The mouse lens fiber-cell intrinsic membrane protein MP19 gene (Lim2) and granule membrane protein GMP-17 gene (Nkg7): Isolation and sequence analysis of two neighboring genes

Ling Zhou, XiaLian Li, Robert L. Church

Emory Eye Center, Department of Ophthalmology, Emory University School of Medicine, Atlanta, GA

Purpose: The lens fiber cell intrinsic membrane protein MP19 appears to play a key role in lens fiber cell structure or communication, and thus cataractogenesis. The goal of this study was to isolate and characterize the entire gene structure of the MP19 gene, termed Lim2, and to investigate gene sequences surrounding this lens-specific gene.

Methods: A 129/SvJ mouse genomic DNA library was screened using radioisotope labeled bovine MP19 cDNA. From this screening, an 11 kb genomic fragment was isolated which contained the entire Lim2 gene, and a neighboring gene, Nkg7, which codes for a 17 kDa granulocyte membrane protein termed GMP-17. The nucleotide sequence of this entire fragment was obtained using double strand automated sequencing techniques. Using CAT and green fluorescent protein reporter constructs, Lim2 5'-upstream promoter sequences were analyzed.

Results: An 11,182 base pair genomic clone containing the entire murine Lim2 gene and another downstream gene, Nkg7, was obtained and completely sequenced. These two genes are only 1,182 base pairs apart, from the poly(A) signal of the Lim2 gene to the published transcriptional start site of Nkg7. Interestingly, the protein coded for by Nkg7, GMP-17, is very similar to the product of the lens Lim2 gene, MP19, in many respects. Both proteins are transmembrane proteins, with each having 4 transmembrane loops. The amino acid sequence of the two proteins is 34% identical, and 49% with respect to similar amino acids. The size of mouse Lim2 is 5,896 base pairs from the transcriptional start site to the poly(A) signal, and contains five exons and four introns. Exons 2-5 of the Lim2 gene encode a polypeptide of 173 amino acids, having over 92% identity to human MP19. Using chloramphenicol acetyltransferase (CAT) and green fluorescent protein (GFP) reporter constructs, it was determined that about 160 bp of sequence upstream from the start of transcription is both necessary and sufficient for efficient expression levels as well as tissue specificity of expression.

Conclusions: The mouse Lim2 gene is very similar to the human LIM2 gene, both having the same number of exons and introns. The coding nucleotide sequences from both species are 88% identical, and 92% identical at the amino acid level. In the immediate 5'-upstream region of these two genes, several highly conserved regions are observed. Due to the similarity of the MP19 and GMP-17 proteins, it is interesting to speculate that the lens MP19 and the lymphocyte-associated GMP-17 may have originated from one primordial gene which, through genetic drift, resulted in two separate proteins having similar functions in two widely separated tissue types.

According to the World Health Organization, cataracts are the leading cause of blindness around the world; an estimated 17 million people suffer loss of vision from cataracts. In the U.S., where surgery has greatly reduced this risk, tens of thousands still lose their sight from this condition, and millions more have poor vision because of cataracts. At present, the exact biochemical events leading to cataract formation are not known, however, it is obvious that many different insults, ranging from environmental to genetic, are involved.

The generation and characterization of To3 (total opacity of lens number 3) cataractous transgenic mice in our laboratory [1,2] has implicated the lens intrinsic membrane protein MP19 as having a critical function in maintaining lens homeostasis during both normal lens development and cataractogenesis. The identification of a single G->T transversion within exon 2 of the Lim2 gene in the To3 mouse mutant [1] established the linkage between Lim2 and cataractogenesis. To date, however, the specific role of MP19 in the lens is not known. The objective of this study is to extend and expand upon our initial knowledge in order to better understand the role of MP19 in normal lens homeostasis and during cataractogenesis.

GMP-17 (Granule Membrane Protein of 17,000 daltons), also called NKG7, TIA-1, G-CSF, and or GIG-1 [3,4], is a cytotoxic granule protein found in activated cytotoxic lymphocytes (CTLs) including natural killer (NK) cells and a subpopulation of CD8+ T cells with cytotoxic potential [5].

As a cytotoxic granule-associated protein, the function of GMP-17 has not been firmly established, although some experimental evidence shows that GMP-17 may act as a nucleolysin in the Fas-mediated apoptotic pathway [6-8]. In other instances, target cell-induced NK cell degranulation results in translocation of GMP-17 from granules to the effector cell surface after fusion of the granules with plasma membrane, suggesting a possible role for GMP-17 in regulating the effector function of lymphocytes and neutrophils [5]. The
Figure 1. DNA sequence of mouse Lim2 and Nkg7. Approximately 11,100 bp of mouse DNA were sequenced to obtain the entire genes for Lim2 and Nkg7. Nucleotide sequences marked in red are exon regions of the two genes. Dinucleotides marked in green indicate splice acceptor and donor boundaries. Translated amino acid sequence for each exon are indicated in blue.
Figure 1 continued. See the legend on the previous page.
gene for human GMP-17 was named **NKG7** after the first publication of the human gene [9].

**METHODS**

Isolation and sequence analysis of genomic DNA: An ~11 kb genomic clone was isolated previously from a 129/SvJ mouse genomic DNA library [2] which contained the entire **Lim2** gene, including about 400 bp upstream from the **Lim2** transcription start site. This DNA was used as a template in cycle sequencing reactions using Cy5-labeled oligonucleotide primers and the Cy5 AutoCycle Sequencing Kit (Pharmacia Biotech, Piscataway, NJ). These reactions were then electrophoresed and analyzed using an ALFexpress automated DNA sequencer (Pharmacia Biotech). Various sequencing primers were synthesized from newly obtained DNA sequence of this clone. Another 1,500 bp fragment was further isolated from the 129/SvJ mouse genomic DNA library using the procedures outlined in earlier [2] and a radioactive probe constructed from the 5'-upstream 100 bp from the above 11 kb clone. This 1,500 bp fragment contained further 5'-upstream DNA sequence.

**Functional analysis of the mouse **Lim2** promoter:** PCR engineering was used to generate **Lim2** 5'-upstream fragments of various length, which were then used to prepare both chloramphenicol acetyltransferase (CAT) and green fluorescent protein (GFP) reporter constructs. These were transiently transfected into chicken embryo (E12) primary lens epithelial or fibroblast cell cultures and assayed for promoter activity.

For CAT analysis, the DNA fragments were cloned into the promoterless CAT reporter vector pSVOATCAT [10]. CAT assays were carried out using the general procedures of Chepelinsky [11] and Klement et al. [12], with the modification of Cassinotti and Weitz [13], to improve sensitivity. Transient transfections were carried out as described by Klement et al. [12], using lipofectamine (Life Technologies,

---

**Figure 2.** Mouse and human promoter sequence. DNA sequence of about 400 bp upstream of the transcription start site (+1) of both mouse and human **Lim2**. The noncoding exon 1 is marked in brown. Sequences having high homology between mouse and human are marked in red, dark purple, green, purple, and blue.

**Figure 3.** CAT expression with various promoter constructs. The X-axis indicates relative promoter activity as indicated by CAT expression. The Y-axis indicates the promoter fragment size contained in the CAT reporter construct. A: CAT activity results obtained from primary chick lens epithelial cell cultures. B: CAT activity results obtained from primary chick fibroblast cultures. All results are shown as a percent of the SV40 control promoter activity.
Gaithersburg, MD) instead of calcium phosphate. The chick embryo lens epithelial cell cultures were used as described by Borras et al. [14]. Embryonic chick lenses (day 12) were cultured in 35 mm collagen-coated dishes (8 lenses per dish) for about 1 week, or until full confluence was achieved. CAT construct DNA (10 µg) was added to the cultures using lipofectamine (30 µl per dish). Cultures were incubated with the transfection mixture for 5 h, after which the cells were washed with serum-free culture medium and complete medium (D-MEM medium with 10% fetal calf serum) was added for further incubation. Cells were co-transfected with the plasmid pIC 409 containing the β-galactosidase (β-Gal) gene to monitor transfection efficiency. Seventy-two h after transfection, the cells were harvested and assays for CAT and β-Gal activity were carried out. Each complete promoter experiment (including all of the promoter fragment sizes) was repeated.

Figure 4. GFP expression with various promoter constructs. Various sized 5’-upstream Lim2 fragments were cloned into the green fluorescent protein (GFP) reporter vector and then transfected into chick embryo lens epithelial cells. A: Epithelial cells transfected with the empty vector (pro-). B: Epithelial cells transfected with vector containing the CMV promoter as a positive control (CMV). C: Epithelial cells transfected with Lim2 upstream fragment -140/+56 (-140). D: Epithelial cells transfected with Lim2 upstream fragment -160/+56 (-160). E: Epithelial cells transfected with Lim2 upstream fragment -200/+56 (-200). F: Epithelial cells transfected with Lim2 upstream fragment -400/+56 (-400).
four separate times. The data presented are representative of the results from these experiments.

For GFP expression analysis, the same upstream DNA fragments outlined above were cloned into the multiple cloning site of the promoterless green fluorescent protein expression vector pEGFP-1 (CLONTECH, Palo Alto, CA). Monolayered chick embryo lens epithelial cell were transfected with the vectors using lipofectamine (Life Technologies). Approximately 24 to 48 h following transfection, the cells were observed for fluorescence using an inverted fluorescent microscope.

DNA sequence analysis: DNA sequences were analyzed using DNAsis sequence analysis software for Windows (Hitachi Software, San Bruno, CA), and with the various online programs such as the Protein Hydrophilicity/Hydrophobicity Search and Comparison Server, BLAST at NCBI, and Prediction of Transmembrane Helices in Proteins.

RESULTS & DISCUSSION

DNA sequencing of the ~11 kb and the ~1.5 kb fragments (AF320075) demonstrated the presence of two genes in close approximation to one another (Figure 1). The mouse Lim2 gene contains 5,896 base pairs from the transcriptional start site to the poly(A) signal. The mouse gene is very similar to the human LIM2 gene [15] in that it contains five exons encoding a polypeptide of 173 amino acids. The different exons in the mouse gene are each exactly the same size as those of the human gene. The total length of the mouse Lim2 gene, however, is considerably smaller than that of the human gene, which is 8,056 bp in length [15]. This difference in size is due almost entirely to the presence of Alu repeats in the human gene, there being seven separate Alu repeats in the human gene.

At least 1,500 bp of both mouse and human Lim2 upstream sequence have been obtained. Within the first 400 bp upstream from the transcriptional start site are several stretches of sequence that have very high homology between the two species (Figure 2). Five such stretches of sequence have been identified. One of these stretches -191 to -177 (red) contains an SP1 consensus sequence, completely conserved between the two sequences.

MP19     1  MY---SFMGGGLFCAM-VQ-TILLVTA-TATDHMQYRLSNGS-P-AQQL  50
M     S     LF A  +G T  L  A T TD W+     G  F AH GL
GMP-17  1  MEPFRSLRA---LF-AGSLGLTSSLIALTTDFWIVAT--GPHFSAHGL  50
MP19    51  WRYCLGNKCFLQTESI-A-YWNATRAFMILSA-LCATSGIIMGVLAFPAQ  100
M N Q E A Y + T+F IL A L G + V
GMP-17  51  WPTS-------Q-ETQVAGYIHVTQSFCILAVMLGVSGSFAHQGLWPTSQETQV  100
MP19   101  STFTRLSR-PF-SA-G-------IMFF-A---STFLVALLAITYGVTYS-SF  150
M S F L S F S A G +MF A SL V A+X+T++F IL A L
GMP-17 101  S-FILLSCIPALFAQGRLIVTYMAFSAALSIL-V-MAVYTMRSWHEQ  150
MP19   151  --FGLGR--FGDFMSY-YLQWLLATMSTFAIR-PF-C-CDAYHMH  200
M S F F SHG Y LGV SF I F C A H
GMP-17 151  TPFQSVQTF---FSFNSFY-LGWV--------SF---ILFLAGGLSLGA   200
MP19   201  ECR-BLA---TPR  250
MP19   201  --CRTTRRAEYTL
CR R A T
GMP-17 201  --CRTTRRAEYTL

Figure 5. Amino acid sequence comparison of MP19 and GMP-17. Sequence alignment was carried out using DNAsis. Amino acids which are identical between the two sequences are marked. Similar amino acids are marked with a “+” between the two sequences.

Figure 6. Hydropathic profiles of MP19 and GMP-17. A: The hydropathic profile of MP19. B: The hydropathic profile of GMP-17. Sequences spanning the membrane are marked in red; sequences that are inside the cell are marked in blue; and sequences that are outside the cell are marked in pink.

MP19     MYSPMG
GMP-17    MEPFRS
Transmembrane 1
MP19     GILFCAMVGTILLVATAT
GMP-17    LALFACSLGLTSSLIALTTDFWIVAT
Outside
MP19     DHHMQYRLSGSFAHGLWRYCLGNKCFLQTESMFAGHGLWPTSQETQV
GMP-17    VATGPHFSASOMTLFTSQETQVAGYIHVTQS
Transmembrane 2
MP19     MILDACLTRSGIIMGVLAF
GMP-17    FCILAVMLGVSGSFAHQGLWPTSQETQV
Inside
MP19     AQQSTFTLSRP
GMP-17    LSAAPGR
Transmembrane 3
MP19     FSAGIMFFASTLFLVLLALAIYTG
GMP-17    GLLVSTVMFASALSILVAMAVY
Outside
MP19     VTVSFLGRGFDWR
GMP-17    TSMMWQTTPSFQVQTXF
Transmembrane 4
MP19     FSWSYLGWVALLMTSFARIFYM
GMP-17    WSFYLGWSFILFLFAGGLSLGA
Inside
MP19     CAYMRHECCRRLATFR
GMP-17    HCRTRRAEYTL

Figure 7. Polypeptide fragment length comparison between MP19 and GMP-17. The hydropathic profiles from Figure 6 were used to obtain the length of sequence found inside the cell, the transmembrane sequence, and the sequence found outside the cell. Sequences spanning the membrane are marked in red; sequences that are inside the cell are marked in blue; and sequences that are outside the cell are marked in pink.

© 2001 Molecular Vision
the two species. As in the human upstream sequence [15], no TATA or CAAT boxes were found in the mouse upstream sequence. Interestingly, the first ~190 bp upstream of the transcription start site are much more conserved (75%) than are the next ~190 bp (49%).

To determine possible upstream regulatory sequences in the mouse, both CAT and GFP constructs were prepared and assayed in embryonic chick lens epithelial cells. The empty vector (Pro-) was used as a negative control and the viral promoter of SV40 was used as a positive control (SV40). Different lengths of the mouse upstream sequence, including the first 56 bp of exon 1 were used to determine promoter activity. Chick embryo fibroblasts were used as a tissue-specific control. As can be seen in Figure 3A, cloned sequences shorter than -140/+56 were completely negative for CAT expression, indicating that the basal promoter for mouse Lim2 was at least greater than 140 bp. All sequences having lengths above 160 bp from the transcription start site appeared to be uniformly positive for CAT expression. The same CAT expression constructs were also tested in chick fibroblasts (Figure 3B). Except for the positive control (SV40), all constructs were uniformly negative for expression of CAT. These data indicate that the basal promoter resides in the sequence between -140/+56 and -160/+56. Interestingly, a completely conserved sequence is found between -147 and -129 bp upstream of the transcription start site (Figure 2, sequence in red). This sequence also contains the consensus SP1 binding site. Equivalent data are observed when green fluorescent protein (GFP) is used instead of CAT as the reporter gene (Figure 4). Very good fluorescence is observed in the SV40 positive control and in all Lim2 upstream sequences that include the -147 to -129 conserved sequence.

As can be seen in Figure 1, another complete gene was found just downstream of the Lim2 polyadenylation site.

This gene was identified as Nkg7, which encodes a protein known as GMP-17, a type III integral membrane protein of 165 amino acids, with a relative molecular mass of 17,700 daltons. These two genes are located immediately next each other, having only 1,182 bp separating the two. The mouse Lim2 gene contains 5,896 base pairs from the transcriptional start site to the polya(A) signal. Lim2 contains five exons encoding a polypeptide of 173 amino acids. The exons range in size from 59 bases (exon 1) to 182 bases (exon 2). The total sequence contains an additional 1,184 bp of upstream 5' flanking sequence, containing putative promoter sequence.

In contrast, the entire Nkg7 gene contains 1,074 bp with 4 exons encoding a protein of 165 amino acids. The exons range in size from 60 bases (exon 4) to 157 bases (exon 1). GMP-17 contains four hydrophobic, possibly membrane-spanning, regions. This granule membrane protein is thought to function in granule targeting or signaling. Interestingly, when the product of the lens Lim2 gene, MP19, is compared with GMP-17, the two proteins share several similar structural aspects. Both proteins are quite similar with respect to amino acid sequence (Figure 5), having a 34% identity and 49% similarity. Both proteins appear to span the membrane four times and the hydrophobic profile is very nearly identical (Figure 6). The first extracellular loop in MP19 is about 11 amino acids larger that the first extracellular loop in GMP-17 (Figure 6, Figure 7). Otherwise, the structures of the two proteins is are almost identical.

The interesting aspects of these observations are the many similarities between GMP-17 and MP19. Both proteins are integral membrane proteins and both proteins have four transmembrane domains. GMP-17 is composed of 165 amino acids; MP19 is composed of 173 amino acids. The amino acid sequence identity between the two proteins is 34%, and the amino acid similarity is 49%. The two genes are literally “next door neighbors,” with only about 1,200 bp separating them in the mouse genome. The GMP-17 gene contains 4 short coding exons, as does Lim2, and the entire GMP-17 gene is contained within 1,074 bp. The overall size of the Nkg7 gene is much smaller than the Lim2 gene (~6,000 bp).

As is true for MP19, there is, as yet, no known function for GMP-17. Since GMP-17 is relocalized from granules to the plasma membrane, it is thought that this protein contributes to the formation of junctions between effector cells and target cells following granule exocytosis, or that it regulates ion channels required for cytotoxic effector function. Very similar functions have been postulated for MP19 in the lens fiber cell. It is interesting to postulate that Lim2 and Nkg7 may be genes which originally had a common ancestor, and that, through gene duplication, became two different genes serving a very similar or identical function in different cell types. In birds, it appears that the Lim2 gene does not exist. There has been much speculation on just what the chicken lens uses to serve the function of MP19. Perhaps the chicken, and lower animals, use a protein much like GMP-17, or an early ancestor of both proteins. The fact that the mouse Nkg7 gene is just 1,183 bases downstream from Lim2 suggests that both genes

Figure 9. Larger proteins that share the MP19 hydrophatic profile as the n-terminal end of the molecule. A: A comparison of MP19 (red) with rat 3-oxo-5-a-steroid 4-dehydrogenase 1 (blue). B: A comparison of MP19 (red) with Escherichia coli maltose transport system permease (blue). C: A comparison of MP19 (red) with Pyrococcus horikoshii abc transporter permease (blue).
may be closely related in function, though they are expressed with different tissue-specificity.

Very similar data were found with respect to the human LIM2 and NKG7 genes. We have determined that the same gene arrangement exists in the human AF305941, with the distance between the two genes being much greater in the human (7,210 bp) than in the mouse (1,182 bp). In the human, there are nine Alu family DNA repeats between the two genes, taking up 2,383 bp. However, even with removing the Alu repeats, there are still approximately 4,800 bp separating the human genes. There are thus approximately three times more nucleotides separating the human genes compared to the mouse genes. No explanation for this difference in the separation of these two genes is known.

To further investigate the potential membrane spanning nature of mouse MP19 and to possibly obtain further information concerning the function of this lens membrane protein, we carried out a hydropathic profile search using the Protein Hydrophilicity/Hydrophobicity Search and Comparison Server which calculates the hydropathic profile of a protein and then searches the protein sequence library for proteins with similar profiles. This program compares only the hydropathic profile of proteins, not the amino acid sequence or the size of the proteins being compared. The results were quite striking. Figure 8A shows the comparison of mouse MP19 (red line) and GMP-17 (blue line). The hydropathic profiles are essentially identical. Another protein of approximately the same size as MP19 and GMP-17, the Chlamydia pneumoniae lipoprotein signal peptidase (LSPA, Q9Z817) also displays a hydropathic profile that is quite similar to that of MP19 and GMP-17 (Figure 8B). This protein is an integral membrane protein that specifically catalyzes the removal of signal peptides from prolipoproteins. It has been postulated that MP19 and GMP-17 are associated with junctional complexes in their respective cell types [5,16]. However, since LSPA apparently serves as a leader sequence peptidase, perhaps MP19 and GMP-17 might serve a similar function. Several much larger proteins appear to have the MP19 hydropathic profile as part of their C-terminus structure (Figure 8C-F). These proteins are widely diverse throughout the animal kingdom and have considerably different functions. The proteins range from mouse n-formyl peptide receptor (Figure 8C), through goat naphthoquinone oxidoreductase chain 2 (Figure 8D), human g protein-coupled receptor (Figure 8E), Salmonella typhimurium regulatory protein (Figure 8F), Bacillus subtilis hexuronate transporter (Figure 8G), and Escherichia coli permease (Figure 8H). There are also larger proteins which share the MP19 hydropathic profile as the n-terminal end of the molecule. These include rat 3-oxo-5-α-steroid 4-dehydrogenase 1 (Figure 9A), Escherichia coli maltose transport system permease (Figure 9B), and Pyrococcus horikoshii abc transporter permease (Figure 9C). It can be argued that the results shown are relatively uninteresting in light of the common length and necessary hydrophobicity of transmembrane helices. However, the searches carried out include all proteins from the SwissProt database, which includes literally hundreds of transmembrane-type proteins. That only about 10 different proteins possess the striking similarity of transmembrane profile while having literally no amino acid similarity would seem to be of some interest, and a possible method of determining the function of MP19 within the lens. It is possible that the transmembrane structure of MP19 is utilized as a purely structural anchor with no active function. However, in general, each of the proteins shown in Figure 8 and Figure 9 has a transport or enzymatic function. Therefore, it is also possible that MP19 serves to allow the transport of macromolecules across the fiber cell membrane. Further studies will hopefully shed light on the function of lens MP19.

**ACKNOWLEDGEMENTS**

This work was supported in part by National Institutes of Health grants R01 EY11516 (RLC), R01 EY12301 (RLC), F32 EY07022 (LZ), T32 EY07092, P30 EY06360, C06 EY06307, the Knights Templar Educational Foundation of Georgia, Inc., and a Departmental Grant from Research to Prevent Blindness, Inc. RLC is presently a Research to Prevent Blindness Senior Scientific Investigator.

**REFERENCES**


10. Lok S, Stevens W, Breitman ML, Tsui LC. Multiple regulatory


