AIM1, a novel non-lens member of the $\beta\gamma$ -crystallin superfamily, is associated with the control of tumorigenicity in human malignant melanoma

MICHAEL E. RAY*[†], GRAEME WISTOW[‡], YAN A. SU*, PAUL S. MELTZER*, AND JEFFREY M. TRENT*§

*Laboratory of Cancer Genetics, National Human Genome Research Institute, National Institutes of Health, 9000 Rockville Pike, Building 49, Room 4A22, Bethesda, MD 20892; [†]Department of Human Genetics, University of Michigan Medical School, Medical Science II M4708, Ann Arbor, MI 48109; and [‡]Section of Molecular Structure and Function, National Eye Institute, Building 6, Room 222, National Institutes of Health, Bethesda, MD 20892

Contributed by Francis S. Collins, December 27, 1996

ABSTRACT AIM1 is a novel gene whose expression is associated with the experimental reversal of tumorigenicity of human malignant melanoma. The predicted protein product of the major 4.1-kb transcript shows striking similarity to the $\beta\gamma$ -crystallin superfamily. All known members of this superfamily contain two or four characteristic motifs arranged as one or two symmetrical domains. AIM1, in contrast, contains 12 $\beta\gamma$ motifs, suggesting a 6-domain structure resembling a trimer of β - or γ -crystallin subunits. The structure of the AIM1 gene shows remarkable similarity to β -crystallin genes, with homologous introns delineating equivalent protein structural units. AIM1 is the first mammalian member of the $\beta\gamma$ superfamily with a primarily non-lens role. Other parts of the predicted AIM1 protein sequence have weak similarity with filament or actin-binding proteins. AIM1 is a good candidate for the putative suppressor of malignant melanoma on chromosome 6, possibly exerting its effects through interactions with the cytoskeleton.

AIM1 (absent in melanoma) is a novel gene whose expression is altered in association with tumor suppression in a model of human melanoma (1). The AIM1 gene is localized to 6q21, within the putative chromosome 6 tumor suppressor region for human melanoma (2). Characterization of the AIM1 gene and its major transcripts reveals, unexpectedly, that AIM1 belongs to the $\beta\gamma$ -crystallin ($\beta\gamma$) superfamily.

The optical properties of the eye lens largely depend on abundant soluble structural proteins, the crystallins (3–5). Although some crystallins are enzymes, the result of relatively recent gene recruitment events in distinct evolutionary lineages, other crystallins, related to stress-inducible proteins, have more ancient origins and are represented ubiquitously in all vertebrates (6, 7). This ubiquitous class includes the β - and γ -crystallins, which are closely related in sequence and in gene and protein structure (8).

Three non-lens members of the $\beta\gamma$ superfamily have been identified through a conserved sequence signature: Protein S of *Myxococcus xanthus*, a sporulating bacterium (9, 10); spherulin 3a of the slime mold *Physarum polycephalum* (11, 12); and an epidermis differentiation-specific protein (EDSP or ep37) of the amphibian *Cynops pyrrhogaster* (13–15).

The structure of Protein S has been confirmed by NMR analysis (16). The structure of a yeast killer toxin (WmKT) has also revealed unexpected similarity to the folding pattern of the $\beta\gamma$ superfamily (17). Although an ancestral relationship has been discussed, WmKT may represent an interesting example of evolutionary convergence.

Copyright @ 1997 by The National Academy of Sciences of the USA 0027-8424/97/943229-6\$2.00/0

PNAS is available online at http://www.pnas.org.

Except for spherulin 3a, and possibly WmKT, with two motifs, all other identified members of the $\beta\gamma$ superfamily have a 4-fold repeat of the characteristic $\beta\gamma$ motif. In contrast, *AIM1* contains 12 $\beta\gamma$ motifs. Furthermore, the exon/intron structure of the *AIM1* gene shows remarkable similarity with β -crystallin genes, with each motif coded by a separate exon.

Protein S, a calcium-binding protein of the bacterial spore coat, and spherulin 3a, a development-specific protein, are both induced by stresses that trigger sporulation or encystment (18). EDSP is expressed specifically in differentiating ectoderm (13). In lens, β - and γ -crystallins are expressed specifically in differentiating and elongating fiber cells. The connections of various members of the $\beta\gamma$ superfamily with processes of stress response, differentiation, and changes in cell morphology may provide clues for the possible function of *AIM1* in suppressing melanoma tumor formation and in normal tissues.

MATERIALS AND METHODS

cDNA Cloning. A 554-bp cDNA fragment of *AIM1* was originally isolated from a subtracted cDNA library (2). Using driver-driven cDNA subtraction of microcell-suppressed sublines provides a significant improvement in detecting differentially expressed clones over other strategies attempted (e.g., differential display; ref. 2). Based on the sequence of this fragment, an oligonucleotide was designed (AIM1-R1: 5'-GCAACAGGGTTATTCATGTC-3') for hybridization capture of longer *AIM1* cDNA clones from a human liver library using the GIBCO/BRL GeneTrapper cDNA System (following manufacturer's instructions). Multiple clones isolated by this procedure were selected for sequencing.

Northern Blot Analyses. RNA was isolated from several human melanoma cell lines: A375, JT 88, M91–054, UACC 091, UACC 383, UACC 827, UACC 1273, and UACC 1308. Expression in human tissues and other tumor was examined using Northern blots from CLONTECH. Probes spanning the length of the *AIM1* cDNA were generated by PCR. RNA isolation and Northern blot analyses were performed as described (2).

Genomic Sequencing and Primer Extension. A genomic P1 clone (GS 5245) positive for *AIM1* was isolated (2) and used as template for genomic sequencing using the custom *AIM1* cDNA sequencing primers. The 5' end of the gene was identified by primer extension of total cellular RNA from cell line UACC-903(+6) with specific antisense primers AIM1-32 (5'-ATCACTACCTTTCCAGGTCG-3') and AIM1-56 (5'-

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

Abbreviation: $\beta\gamma$ superfamily, $\beta\gamma$ -crystallin superfamily.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. U83115 and U83116).

[§]To whom reprint requests should be addressed at: National Human Genome Research Institute, National Institutes of Health, 49 Convent Drive, MSC 4470, Bethesda, MD 20892-4470.

ACTAATCACGACCGGCTCGT-3') by using standard methods (19). The GenBank accession number for the AIM1 cDNA is U83115. The genomic sequence of the 5' end that gives the first ≈ 80 codons of exon 1 and the putative promoter region is accession number U83116.

Sequence Comparison and Molecular Modeling. The sequence of AIM1 was compared with nucleotide and protein sequence databases using BLAST programs (20). Divergent $\beta\gamma$ motifs were revealed by inspection. Structural models for AIM1 were investigated using QUANTA (Molecular Simulations, Waltham, MA) running on a Silicon Graphics workstation.

RESULTS

AIM1 Is Associated with Suppressed Melanoma. Based on sequence from the original AIM1 cDNA fragment isolated from the suppressed melanoma cell line UACC-903(+6)(1,2), 14 clones with cDNA inserts ranging from 2.8 kb to 6.8 kb were isolated from a human liver library. Expression of AIM1 in melanoma, normal tissues, and other tumors was examined by Northern blotting. UACC-903 is a highly tumorigenic human melanoma cell line, while UACC-903(+6) and D151JA are both microcell hybrid derivatives of UACC-903 that show marked suppression of tumorigenicity due to complete or partial chromosome 6 introduction (1). As shown in Fig. 1A, two major AIM1 transcripts of 4.1 and 7.2 kb were detected in UACC-903(+6) and D151JA, whereas neither could be detected in the parental line UACC-903². Thus, AIM1 expression is associated with chromosome 6-mediated tumor suppression of human melanoma. AIM1 expression was also



FIG. 1. Expression of AIM1 in tumor cells and normal tissues. Northern blot analyses of: human melanoma cell line UACC-903, its suppressed hybrid, UACC-903(+6), and its radiation-reduced chromosome 6 hybrid, D151JA (A); eight additional human melanoma cell lines (B); normal human tissues (C); human tumor cell lines (promyelocytic leukemia HL-60, cervical carcinoma HeLa S3, chronic myeloid leukemia K-562, lymphoblastoid leukemia MOLT-4, Burkitt's lymphoma Raji, colorectal adenocarcinoma SW 480, lung carcinoma A549, and melanoma G361 cells) (D).

undetectable in eight other human melanoma cell lines (Fig. 1*B*).

In normal human tissues AIM1 showed tissue-preferred expression (Fig. 1C). The 4.1-kb AIM1 transcript was abundant in placenta. The 7.2-kb transcript was also present at lower levels in placenta and was detectable in heart, lung, and pancreas. No detectable expression was seen in brain or skeletal muscle, and expression was barely detectable in liver and kidney. In other human tumors, AIM1 expression was significant in the promyelocytic leukemia cell line HL-60, cervical carcinoma cell line HeLa S3, chronic myeloid leukemia cell line K-562, and lymphoblastoid leukemia cell line MOLT-4 (Fig. 1D). Burkitt's lymphoma Raji cells showed high expression, while none was detectable in colorectal adenocarcinoma cell line SW480, lung carcinoma A549, or melanoma G361. Both 7.2-kb and 4.1-kb transcripts were detected in these tumor cells, with highest expression of the 7.2-kb AIM1 transcript in Burkitt's lymphoma cells. Variant transcript sizes were detected in several tissues and lines, possibly the result of alternative transcription or posttranscriptional events.

Sequence Analysis of *AIM1*. A 6.8-kb sequence cDNA was assembled from 14 clones. One clone revealed an alternative polyadenylation signal (position 5528–5532 in the complete 7.2-kb transcript sequence), which could account for some of the minor transcripts observed (Fig. 1). The cDNA sequence contained a long ORF with a 1,906-bp 3' untranslated region (UTR) including a single Alu element. A BLASTN search (20) revealed that UTR sequences are closely similar to several ESTs (GenBank accession numbers G06991, H12321, H40510, H52190, R67842, R83261, R88722, R90978, and T53644); however, the rest of the *AIM1* sequence revealed no close identity with known sequences.

The sequence of AIM1 (Fig. 24) was completed by sequencing a P1 genomic clone (GS 5245). Intron positions were identified and the likely 5' end of the AIM1 7.2-kb transcript ORF was located. The 5' termini of both the 7.2- and 4.1-kb AIM1 transcripts were determined by primer extensions of RNA from UACC-903(+6). The start site for the 7.2-kb transcript was identified 120 bp upstream of the translation start site. The 5' end of the 4.1-kb transcript was found to be 20–30 bp upstream of the second intron position (Fig. 2B). The 5' end of the 4.1-kb transcript thus may originate from a presently unidentified alternative first exon.

AIM1 Encodes a Novel Gene Product Related to β - and γ -Crystallins. The long ORF of the 7.1-kb AIM1 transcript codes for 1,723 amino acid residues with a molecular mass of 189 kDa, whereas the predicted product of the 4.1-kb transcript, possibly including a small unidentified first exon, would correspond to a protein 717–727 amino acids with molecular mass of 81–83 kDa. BLAST database searches showed that the 700-residue C-terminal region of AIM1, corresponding closely to the product of the 4.1-kb mRNA, has extensive similarity with members of the $\beta\gamma$ superfamily.

The multimeric β - and monomeric γ -crystallins are closely related proteins that are major components of the vertebrate eye lens (7, 21). Both families consist of two highly similar domains, joined by a connecting peptide. Each domain consists of two similar structural motifs of a modified greek key type (22), composed of four antiparallel β -strands: *a*, *b*, *c*, and *d* (8). In β - and γ -crystallins (8, 23, 24) and Protein S (16), the paired β -sheets pack closely together forming a symmetrical wedgeshaped domain.

 $\beta\gamma$ motifs may be further classified as A- or B-type. Vertebrate members of the $\beta\gamma$ superfamily conform to an ABAB motif pattern (7). B-type motifs are the most highly conserved and form most of the contacts between domains (8, 23, 24). In Protein S, the order of motifs is reversed to BABA, suggesting a separate history of duplication events (10, 16).

The folding and packing requirements of $\beta\gamma$ motifs define a characteristic sequence signature. Using motif 1 of bovine

Α



FIG. 2. (A) The sequence of human AIM1. Protein sequence is shown in single-letter code. Relative positions of introns in the AIM1 gene are marked by vertical lines. The 12 $\beta\gamma$ motifs are numbered and indicated by colored boxes. A-type motifs are blue and B-type motifs are red. Short underlines indicate connecting peptides between predicted structural domains. Gray shading indicates the position of the two novel linker peptides. (B) Schematic of the complete AIM1 gene is shown with shaded portions representing coding sequence and splicing as indicated. For the 7.2-kb AIM1 transcript, a large first exon is followed by 17 smaller exons. The 4.1-kb AIM1 transcript begins just upstream of the second exon, probably beginning with an unidentified, 20- to 30-bp first exon containing a translation start codon.

 γ B-crystallin (8, 23, 25) as an archetype, the most important residues are G13, F or Y at positions 6 and 11, and S34. Other positions are also well conserved, particularly in B-type motifs, with their characteristic Tyr and Trp corners (26).

AIM1 Protein Structure. Alignment of 12 AIM1 amino acid segments (a1–a12) with the $\beta\gamma$ motifs of bovine γ B-crystallin, Protein S (9), and EDSP (7, 13) reveals clear sequence similarity (Fig. 3*A*). The 12 AIM1 motifs follow the ABAB pattern with highest conservation of B-type motifs. However, some A-type motifs are quite divergent. Motifs a1 and a7 lack the important G13 of the hairpin between strands *a* and *b*. In a1, there is a second unusual substitution of leucine for the conserved S34, which could compensate for the less tightly folded hairpin and for the less bulky valine at position 6. In spite of these differences, a generally similar $\beta\gamma$ motif structure is possible for these motifs.

The other major differences between AIM1 motifs and those of other superfamily members are increased length in the *c*-*d* loops in motifs a1 and a2. Secondary structure prediction suggests that these extensions may form α -helical structures, reminiscent of those in motifs 1 and 3 of Protein S (16). Overall, the first domain of AIM1 (motifs a1 and a2) is the most divergent in structure. In crystallins, stability is so important that structural compromises are probably strongly disfavored. In AIM1, functionality is likely to be more important than stability, so major modifications of the ancestral fold may have occurred to enhance function. Thus, the divergent a1/a2 domain may have an important role in AIM1 function.

Connections Between Motifs. The $\beta\gamma$ motifs of AIM1 are linked in three different ways. First, the connections between A/B pairs of motifs, a1/a2, a3/a4, a5/a6, a7/a8, a9/a10, and a11/a12, are of zero length, so that these motifs must associate directly to form domains. Next, there are three connecting peptides between motifs a2/a3, a6/a7, and a10/a11 joining domains. These connecting peptides have a special significance in β - and γ -crystallins.

 γ -Crystallins are monomers and have "bent" connecting peptides that allow the B-type motifs of the two domains in the same polypeptide to associate (8, 23). β -crystallins, in contrast, have extended connecting peptides (24). Because domains in the same monomer cannot associate in this conformation, dimerization occurs through interdomain interactions between N- and C-terminal domains of separate polypeptides. In some А

		hin ou	6	11 1	3										34			
ps-2	VKAILYQNDGFACDQIEVVANAEELGPLNNNVSSIRVIS														IS			
ps-4	\mathbf{L}	LAVVINKNDN SEDTLPVNSDAPTLGAMNNNTSSIRIS.																
g-1	G	KITF	YEDR	.GFQC	HCYE	CSSE	CPNL		QPY	ζF	• • •	• • • •	• • •	SRC	INS 1	RVI	DS	
g-3	F	RMRI	YERD	DFRG	QMSE	ITDD	CPSL	. Q	DRFH	łL	• • •	• • • •		TE	THSI	JNVI	E	
e-1	N'	TITV	YEHS	NFQC	THKI	FTSD	VPNL	.v	NESF	M	•••	• • • •	•••	DC1	SSI	KI	/G "A	"
e-3	PQITVMEHVNCCCKALVLTEE.TNLAFGNMHDNISSHRVQR t														QR ty	pe		
a-1	Gl	GKVVINSEPDWSEKCIEVFSDIQDCSSWSLSPVI IKVVR															JR	
a-3	SI	SHIDLETEPECLCILSSYFDDTEEMQGFGVMQKTCSMKVHW																
a-5	Pl	PKVVVVEKPFEEKCVELETGMCSFVMEGGETEEATGDDHLPFTSVGSMKVLR																
a-7	AHMIMYSEKNEGSKGSSIDVLGIVANLKETGYGVKTQSINVLS																	
a-9	N	QIHL	SEP	QEQC	HSQS	FEET	TSQI	DDS.						FST	KSC	RV	SG	
a-11	a-11 PTIILPEREDRKCKKIELNAETVNLRSLGFN																	
ps-1	1 ANITYFYNEDFQCKQVDLPPGNYTRAQLAALGIENNTISSVKVPP																	
ps-3	Pl	RARF	YKE	QFDC	KEVD	LPPG	QYTQ	AELE	ERYGI	DN				NTI	SST	κp	QG	
g-2	G	CWML	ERP	NYQC	HQYF	LRRG	DYPD	YQQW	MGFN	J				DSI	RSC	RL	ΓP	
g-4	G:	SWVL	YEMP	SYRG	RQYL	LRPG	EYRR	YLDW	IGAMN	J			• • •	AKV	'GSI	RRV	/м "в	п
e-2	2 QPWILLQDINVSCQCLPLEEGEYSGISMNDGASSIRLIT													T ty	pe			
e-4 GAWALYEHINRGCRCIVARAGEYLANYCTIGFNDOV.SHVYPL														PL				
a-2 GCWILYEQPNFECHSIPLEEGELELSGLWGIEDILERHEEAESDKPVVIGSIRHVV															<i>I</i> V			
a-4 GTWLIYEEPGPQCVPFILEPGEYPDLSFWDTEAAYICSMRPLK																		
a-6 GIWVAWEKPGFTCHQYLLEEGEYRDWKAWGGYNGELCGLRFIL														L				
a-8 GVWVAYENPDFTGEQYILDKGFYTSFEDWGGKNYKISSVOPIC														C				
a-10 GSWVVYDGENFTCNQYVLEEGHYPCLSAMGCPPGATFKSLRFID																		
a-12	G.	IWVT	EYG	SYRG	RQFL	LSPA	EVPN	WYEF	SGC.					RQI	GSI	RPI	v	
					. h													
<-d-> <-d->																		
		trp-c	OLU	er		1	YI-C	orne	iT.									
-																		
в																		
	.B-	type	mot	if		Con	inect	ing	pept	ide	A	-ty	pe	mot	if.	• •		
34 3	35	36	37	38	39	p1	p2	p3	p4	p5	1	2	3	4	5	б		
s	т	R	н	v	v	0	D	v	R	v	S	н	т	D	т.	F	A2/A3	
s	L	R	P	ī	ī.	Ğ	D	F	s	N	A	н	м	т	м	Ŷ	A6/A7	
s	L	R	F	I	D	v	E	F	s	E	P	т	т	Ť	Т.	F	A10/A	11
s	т	R	P	T	ĸ	v	D	ŝ	0	E	н	ĸ	т	T	т.	v	BB2	
c	ĉ	P	T	т Т	D	ò	ц ц		Š		11	5	1 M	'n		1	100	

FIG. 3. (A) Alignment of the 12 $\beta\gamma$ motifs of AIM1 (a1-a12) with A- and B-type $\beta\gamma$ motifs from Protein S (ps-1 through ps-4), bovine γ B-crystallin (g-1 through g-4), and EDSP (e-1 through e-4) (7). Numbers above indicate the positions of structurally critical residues in motif 1 of bovine γ B-crystallin. The positions of β -strands a, b, c, and d are indicated below, along with the characteristic Trp and Tyr corners of B-type motifs. Important residues highlighted in black are glycine at position 34. Residues highlighted in gray are generally hydrophobic and contribute to packing between β -sheets. (B) Alignment of AIM1-connecting peptide sequences (a2/a3, a6/a7, and a10/a11) with those of bovine β B2- and γ B-crystallin. For comparison, motif sequences are numbered according to the scheme for motif 1 of γ B. The extra residues of the connecting peptide are numbered p1-p5.

experiments, engineering of a γ -crystallin connecting peptide into a β -crystallin leads to monomer instead of dimer formation (27, 28).

In AIM1 the connecting peptides are slightly more β -like than γ -like (Fig. 3*B*). The a6/a7 connecting peptide sequence contains P 37, typical of β -crystallins, whereas none of the AIM1 peptides contain P 39, typical of γ -crystallins (29). The AIM1 peptides also contain acidic residues at position p2, and aromatics at p3.

AIM1 also contains two 9–10 residue peptides linking motifs a4/a5 and a8/a9. Inspection of the structure of the β B2 dimer (24) shows that peptides of this length would be sufficient to link the C-terminal domain of one β B2 monomer to the N-terminal domain of its dimer partner. These linker peptides would also be long enough to link up monomers in a γ B conformation.

Models of AIM1 Domain Assembly. The 12 $\beta\gamma$ motifs would form a structure resembling a trimer of β - or γ -crystallins. A possible model for AIM1 assembly with pseudo-threefold symmetry is illustrated in Fig. 4*A*. In this model the connecting peptides between motifs a2/a3, a6/a7, and a10/a11 adopt an extended, β -like conformation. B-type motifs form interdomain contacts in the order a2/a12, a4/a6, and a8/a10 with the longer a4/a5 and a8/a9 linker peptides joining pairs of domains with interdomain contacts. This would be a relatively constrained structure with each domain linked covalently or



FIG. 4. Modeling the $\beta\gamma$ motifs of AIM1. (A and B) schematics of pseudo-threefold AIM1 structures. (A) β -like conformation with "straight" connecting peptides shown as dotted line. (B) γ -Like conformation with "bent" interdomain-connecting peptides. (C) A conceptual illustration of the pseudo-threefold structure. Pairs of domains from the β B2 structure [Bax *et al.* (24)] were assembled at separations consistent with the lengths of connecting peptides. Non-hydrogen atoms are shown in a van der Waals representation.

noncovalently to two others. There are various possibilities for the orientations of the paired domains, but one option would be a three-bladed propeller with a twist about each of the connecting peptides. A conceptual illustration of this idea is shown in Fig. 4C.

Another threefold arrangement could be formed with the connecting peptides in a γ -like, bent conformation (Fig. 4*B*). This arrangement would give no covalent or noncovalent link between the first (a1/a2) and last (a11/a12) domains and would probably be more flexible than the β -like model. Finally, there is the possibility that AIM1 domains form inter- rather than intramolecular contacts, perhaps dimerizing with one or more additional AIM1 molecules or other $\beta\gamma$ superfamily members.

Exon/Intron Structure. AIM1 provides the first example of exon/intron structure for a non-lens member of the $\beta\gamma$ superfamily and shows remarkable similarity to β -crystallin genes (7, 21, 30). In these genes each structural motif is encoded in a separate exon (Fig. 5). The two "intradomain" introns are in phase 2 and both fall within the codon of the second residue of a B-type motif. These introns are missing from the otherwise very similar γ -crystallin genes (7, 21, 31). The N-terminal "arm" of a β -crystallin is encoded in one or two exons upstream of the first A-type motif and separated from it by an intron in phase 0. A similar intron falls near the beginning of sequences coding for the interdomain connecting peptide and thus separates exons encoding motifs 2 and 3.



FIG. 5. The correspondence of gene and protein structure in the $\beta\gamma$ superfamily. Structures of "generic" β - and γ -crystallin genes and of AIM1 are shown with boxes representing exons. Coding sequences are shown in black. Exons are numbered according to the motif(s) they encode; upstream exons in the crystallins have negative numbers, while exons coding for linker peptides in AIM1 are labeled L1 and L2. The phases of introns are shown in parentheses. The fourfold pattern of the greek key motifs in a crystallin is indicated.

Precisely the same pattern and organization is seen in AIM1 (Figs. 2 and 5). The additional linker sequences L1 and L2 are coded separately in exons, flanked by phase 0 introns, reminiscent of the exons for the N-terminal arms of β -crystallins and suggesting possible common evolutionary origins. In β B-crystallins, C-terminal extensions are coded in the same exon as motif 4, the final B-type motif. In AIM1, however, the exons for every B-type motif end at the same position with phase 0 introns and include no linker or arm sequences.

N- and C-terminal Sequences of AIM1. Flanking the 12 $\beta\gamma$ motifs are novel AIM1. The 1,000-residue N-terminal region of the 7.2-kb mRNA product shows weak similarity with neurofilaments, and secondary structure prediction suggests that this region is predominantly α -helical. It may form an extended filament, capable of interaction with cytoskeleton. C-terminal to the 12 $\beta\gamma$ motifs is a region of about 130 residues that shows local similarities with members of the gelsolin family, particularly one from *Caenorhabditis elegans* (GenBank U01183) (not shown), focusing on a twofold repeat of a DXXXDXXXW motif. Gelsolins are actin-modulating proteins (32).

DISCUSSION

AIM1 maps closely to a locus for suppression of malignant melanoma on chromosome 6. It is not expressed in tumorigenic but is expressed at high levels in suppressed melanoma cells. As such, it is a strong candidate for a functional role in suppression of malignancy. The *AIM1* gene gives rise to two major transcripts, both encoding sequences showing membership in a widespread superfamily, the $\beta\gamma$ -crystallins.

All members of the $\beta\gamma$ superfamily contain repeats of a characteristic structural motif. It has been suggested that gene duplication gave rise to two similar ancestral motifs. Gene fusion resulted in a symmetrical one-domain protein, perhaps resembling spherulin 3a of *P. polycephalum* (11, 12). Subsequent gene duplication and fusion led to the four-motif members of the superfamily, with an ABAB pattern, which later proliferated into multigene families (3, 7, 21).

AIM1 represents the most extreme case of this evolutionary process. It contains $12 \beta \gamma$ motifs arranged in an (AB)₆ pattern. Each AB pair shows characteristics of a single domain, each ABAB unit resembles a β or γ monomer with a connecting peptide linking two domains, and the ABAB units are joined with novel linker peptides. Overall, AIM1 resembles a trimer of β - or γ -crystallins. There are several possibilities for the subunit arrangement of this protein, but one interesting possibility is a pseudosymmetrical threefold arrangement with each domain covalently or noncovalently bonded to two other domains.

The β - and γ -crystallins have provided one of the most striking examples of mapping between genomic and protein structures. AIM1 emphasizes this mapping even further, with multiple structural motifs and distinct linker peptides coded in separate exons. It also shows that the arrangement of exons and introns in the vertebrate members of the superfamily must be older than the origins of the β - and γ -crystallins themselves. In β -crystallin genes, each structural motif is encoded by a separate exon, whereas γ -crystallin genes lack introns between motifs within each domain (7, 21). AIM1 provides the first examples of intron positions in a noncrystallin member of the superfamily. Remarkably, the introns defining each ABAB unit of AIM1 are precisely homologous to those of β -crystallin genes, suggesting that β -crystallin genes retain the ancestral structure of the family. γ -Crystallin genes must have lost the ancestral intradomain introns, perhaps as a result of gene conversion with a processed pseudogene. Because both A/B introns are missing in these genes, it may have taken two such events to eliminate both introns. However, because γ -crystallins retain the B/A intron between domains, common to β -crystallins and AIM1, it seems likely that this intron was present ancestrally and that the last common ancestor of this superfamily in vertebrates had the four-motif structure.

The β - and γ -crystallins are highly stable structural proteins of the vertebrate eye lens (3, 5, 7). Their sheer abundance in lens contributes to the refractive index and optical properties of the tissue. However, it seems likely that they also have roles beyond that of passive "filler." Although mammalian y-crystallins may be truly lens-specific, some β -crystallins are expressed at lower levels in several non-lens tissues (33, 34). A relative of the β - and γ -crystallins, EDSP/ep37 of Cynops (14), is expressed in differentiating epidermis and shows an intracellular localization consistent with an association with cytoskeleton and a possible role in cell morphology (13, 15). Indeed, it has been suggested that the recruitment of many crystallins, including some of the enzyme crystallins, may have connections with processes in cell elongation, a key feature of lens cell differentiation (6, 7). β - and γ -crystallins are expressed specifically in elongating lens fiber cells that are undergoing large changes in cytoskeletal architecture and composition.

This suggests the possibility that $\beta\gamma$ superfamily members may have roles in management of cell morphology and shape. This may also apply to AIM1, which, in addition to its similarity to the $\beta\gamma$ superfamily, also contains regions with weak similarity to filaments or to a family of actin-binding proteins. Indeed, suppressed melanoma cells [UACC-903(+6)] expressing AIM1 have a flattened and stellate morphology, distinct from the irregular, "grape-like" morphology characteristic of the parental tumorigenic melanoma cells (UACC-903) (1). Lack of mediators of normal cell morphology could contribute to cells becoming anchorage-independent and invasive, resulting in a higher tumorigenic and metastatic potential. In this regard, it is intriguing to note that there are no natural tumors of vertebrate lens fiber cells.

- Trent, J. M., Stanbridge, E. J., McBride, H. L., Meese, E. U., Casey, G., Araujo, D. E., Witkowski, C. M. & Nagle, R. B. (1990) *Science* 247, 568–571.
- Ray, M. E., Su, Y. A., Meltzer, P. S. & Trent, J. M. (1996) Oncogene 12, 2527–2533.
- Wistow, G. J. & Piatigorsky, J. (1988) Ann. Rev. Biochem. 57, 479–504.

- Tardieu, A., Veretout, F., Krop, B. & Slingsby, C. (1992) *Biophys.* J. 21, 1–12.
- de Jong, W. W., Lubsen, N. H. & Kraft, H. J. (1994) Molecular Evolution of the Eye Lens in Progress in Retinal and Eye Research, eds. Chader, G. & Osbourne, N. (Elsevier Science, London), Vol. 13, pp. 391–442.
- 6. Wistow, G. (1993) Trends Biochem. Sci. 18, 301-306.
- Wistow, G. J. (1995) Molecular Biology and Evolution of Crystallins: Gene Recruitment and Multifunctional Proteins in the Eye Lens, Molecular Biology Intelligence Series (R. G. Landes Company, Austin, TX).
- Blundell, T., Lindley, P., Miller, L., Moss, D., Slingsby, C., Tickle, I., Turnell, B. & Wistow G. (1981) Nature (London) 289, 771–777.
- Inouye, S., Franceschine, T. & Inouye, M. (1983) Proc. Nat. Acad. Sci. USA 80, 6829–6833.
- 10. Wistow, G., Summers, L. & Blundell, T. (1985) *Nature (London)* 315, 771–773.
- 11. Bernier, F., Lemieaux, G. & Pallotta, D. (1987) Gene 59, 265-277.
- 12. Wistow, G. (1990) J. Mol. Evol. 30, 140-145.
- Takabatake, T., Takahashi, T. C., Takeshima, K. & Takata, K. (1991) Dev. Growth Differ. 33, 277–282.
- 14. Wistow, G., Jaworski, C. & Rao, P. V. (1995) *Exp. Eye Res.* 61, 637–639.
- Takahashi, T. C., Ogawa, M., Takabatake, T. & Takeshima, K. (1996) New Members of the β/γ Crystallin Superfamily in Vertebrate Not Expressed in the Eye Lens. Taniguchi Symposium on Developmental Biology VII: Molecular Basis of Lens Induction and Lens Development (Senri Life Science Center, Osaka).
- Bagby, S., Harvey, T. S., Eagle, S. G., Inouye, S. & Ikura, M. (1994) Structure 2, 107–122.
- 17. Antuch, W., Guntert, P. & Wuthrich, K. (1996) *Nat. Struct. Biol.* **3**, 662–665.
- Raub, T. J. & Aldrich, H. A. (1982) in Aldrich, H. C. & Daniel, J. W., eds. *Cell Biology of Physarum and Didymium* (Academic, New York), Vol. 2.

- Triezenberg, S. J. (1992) in *Current Protocols in Molecular Biology*, ed. Ausubel, F. M. (Greene Publishing Associates and Wiley, New York), Vol. 1, Suppl. 20, pp. 4.8.1–4.8.5.
- New York), Vol. 1, Suppl. 20, pp. 4.8.1–4.8.5.
 Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990) J. Mol. Biol. 215, 403–410.
- Lubsen, N. H. M., Aarts, H. J. & Schoenmakers J. G. G. (1988) Prog. Biophys. Mol. Biol. 51, 47–76.
- 22. Richardson, J. S. (1977) Nature (London) 268, 495-500.
- Wistow, G., Turnell, B., Summers, L., Slingsby, C., Moss, D., Miller, L., Lindley, P. & Blundell, T. (1983) *J. Mol. Biol.* 170, 175–202.
- Bax, B., Lapatto, R., Nalini, V., Driessen, H., Lindley, P. F., Mahadevan, D., Blundell, T. L. & Slingsby, C. (1990) Nature (London) 347, 776–780.
- Hay, R. É., Woods, W. D., Church, R. L. & Petrash, J. M. (1987) Biochem. Biophys. Res. Commun. 146, 332–338.
- Hemmingsen, J. M., Gernert, K. M., Richardson, J. S. & Richardson, D. C. (1994) *Protein Sci.* 3, 1927–1937.
- Mayr, E. M., Jaenicke, R. & Glockshuber, R. (1994) J. Mol. Biol. 235, 84–88.
- 28. Trinkl, S., Glockshuber, R. & Jaenicke, R. (1994) Protein Sci. 3, 1392–1400.
- Hope, J. N., Chen, H. C. & Hejtmancik, J. F. (1994) J. Biol. Chem. 269, 21141–21145.
- den Dunnen, J. T., Moormann, R. J. M., Lubsen, N. H. & Schoenmakers, J. G. G. (1986) Proc. Natl. Acad. Sci. USA 83, 2855–2859.
- Moormann, R. J. M., den Dunnen, J. T., Mulleners, L., Andreoli, P., Bloemendal, H. & Schoenmakers, J. G. G. (1983) *J. Mol. Biol.* 171, 353–368.
- 32. Schafer, D. A. & Cooper, J. A. (1995) Annu. Rev. Cell Biol. 11, 497–518.
- Head, M. W., Peter, A. & Clayton, R. M. (1991) Differentiation 48, 147–156.
- Head, M. W., Sedowofia, K. & Clayton, R. M. (1995) *Exp. Eye Res.* 61, 423–428.