Arch. Clin. Exp. Ophthalmol. 236, 146–150.
- Dahlman, J.M., Margot, K.L., Ding, L., Horwitz, J., Posner, M., 2005. Zebrafish alpha-crystallins: protein structure and chaperone-like activity compared to their mammalian orthologs. Mol. Vis. 11, 88–96.
- de Jong, W.W., Leunissen, J.A., Voorter, C.E., 1993. Evolution of the alphacrystallin/small heat-shock protein family. Mol. Biol. Evol. 10, 103–126.
- Delaye, M., Clark, J.I., Benedek, G.B., 1982. Identification of the scattering elements responsible for lens opacification in cold cataracts. Biophys. J. 37, 647–656.
- den Dunnen, J.T., Szpirer, J., Levan, G., Islam, Q., Schoenmakers, J.G., 1987. All six rat gamma-crystallin genes are located on chromosome 9. Exp. Eye Res. 45, 747–750.
- den Dunnen, J.T., van Neck, J.W., Cremers, F.P., Lubsen, N.H., Schoenmakers, J.G., 1989. Nucleotide sequence of the rat gamma-crystallin gene region and comparison with an orthologous human region. Gene 78, 201–213.
- Detrich III, H.W., 1991. Cold-stable microtubules from Antarctic fish. In: di Priscu, G. (Ed.), Life Under Extreme Conditions: Biochemical Adaptations. Springer-Verlag, Berlin, pp. 35–49.

- DeVries, A.L., 1988. The role of antifreeze glycopeptides and peptides in the freezing avoidance of antarctic fishes. Comp. Biochem. Physiol. B 90, 611–621.
- di Prisco, G., Eastman, J.T., Giordano, D., Parisi, E., Verde, C., 2007. Biogeography and adaptation of Notothenioid fish: hemoglobin function and globin-gene evolution. Gene 398, 143–155.
- Eastman, J.T., 1993. Antarctic Fish Biology: Evolution in a Unique Environment. Inc., New York, Academic Press.
- Franck, E., Madsen, O., van Rheede, T., Ricard, G., Huynen, M.A., de Jong, W.W., 2004. Evolutionary diversity of vertebrate small heat shock proteins. J. Mol. Evol. 59, 792–805.
- Gellman, S.H., 1991. On the role of methionine residues in the sequenceindependent recognition of nonpolar protein surfaces. Biochemistry 30, 6633–6636.
- Hall, T., 1999. BioEdit: a user friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. 41, 95–98.
- Hall, B.G., 2004. Phylogenetic Trees Made Easy: A How-to Manual. Sunderland, Mass, Sinauer Associates.
- Hejtmancik, J.F., Thompson, M.A., Wistow, G., Piatigorsky, J., 1986. cDNA and deduced protein sequence for the beta B1-crystallin polypeptide of the chicken lens. Conservation of the PAPA sequence. J. Biol. Chem. 261, 982–987.
- Kiss, A.J., 2005. Functional, biochemical and molecular analyses of the cold stable eye lens crystallins from the Antarctic toothfish *Dissostichus mawsoni*. (Dissertation). Ecology, Ethology & Evolution. University of Illinois. Urbana-Champaign.
- Kiss, A.J., Mirarefi, A.Y., Ramakrishnan, S., Zukoski, C.F., Devries, A.L., Cheng, C.H., 2004. Cold-stable eye lens crystallins of the Antarctic nototheniid toothfish *Dissostichus mawsoni* Norman. J. Exp. Biol. 207, 4633–4649.
- Kroger, R.H.H., Campbell, M.C.W., Munger, R., Fernald, R.D., 1994. Refractive index distribution and spherical aberration in the crystalline lens of the African cichlid fish haplochromis burtoni. Vis. Res. 34, 1815–1822.
- Liu, C., Asherie, N., Lomakin, A., Pande, J., Ogun, O., Benedek, G.B., 1996. Phase separation in aqueous solutions of lens gamma-crystallins: special role of gS. Proc. Natl. Acad. Sci. U. S. A. 93, 377–382.
- Lu, S.F., Pan, F.M., Chiou, S.H., 1996. Characterization of gamma-crystallin from the eye lens of bullfrog: complexity of gamma-crystallin multigene family as revealed by sequence comparison among different amphibian species. J. Protein. Chem. 15, 103–113.
- Marchetti, M.A., Pizarro, G.O., Sagher, D., Deamicis, C., Brot, N., Hejtmancik, J.F., Weissbach, H., Kantorow, M., 2005. Methionine sulfoxide reductases B1, B2, and B3 are present in the human lens and confer oxidative stress resistance to lens cells. Invest. Ophthalmol. Vis. Sci. 46, 2107–2112.
- McDermott, J.B., Cvekl, A., Piatigorsky, J., 1997. A complex enhancer of the chicken beta A3/A1-crystallin gene depends on an AP-1-CRE element for activity. Invest. Ophthalmol. Vis. Sci. 38, 951–959.
- Miyake, T., Amemiya, C.T., 2004. BAC libraries and comparative genomics of aquatic chordate species. Comp. Biochem. Physiol. C 138, 233–244.
- Narberhaus, F., 2002. Alpha-crystallin-type heat shock proteins: socializing minichaperones in the context of a multichaperone network. Microbiol. Mol. Biol. Rev. 66, 64–93.
- Nelson, W., Soderlund, C., 2005. Software for restriction fragment physical maps. In: Meksem, K., Kahl, G. (Eds.), The Handbook of Plant Genome Mapping: Genetic and Physical Mapping, vol. 1. Wiley-VCH, Hoboken, NJ, pp. 285–306.
- NRC, 2003. Frontiers in Polar Biology in the Genomic Era. Washington, D.C, National Academies Press.
- Nylander, J.A.A., 2004. MrModeltest. vol. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.
- Pan, F.-M., Chuang, M.-H., Chiou, S.-H., 1997. Characterization of gS-crystallin isoforms from lip shark (*Chiloscyllium colax*): Evolutionary comparison between gS and b/g crystallins. Biochem. Biophys. Res. Commun. 240, 51–56.
- Parker, R.W., Paige, K.N., DeVries, A.L., 2002. Genetic variation among populations of the Antarctic toothfish: evolutionary insights and implications for conservation. Polar Biol. 25, 256–261.

- Piatigorsky, J., 1998. Multifunctional lens crystallins and corneal enzymes. More than meets the eye. Ann. N. Y. Acad. Sci. 842, 7–15.
- Pointer, M.A., Cheng, C.-H.C., Bowmaker, J.K., Parry, J.W.L., Soto, N., Jeffery, G., Cowing, J.A., Hunt, D.M., 2005. Adaptations to an extreme environment: retinal organisation and spectral properties of photoreceptors in Antarctic notothenioid fish. J. Exp. Biol. 208, 2363–2376.
- Posner, M., Kantorow, M., Horwitz, J., 1999. Cloning, sequencing and differential expression of alphaB-crystallin in the zebrafish, *Danio rerio*. Biochim. Biophys. Acta 1447, 271–277.
- Privalov, P.L., 1990. Cold denaturation of proteins. Crit. Rev. Biochem. Mol. Biol. 25, 281–305.
- Rajaraman, K., Raman, B., Ramakrishna, T., Rao, C.M., 2001. Interaction of human recombinant alphaA- and alphaB-crystallins with early and late unfolding intermediates of citrate synthase on its thermal denaturation. FEBS Lett. 497, 118–123.
- Riyahi, K., Shimeld, S.M., 2007. Chordate betagamma-crystallins and the evolutionary developmental biology of the vertebrate lens. Comp. Biochem. Physiol. B 147, 347–357.
- Runkle, S., Hill, J., Kantorow, M., Horwitz, J., Posner, M., 2002. Sequence and spatial expression of zebrafish (*Danio rerio*) alphaA-crystallin. Mol. Vis. 8, 45–50.
- Sagher, D., Brunell, D., Hejtmancik, J.F., Kantorow, M., Brot, N., Weissbach, H., 2006. Thionein can serve as a reducing agent for the methionine sulfoxide reductases. Proc. Natl. Acad. Sci. U. S. A. 103, 8656–8661.
- Sambrook, J., Russell, D.W., 2001. Protocol 3: selection of poly(A)+ RNA by oligo(dT)-cellulose chromatography. In: Irwin, N., Janssen, K.A. (Eds.), Molecular Cloning: A Laboratory Manual, vol. 1. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, pp. 7.13–17.17.
- Shimeld, S.M., Purkiss, A.G., Dirks, R.P., Bateman, O.A., Slingsby, C., Lubsen, N.H., 2005. Urochordate betagamma-crystallin and the evolutionary origin of the vertebrate eye lens. Curr. Biol. 15, 1684–1689.
- Siezen, R.J., Benedek, G.B., 1985. Controlled modulation of the phase separation and opacification temperature of purified bovine gamma IV-crystallin. Curr. Eye Res. 4, 1077–1085.

- Simmons, M.P., Ochoterena, H., Freudenstein, J.V., 2002. Amino acid vs. nucleotide characters: challenging preconceived notions. Mol. Phylogenet. Evol. 24, 78–90.
- Skow, L.C., Donner, M.E., Huang, S.M., Gardner, J.M., Taylor, B.A., Beamer, W.G., Lalley, P.A., 1988. Mapping of mouse gamma crystallin genes on chromosome 1. Biochem. Genet. 26, 557–570.
- Slingsby, C., Clout, N.J., 1999. Structure of the crystallins. Eye 13, 395-402.
- Smith, A.A., Wyatt, K., Vacha, J., Vihtelic, T.S., Zigler Jr., J.S., Wistow, G.J., Posner, M., 2006. Gene duplication and separation of functions in alphaBcrystallin from zebrafish (*Danio rerio*). FEBS J. 273, 481–490.
- Soderlund, C., Humphray, S., Dunham, A., French, L., 2000. Contigs built with fingerprints, markers, and FPC V4.7. Genome Res. 10, 1772–1787.
- Van Montfort, R.L., Bateman, O.A., Lubsen, N.H., Slingsby, C., 2003. Crystal structure of truncated human betaB1-crystallin. Protein Sci. 12, 2606–2612.
- Waddell, P.J., Steel, M.A., 1997. General time-reversible distances with unequal rates across sites: mixing gamma and inverse Gaussian distributions with invariant sites. Mol. Phylogenet. Evol. 8, 398–414.
- Willard, H.F., Meakin, S.O., Tsui, L.C., Breitman, M.L., 1985. Assignment of human gamma crystallin multigene family to chromosome 2. Somat. Cell Mol. Genet. 11, 511–516.
- Wistow, G., 2002. A project for ocular bioinformatics: NEIBank. Mol. Vis. 8, 161–163.
- Wistow, G., Wyatt, K., David, L., Gao, C., Bateman, O., Bernstein, S., Tomarev, S., Segovia, L., Slingsby, C., Vihtelic, T., 2005. gN-crystallin and the evolution of the bg-crystallin superfamily in vertebrates. FEBS J. 272, 2276–2291.
- Wu, Z., Delaglio, F., Wyatt, K., Wistow, G., Bax, A., 2005. Solution structure of (gamma)S-crystallin by molecular fragment replacement NMR. Protein Sci. 14, 3101–3114.
- Yu, C.M., Chang, G.G., Chang, H.C., Chiou, S.H., 2004. Cloning and characterization of a thermostable catfish alphaB-crystallin with chaperone-like activity at high temperatures. Exp. Eye Res. 79, 249–261.