Mutation of a Gene Encoding a Protein with Extracellular Matrix Motifs in Usher Syndrome Type IIa

James D. Eudy,* Michael D. Weston,* SuFang Yao, Denise M. Hoover, Heidi L. Rehm, Manling Ma-Edmonds, Denise Yan, Iqbal Ahmad, Jason J. Cheng, Carmen Ayuso, Cor Cremers, Sandra Davenport, Claes Moller, Catherine B. Talmadge, Kirk W. Beisel, Marta Tamayo, Cynthia C. Morton, Anand Swaroop, William J. Kimberling, Janos Sumegi†

Usher syndrome type IIa (OMIM 276901), an autosomal recessive disorder characterized by moderate to severe sensorineural hearing loss and progressive retinitis pigmentosa, maps to the long arm of human chromosome 1q41 between markers AFM268ZD1 and AFM144XF2. Three biologically important mutations in Usher syndrome type IIa patients were identified in a gene (*USH2A*) isolated from this critical region. The *USH2A* gene encodes a protein with a predicted size of 171.5 kilodaltons that has laminin epidermal growth factor and fibronectin type III motifs; these motifs are most commonly observed in proteins comprising components of the basal lamina and extracellular matrixes and in cell adhesion molecules.

Usher syndrome, an autosomal recessive disorder, is the most frequent cause of combined deafness and blindness in adults and affects 3 to 6% of children born with hearing impairment. The frequency of the syndrome in the United States is ~4.4/100,000 persons. Affected individuals have sensorineural hearing deficiencies at birth and later develop progressive retinitis pigmentosa (RP). Usher syndrome is both clinically and genetically heterogeneous (*1*). Three forms of Usher syndrome, types I, II, and III, can be distinguished from one another clinically. Type I differs from types II and III with respect to severity of hearing loss and vestibular involvement. Type I patients are profoundly deaf, whereas type II patients have mild hearing impairment. The hearing loss of type III patients is progressive. Type I patients have no vestibular responses, and type III individuals have variable vestibular dysfunction, whereas type II patients are normal in this regard. In all cases of Usher syndrome, deafness is associated with RP. The frequency of hearing impairment within the RP population is estimated to range between 8.0 and 33.3% (*2*), a substantial proportion of the hearing loss being due to the Usher syndromes. In fact, of the ~16,000 deaf and blind people in the United States, more than half are believed to have Usher syndrome. Considering the tremendous burden imposed by the loss of both major senses and the fact that Usher syndrome is the major cause of deafness and blindness in technologically developed countries, progress in understanding the underlying pathological basis of this disorder will affect a large segment of the population.

Usher syndrome type II is the most common of the three Usher syndromes. Although Hallgren originally observed that Usher type II made up only ~10% of all Usher cases that he encountered (*3*), more recent research shows that type II actually accounts for over half of all Usher cases (*4*). In our series of 560 families with Usher syndrome, 59% have Usher syndrome type II.

At least nine distinct genetic loci have been identified as being associated with the three clinical types of Usher syndrome (*5*); six loci correspond to the type I phenotype and two to the type II phenotype, and one locus has been mapped for Usher type III. At present, only the *USH1B* gene has been identified (*6*), a gene that codes for the unconventional myosin VIIa. Preliminary studies have suggested a role for myosin VIIa in trafficking of vesicles in photoreceptors and cochlear hair cells (*7*).

Recently, we localized the *USH2A* gene to a 1-mega–base pair interval between markers D1S474 and AFM144XF2 and accordingly generated a yeast artificial chromosome (YAC) map to encompass the region (*8*). Haplotype analysis to narrow the *USH2A* gene critical region placed the gene between markers AFM268ZD1 and AFM144XF2 (Fig. 1) (*9*). Before refinement of the *USH2A* region, we isolated two genes from within the D1S474 to AFM144XF2 interval, neither of which exhibited mutations in their coding regions. To identify a candidate gene for Usher syndrome type IIa, we constructed a bacterial artificial chromosome (BAC) contig of ~300 kb between AFM268ZD1 and AFM144XF2 (*10*). During the process of generating sequence tagged sites (STSs) from the ends of the BAC clones, GenBank

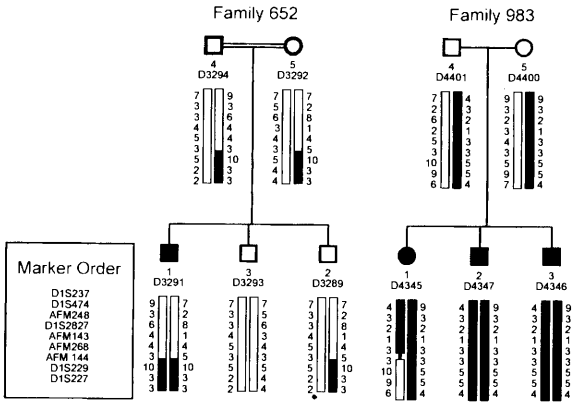

Fig. 1. Refinement of the *USH2A* gene critical region. Nine polymorphic markers known to lie in the region were typed and examined in USH2A families. Homozygosity mapping in family 652 excluded the *USH2A* gene from the region centromeric to AFM268ZD1 in family 652, whereas in family 983 the *USH2A* gene was excluded from the telomeric side of AFM144XF2.

J. D. Eudy, S. F. Yao, M. Ma-Edmonds, J. J. Cheng, C. B. Talmadge, J. Sumegi, Department of Pathology and Microbiology, University of Nebraska Medical Center, Omaha, NE 68198, USA.
M. D. Weston, D. M. Hoover, K. W. Beisel, W. J. Kimberling, Department of Genetics, Boystown National Research Hospital, Omaha, NE 68131, USA.
H. L. Rehm, Department of Genetics, Harvard Medical School, Boston, MA 02115, USA.
D. Yan and A. Swaroop, Department of Ophthalmology, W. K. Kellogg Eye Center, University of Michigan, Ann Arbor, MI 48105, USA.
I. Ahmad, Department of Cell Biology and Anatomy, University of Nebraska Medical Center, Omaha, NE 68198, USA.
C. Ayuso, Fundacion Jimenez Diaz, Avenida de los Reyes Catolicos, 2, 28003 Madrid, Spain.
C. Cremers, Katholieke Universiteit Nijmegen, Sint Radboudziekenhuis, Philips van Leydeniaan 15, 6500 NB Nijmegen, The Netherlands.
S. Davenport, 5801 Southwood Drive, Bloomington, MN 55437–1739.
C. Moller, Ear Nose and Throat Clinic, University Hospital (RIL), S-58180 Linkoping, Sweden.
M. Tamayo, Unidad de Genetica Clinica, Universidad Javeriana, Carrera 7 # 40-62 (2o.Piso), Bogota, Colombia.
C. C. Morton, Department of Obstetrics, Gynecology and Reproductive Biology and Pathology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: jsumegi@mail.unmc.edu

BLAST search analysis revealed that BAC 133c8 contained an open reading frame (ORF) on the centromeric end that showed substantial similarity to epidermal growth factor motifs present in the laminin family of proteins (*11, 12*). To identify additional potential exons, the BAC 133c8 was digested with Hind III, the fragments were subcloned into pBluescript KS^+ plasmid (Stratagene), and their sequences were determined. One of the subclones also contained an ORF with a sequence similar to that of laminin epidermal growth factor (LE) motifs, suggesting that the two ORFs originated from the same gene. Polymerase chain reaction (PCR) primer pairs generated from the DNA sequences of the ORFs were used to screen a lambda gt10 human retina cDNA library by a combination of PCR and hybridization techniques. Screening more than 5×10^6 plaques, we isolated three overlapping cDNA clones. In addition to cDNA library screening, 5′ rapid amplification of cDNA ends (5′ RACE) was performed on human retinal RACE-ready cDNA from Clontech. The library screening and 5′ RACE procedures allowed us to generate a contiguous cDNA sequence of 4782 base pairs (bp) (*13*). Subsequent BLAST search analysis identified a GenBank expressed sequence tag (EST) obtained from a retina cDNA library (GenBank accession number AF017021) that overlapped and extended our cDNA contig to 6330 bp. The consensus sequence of the cDNA contig has a 5′ untranslated region (UTR) sequence of 370 bp and a large ORF from nucleotides 371 to 5023, followed by a 3′ UTR of 1307 bp (GenBank accession number AF055580).

To prove that the isolated gene is responsible for Usher syndrome type IIa, we designed PCR primers to screen DNA from 96 unrelated type II patients who showed linkage to 1q and from 96 control subjects without previous history of hearing or visual impairment. Heteroduplex analysis identified three different mutations—2314delG, 2913delG, and 4353-54delCT—which were found exclusively among the Usher type IIa patients (*14*) (Fig. 2A). All mutations were observed to segregate with the Usher II phenotype and resulted in frameshift mutations with premature terminations. 2314delG causes a frameshift at codon 772, after which the ORF continues for 20 codons and ends as TAG; 2913delG starts a frameshift at codon 972, continues for 43 codons, and ends as TGA; and 4353-54delCT causes a frameshift at codon 1452 and ends as TAA 28 codons downstream. The most frequent of the three mutations was the 2314delG (Fig. 2B). Of the 96 probands, 21 tested positive for the 2314delG mutation, 8 of which were homozygous and 13 heterozygous. All but two of these individuals had northern European ancestry (Swedish, Dutch, German, or English). The two non–northern European patients were both homozygous for 2314delG; one was from Spain and the other was an African American from Nebraska, USA. Examination of various haplotypes for markers AFM143XF10, AFM268ZD1, and AFM144XF2 failed to reveal any substantial disequilibrium with the 2314delG mutation. These data suggest that the 2314delG mutation did not arise in

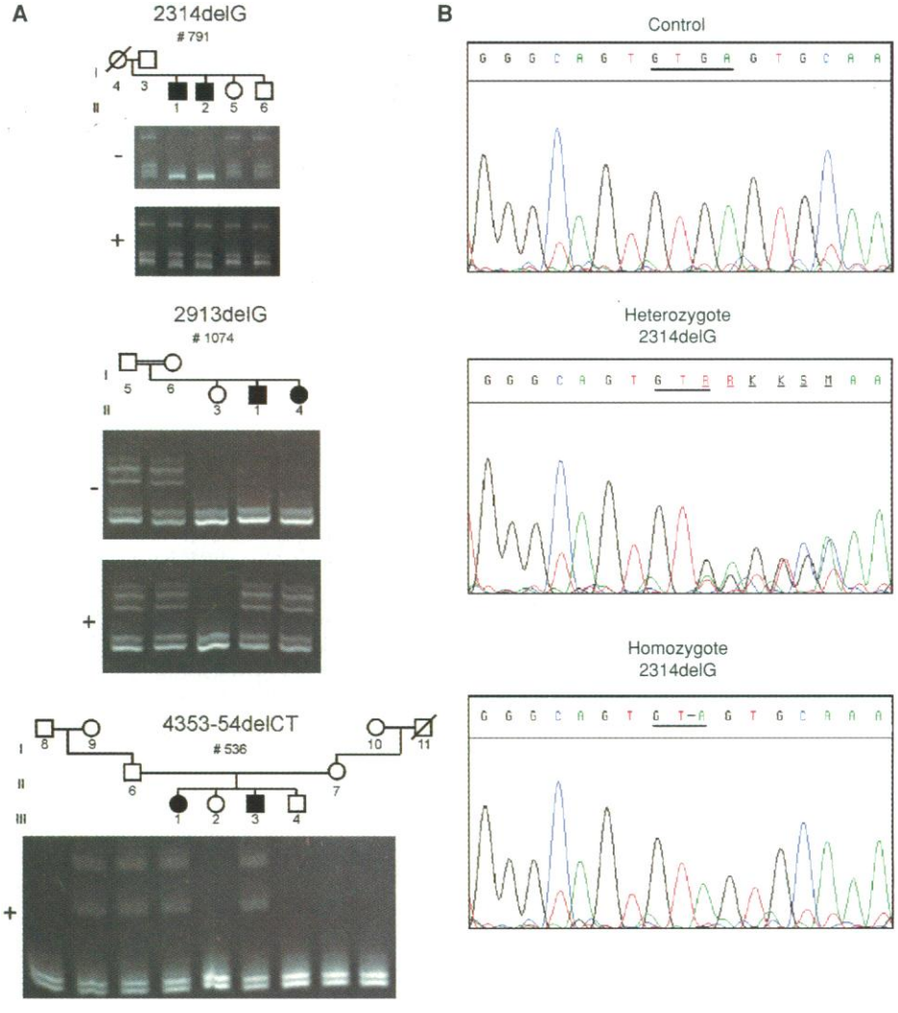

Fig. 2. Heteroduplex analysis and DNA sequencing of *USH2A* mutations. (**A**) Heteroduplex gels of three families segregating deletions of the *USH2A* gene. Family 791 segregates for the 2314delG mutation, 1074 segregates for the 2913delG mutation, and 536 segregates for the 4353-54delCT mutation. Gel pictures are labeled minus (−) and plus (+) to indicate inclusion of an amplification product from a control CRT-1 cDNA or 133c8 BAC templates for the *USH2A* gene, respectively. Inclusion of control PCR product allowed detection of mutations present in the homozygous state. Note the lack of heteroduplex formation in affected individuals II-1 and II-2 from family 791 and II-1 and II-4 from family 1074 when the PCR product from the cDNA or BAC template was not added before the heteroduplex formation. (**B**) DNA sequences surrounding the 2314delG mutation. The top panel shows the sequence of the PCR product from the cDNA clone, and the middle and bottom panels are sequences for PCR products from I-3 and II-1 of family 791. Note the deletion of guanine in the bottom panel. The 2314delG mutation was observed by both heteroduplex analysis and direct sequencing with both forward and reverse primers in 21 USH2A probands (8 homozygotes and 13 heterozygotes) and verified by heteroduplex analysis in all available family members (*15*).

a common ancestor, although further studies of the origins of the 2314delG mutation are warranted. The 4353-54delCT was observed in a heterozygous Usher type IIa patient, one of the few found in the Louisiana Acadian population in our series. In total, the DNA tested for mutations accounted for 18% of the ORF, whereas the Usher type IIa mutations identified above accounted for 16% of the mutations expected to be found in the 96 Usher II cases.

To investigate the expression pattern of *USH2A*, we performed Northern blot analysis of RNA from various tissues, using the cDNA clone CRT-1 as a hybridization probe (*15*). Three transcripts of ~6.5, 5.0, and 1.9 kb were detected in the human retina, but no discernable signal was found in other tissues (Fig. 3A). We confirmed expression of *USH2A* in human fetal cochlea, eye, brain, and kidney and in adult retina and Y79 retinoblastoma cells, using reverse transcriptase–PCR (RT-PCR) with PCR primer pairs (Fig. 3, C and D). The PCR primer 1STS 1336 also amplified a product from monkey retina RNA (Fig. 3D). The sequence of the RT-PCR product from monkey RNA showed high sequence similarity with the corresponding human product, only two conservative changes being found within the 109 nucleotide sequence (*16*).

The spectrum of mutations observed in the ORF and the expression of the gene in the target tissues establish this gene as the *USH2A* gene. Conceptual translation of the *USH2A* ORF results in a polypeptide of 1551 amino acid residues with a predicted size of 171.5 kD and an isoelectric point of 7.45 (Fig. 4A). A BLAST search of GenBank with the deduced *USH2A* protein sequence revealed 32% identity and 47% similarity between amino acid residues 300 and 1050 to all laminin family members. The polypeptide chain contains 10 LE domains, each with ~50 amino acid residues arranged in tandem (Fig. 4B). Laminins are one of the major components forming the extracellular matrix of basement membranes in all tissues, and the LE motif is present in other extracellular matrix proteins. Sequence similarity between the USH2A protein and the laminins ends at amino acid residue 1050, and analysis of the COOH-terminal region from residues 1050 to 1551 with the Paircoil program did not identify the characteristic coiled-coil domains present in all laminins identified thus far (*17*). From amino acid residues 1090 to 1500, however, the USH2A protein has four tandem repeats of ~100 amino acids residues with a sequence similar to those of a variety of proteins containing fibronectin type III (F3) repeats. The greatest similarity is with several cell adhesion proteins and receptors, including neogenin and the leukocyte common antigen-related protein (*18*). Analysis of the F3 repeat region of the USH2A protein with the Transmembrane prediction program (TMpred) (*17*) identified two statistically significant potential transmembrane helices at amino acid residues 1366 to 1383 (outside to inside helix) and 1447 to 1465 (inside to outside helix). The first 20 amino acid residues of the USH2A protein are highly hydrophobic with characteristics of a signal peptide and may represent a signal for secretion. In addition, this protein contains 18 potential N-glycosylation sites.

Collectively, sequence similarity of the USH2A protein with the sequences of laminins and cell adhesion molecules, the hydrophobic NH_2-terminal 20 residues, and

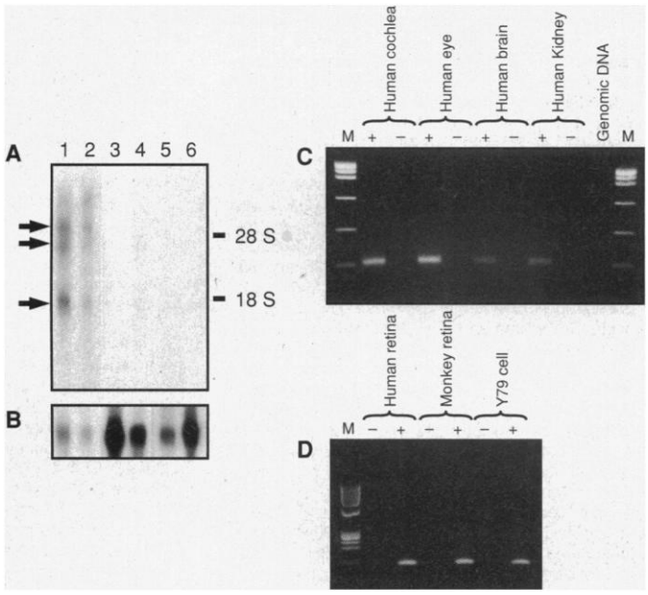

Fig. 3. Expression pattern of USH2A in fetal and adult tissues. (**A**) Northern blot analysis, indicating the positions of 18S and 28S rRNA. (**B**) Hybridization of human β-actin probe. (**C**) Expression of USH2A in fetal human cochlea, eye, brain, and kidney. (**D**) Expression of USH2A in adult human retina, monkey retina, and Y79 human retinoblastoma cells (+, reactions with RT; −, reactions without RT). The PCR reaction products were run on 4% NuSieve 3:1 agarose gel (FMC) with 100-bp markers (M) from GIBCO-BRL.

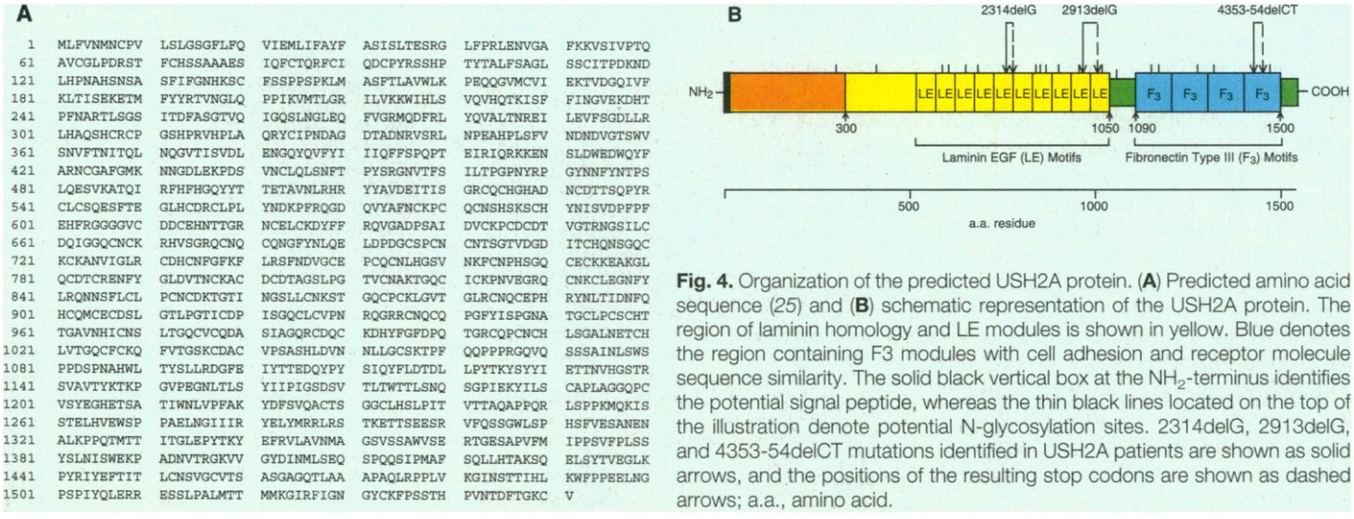

Fig. 4. Organization of the predicted USH2A protein. (**A**) Predicted amino acid sequence (*25*) and (**B**) schematic representation of the USH2A protein. The region of laminin homology and LE modules is shown in yellow. Blue denotes the region containing F3 modules with cell adhesion and receptor molecule sequence similarity. The solid black vertical box at the NH_2-terminus identifies the potential signal peptide, whereas the thin black lines located on the top of the illustration denote potential N-glycosylation sites. 2314delG, 2913delG, and 4353-54delCT mutations identified in USH2A patients are shown as solid arrows, and the positions of the resulting stop codons are shown as dashed arrows; a.a., amino acid.

the potential N-glycosylation sites all suggest that the USH2A protein is either a novel tissue-specific extracellular matrix protein or a cell adhesion molecule. Of potential relevance, two regions of the human retina are particularly rich in extracellular matrix proteins (*19*); these areas include Bruchs membrane, which is the specialized basement membrane underlying the retinal pigment epithelium (RPE), and the interphotoreceptor cell matrix (IPM). The IPM is in direct contact with both photoreceptors and the RPE and provides structural and trophic support for development and maintenance of the neural retina. Given that other extracellular matrix proteins and cell adhesion molecules are known to influence neural-glial interactions and to be involved in synapse development and stabilization, the USH2A protein may be involved in similar processes (*20*). Extracellular matrix proteins also play a fundamental role in the cochlea; several collagen genes, including *COL1A2*, *COL2A1*, and *COL3A1*, are highly expressed in the membranous labyrinth of the cochlea, both in connective tissue elements and other cell types, including inner hair cells (*21*). In addition, both autosomal and X-linked forms of Alport syndrome are due to mutations in collagen type IV genes (*22*). Furthermore, patients with Kallmann syndrome, a disorder characterized by anosmia and hypogonadism, often develop sensorineural deafness and occasionally ocular abnormalities (*23*). The gene for the X-linked form of this syndrome, *KAL1*, is thought to encode an adhesion molecule that contains four F3 repeats and may direct neuronal migration during development (*24*).

The *USH2A* gene, which expresses a putative extracellular matrix protein, and the *USH1B* gene, responsible for an unconventional myosin, do not appear to have obvious functional correlations; however, little is known about the physiological roles of either protein, and a common molecular pathological process could conceivably be etiologic in both forms of the syndrome.

Identification of the *USH2A* gene will lead to the development of a differential diagnostic tool for patients with Usher syndrome. This is particularly important in that one mutation, G2314del, seems to be especially frequent. Studies are currently under way to determine the specific tissue distribution of the USH2A protein in the cochlea and retina. Defining the molecular events that result in deafness and blindness may provide further insight into processes involved in the development and maintenance of the retina and cochlea, as well as into the physiology of vision and hearing.

REFERENCES AND NOTES

1. W. J. Kimberling and C. Moller, *J. Am. Acad. Audiol.* **6**, 63 (1995).
2. J. Boughman and G. Fishman, *Br. J. Ophthalmol.* **67**, 449 (1983).
3. B. Hallgren, *Acta Psychiatr. Scand. Suppl.* **138**, (1959).
4. L. Piassa *et al.*, *Arch. Ophthalmol.* **1041**, 1336 (1986); T. Rosenberg, *et al.*, *Clin. Genet.* **5**, 314 (1997).
5. J. Kaplan *et al.*, *Genomics* **14**, 979 (1992); W. J. Kimberling *et al.*, *ibid.*, p. 988; R. J. Smith *et al.*, *ibid.*, p. 995; S. Wayne *et al.*, *Hum. Mol. Genet.* **5**, 1689 (1996); H. Chaib *et al.*, *ibid.*, **6**, 27 (1997); W. J. Kimberling *et al.*, *Genomics* **7**, 245 (1990); S. Pieke Dahl *et al.*, *J. Med. Genet.* **30**, 843 (1993); E. M. Sankila *et al.*, *Hum. Mol. Genet.* **4**, 93 (1995).
6. D. Weil *et al.*, *Nature* **374**, 60 (1995).
7. A. el-Amraoui *et al.*, *Hum. Mol. Genet.* **5**, 1171 (1996).
8. J. Sumegi *et al.*, *Genomics* **35**, 79 (1996).
9. Fluorescent (6-carboxyfluorescein, tetrachlorofluorescein, or hexachlorofluorescein) oligonucleotide primers were used to amplify polymorphic markers around the *USH2A* locus, as described earlier (*8*). PCR products were pooled, denatured, and separated on 4.25% polyacrylamide gels with an ABI 377 automated sequencer running GENESCAN2.1 software. Genotyping data were collected and analyzed with GENEOTYPER 2.0. The order of the markers was obtained from the Genome Database at Johns Hopkins University (gdbwww.gdb.org/gdb/regionSearch.html) and the database at the Whitehead Institute for Biomedical Research/Massachusetts Institute of Technology (MIT) Center for Genome Research (www.genome.wi.mit.edu).
10. Screening of BAC libraries from Research Genetics and Genome Systems (St. Louis, MO) was performed with PCR primer pairs from the STSs (AFM248NC1, AFM143XF10, AFM268ZD1, AFM144XF2, WI-3484, WI-3128, WI-6236, and WI-9496) listed at Whitehead Institute for Biomedical Research/MIT Center for Genome Research. BACs were propagated and DNA extracted according to the manufacturer's recommendation. Rescue of the BAC insert end was performed by digestion of 1 μg of BAC DNA with Nhe I, followed by phenol-chloroform extraction and precipitation. DNA was religated and used to transform *Escherichia coli* DH5alpha. Colonies were propagated, the DNA was extracted and sequenced with IRD700-labeled T7 and SP6 primers, and the SequiTherm EXCEL II Long-Read Premix DNA Sequencing Kit (Epicentre Technologies) was optimized for use with the LI-COR 4000LS automated sequencer. We used the following PCR primer pairs, designed by the Oligo 5.0 primer analysis software program (NBI): 1STS4277: 5′-GTCCCTTCAGGAATGGATCTCC-3′/5′-TGCCTGTGACAAAGTCTGAGAACTG-3′ (product size 292 bp), 1STS1336: 5′-GGTCTTTGCATTGGTCACAACG-3′/5′-TCATGCCATACAACTGGTGCAG-3′ (product size 109 bp), 1STS526: 5′-CAATGAATTGCCTTCTTGTGCC-3′/5′-CTCTTTGGAGAACAGGACTGATCC-3′ (product size 267 bp), 1STS1337: 5′-TGGACTGGACGTTGTCTCAGCTTC-3′/5′-AGTTTCTGTCATTTCTGTCCTCGC-3′ (product size 164 bp), 1STS527: 5′-ATGGACCCTTAGTGCAGCCTTAGG-3′/5′-TGATTCTCTTTCGGTAGGAGTCTGC-3′ (product size 137 bp), 1STS1976: 5′-GGAACTGGATTTATTCTGCTCCTG-3′/5′-AATTGGGGGAATTTCGGGG-3′, (product size 100 bp), 1STS1977: 5′-CTCATCAAGAGCTGTCAGGGATTAG-3′/5′-GCACATGGGACAATCTTCAGATCAC-3′ (product size 288 bp). All PCR reactions were carried out in the same standard buffer condition at 55°C annealing temperature. BACs were sized by digestion of 1 μg of DNA with Not I followed by field inversion electrophoresis (FIGE).
11. R. Timpl and J. C. Brown, *Bioessays* **18**, 123 (1995).
12. BLAST analyses were performed with the use of the National Center for Biotechnology database (www.ncbi.nlm.nih.gov/BLAST/).
13. SM buffer (10 ml) was added to each of 100 agar plates (150 mm) containing 50,000 phages from a Clontech human retinal lambda gt10 library. Phages were eluted by gentle shaking on a rotary shaker overnight at 4°C. Portions (2 μl) from each plate lysate were analyzed by PCR with primer pairs 133hmr (5′-GGTCTTTGCATTGGTCACAACG-3′/5′-TCATGCCATACAACTGGTGCAG-3′) and 133hm2 (5′-GGAGGACAGTGTAATTGTAAGAGACACG-3′/5′-TGGTGACAGGTAATATCTCCATCCAC-3′). Samples were denatured at 95°C for 5 min and subjected to 35 cycles of 94°C for 30 s, 55° for 30 s, and 72°C for 1.5 min. Buffer conditions were those suggested by GIBCO-BRL. Individual clones were isolated from positive pools by standard Southern blot (DNA) hybridization. cDNAs were subcloned into the pBluescript KS$^+$ plasmid (Stratagene) and sequenced. 5′ RACE was performed with the 133hm B primer and human retinal RACE-ready cDNA from Clontech, following the manufacturer's suggestions. Products were subcloned into pCR2.1 (Invitrogen) and sequenced. cDNA clones and 5′ RACE products were mapped back to the BAC contig by Southern hybridization to ensure their authenticity. The cDNA sequence has been submitted to GenBank (accession number AF055580).
14. The following primer pairs were used to amplify products for mutation detection: 2186F 5′-TTAGGTGTGATCATTGCAATTTTGG-3′ and 2569B 5′-TCTTATCACAGTTGCAAGGCAGACA-3′ (product size 384 bp), 2869F 5′-TCATGCCATACAACTGGCGCAG-3′ AND 2977R 5′-GGTCTTTGCATTGGTCACAACGTTG-3′ (product size 109 bp), and 4266F 5′-GCTGTTGCACACTGCTAAATCCC-3′ and 4379R 5′-CTGGTCACACAACCAACTGAATTG-3′ (product size 114 bp). Thirty-seven cycles of 95°C for 30 s, 52°C for 30 s, and 72°C for 40 s were used to amplify genomic DNA from patients and controls under standard PCR buffer conditions. Products were subjected to heteroduplex analysis by means of previously established protocols and loaded onto vertical Mutation Detection Enhancement gels (FMC). Gels were stained with ethidium bromide and visualized under ultraviolet light. DNA fragments with altered mobility were sequenced directly by standard cycle sequencing reactions with ABI dye terminators and were analyzed with an ABI 377 sequencer.
15. Total RNA for the Northern blot analysis was isolated from human tissues (acquired from the National Disease Research Interchange, Philadelphia, PA) with Trizol reagent (GIBCO-BRL), separated on 1.2 to 1.4% agarose gels, and transferred to Hybond-N (Amersham). Northern blots were hybridized to ^{32}P-labeled CRT-1 cDNA probe, generated by Rediprime DNA-labeling kit (Amersham), in Express HYB solution (Clontech) according to standard procedures and the manufacturers' instructions. Total RNA from human fetal cochlea, eye, brain, and kidney and from adult retina, monkey retina, and Y79 retinoblastoma cells was extracted by the guanidine isothiocyanate method (J. M. Chirgwin, A. E. Przybyla, R. J. MacDonald, W. Rutter, *Biochemistry* **18**, 5294 (1979). Some (3 μg) of the total RNA was reverse transcribed with the SuperScript II kit according to the manufacturer's recommendations (GIBCO-BRL). To evaluate genomic contamination, we performed parallel PCR reactions with and without RT. Portions from the +/−RT reactions were used as the templates for PCR. The following PCR primers were used: F11B11 5′-GGGATTTAGCAGTGTGCAACAGC-3′/5′-GGAAAAGTTGTGGGGTATGACATC-3′ (product size 100 bp) 1STS 1336 5′-GGTCTTTGCATTGGTCACAAGG-3′/5′-TCATGCCATACAACTGGTGCAG-3′ (product size 109 bp).
16. J. D. Eudy *et al.*, unpublished data.
17. The Paircoil program (MIT) and TMpred (ISREC, Lausanne, Switzerland) were accessed through the Baylor College of Medicine Search Launcher (http://gc.bcm.tmc.edu:8088/search-launcher/launcher.html).
18. J. Vielmetter *et al.*, *Genomics* **41**, 414 (1997); M. Streuli, N. X. Krueger, A. Y. Tsai, H. Saito, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 8698 (1989).
19. T. E. Ogden, *Retina* (Mosby, St. Louis, MO, ed. 2, 1994), pp. 58–71.
20. U. Bartsch, F. Kirchhoff, M. Schachner, *J. Neurocy-*

tol. **19**, 550 (1990).
21. N. G. Robertson *et al.*, *Genomics* **23**, 42 (1994).
22. D. F. Barker *et al.*, *Science* **248**, 1224 (1990); T. Mochizuki *et al.*, *Nature Genet.* **8**, 77 (1994).
23. R. Legouis *et al.*, *Cell* **67**, 423 (1991).
24. N. Soussi-Yanicostas *et al.*, *J. Cell Sci.* **109**, 1749 (1996).
25. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
26. This work was supported by a grant from NIH-NIDCD 5PO1 DC01813-05, by the Nebraska Research nitiative Fund, and by a grant from the Foundation Fighting Blindness. We extend our sincere gratitude to the patients and their families involved in the study for their assistance and cooperation. We thank M. Gillett for assistance. A.S. is a recipient of a Research to Prevent Blindness Lew R. Wasserman Merit Award and supported by EY07003 (CORE). C.C.M. is supported by the John Alden Trust and NIH-NIDCD DC03402. The retina EST was provided by the IMAGE Consortium. We thank J. Edwards for artwork and V. Wrobleski for preparation of the manuscript.

29 December 1997; accepted 1 May 1998

Promotion of Met-tRNA$_i^{Met}$ Binding to Ribosomes by yIF2, a Bacterial IF2 Homolog in Yeast

Sang Ki Choi, Joon H. Lee, Wendy L. Zoll, William C. Merrick, Thomas E. Dever*

Delivery of the initiator methionine transfer RNA (Met-tRNA$_i^{Met}$) to the ribosome is a key step in the initiation of protein synthesis. Previous results have indicated that this step is catalyzed by the structurally dissimilar translation factors in prokaryotes and eukaryotes—initiation factor 2 (IF2) and eukaryotic initiation factor 2 (eIF2), respectively. A bacterial IF2 homolog has been identified in both eukaryotes and archaea. By using a combination of molecular genetic and biochemical studies, the *Saccharomyces cerevisiae* IF2 homolog is shown to function in general translation initiation by promoting Met-tRNA$_i^{Met}$ binding to ribosomes. Thus, the mechanism of protein synthesis in eukaryotes and prokaryotes is more similar than was previously realized.

A universally conserved step in gene expression is the initiation of protein synthesis at an AUG codon by using a specific initiator tRNA^{Met} (tRNA$_i^{Met}$). In prokaryotes, the methionine linked to the tRNA$_i^{Met}$ is formylated (fMet-tRNA$_i^{Met}$), and the translation factor IF2 facilitates the AUG codon-dependent binding of fMet-tRNA$_i^{Met}$ to the small (30S) ribosomal subunit (*1*). Selection of the correct AUG initiation codon is mediated by base-pairing interactions between 16S ribosomal RNA (rRNA) and the Shine–Dalgarno sequence of complementary nucleotides located 5′ of the start site. In eukaryotes, the heterotrimeric translation factor eIF2 forms a stable ternary complex with guanosine triphosphate (GTP) and Met-tRNA$_i^{Met}$, which can then stably bind to 40S ribosomal subunits (*2*). This 43S preinitiation complex binds near the 5′ end of an mRNA and then scans in a 3′ direction and selects the AUG start site through base-pairing interactions between the AUG codon of the mRNA and the anticodon of the tRNA$_i^{Met}$ (*3*). Several translation factors, including eIF2, are also thought to play an important role in AUG start site selection (*4*).

IF2 from *Escherichia coli* is a single polypeptide chain of 97.3 kD and contains a consensus GTP binding domain near the center of the protein. *Saccharomyces cerevisiae* has an open reading frame (ORF) on chromosome I (YAL035w) that is similar to prokaryotic IF2 proteins (*5*). We refer to the protein product of this gene, previously designated *FUN12* (*6*), as yIF2 for yeast IF2. The full-length 1002–amino acid yIF2 protein sequence shows 27% identity and 48% similarity with IF2 from *E. coli*. An IF2 homolog, as well as homologs of all three eIF2 subunits, were also identified in the archaeon *Methanococcus jannaschii* (*7*). The archaeal protein is truncated at its NH_2-terminus; however, the NH_2-terminal regions of *E. coli* IF2 and yIF2 are dispensable for cell viability (*8, 9*). The discovery of IF2 homologs in all three kingdoms suggests a greater conservation in the mechanism of translation initiation than had been anticipated. We disrupted one copy of *FUN12* in a diploid yeast strain (*10*) and identified, by tetrad analysis, two spores with wild-type growth rates and two *fun12*Δ spores with a severe slow-growth phenotype on rich medium. Therefore, *FUN12* is critically important but nonessential for growth in yeast. Mitochondrial protein synthesis is very similar to prokaryotic translation, and null mutations in the yeast gene *IFM1* encoding mitochondrial IF2 cause a respiratory-deficient phenotype (*11*). In contrast, the slow-growing *fun12*Δ strains were able to grow on nonfermentable carbon sources (*9*), demonstrating that *FUN12* is not required for mitochondrial protein synthesis. To determine whether the slow-growth phenotype in the *fun12*Δ strains resulted from a defect in cellular translation initiation, we analyzed polysome profiles by velocity sedimentation of cell extracts in sucrose gradients (*12*). As expected, distinct 40S and 60S ribosomal subunit peaks, as well as 80S monosome and polysome peaks, could be detected in the wild-type *FUN12* extract (Fig. 1A, upper panel). The isogenic *fun12*Δ strain showed a dramatic reduction in the larger polysomes and a corresponding increase in the 80S monosome peak (Fig.1A, lower panel). The polysome-to-monosome ratio in the *fun12*Δ strain was reduced by 70% compared with the ratio in the isogenic wild-type strain (Fig. 1A). The 80S monosomes that accumulated in the *fun12*Δ strain appear to be inactive 80S particles composed of 40S and 60S subunits but lacking mRNA because, when these extracts were sedimented in high-salt sucrose gradients, the 80S particles dissociated into 40S and 60S subunits (*9*). In contrast, the 80S monosomes in the wild-type *FUN12* extract were relatively stable in the high-salt gradients (*9*), as expected for translationally active 80S ribosomes (*13*). The reduction in large polysomes and the increase in 80S monosomes in the *fun12*Δ strain are indicative of a translation initiation defect. Consistent with yIF2 functioning in general translation initiation, indi-

Table 1. Stimulation of first peptide bond synthesis by yIF2 and eIF2 in a reconstituted system from rabbit reticulocytes. The AUG-directed synthesis of methionyl-puromycin was assayed as described in (*16*). Each reaction mixture contained ribosomes and, where indicated, translation factors (eIF1A, eIF5, and eIF5A) purified from rabbit reticulocytes. The GST-yIF2 fusion protein (5 μg) and the GST control protein (5 μg) were purified from yeast crude extracts (*15*); eIF2 (5 μg) was purified from rabbit reticulocytes. The GST-yIF2 fusion protein lacks the NH_2-terminal region of yIF2 with the GST protein fused directly to the GTP binding domain of yIF2. Results shown are typical of three independent assays.

Additions	Methionyl-puromycin synthesis (pmol)
eIF2 (alone)	0.05
eIF2 + factors	0.59
GST-yIF2 (alone)	0.08
GST-yIF2 + factors	0.54
GST + factors	0.02

S. K. Choi, J. H. Lee, T. E. Dever, Laboratory of Eukaryotic Gene Regulation, National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD 20892–2716, USA.
W. L. Zoll and W. C. Merrick, Department of Biochemistry, Case Western Reserve University, Cleveland, OH 44106, USA.

*To whom correspondence should be addressed: E-mail: tdever@box-t.nih.gov