J Biol Chem. Author manuscript; available in PMC 2011 August 18.

Published in final edited form as:
J Biol Chem. 2006 December 15; 281(50): 38172–38180. doi:10.1074/jbc.M606010200.

Mammalian Transcription in Support of Hybrid mRNA and Protein Synthesis in Testis and Lung[*,S]

Carolyn Fitzgerald[‡], **Curtis Sikora**[‡], **Vannice Lawson**[§], **Karen Dong**[‡], **Min Cheng**[‡], **Richard Oko**[§], and **Frans A. van der Hoorn**[‡,1]

[‡]Department of Biochemistry and Molecular Biology, University of Calgary, Calgary, Alberta T2N 4N1, Canada

[§]Department of Anatomy and Cell Biology, Queen's University, Kingston, Ontario K7L 3N6, Canada

Abstract

Post-transcriptional mechanisms including differential splicing expand the protein repertoire beyond that provided by the one gene-one protein model. Trans-splicing has been observed in mammalian systems but is low level (sometimes referred to as noise), and a contribution to hybrid protein expression is unclear. In the study of rat sperm tail proteins a cDNA, called 1038, was isolated representing a hybrid mRNA derived in part from the ornithine decarboxylase antizyme 3 (*Oaz3*) gene located on rat chromosome 2 fused to sequences encoded by a novel gene on chromosome 4. Cytoplasmic *Oaz3* mRNA is completely testis specific. However, in several tissues *Oaz3* is transcribed and contributes to hybrid 1038 mRNA synthesis, without concurrent *Oaz3* mRNA synthesis. 1038 mRNA directs synthesis of a hybrid 14-kDa protein, part chromosome 2- and part chromosome 4-derived as shown *in vitro* and in transfected cells. Antisera that recognize a chromosome 4-encoded C-terminal peptide confirm the hybrid character of endogenous 14-kDa protein and its presence in sperm tail structures and 1038-positive tissue. Our data suggest that the testis-specific *OAZ3* gene may be an example of a mammalian gene that in several tissues is transcribed to contribute to a hybrid mRNA and protein. This finding expands the repertoire of known mechanisms available to cells to generate proteome diversity.

The mammalian sperm tail contains unique structures not present in cilia, viz. the outer dense fibers (ODF)[2] and the fibrous sheath (FS). In the midpiece region of spermatozoa ODF are each associated with one microtubule doublet of the axoneme. In the sperm tail principal piece, two ODF are replaced by the longitudinal columns of the FS that are connected throughout the principal piece by ribs (1–3). ODF are thought to provide elastic recoil for tail movement (4), and FS acts as scaffold for components involved in signaling pathways (5). Several major ODF proteins, including Odf1, Odf2, and Spag4, have been cloned by us (6–8) and others (9–11), and we proposed that many of them specifically

---

[*]This work was supported by grants from the Canadian Institutes of Health Research (to F. A. v. d. H. and R. O.).

[S]The on-line version of this article (available at http://www.jbc.org) contains supplemental Figs. S1–S3.

[1]To whom correspondence should be addressed: Dept. of Biochemistry and Molecular Biology, University of Calgary, 3330 Hospital Dr. NW, Calgary, Alberta T2N 4N1, Canada. Tel.: 403-220-4243; Fax: 403-210-8109; fvdhoorn@ucalgary.ca.

[2]The abbreviations used are: ODF, outer dense fiber; FS, fibrous sheath; RT-PCR, reverse transcriptase PCR; HA, hemagglutinin.

interact through dimerization motifs (7). For example, we showed that Spag4 precedes synthesis of other ODF proteins, binds to the forming axoneme, and likely acts as starting point for ODF development by attracting Odf1, followed by several other ODF proteins (12). ODF in turn bind ODF-associated proteins, including Spag5, which plays multiple roles both in germ cells and in somatic cells (13, 14), and KLC3, a kinesin light chain motor protein (15, 16). Analysis of major ODF proteins indicated that several appear related and are encoded by differentially spliced mRNAs. For example, differentially spliced mRNAs encoding Spag5 (14) and Odf2 (17) variants were detected. Clearly, differential splicing expands the protein repertoire available to spermatids and can result in production of functionally different but related proteins. Indeed, it has been estimated that 40–60% of human genes are alternatively spliced (18).

Few reports have been published demonstrating mammalian trans-splicing, and results from these data support the notion that trans-splicing may be a rare occurrence (19–22); for example, the amount of trans-spliced *Cyp3A43/3A4* transcript is estimated at 0.15% of that of *Cyp3A43* mRNA (23). In most instances trans-spliced mRNAs are predicted to encode a regular, not hybrid, protein (20, 24). Recently, trans-spliced human *ACAT1* mRNA was shown to direct synthesis of a hybrid protein under experimental transfection conditions (25). Elements and exonic splicing enhancer sequences that play a role in trans-splicing have been identified (26), and trans-splicing was shown to involve both regular and cryptic splice sites. Based on current data on mammalian trans-splicing it has been suggested that this process might resemble a level of noise associated with normal nuclear RNA processing (27) of uncertain functional importance.

Here we report that an integral component of sperm tail ODF and FS, a 14-kDa protein, is a hybrid protein encoded by a hybrid mRNA. The hybrid mRNA has contributions from the spermatid-specific *Oaz3* gene (28) located on rat chromosome 2 and a novel gene on rat chromosome 4. *Oaz3* (previously called *Oaz-t*) (29) belongs to the family of ornithine decarboxylase antizyme genes that also include ubiquitous somatic *Oaz1* and *Oaz2* genes. Translation of all OAZ mRNAs is induced by polyamines and depends on translational frameshift at the same TCCTGA sequence (30). In addition, we discovered that other tissues, including lung, transcribe these two genes to give rise to detectable levels of the hybrid mRNA and its hybrid protein, but not of normal spliced *Oaz3* mRNA.

EXPERIMENTAL PROCEDURES

Male Germ Cell and Component Isolation

Rat male germ cells, including pachytene spermatocytes, round spermatids, and elongating spermatids, were isolated by centrifugal elutriation as described previously (31). Rat ODF and FS were isolated from tails of rat epididymal spermatozoa as described previously (1).

Antibodies

Polyclonal antibodies were raised against peptide YQNQQLEKESRSSQEHLC, located in the chromosome 4-specific part of 1038 protein. Affinity purification of anti-1038 antibodies was done using peptide covalently coupled to agarose beads. Polyclonal antibodies were also

raised against total rat ODF and FS proteins, and anti-14 antibodies were isolated by affinity purification using filters containing immobilized 14-kDa protein as described previously (12).

Molecular Cloning of 1017 and 1038 cDNAs

Affinity-purified anti-14 antiserum, which recognizes the 14-kDa polypeptide in ODF and FS, was used to screen a rat testicular λ Zap II cDNA library (Stratagene, La Jolla, CA) following the method of Young and Davis (32). The phagemids of six positive clones were excised by transfecting them with helper phage into XL-1 Blue cells, and the resulting secreted and circularized plasmids were transformed into SOLAR cells with ampicillin selection. The selected colonies were plasmid isolated and sequenced by Cortec Service (Queen's University, Kingston, ON) using the ABI PRISM™ Dye Terminator Cycle Sequencing kit with AmpliTag® DNA polymerase. Representative clones of 1017 and 1038 were chosen for second and third round sequencing analysis in which both strands were sequenced completely.

Rat-Hamster Radiation Hybrid Mapping

Radiation hybrid mapping was PCR based using genomic DNA from hybrid cells of the T55 panel (5255 genetic markers) as template at the Canadian Institutes of Health Research Genome Resource Facility, The Centre for Applied Genomics at the Hospital for Sick Children in Toronto, Ontario, Canada. The T55 panel contains a map of the rat genome using 5255 genetic markers. Genomic location of the 3′-end of 1038 cDNA was mapped using primers 5′-GGAGAGAATGTGGGGAATACCAG-3′ and 5′-GCCTCTAGCCATGTTCCTCAAGC-3′, which generate a 346-bp PCR fragment. The 5′-1038 cDNA region was mapped using primers 5′-GAGAAACTGCCTTGTACCAG-GTCC-3′ and 5-′CCCTGCCACTTAGGAAATTCTCTC-3′, which generate a 336-bp PCR fragment. Twenty-five nano-grams of hybrid DNA was used in PCR analysis. For linkage analysis, PCR typing results were submitted to the Otsuka GEN Research Institute data base (ratmap.ims.u-tokyo.ac.jp/menu/RH.html).

RNA Isolation

RNA was isolated from 100 mg of rat tissues and from elutriated male germ cells. Tissue was homogenized at 4 °C in 1 ml of homogenization buffer (10 mM HEPES, pH 8, 50 mM NaCl, 0.5 M sucrose, 1 mM EDTA, pH 8, 0.5 mM spermidine, 0.15 mM spermine tetrachloride). The extract was filtered and centrifuged at 800 × *g* for 10 min at 4 °C to pellet nuclei. The pellet was washed three times in 1 ml of homogenization buffer and resuspended in 1 ml of TRIzol (Invitrogen) to isolate nuclear RNA. The supernatant was used for cytoplasmic RNA extraction using TRIzol. All nuclear RNA preparations were incubated with RNase-free DNase to remove any contaminating DNA.

RT-PCR

RNA was transcribed into cDNA using the M-MLV RT kit (Invitrogen); for each tissue, 2 μg of cytoplasmic or nuclear RNA was used with 100 ng of random primers to generate cDNA. For each primer set, 35 cycles of PCR were carried out under the following conditions: a

denaturing step for 30 s at 94 °C, an annealing step for 30 s at 54 °C, and an elongation step for 1 min at 72 °C. For PCR the following primers were used: pr678, 5′-GAGAAACTGCCTTGTACCAGGT-CC-3′; pr438, 5′-GCCTCTAGCCATGTTCCTCAAGC-3′; pr117a, 5′-GTGGAAGTCTAACTGAACTGAACTGGGTC-TTG-3′; β-actin, 5′-CAACACCCCAGCCATGTACG-3′ and 5′-AGGAAGAGGATGCGGCAGTGG-3′; Odf1, 5′-CATCC-CGGGTCTATTTGTCGTCCTTCTGAG-3′ and 5′-CATGT-CGACTAAATGGCCGCACTGAGTTGTCTTTTG-3′. PCR fragments were analyzed at the University of Calgary Core DNA Services.

Mass Spectrometry

Gel-separated proteins were submitted to the Southern Alberta Mass Spectrometry facility (University of Calgary). Proteins were trypsinized, treated with or without iodoacetamide, and initially analyzed by matrix-assisted laser desorption ionization on a Voyager-DE STR (Applied Biosystems). Selected peptide sequences were obtained by liquid chromatography-tandem mass spectroscopy using an Agilent Nano LC chromatographic system and a QSTAR Pulsar iHybrid quadrupole time-of-flight mass spectrometer (Applied Biosystems).

Protein Analysis

Western blot analysis was carried out essentially as described previously (12). In short, proteins were boiled in loading buffer, separated on 10–18% acrylamide SDS-PAGE gradient gels, transferred onto a polyvinylidene fluoride membrane (Amersham Biosciences), blocked overnight at 4 °C in blocking buffer (54 mM Tris, pH 7.5, 150 mM NaCl, 0.1% Nonidet P-40, 0.1% Tween 20, 5% dry nonfat milk), and analyzed using primary antibodies (anti-1038 dilution, 1/100; anti-14 dilution, 1/200) followed by goat anti-rabbit IgG antibody (Sigma) diluted 1/12,000. Prestained Protein Ladder SM0671 (Fermentas) was used as a size marker. LumiGLO substrate (Kirkegaard & Perry Laboratories, Inc.) was used to develop the blot. The image was captured using a 3000 Versa-Doc Imaging System (Bio-Rad).

In Vitro Translation and Immunoprecipitations

HA-tagged 1038 protein was produced *in vitro* using the TNT system (Promega) in the presence of [^{35}S]methionine as described (12), and 25 μl of *in vitro* translated HA or HA-1038 fusion protein were incubated with anti-1038, anti-14, and anti-HA antibodies. Immune complexes were precipitated using Protein A-Sepharose CL-4B beads (Pharmacia Biotech), washed three times with wash buffer (50 mM Tris, pH 8), and analyzed by SDS-PAGE and autoradiography as described (12).

Transfection

Constructs expressing HA or HA-1038 fusion protein were transfected into human 293 cells and mouse 3T3 cells using the calcium phosphate co-precipitation method. 36 h after transfection proteins were extracted from these cells and complexes were immunoprecipitated using anti-1038, anti-HA, and anti-14 antibodies. Immunoprecipitates were analyzed by Western blotting using the indicated antibodies as described above.

Immunoelectron Microscopy

The LR-white-embedded rat testicular and epididymal tissues were processed for immunogold labeling by using previously described procedures (33). Briefly, the ultra thin sections were mounted on formvar-coated nickel grids and blocked with 10% normal goat serum prior to overnight incubation of the affinity-purified anti-1038 antibodies at 4 °C. The sections were then washed and incubated with goat anti-rabbit secondary antibody conjugated to 10-nm gold particles (1:20; Sigma) followed by counterstaining with uranyl acetate and lead citrate. In indicated instances anti-1038 antibody was preincubated with C-terminal 1038-specific peptide bound to Sepharose beads before use in electron microscopy. The final sections were analyzed by transmission electron microscope (Hitachi 7000).

RESULTS

Rat ODF and FS contain several major proteins, one of which is 14,000 in size. We raised an antiserum against all major ODF proteins from which an anti-14 antibody was affinity purified with specificity for the 14-kDa protein. Screening a rat testicular cDNA expression library with affinity-purified anti-14 antibodies resulted in isolation of two cDNAs, 1017 (GenBank™ DQ431008) and 1038 (GenBank™ DQ431007). Sequence analysis demonstrated that they have identical 5′, but different 3′, sequences. We determined that 1017 cDNA is encoded by the rat *Oaz3* (28) gene, located on chromosome 2 (Fig. 1*A*). Analysis of 1038 cDNA demonstrated that it is composed of sequences derived from *Oaz3* exon 1 and part of exon 2 (Fig. 1*A*, indicated as *exon* Δ*2*) and from sequences on rat chromosome 4 (Fig. 1*A*). No genes have yet been described or predicted for this region. This suggested that 1038 mRNA is a hybrid mRNA. Northern blotting using 1017- and 1038-specific probes confirmed the presence of the two mRNAs in both rat and mouse (supplemental Fig. S1).

To verify the 1038 cDNA bioinformatics result and exclude cloning artifacts, we screened genomic DNA from rat-hamster radiation hybrid cell panel T55 using primer pairs that specifically detect rat sequences representing either the 5′-region shared by 1017 and 1038 cDNAs or the 1038-specific 3′-region. Fig. 1*B* shows that 1038-specific genomic sequences are located in close proximity to four rat chromosome 4 markers (*4.020-D4Rat22*, *4.019-D4Rat20*, *4.021-D4Rat96*, and *4.018-D4Got39*) that are at the location predicted for 1038 (position 61.437). As expected, the 5′-sequences shared between 1017 and 1038 mapped to the known location of the rat *Oaz3* gene on chromosome 2 (Fig. 1*B*, markers *2.039-D2Rat134*, *2.040-D2Mgh9*, and *8.022-D8Got54*). We conclude that 1038 mRNA is a genuine hybrid transcript. Analysis of the crossover point between chromosome 2 and 4 sequences in 1038 mRNA and comparison to the corresponding chromosome 2 and chromosome 4 genomic sequences showed that formation of the hybrid 1038 mRNA appears to involve a cryptic 5′-splice site on chromosome 2 (AG CT) and a canonical 3′-splice site on chromosome 4 (CCAGCCAG CA). The *Oaz3* exon 2 5′-splice site (AG TA) is used in formation of 1017 (*Oaz3*) mRNA. These results suggest, but do not prove, that trans-splicing may be involved in 1038 hybrid mRNA production.

Hybrid 1038 mRNA Expression in Postmeiotic Male Germ Cells

To demonstrate expression of the novel, hybrid 1038 mRNA in male germ cells we used Northern blot and RT-PCR approaches. First, developmental Northern blots using specific riboprobes showed that trans-spliced 1-kb 1038 mRNA and normal spliced 1-kb 1017 mRNA can first be detected in rat testis starting at postnatal day 31 (Fig. 2*A*, *lane 2*, *panels I* and *II*), a time corresponding to the first appearance of spermatids. No 1038 mRNA is expressed at day 27 (*lane 1*), corresponding to late pachytene spermatocytes. Second, hybrid 1038 mRNA spermatid expression was demonstrated by RT-PCR using cytoplasmic RNA from fractionated pachytene spermatocytes, round spermatids, and elongating spermatids (Fig. 2*B*, *lanes 2* and *3*). PCR products observed using spermatocyte RNA (*lane 1*) result from the presence of small numbers of spermatids in the spermatocyte fraction, because primers specific for *Odf1*, a gene exclusively expressed in spermatids (12, 34), generated *Odf1* PCR fragment using spermatocyte RNA. Thus hybrid 1038 mRNA expression in testis appears restricted to postmeiotic cells.

OAZ3 Transcription in Support of Hybrid 1038 mRNA Synthesis

Oaz3 expression had been suggested to be testis specific (28). We analyzed 1017 (*Oaz3*) mRNA and 1038 mRNA expression in various tissues by RT-PCR of nuclear and cytoplasmic RNA using 1038-specific primers and 1017-specific primers (locations as shown in Fig. 3). The result of cytoplasmic RNA analysis shows that 1017 (*Oaz3*) mRNA is exclusively expressed in testis (Fig. 3, *panel cytoplasmic RNA*, *lane 5*), confirming the above suggestion (28). Unexpectedly however, we observed cytoplasmic hybrid 1038 RNA in kidney, small intestine, lung, brain, and ovary (*panel cytoplasmic RNA*, *lanes 1*, *4*, *6*, *8*, and *9*). Because 1038 hybrid mRNA derives in part from *Oaz3* gene sequences, this result indicates that *Oaz3* must be actively transcribed in these tissues. Analysis of nuclear RNA using the primers confirmed that indeed *Oaz3* nuclear pre-mRNA can be detected in all 1038 mRNA-positive tissues (Fig. 3, *nuclear RNA panel*, *1248 bp DNA*, *lanes 1*, *4*, *6*, *8*, and *9*) indicative of transcription of the *Oaz3* gene. Several tissues express neither transcript.

1038 mRNA Encodes a Hybrid 14-kDa Protein in Transfected Cells

Sequence analysis and comparison of 1017 and 1038 cDNAs show that at the position where 1017 mRNA has 5 adenosine residues (position 68–72), 1038 mRNA has 6 adenosine residues (position 68–73) (Fig. 1*D*, *boxed A residue*). We confirmed by genomic PCR that the *Oaz3* gene encodes 5 adenosine residues at that position, not 6 (in agreement with the *Oaz3* GenBank™ NM_016901 entry). We also determined by RT-PCR that whereas nuclear 1017 pre-mRNA has 5 adenosine residues, nuclear trans-spliced 1038 RNA has 6 adenosine residues. This suggests that an unknown nuclear process adds 1 adenosine residue to hybrid 1038 RNA before transport to the cytoplasm, which addition affects the reading frame from that point on (see Fig. 1*D*). We also determined by N-terminal sequence analysis of purified rat sperm tail 14-kDa protein that the N-terminal sequence is PCTR (not shown), indicative of a translation start at the Leu codon (Fig. 1*D*, *boxed CTG codon*). Translation starts at the Leu codon and other non-AUG codons have been reported in several instances (reviewed in Ref. 35). Based on translation start at the Leu codon, the additional adenosine residue and the (previously published; Ref. 28) occurrence of a translational frameshift at the sequence

UCCUGA present in both 1038 and 1017 mRNA (Fig. 1*D*, *boxed TCCTGA sequence*), hybrid 1038 mRNA is predicted to encode a hybrid 14-kDa protein whose N-terminal 105 amino acids are encoded on chromosome 2 and its C-terminal 33 amino acids by the novel gene on chromosome 4 (Fig. 1*E*).

We sought first to establish that under experimental conditions hybrid 1038 mRNA can direct synthesis of the predicted 14-kDa hybrid protein. For this we generated antibodies (anti-1038) raised against a chromosome 4 1038-specific peptide (Fig. 1*E*, *boxed residues in bold*). In indicated experiments we also employed anti-14 antibodies used to isolate the 1038 cDNA clones. First, HA-1038 cDNA (encoding a HA-tagged 1038 fusion protein of 15 kDa) directs synthesis of a 15-kDa protein as shown by immunoprecipitation using anti-1038 and anti-14 antisera (Fig. 4*A*, *lanes 1* and *4*). Preincubation of anti-1038 antiserum with peptide blocks immunoprecipitation (*lanes 2* and *3*). Second, we transfected human 293 and mouse 3T3 cells with an HA-1038 expression construct and with an HA construct lacking the insert. Extracts from transfected cells were investigated by immunoprecipitation-Western blot analysis. Fig. 4*B* shows that HA-1038 cDNA directs synthesis of a hybrid 15-kDa protein recognized by anti-HA and by anti-1038 antibodies in both 293 cells (*lanes 2* and *6*) and 3T3 cells (*lane 4*). In addition, as expected, the HA-1038-directed 15-kDa protein is recognized by anti-14 antiserum (*lane 8*). HA controls lack this protein (*lanes 1*, *3*, *5*, and *7*).

We observed that the translational frameshift does not, however, occur in bacteria; a construct linking 1038 mRNA to maltose-binding protein (MBP) coding sequences, called MBP-1038, was expressed in bacteria and the resulting protein isolated and its sequence analyzed by tandem mass spectrometry. The mass spectrometry sequencing results (supplemental Fig. S2) established that the amino acid sequence of bacterially produced 1038 was in accordance with the sequence predicted based on the presence of 6 adenosine residues, not 5, but that the translational frameshift did not occur: maltose-binding protein 1038 was shorter than expected and is recognized by anti-14, but not by anti-1038, antibodies.

Hybrid 14-kDa Protein Localizes to ODF and FS

We find that in testis hybrid 1038 mRNA is a postmeiotic transcript, and we sought to detect expression of the corresponding hybrid 14-kDa protein in rat ODF and FS by immunoelectron microscopy of rat epididymal sperm and by Western blotting of purified ODF and FS preparations using affinity-purified anti-1038 antiserum. The results show that in the midpiece and principal piece of spermatozoa gold label was detected over ODF and FS. Fig. 5*A* shows staining of several principal piece (*PP*) sections and one midpiece (*MP*) section. Fig. 5*B* shows two midpiece sections with gold label predominantly over ODF as well as two endpiece sections. Additional higher power images are presented in supplemental Fig. S3). Preincubation of anti-1038 antiserum with peptide abolished gold label (Fig. 5*C* shows two examples), and use of preimmune serum did not result in any gold labeling of these structures, as expected. Western blotting demonstrated that anti-1038 antibodies recognize the 14-kDa (1038) protein in both ODF and FS (Fig. 5*D*, *lanes 1* and *2*, respectively), more so in ODF. Preimmune sera did not detect 14-kDa protein when used in

Western blotting experiments (not shown). Relative intensity on Westerns and abundance of gold label over ODF compared with FS are in agreement.

Hybrid 14-kDa Protein Expression in Lung

Several tissues, including lung, express hybrid 1038 mRNA, but not normal spliced 1017 (*Oaz3*) mRNA. First, we used Western blot analysis of lung, testis, and liver protein extracts and showed that lung contains hybrid 14-kDa 1038 protein (Fig. 6*A*, *lane 3*). As expected from the immunoelectron microscopic results, testis expresses the hybrid 14-kDa protein (Fig. 6*A*, *lane 2*). Reprobing these blots, after complete stripping, with anti-14 antiserum detected the same protein band, indicating that the 14-kDa protein contains epitopes recognized by both anti-14 and anti-1038 sera as expected for the hybrid protein. 1038 protein is not expressed in liver (Fig. 6*A*, *lane 1*), a tissue negative for 1038 mRNA (Fig. 3, *lane 7*). Second, we covalently coupled anti-1038 antibodies to Sepharose beads, incubated these beads with protein extracts from testis, lung, and liver, and analyzed the bound complexed proteins by Western blotting using anti-14 antiserum. The results (Fig. 6*B*) show that the 14-kDa hybrid protein is complexed to immobilized anti-1038 antibodies and detected with anti-14 antiserum in testis and lung, but not in liver.

DISCUSSION

In this work we have documented the expression of a hybrid 1038 mRNA and protein in postmeiotic male germ cells. In addition we observed that several rat tissues express hybrid 1038 mRNA without concurrent accumulation of 1017 (*Oaz3*) mRNA. This indicates that in these latter tissues, the *Oaz3* gene is transcribed for the purpose of contributing to a hybrid mRNA, rather than expression of normal spliced *Oaz3* mRNA. To our knowledge this is the first demonstration that a mammalian gene can be transcribed in one tissue (testis) to give rise to both normal and hybrid mRNA species but in other tissues only contributes to the hybrid mRNA.

We have demonstrated that 1038 mRNA is a fusion of sequences originating from *Oaz3* on rat chromosome 2 and a novel gene on chromosome 4. In addition, we showed that the hybrid mRNA encodes a true hybrid 14-kDa protein that is a significant constituent of two important sperm tail structures, ODF and FS. The role of this hybrid protein, as indeed of other ODF proteins, remains to be determined, but ODF have been suggested to play a role in elastic recoil (4). Of note, *Drosophila* sperm contains structures similar to ODF: products of the testis-specific *Mst(3)CGP* family contribute to those structures and mutation of *Mst(3)CGP* genes renders animals infertile (36, 37).

Mammalian trans-splicing has been documented in a number of instances. In these reports the trans-spliced transcripts are rare (sometimes described as noise) (27) and usually involve fusion of an untranslated region derived from one gene spliced to a region containing a complete coding sequence derived from another gene. In addition, in previous reports cells displaying low levels of trans-spliced mRNAs invariably express higher levels of corresponding normal spliced mRNAs (23). A recent review suggested that trans-splicing is perhaps an aberrant process not intended to produce relevant hybrid proteins (27), and new bioinformatics algorithms to predict alternative splicing also suggest trans-splicing results

from gene level abnormalities (38). We do not know the underlying mechanisms involved in synthesis of hybrid 1038 mRNA. We do not believe it involves recombination of genomic DNA in, for example, a subset of germ cells, because hybrid protein is present in all elongating spermatids (as shown in Fig. 5) and genomic PCR using appropriate primer combinations failed to reveal evidence for genomic DNA rearrangements (not shown). Sequence analysis of the involved genomic regions on chromosomes 2 and 4 and comparison to the 1038 mRNA sequence suggested the use of a cryptic 5′-splice site on chromosome 2 (5′-AG CT-3′) and a consensus 3′-splice site on chromosome 4 (5′-CCAGCCAG CA-3′) to generate 1038 hybrid mRNA (Fig. 1, *C* and *D*). This suggested the involvement of trans-splicing mechanisms, but we cannot rule out other RNA recombination events. Use of non-canonical splice sites in formation of trans-spliced mRNA had been reported previously (20, 24). We also showed here that the normal exon 2 donor splice site (5′-AG TA-3′) is used in generation of normal spliced 1017 (*Oaz3*) mRNA (Fig. 1*C*).

Several interesting implications emerge from our findings. First, the enzymatic activity of OAZ3 protein, an exclusively postmeiotic male germ cell-specific protein, is known. Because the ~100 amino acid residues in the C-terminal part of OAZ3 are replaced by 33 amino acids derived from chromosome 4, one can predict a novel (currently unknown) function for the hybrid 14-kDa protein. Of note, however, this replacement did not alter the subcellular localization; both 1017 (OAZ3) protein (not shown) and 1038 (14 kDa) protein localize to ODF and FS. We had demonstrated previously that many ODF proteins specifically interact using dimerization motifs resembling leucine zippers (reviewed in Ref. 39); it is interesting to note that the 1017-encoded 14-kDa protein contains a motif resembling a leucine zipper (L-N_6-I-N_6-L-N_6-IN_6-L) that may well interact with the functional Odf1 or Odf2 leucine zippers. In contrast, the 1038-encoded 14-kDa protein does not encode such a predicted motif. In this regard it is also noteworthy that those tissues that only express hybrid 1038 mRNA have a common endodermal origin, except brain. More importantly, 1038 mRNA-positive tissues contain ciliated cells, and it will be interesting to determine the subcellular localization of the hybrid 14-kDa protein to ascertain a possible role during development or in cilia. Be that as it may, the characterization of the hybrid 14-kDa protein is a first example showing that spermatids are able to expand their protein repertoire by using hybrid mRNA.

Second, our data indicate clearly that the testis-specific characteristic of the *Oaz3* gene, as evidenced by the presence of cytoplasmic *Oaz3* mRNA, is the result of multiple contributions: (i) *Oaz3* is actively transcribed in several tissues, including testis, but not in many others, indicating a distinct regulation at the transcription initiation stage, and (ii) *Oaz3* nuclear pre-mRNA undergoes normal splicing exclusively in spermatids. These data thus demonstrate that tissue specificity is not necessarily only a consequence of transcription regulation. A corollary to our findings is that care should be exercised when interpreting results from global gene expression analyses. Clearly, the use of RNA or corresponding expressed sequence tags from entire tissues and whole cells would result in inclusion of nuclear transcripts that, in the case documented here for *Oaz3*, do not necessarily appear as mature cytoplasmic transcripts. Our analysis of the hybrid 1038 mRNA transcript and corresponding hybrid protein, which localizes to two essential sperm tail structures, documents a novel way to increase the protein repertoire of a cell.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

1. Oko R. Biol Reprod. 1988; 39:169–182. [PubMed: 3207795]
2. Vera JC, Brito M, Zuvic T, Burzio LO. J Biol Chem. 1984; 259:5970–5977. [PubMed: 6715381]
3. Irons MJ, Clermont Y. Anat Rec. 1982; 202:463–471. [PubMed: 7200337]
4. Baltz JM, Williams PO, Cone RA. Biol Reprod. 1990; 43:485–491. [PubMed: 2271730]
5. Eddy EM, Toshimori K, O'Brien DA. Microsc Res Tech. 2003; 61:103–115. [PubMed: 12672126]
6. Schalles U, Shao X, van der Hoorn FA, Oko R. Dev Biol. 1998; 199:250–260. [PubMed: 9698445]
7. Shao X, Tarnasky HA, Schalles U, Oko R, van der Hoorn FA. J Biol Chem. 1997; 272:6105–6113. [PubMed: 9045620]
8. van der Hoorn FA, Tarnasky HA, Nordeen SK. Dev Biol. 1990; 142:147–154. [PubMed: 1699827]
9. Petersen C, Aumuller G, Bahrami M, Hoyer-Fender S. Mol Reprod Dev. 2002; 61:102–112. [PubMed: 11774381]
10. Hoyer-Fender S, Petersen C, Brohmann H, Rhee K, Wolgemuth DJ. Mol Reprod Dev. 1998; 51:167–175. [PubMed: 9740324]
11. Burmester S, Hoyer-Fender S. Mol Reprod Dev. 1996; 45:10–20. [PubMed: 8873064]
12. Shao X, Tarnasky HA, Lee JP, Oko R, van der Hoorn FA. Dev Biol. 1999; 211:109–123. [PubMed: 10373309]
13. Shao X, Xue J, van der Hoorn FA. Mol Reprod Dev. 2001; 59:410–416. [PubMed: 11468777]
14. Fitzgerald CJ, Oko RJ, van der Hoorn FA. Mol Reprod Dev. 2006; 73:92–100. [PubMed: 16211599]
15. Zhang Y, Oko R, van der Hoorn FA. Dev Biol. 2004; 275:23–33. [PubMed: 15464570]
16. Bhullar B, Zhang Y, Junco A, Oko R, van der Hoorn FA. J Biol Chem. 2003; 278:16159–16168. [PubMed: 12594206]
17. Hoyer-Fender S, Neesen J, Szpirer J, Szpirer C. Cytogenet Genome Res. 2003; 103:122–127. [PubMed: 15004474]
18. Matlin AJ, Clark F, Smith CW. Nat Rev Mol Cell Biol. 2005; 6:386–398. [PubMed: 15956978]
19. Song MS, Chang JS, Song SJ, Yang TH, Lee H, Lim DS. J Biol Chem. 2005; 280:3920–3927. [PubMed: 15546880]
20. Zhang C, Xie Y, Martignetti JA, Yeo TT, Massa SM, Longo FM. DNA Cell Biol. 2003; 22:303–315. [PubMed: 12941158]
21. Tasic B, Nabholz CE, Baldwin KK, Kim Y, Rueckert EH, Ribich SA, Cramer P, Wu Q, Axel R, Maniatis T. Mol Cell. 2002; 10:21–33. [PubMed: 12150904]
22. Caudevilla C, Silva-Azevedo L, Berg B, Guhl E, Graessmann M, Graessmann A. FEBS Lett. 2001; 507:269–279. [PubMed: 11696354]
23. Finta C, Zaphiropoulos PG. J Biol Chem. 2002; 277:5882–5890. [PubMed: 11726664]
24. Hirano M, Noda T. Gene. 2004; 342:165–177. [PubMed: 15527976]
25. Yang L, Lee O, Chen J, Chen J, Chang CC, Zhou P, Wang ZZ, Ma HH, Sha HF, Feng JX, Wang Y, Yang XY, Wang L, Dong R, Ornvold K, Li BL, Chang TY. J Biol Chem. 2004; 279:46253–46262. [PubMed: 15319423]
26. Caudevilla C, Codony C, Serra D, Plasencia G, Roman R, Graessmann A, Asins G, Bach-Elias M, Hegardt FG. Nucleic Acids Res. 2001; 29:3108–3115. [PubMed: 11452036]
27. Takahara T, Tasic B, Maniatis T, Akanuma H, Yanagisawa S. Mol Cell. 2005; 18:245–251. [PubMed: 15837427]
28. Ivanov IP, Rohrwasser A, Terreros DA, Gesteland RF, Atkins JF. Proc Natl Acad Sci U S A. 2000; 97:4808–4813. [PubMed: 10781085]

29. Tosaka Y, Tanaka H, Yano Y, Masai K, Nozaki M, Yomogida K, Otani S, Nojima H, Nishimune Y. Genes Cells. 2000; 5:265–276. [PubMed: 10792465]
30. Matsufuji S, Matsufuji T, Miyazaki Y, Murakami Y, Atkins JF, Gesteland RF, Hayashi S. Cell. 1995; 80:51–60. [PubMed: 7813017]
31. Junco A, Bhullar B, Tarnasky HA, van der Hoorn FA. Biol Reprod. 2001; 64:1320–1330. [PubMed: 11319135]
32. Young RA, Davis RW. Proc Natl Acad Sci U S A. 1983; 80:1194–1198. [PubMed: 6219389]
33. Tovich PR, Sutovsky P, Oko RJ. Biol Reprod. 2004; 71:1182–1194. [PubMed: 15189827]
34. Oosterhuis JH, van der Hoorn FA. J Biol Chem. 1999; 274:11708–11712. [PubMed: 10206985]
35. Touriol C, Bornes S, Bonnal S, Audigier S, Prats H, Prats AC, Vagner S. Biol Cell. 2003; 95:169–178. [PubMed: 12867081]
36. Schafer M, Borsch D, Hulster A, Schafer U. Mol Cell Biol. 1993; 13:1708–1718. [PubMed: 8441407]
37. Kuhn R, Kuhn C, Borsch D, Glatzer KH, Schafer U, Schafer M. Mech Dev. 1991; 35:143–151. [PubMed: 1684716]
38. Shao X, Shepelev V, Fedorov A. Bioinformatics. 2006; 22:692–698. [PubMed: 16308355]
39. Kierszenbaum AL. Mol Reprod Dev. 2002; 61:1–2. [PubMed: 11774369]

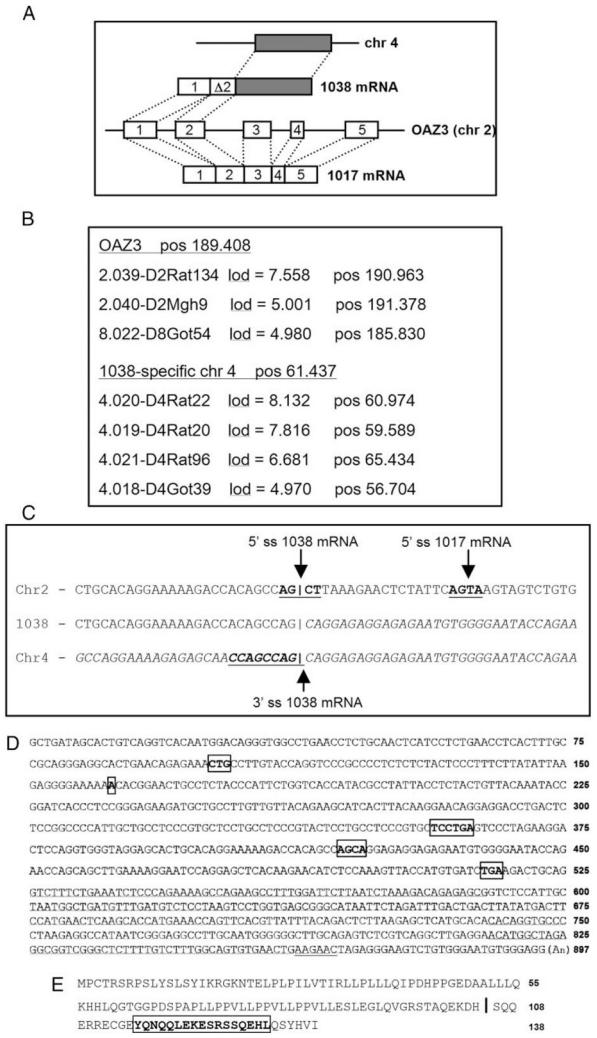

FIGURE 1. Hybrid 1038 mRNA is encoded by rat chromosomes 2 and 4

A, sequence analysis of 1038 mRNA shows it is a fusion of *Oaz3* sequences (*Oaz3 rat chr 2*) and a novel transcription unit on chromosome 4 (*chr 4*). Indicated are the chr 4 sequence in 1038 mRNA (*gray box*), the 5 exons that contribute to 1017 (*Oaz3*) mRNA, and exon 1 and part of exon 2 (Δ*2*) from *Oaz3* that contribute to 1038 mRNA. *B*, screening rat-hamster radiation hybrid genomic DNAs with 1017-and 1038-specific primers using genomic PCR and comparison to known locations of markers confirms the location of exon 1 and 2 *Oaz3* sequences in 1038 mRNA on chromosome 2 and 1038-specific sequences on chromosome 4. We found that three chromosome 2 markers (*2.039-D2Rat134*, *2.040-D2Mgh9*, and *8.022-D8Got54*) that colocalize with *Oaz3* genomic sequences as evidenced by their high lod scores (*lod*) have a chromosomal position (*pos*) close to the *Oaz3* gene. Similarly, we found four chromosome 4 markers (*4.020-D4Rat22*, *4.019-D4Rat20*, *4.021-D4Rat96*, and *4.018-D4Got39*) with high lod scores indicative of significant likelihood of colocalization with the novel 1038 sequences that have locations close to the predicted genomic 1038-specific chr 4 sequences. *Lod*, logarithm of the odds; *pos*, position on the chromosome. *C*, alignment of 1038 mRNA (*1038*) and the corresponding chromosome 2 (*Chr 2*) and 4 sequences (*Chr 4*).

1038 mRNA sequences derived from chromosome 4 are in *italics*. Indicated are the cryptic 5′-splice site in the chr 2 *Oaz3* gene sequence used to generate 1038 mRNA (AG CT: *5′-ss 1038 mRNA*), the 5′-splice site used for *Oaz3* testicular mRNA (AGTA: *5′-ss 1017 mRNA*), and the 3′-splice site on chr 4 used to generate 1038 hybrid mRNA (CCAGCCAG CA: *3′-ss 1038 mRNA*). *D*, nucleotide sequence of 1038 hybrid cDNA. Indicated are important features in *bold* and surrounded by a *box*, discussed under “Results”, including the CUG translation start site (*CTG*), the additional adenosine residue only present in mRNA not the genomic sequence (*A*), the translational frameshift sequence common to all *Oaz* genes (*TCCTGA*), the link between sequences derived from chromosomes 2 and 4 (*AGCA*), and the stop codon (*TGA*). Numbering of nucleotides is given on the *right*. *E*, predicted amino acid sequence from 1038 hybrid mRNA. Indicated are the point where sequences encoded by chromosomes 2 and 4 are linked (l) and the peptide used to raise a 1038-specific antibody (*bold*, *box*). Numbering of amino acid residues is on the *right*.

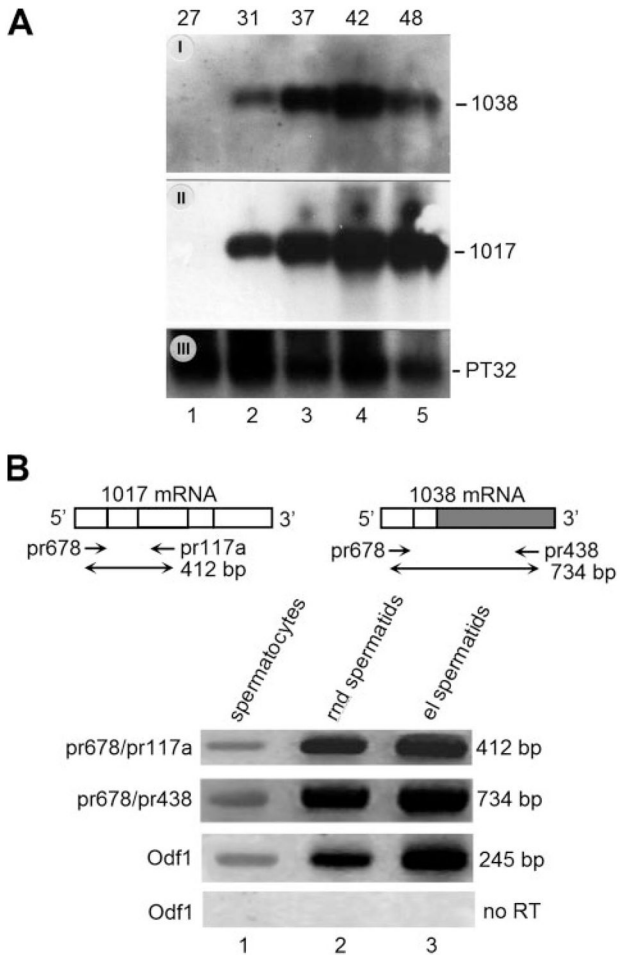

FIGURE 2. Hybrid 1038 mRNA is expressed in postmeiotic male germ cells

A, developmental Northern blot analysis of gene expression using RNA isolated from 27-, 31-, 37-, 42- and 48-day-old rats. *Blot I* was probed with a 1038-specific cDNA probe. *Blot II* was hybridized using a 1017-specific cDNA probe. *Blot III* shows results of reprobing a stripped filter using a perinuclear theca *Pt32* cDNA probe as internal control. *B*, RT-PCR analysis of gene expression in elutriated, fractionated male germ cells. Indicated primers were used in RT-PCR of cytoplasmic RNA isolated from fractionated spermatocytes (*lane 1*), round spermatids (*lane 2*), and elongating spermatids (*lane 3*). Primer pr438 is specific for the chromosome 4 sequence in 1038 mRNA. Controls included PCR experiments done with primers for the spermatid-specific *Odf1* mRNA (generating a 245-bp fragment) and reactions without addition of RT (*no RT*).

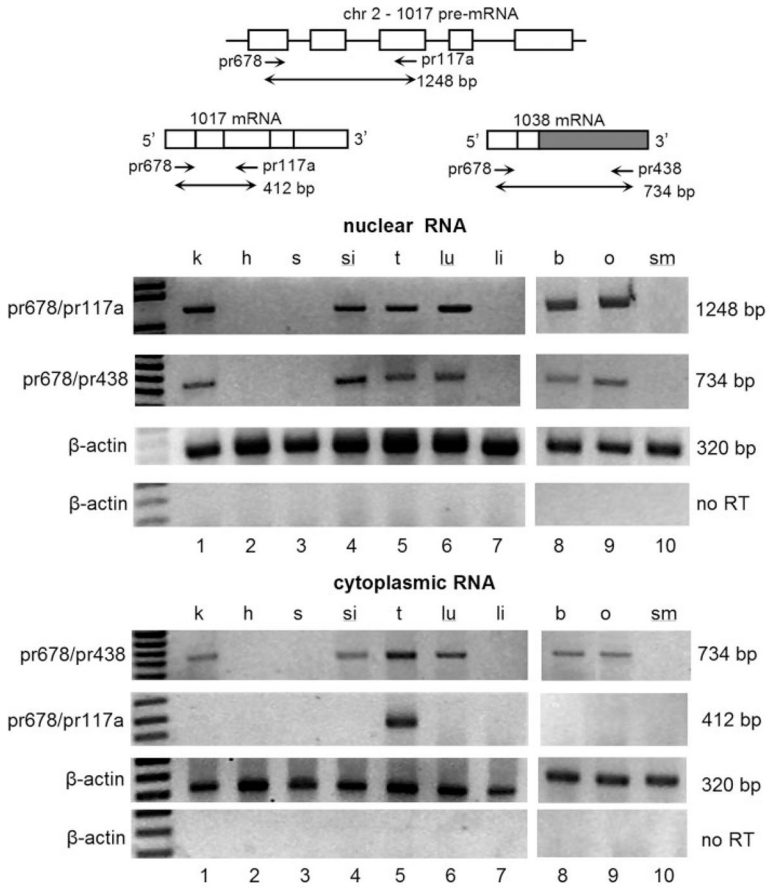

FIGURE 3. 1017 RNA and 1038 RNA expression patterns in different tissues

RT-PCR using RNA isolated from nuclei or from cytoplasm of kidney (*k*), heart (*h*), spleen (*s*), small intestine (*si*), testis (*t*), lung (*lu*), liver (*li*), brain (*b*), ovary (*o*), and smooth muscle (*sm*). Primer pr438 is specific for the chromosome 4 sequence in 1038 mRNA, and primer pr117a detects *Oaz3* pre-RNA in nuclear RNA preparations and mature *Oaz3* mRNA in cytoplasmic preparations. Control reactions were done with *β*-actin-specific primers and without RT. Note that 1017 pre-RNA is detected in all 1038 mRNA-positive tissues (*lanes 1*, *4 – 6*, *8*, and *9*) but mature cytoplasmic 1017 mRNA is only observed in testis.

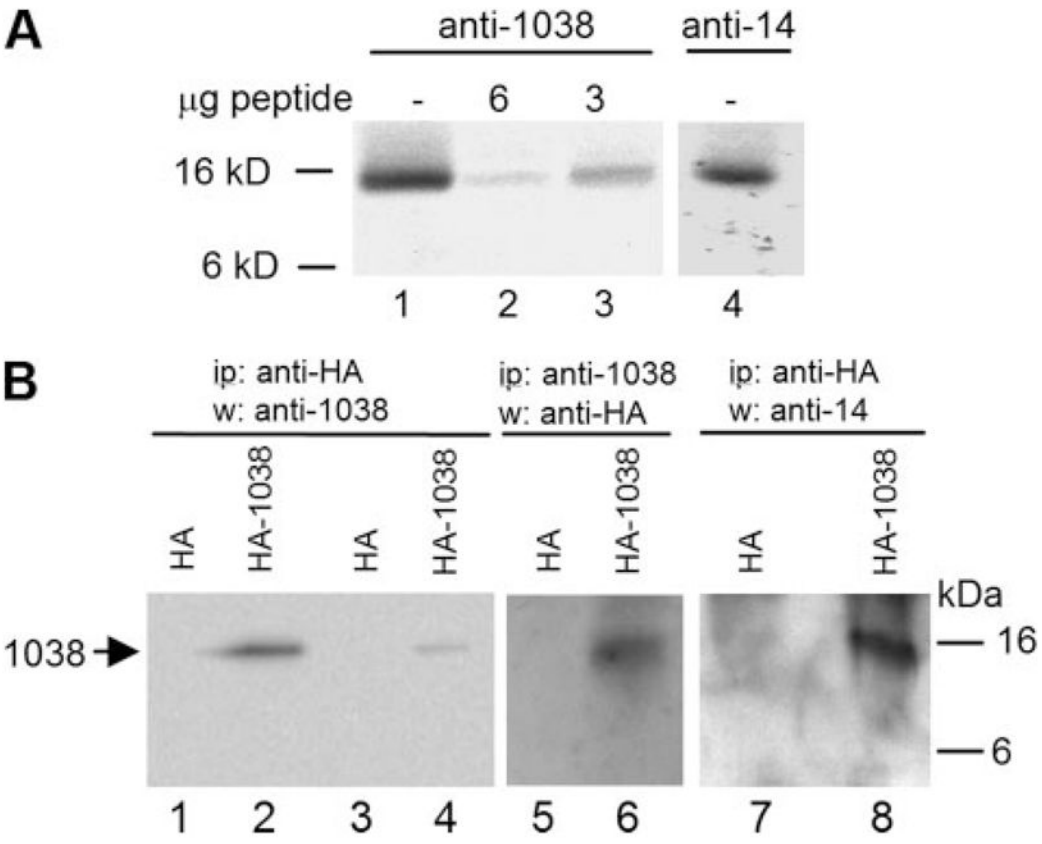

FIGURE 4. Chromosome 4-encoded amino acid residues are present in 1038 mRNA-directed protein

A, anti-1038 peptide antibodies were used in immuno-precipitation assays of *in vitro* translated HA-tagged 1038 protein. Antibodies were used direct (*lane 1*) or after preincubation with indicated amounts of peptide (*lanes 2* and *3*). In control immunoprecipitation assays the product was incubated with anti-14 antiserum (*lane 4*). *B*, human 293 cells (*lanes 1,2,5–8*) or mouse 3T3 cells (*lanes 3* and *4*) were transfected with constructs expressing HA tag (*lanes 1*, *3*, *5*, and *7*) or HA-tagged 1038 protein (*lanes 2*, *4*, *6*, and *8*). After transfection proteins were immunoprecipitated with indicated antisera (*ip*) and analyzed by Western blotting using indicated antisera (*w*).

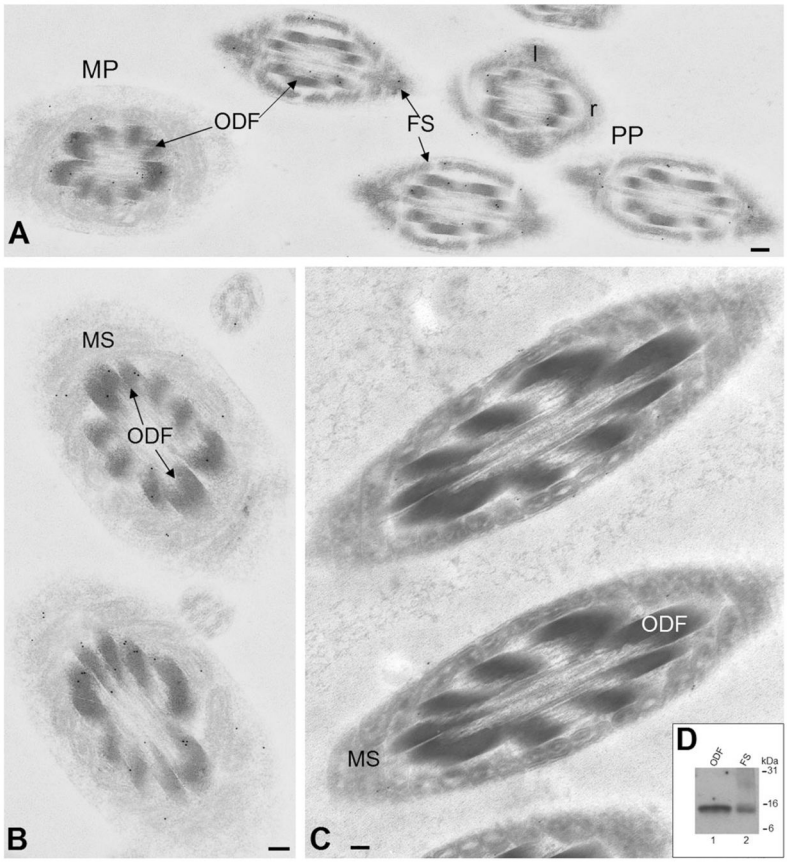

FIGURE 5. Hybrid 14-kDa 1038 protein is present in rat ODF and FS

A–C, electron micrographs of sections through the tails of rat spermatozoa immunogold labeled with affinity-purified anti-1038 serum. *A*, cross-section through one midpiece (*MP*) and several principal pieces (*PP*) showing gold label over ODF and FS. *B*, gold label is present over ODF in the midpiece. *MS*, mitochondrial sheath. *C*, specific gold label is quenched on preincubation of anti-1038 serum with C-terminal 1038-specific peptide. *D*, Western blot analysis of purified ODF (*lane 1*) and FS (*lane 2*) using affinity-purified anti-1038 serum. *Bars*, 0.1 μm.

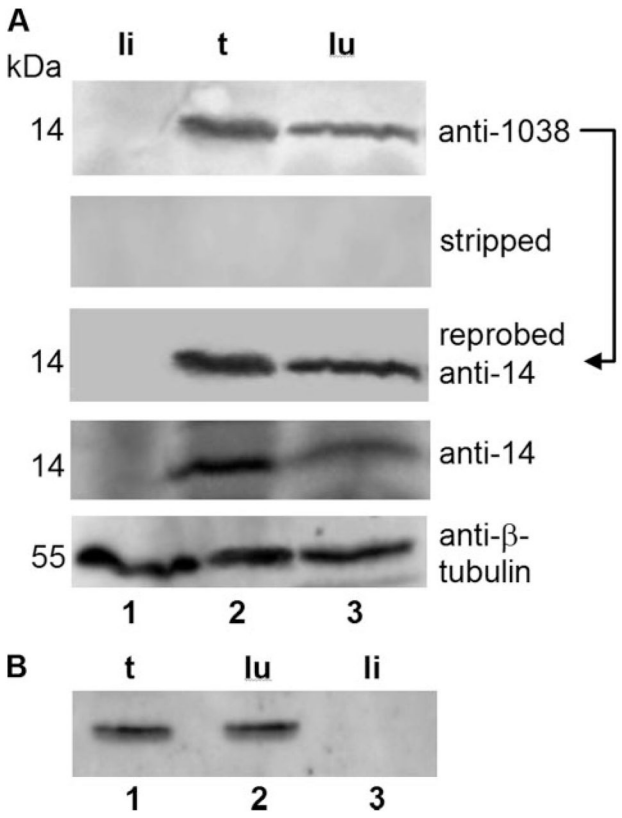

FIGURE 6. Hybrid 14-kDa 1038 protein is expressed in lung

A, proteins extracted from liver (*lane 1; li*), testis (*lane 2; t*), and lung (*lane 3; lu*) were analyzed by Western blotting using affinity-purified anti-1038 antiserum (*anti-1038*). After development, blots were stripped, examined for absence of any signal (*stripped*), and incubated with anti-14 antiserum (*reprobed anti-14*). Western blots were also probed directly with anti-14 antiserum (*anti-14*). Controls included anti-*β*-tubulin antibodies (*anti-β-tubulin*). *B*, proteins extracted from testis (*lane 1; t*), lung (*lane 2; lu*), and liver (*lane 3; li*) were incubated with anti-1038 antibodies that were covalently coupled to beads. Complexed proteins were analyzed by Western blotting using anti-14 antiserum.