CMLS, Cell. Mol. Life Sci. 60 (2003) 1705–1715
1420-682X/03/081705-11
DOI 10.1007/s00018-003-3134-4

CMLS Cellular and Molecular Life Sciences

Research Article

Evidence for posttranscriptional regulation of the multi K homology domain protein vigilin by a small peptide encoded in the 5′ leader sequence

J. Rohwedel[a, *], S. Kügler[b], T. Engebrecht[a], W. Purschke[c], P. K. Müller[a] and C. Kruse[a]

[a] Department of Medical Molecular Biology, Medical University of Lübeck, Ratzeburger Allee 160, 23538 Lübeck (Germany), Fax: + 49 451 5003637, e-mail: rohwedel@molbio.mu-luebeck.de
[b] Department of Neurology, University of Göttingen, 37073 Göttingen (Germany)
[c] NOXXON Pharma AG, 13355 Berlin (Germany)

Received 7 April 2003; received after revision 12 May 2003; accepted 16 May 2003

Abstract. Vigilin, a K homology (KH) protein has been found in all eukaryotic species studied. It has a unique structure of 14–15 consecutively arranged KH domains which apparently mediate RNA-protein binding. Cloning and sequencing of the mouse *vigilin* cDNA confirmed that the amino acid sequences of vertebrate vigilins are highly conserved and contain conserved sequence motifs of nuclear import and export sequences. The human and murine *vigilin* mRNAs carry two alternatively spliced 5′ exons. In the 5′ leader region of one of the splice variants, variant 1A, we found an upstream open reading frame (uORF) highly conserved between mouse and human. Here we present for the first time evidence that a 13 amino acid long peptide encoded by this uORF is an inhibitor of vigilin expression operating on a posttranscriptional level. We propose that the two structurally different 5′ leader sequences of the human *vigilin* mRNA are involved in the regulation of vigilin biosynthesis.

Key words. Translation; uORF; RNA-protein interaction; KH protein; gene expression.

Proteins of the K homology (KH) domain family are RNA-binding proteins whose KH domain plays a crucial role in RNA-protein interaction [1, 2]. The KH RNA-binding motif was first described in the human heterogeneous nuclear ribonucleoprotein K [3] and was subsequently found in numerous other proteins among which are FMR1 [4], NusA [5], Mer1 [6], PSI [7] and vigilin [8]. Because the latter exhibits an accumulation of KH domains and is ubiquitously expressed in heterotrophic eukaryotic species, it has been proposed to form a separate protein subfamily [9]. Using nuclear magnetic resonance (NMR)-spectroscopy, the structure of an individual vigilin KH domain has been resolved to high resolution [2].

* Corresponding author.

The *vigilin* sequence was first identified as a cDNA isolated from human cells [10] and from mesenchymal cells of chicken embryos [8]. The full-length protein carries 14 complete plus one degenerated, reiterative RNA-binding domains of the KH type [2, 11]. The observation that vigilin is part of a nuclear as well as a cytoplasmic multi-protein complex containing tRNA [12, 13] suggested that one cellular function of the protein might be associated with tRNA nucleocytoplasmic export. This functional concept was also in line with the observation that cellular levels of vigilin were always highest under conditions of increased protein synthesis [14], an elevated demand for tRNA molecules being presumed to be a prerequisite for stimulated protein output. Accordingly, by nuclear injection of vigilin into human Hep-2 cells, tRNA export from the nucleus could be accelerated [13]. According to our

hypothesis, quiescent cells do not require high levels of vigilin. To cope with such a situation, a mechanism might be favored by which downregulation of vigilin synthesis is achieved without wasting its own RNA. In this way, vigilin translation may be turned on again immediately when there is a demand for high protein synthesis. Genes are regulated at multiple steps such as transcription, post-transcriptional processing or translation by diverse mechanisms [15, 16]. Specifically, in about 10% of eukaryotic mRNAs, AUG codons located upstream of the cognate start codon are apparently involved in translational regulation mainly of genes that play a role in cell growth and maintenance [17, 18]. In some rare cases, such upstream AUGs are associated with an upstream open reading frame (uORF) which represses downstream translation [19]. Structural analysis of the human *vigilin* gene showed the presence of two additional 5′ exons compared to the chicken gene, designated exon 1A and 1B, which are alternatively spliced [20], though their function, if any, has yet to be elucidated. Here, we present the complete coding sequence of the mouse *vigilin* gene and further corroborate that vigilin is highly conserved from humans to yeast. In particular, an uORF located at the 5′ end of exon 1A is also conserved between human and mouse. We provide the first evidence that a 13-amino-acid long peptide encoded by this uORF of the *vigilin* mRNA may interfere with vigilin translation.

Materials and methods

Cloning and sequencing of mouse *vigilin* cDNAs

Total RNA was isolated from mouse liver using the Rneasy Midi kit (Qiagene) and 500 ng of total RNA was reverse transcribed using oligo-dT primer and Superscript II reverse transcriptase according to the manufacturer's recommendations (Life Technologies). To amplify the complete *vigilin* splice variants carrying either exon 1A or 1B, specific oligonucleotide primer sequences were deduced from the published human sequences [20] as follows (designations and sequences of oligonucleotides are given in parentheses followed by the length of the amplified fragment): exon 1A sense primer (primer 1: 5′ CAAGCGTAGCCTCTTCTCTTTACCAAG 3′; 4331 bp), exon 1B sense primer (primer 2: 5′ CCTCGGAGCGTCCCGGCTTC 3′; 4301 bp) and the same antisense primer (primer 3: 5′ CGTGATGAAGGCCAGAGTGCTGACT 3′) were used to amplify the two fragments using the Expand Long Template polymerase chain reaction (PCR) system (Roche). Reverse transcription (RT)-reactions were denatured for 2 min at 95 °C, followed by 40 cycles of 40 s denaturation at 95 °C, 40 s annealing at 50 °C and 5 min elongation at 68 °C. Amplified fragments were isolated from agarose gels following electrophoretic separation and cloned into the pCR-Blunt II-TOPO plasmid vector using the Zero Blunt Topo PCR cloning kit (Invitrogen) generating the clones pMvig-1A carrying exon 1A and pMvig-1B carrying exon 1B. To verify sequence identity, both strands of the cloned fragments were sequenced by a customer sequence service (MWG-Biotech).

Cloning of two alternatively spliced 5′ ends of human vigilin cDNA

Each of the two alternatively spliced 5′ ends of the human *vigilin* gene were cloned in front of the complete human *vigilin* cDNA consisting of the AUG in exon 2 down to the stop codon in exon 27. To this end the complete human *vigilin* cDNA was cloned into the *Hin*dIII and *XhoI* sites of the plasmid vector pAD-CMV (Stratagene). One microgram of total human RNA, isolated from the cell line MG63, was reverse transcribed with Superscript Reverse Transcriptase (Life Technologies) according to the manufacturer's recommendations with 10 pmol of *vigilin*-specific primer E16as (5′ CATGATGGAGCGGATCAGAC 3′). Sense primers for the following PCR were 1A-HIII (5′ ATCGATAAGCTTCAAGCGTAGCCTCTTCTCTTTAC 3′) for the human *vigilin* splice variant with exon 1A (VIG-1A) and 1B-HIII (5′ ATCGATAAGCTTCCTCGGAGCGTCCCGGCTTCT 3′) for the human *vigilin* splice variant with exon 1B (VIG-1B); the antisense primer for both reactions was E13as (5′ CTGCCACCATCTTCTGCATG 3′). The resulting fragments VIG-1A and VIG-1B were cut with *Hin*dIII and *Bsm*BI, and after gel purification cloned into the appropriate sites of the pAD-CMV vector containing the complete human *vigilin* cDNA. The correct orientation of the fragments was confirmed by sequencing.

Site-directed mutagenesis

To eliminate the first AUG of the uORF in VIG-1A , site-directed mutagenesis was performed on the VIG-1A construct with primers GUG-s (5′ CTTCTCCTTTACCAAGGTGGCGGCTTGTCCCTG 3′) and GUG-as (5′ CAGGGACAAGCCGCCACCTTGGTAAAGGAGAAG 3′) using the Quickchange kit (Promega) according to the manufacturer's instructions. The result was confirmed by sequencing. The mutant construct will be referred to as VIG-1A-GUG.

Cloning of the *β-Gal* reporter constructs

For generation of reporter gene constructs, the 5′ ends of VIG-1A and VIG-1B from the respective nucleotide 1 down to the AUG in exon 2 were cloned 5′ to *β-galactosidase* into the pAD-*β*GAL vector. To achieve this, PCR was performed on the VIG-1A-, VIG-1B- and VIG-1A-GUG constructs with sense primers 1AHIII and 1BHIII, and with the antisense primer E2as-Kpn (5′ ACTGCAGGTACCCTCATGGTTGATCTCACACCTAC 3′). This primer contains an artificially introduced *KpnI*

site one codon downstream of the authentic vigilin AUG. The *Hin*dIII/*Kpn*I-digested and gel-purified PCR fragments were cloned into the pAD-βGAL vector from which the gpt leader sequence had been removed. This cloning procedure leaves the genuine environment of the *vigilin* mRNAs exon 2 AUG intact and results in the constructs VIG-1A-βGAL, VIG-1B-βGAL and VIG-1A-GUG-βGAL. Again, the sequences of the resulting constructs were confirmed by sequencing.

Cell-free translation assay

Generation of ^{35}S-labeled protein was performed in the reticulocyte translation system TnT (Promega). The system was primed with RNA (up to 6 ng/µl) which had been transcribed from the respective linearized plasmids by T7 RNA polymerase using the MaxiScribe kit (Promega). The RNA treated with DNase and precipitated by two rounds of ethanol addition was subsequently assayed for full-length transcripts on denaturing formaldehyde gels. The concentration was determined spectrophotometrically. Aliquots of the translation reactions were removed after 30 and 60 min and seperated by polyacrylamide gel electrophoresis (PAGE). After electrophoresis, the gels were fluorographed and the band intensities were quantified by video densitometry. The phosphoimager Fuji Bas-1500 (Fujifilm) was used for quantification.

Cell culture translation assay

CV1 cells (ATTC, CCL-70) were transiently transfected by liposome mediated gene transfer using Lipofectamine (Life Technologies) essentially as recommended by the manufacturer for Cos-7 cells, except that the amount of DNA was reduced to 500 ng/transfection of 3×10^5 cells, seeded in 3-cm dishes 12 h before starting the transfection. The transfection mixture (1 ml) was incubated with the cells for 5 h, then 1 ml of Dulbecco's modified Eagle's medium (DMEM, Life Technologies) containing 20% fetal calf serum (Life Technologies) was added and the cells were incubated for another 12 h. Protein and RNA were isolated from the same transfection reaction as follows. The cells were washed with ice-cold phosphate-buffered saline (PBS; 137 mM NaCl, 8.5 mM Na_2HPO_4, 1.5 mM KH_2PO_4, 3 mM KCl, pH 7.3), 500 µl of 0.05% trypsin was added and incubated at 37 °C for 5 min followed by addition of 1000 µl of ice-cold DMEM. The cell suspension (750 µl) was centrifuged for 5 min at 250 g at 4 °C, the supernatant was removed and the pellet resuspended in 350 µl RLT lysis buffer of the total RNA isolation kit Rneasy (Qiagen). The subsequent purification steps were performed according to the manufacturer's instructions. The chemiluminescent signals obtained from Northern blotting were visualized on X-ray films, quantified densitometrically and the values for the *β-galactosidase* signals were normalized to the *glyceraldehyde-3-phosphate dehydrogenase (GAPDH)* gene content of the sample.

The β-galactosidase activity was assayed by washing the remaining 750 µl of the transfected cells twice with ice-cold PBS and subsequent resuspension of the pellet in 500 µl of the lysis buffer of the Chemiluminescent β-galactosidase Detection kit (Boehringer). The subsequent processing of the samples was performed essentially as recommended by the manufacturer. The chemiluminescence was measured in glass vials with a liquid scintillation counter (Beckmann) using the single photon counting mode. Counts were normalized to the *β-galactosidase* mRNA content of the sample.

Cultivation and lysis of cells

An established cell line derived from a human epithelial larynx carcinoma (HEp-II, ATCC CCL23) was grown in Iscove's medium (Sigma), supplemented with 10 % fetal calf serum (Biochrom), 100 U/ml penicillin, 100 µg/ml streptomycin, 0.05 mg/ml L-ascorbic acid sodium salt (Sigma) and 2 mM glutamine. Primary human fibroblasts were cultivated as described previously [21, 22]. Fibroblasts were harvested 2 and 5 days after seeding. For SDS/PAGE analysis, 10^6 cells were lysed for at least 1 h in 50 µl of a lysis buffer consisting of 0.1% $NaHPO_4$ (w/v), 100 mM phenylmethylsulfonyl fluoride and 0.5% Nonidet P-40. The cell extract was centrifuged and the supernatant was immediately used, or stored at –70 °C.

Electrophoresis and immunoblotting

Supernatants from centrifuged cell lysates (derived from 1×10^6 cells) were analyzed by SDS-PAGE using 7% or 10% polyacrylamide. The amount of protein was determined and equal amounts were loaded onto the gels. The gels were blotted onto nitrocellulose membranes and vigilin or the exon-1A-uORF-encoded peptide were visualized by immunodetection with the vigilin-specific antibody FP3 [23] or a specific antiserum commercially generated against the exon-1A-uORF-encoded peptide of the sequence MAACPCFATVPTL in rabbits (Eurogentec). For immunization, the 13-amino-acid-long peptide was synthesized and fused to bovine serum albumin (BSA). As controls, the fusion protein between BSA and the exon-1A-uORF-encoded peptide which was used for immunization of rabbits as well as BSA alone were also separated on the same gel and blotted.

Northern blotting and quantification of *vigilin* mRNA

Total RNA from fibroblasts was prepared by acid guanidine thiocyanate/phenol/chloroform (1:1:0.2, by volume) treatment [24]. Electrophoresis, blotting, hybridization and detection of *vigilin* mRNA were as described previously [20, 25]. As an internal control for quantification of *vigilin* mRNA levels, a human *GAPDH*-specific antisense RNA was used. The amount of *vigilin* mRNA was

corrected for the variation in the levels of *GAPDH* content. Densitometric quantification was performed using whole-band scanning and Optoquant software (Computer & Vision).

Immunostaining of fibroblasts

Briefly, $2-4 \times 10^5$ cells were grown on round cover slips (22 mm diameter) for 2 or 5 days, washed twice with PBS and fixed with methanol/DAPI (Roche, 1 µg DAPI/ml methanol) for 10 min and subsequently with acetone for 1 min at –20 °C. The cells were washed again twice with PBS. The following antibodies were used: FP3, detecting vigilin; an antibody specific for the human ribosomal protein P (Euroimmun), or an antibody directed against the exon-1A-uORF-encoded peptide. Antibodies were allowed to bind for 1 h at 37 °C. They were used at dilutions of 1:500, 1:100 and 1:200, respectively, in PBS containing 0.1% BSA. As secondary antibodies Cy3- or FITC-labeled antibodies (Dianova) were used. Specimens were analyzed with the fluorescence microscope Axioskop and digital images prepared using the Axiovision program (Zeiss).

Results

Vigilin is highly conserved from yeast to humans

The amino acid sequence of mouse vigilin from the start codon in exon 2 (nucleotide 103) downstream to the termination signal in exon 27 (nucleotide 3907) was deduced from a mouse cDNA (see below) and used to screen the EMBL database for homologies. As expected, ,we found that mouse vigilin is highly homologous to sequences in several species with a level of homology ranging from 95.5 to 43.3% at the DNA and 98.5 –21.2% at the protein level from rat to yeast (fig. 1). The identity score between vertebrate species lies between 87.2% (chicken) and 98.5% (rat) when compared with the mouse amino acid sequence. In addition to the 15 tandemly repeated KH domains, several sequence motifs were found to be conserved, in particular between mammals and chicken such as a nuclear localization sequence (NLS) and a nuclear export sequence (NES). Both motifs are presumably necessary for effective molecular shuttling between the nucleus and the cytoplasm [26].

Comparison of vigilins with other KH proteins containig import (NLS) and export (NES) sequences

A comparison of vigilin from different species with other KH proteins revealed common structural characteristics within the KH protein superfamily (fig. 2). In contrast to other KH proteins such as the human fragile X mental retardation protein (FMRP), the mouse coding-region-determinant-binding protein (CRD-BP) and the herpes-simplex-virus-infected cell polypeptide 27 (ICP27), the vigilins are composed exclusively of 15 KH domains framed by 75 and 70-amino-acid-long non-KH sequences at the N and C terminus, respectively. Nuclear import and export sequences are part of the KH domains in vigilin, with the exception of the missing NES in yeast. In contrast in other KH proteins the sequences are localized outside the KH domains. Furthermore, the NLS in vigilin is not located at a fixed position but can be found at varying locations along the molecule. Also noteworthy is that other KH proteins contain additional RNA-binding motifs such as the Arg-Gly-Gly (RGG) motif in human FMRP, mouse CRD-BP and herpes simplex virus ICP27 and the ribonucleoprotein (RNP) motif in mouse CRD-BP while vigilin contains exclusively KH domains (fig. 2).

A highly conserved uORF in the 5′-leader sequence of the human and mouse *vigilin* mRNA

Using oligonucelotide primers specific for the 5′ splice variants of the human *vigilin* gene [10, 20], we were able to amplify by RT-PCR two fragments of 4331 and 4301 bp in length from total mouse liver RNA. Both fragments were cloned into the plasmid pCR-Blunt II-TOPO. Sequencing of the resulting clones pMvig-1A and pMvig-1B showed that both cDNA fragments exhibited an ORF displaying 89.6% homology to the human *vigilin* coding sequence spanning exon 2 to exon 27. The two cDNA fragments differed in length by 30 nucleotides at their 5′ ends and were shown to represent two alternatively spliced variants, one with exon 1A and the other with exon 1B (fig. 3A) as has been shown previously for the human homologues [20]. Whereas the analyzed sequences of exon 1A and 1B did not show any homology with each other in human or mouse, the homology between human and mouse was 90.4% for exon 1A and 62.9% for exon 1B. In particular, a region between nucleotides 28 and 69 of exon 1A was almost identical (fig. 3B) and encompasses an uORF coding for 13 amino acids which is the subject of the hypothesis that it plays a regulatory role in its own expression. A structural hallmark of exon 1B was its relatively high GC content of 59.7% and 78.7% for mouse and human, respectively.

Translation of alternatively spliced *vigilin* mRNAs in a cell-free system and in cells indicates that the exon 1A uORF plays a regulatory role

To test for regulatory functions of the conserved 13- amino-acid-encoding uORF in exon 1A, we analyzed the translational properties of both alternatively spliced vigilin 5′ leader sequences in a cell free transcription/translation assay. Both 5′ leader sequences were isolated by RT-PCR and cloned into an existing vigilin expression construct, generating constructs, VIG-1A and VIG-1B. As a reference construct we generated a point mutation in VIG-1A resulting in the inactivation of the AUG of the 13-amino-acid-encoding uORF, which was the only uORF start codon within a

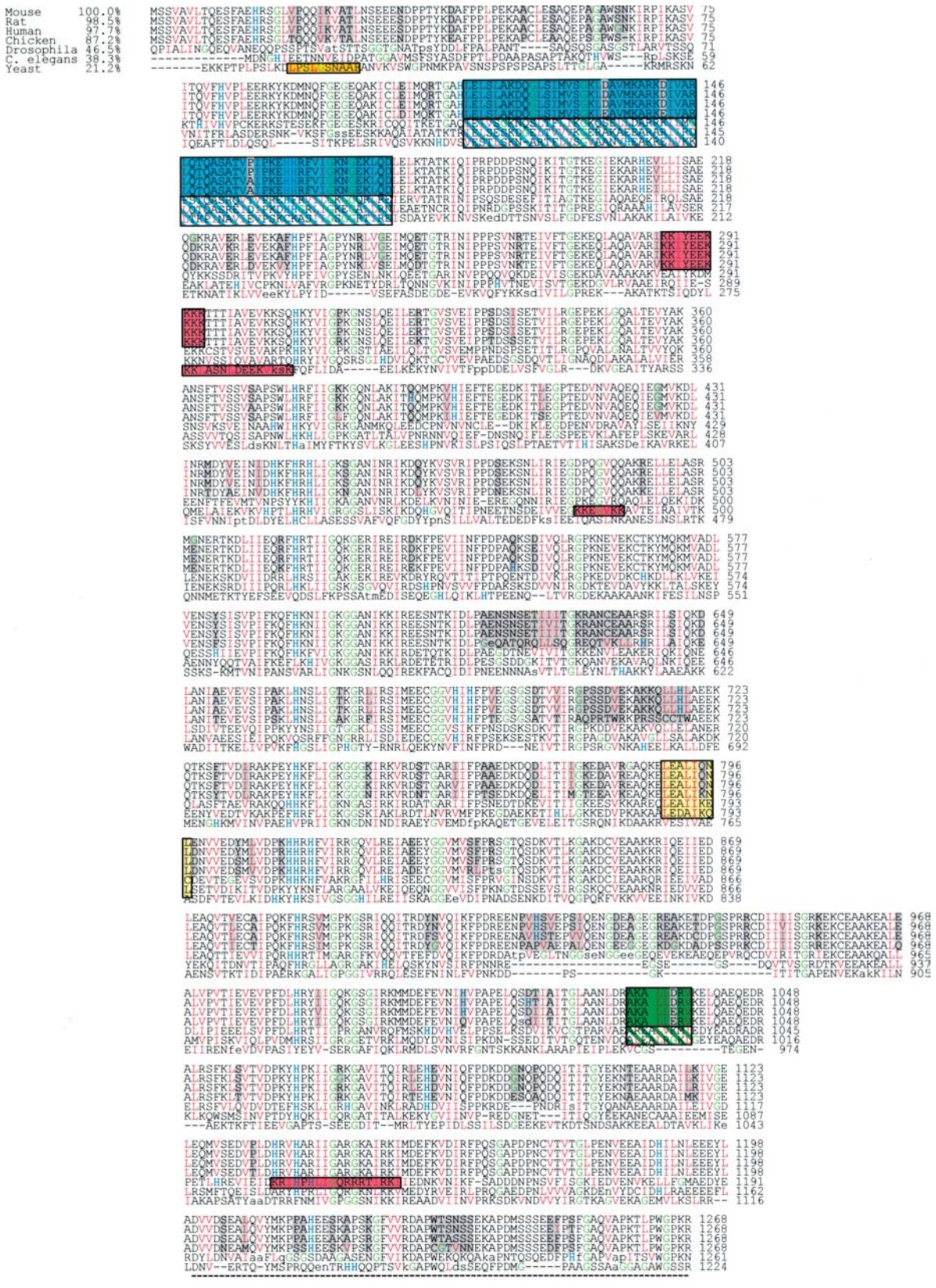

Figure 1. Vigilin is highly conserved among different organisms. The amino acid sequence of mouse vigilin was deduced from the nucleic acid sequence of clone pMvig-1A carrying the complete coding sequence of the mouse *vigilin* cDNA. The EMBL database was screened for similarities to this sequence using the EBI FASTA3 program. Homologous sequences are presented in segments of the consecutive KH domains in addition to the N- and C-terminal sequences. Gray background color indicates sequence differences among mouse, rat, human and chicken. The following conserved sequence motifs are highlighted by colored boxes: a ubiquitin-binding motif (blue), a nuclear localization sequence (NLS; red), a nuclear export sequence (NES; yellow) and a helix-loop-helix motif (green). Low conservation is indicated by hatched boxes.

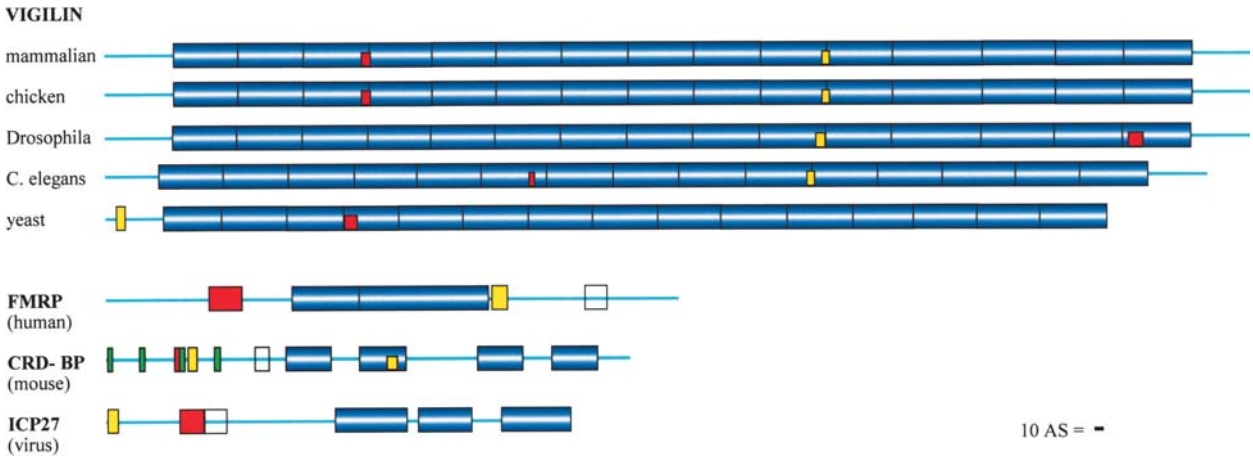

Figure 2. Vigilins represent a unique family among KH proteins. Representative members of the KH protein family are shown including vigilins from different organisms, the human fragile X mental retardation protein (FMRP), the mouse coding-region-determinant-binding protein (CRD-BP) and the herpes-simplex-virus-infected cell polypeptide 27 (ICP27). Homology domains are indicated. Because of the high similarity between mouse, rat and human vigilin they are presented as the mammalian vigilin. Boxes: blue, KH domain; red, NLS; yellow, NES; white, RGG-motif; green, ribonucleoprotein motif.

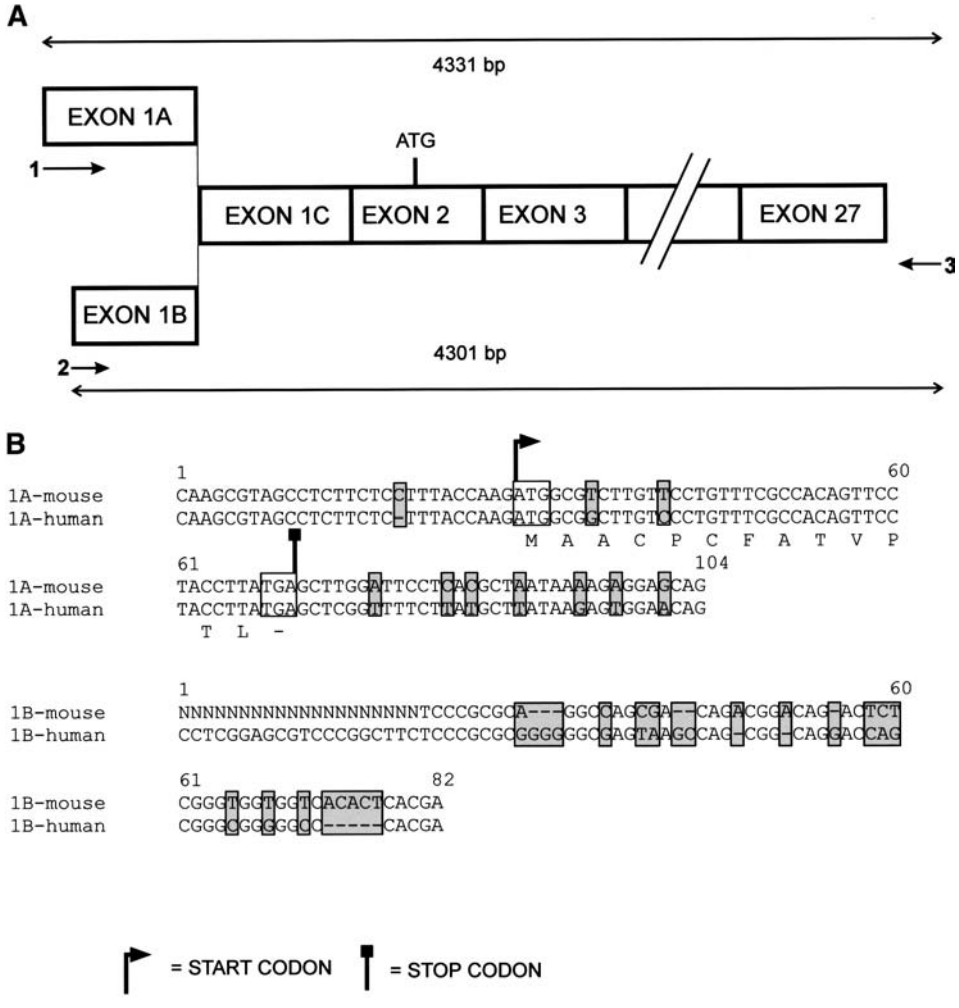

Figure 3. A conserved ORF in the 5′ region of the *vigilin* splice variant 1A. Schematic representation demonstrating that *vigilin* exons 1A and 1B are alternatively spliced to exon 1C followed by exon 2 that contains the translation start codon ATG for the ORF encoding the vigilin protein (*A*). Alignment of the sequences of the exon 1A and 1B splice variants from mouse and human *vigilin* genes demonstrating that an ORF detected in exon 1A is highly conserved between mouse and human (*B*). Sequence differences are boxed. Arrows indicate the location of oligonucleotide primers used for cloning (exon 1A primer, 1, 3; exon 1B primer, 2, 3). The sequence of the primer binding site in exon 1A was corrected according to a published EST sequence (GenBank/EMBL/DDBJ accession no. BG077807) and is not known (N) in exon 1B.

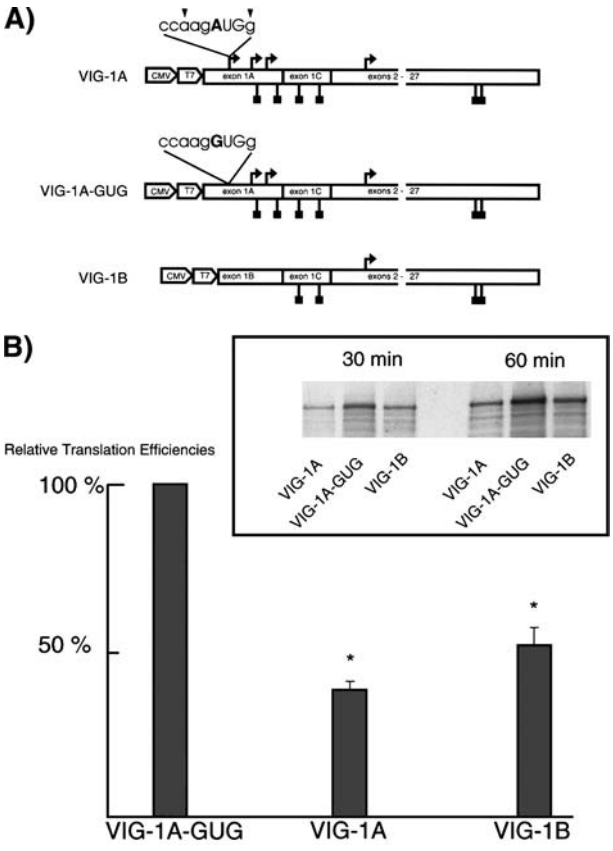

Figure 4. Disruption of the ORF in the 5′ region of a *vigilin* splice variant results in enhanced vigilin translation as demonstrated in a cell-free system. (*A*) Schematic representation of the VIG-1A, VIG-1A-GUG and VIG-1B expression cassettes. Start and stop codons are shown as in figure 3. The nucleotide sequence next to the first AUG in the native VIG-1A mRNA and the mutated sequence in the VIG-1A-GUG transcript are shown in detail. Arrowheads indicate the nucleotides most important for a strong initiation potential (Kozak sequence). CMV, cytomegalovirus promoter; T7, T7 RNA polymerase promoter. (*B*) Translation efficiencies of the cognate VIG-1A and VIG-1B transcripts compared to the mutated VIG-1A-GUG transcripts. Bars represent the mean ± SEM of six translation experiments. Asterisks indicate that the amounts of protein translated from the VIG-1A and VIG-1B transcripts, are statistically significantly different ($p = 0.05$) from VIG-1A-GUG. The inset shows the fluorography of ^{35}S-methionine-labeled protein bands synthesized from the different transcripts.

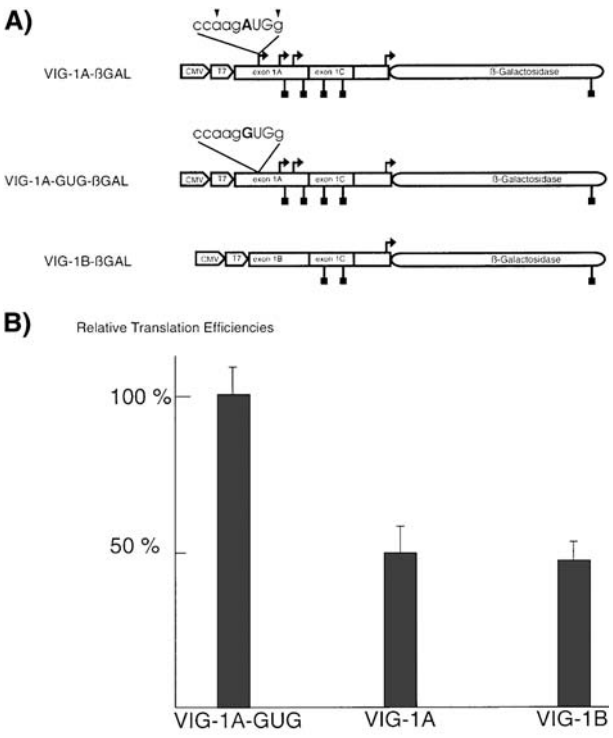

Figure 5. Disruption of the ORF in the 5′ region of a *vigilin* splice variant results in enhanced *vigilin* translation as demonstrated in cell culture. (*A*) Schematic representation of the VIG-1A-βGal, VIG-1A-GUG-βGal and VIG-1B-βGal expression cassettes. Start and stop codons are shown as in figure 3, promoters and sequences as in figure 4. The authentic *vigilin* start codon in exon 2 is used to drive expression of the *β-galactosidase* reporter gene. (*B*) Translation efficiencies of the VIG-1A, VIG-1A-GUG and VIG-1B leader sequence-βGal fusions. Bars represent the mean ± SEM of three independent transfection experiments in CV1 cells.

strong initiation context (clone VIG-1A-GUG) (fig. 4). As mentioned above, the VIG-1B 5′ leader sequence did not contain such an uORF but showed a high GC content of 78.7%. To ensure that the cell-free translation system was not saturated by excess amounts of RNA, aliquots of each reaction were analyzed after 30 and 60 min. The translation rates of both the VIG-1A- and VIG-1B transcripts consistently, only reached about 40–50% of the rate obtained from the mutated VIG-1A-GUG transcript (fig. 4).

To verify the data from the transcription/translation assay, both vigilin 5′ leader sequences and the mutated VIG-1A-GUG leader sequence were cloned into a βGAL reporter vector (fig. 5). These clones contained the 5′ sequences of the VIG-1A, VIG-1A-GUG and VIG-1B transcripts from their very 5′ ends up to the original exon 2 *vigilin* start codon in front of the *β-galactosidase* gene. CV1 cells were transiently transfected with the respective constructs. Saturation effects were avoided by using small amounts of DNA for transfection and also by reducing the transgene expression time to 12 h. The translation efficiency was measured by chemiluminescent detection of β-galactosidase activity in relation to the *β-galactosidase* mRNA content in the respective cell lysate, which was further normalized to the mRNA content of the *GAPDH* housekeeping gene. This procedure allows for a very tight correlation of β-galactosidase activity with the mRNA content in the respective sample.

Consistent with the data obtained in the cell-free transcription/translation assay, very similar protein amounts were synthesized from both naturally occuring vigilin leader sequences, whereas the mutated control transcript was translated with about twice the efficiency of either cognate transcript (fig. 5). Together, our results indicate that by inactivating the uORF, the *vigilin* mRNA splice variant 1A has lost a regulatory module.

Downregulation of *vigilin* translation coincides with expression of the exon 1A-uORF-encoded peptide

Characteristically, genes translationally regulated by uORFs show changes in translational levels without mRNA levels being affected [19]. To test for this effect, we investigated growing fibroblasts for expression at the mRNA and protein levels. Two days after fibroblast seeding, vigilin was detected by immunoblotting but had disappeared completely by day 5 (fig. 6A). In contrast, *vigilin* mRNA was not downregulated, as shown by Northern blot analysis normalized to *GAPDH* mRNA expression (fig. 6B). We then asked, whether the uORF-encoded peptide was expressed at the time when vigilin protein was downregulated. Polyclonal antibodies were generated against the uORF-encoded peptide and their specificity was tested by immunoblotting (fig. 7). Fibroblasts grown for 2 days showed only very weak expression of the 13-amino-acid-long-peptide, whereas vigilin was strongly expressed both in the cytoplasm and in the nucleus (fig. 8). However, at day 4–5 of fibroblast culture, when only residual amounts of vigilin were found in the nucleus, fibroblasts showed high expression of the uORF-encoded peptide, colocalizing in the cytoplasm with ribosomes immunostained with a human ribosomal antibody (fig. 8). These results support our hypothesis that the uORF-encoded peptide controls downstream translation of vigilin.

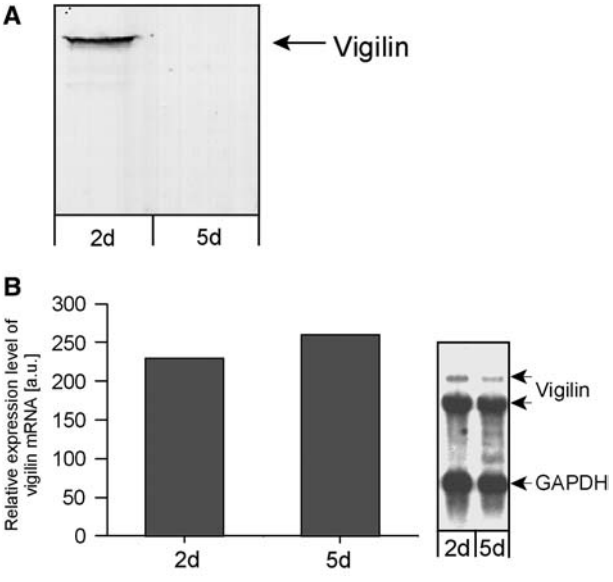

Figure 6. Vigilin protein synthesis is downregulated whereas *vigilin* mRNA transcription is unaffected during cultivation of primary human fibroblasts. (*A*) Immunoblot with anti-vigilin antibodies of cell lysates from growing human skin fibroblasts on day 2 and 5 after seeding following separation on 7% PAGE. Equal amounts of protein were analyzed. (*B*) Northern blot of *vigilin* mRNA from growing human skin fibroblasts on day 2 and 5 after seeding and quantification of these mRNAs corrected for variation in *GAPDH* levels. The two transcripts of different size detected by Northern hybridization are most likely due to usage of alternative polyadenylation sites [8, 20].

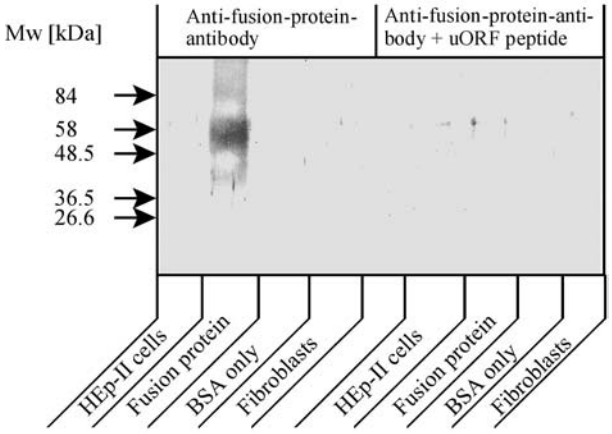

Figure 7. Specificity of an antiserum generated against the exon-1A-uORF-encoded peptide. Four protein samples, a protein lysate from HEp-II cells, a fusion protein between BSA and the exon-1A-uORF-encoded peptide, BSA alone and a protein lysate from primary human skin fibroblasts, were seperated by 10% PAGE and blotted onto nitrocellulose membranes. The blot was incubated with either antiserum generated against the exon-1A-uORF-encoded peptide fused with BSA (anti-fusion protein antiserum) or the same antiserum preabsorbed with the exon-1A-uORF-encoded peptide (anti-fusion protein antiserum + uORF peptide). The antiserum only detects the fusion protein, which is almost completely abolished after preabsorption.

Discussion

Vigilins, a unique protein family with cumulative KH domains and yet with a host of additional sequence motifs

Sequence comparison of *vigilin* from several species revealed a high degree of homology both at the nucleotide and the amino acid level. As shown here, proteins of the vigilin type contain 14–15 tandemly arranged KH domains, an NLS and an NES, in all species investigated to date from yeast to human. While the functional utilization of the NLS was demonstrated a few years ago [23], experimental proof for the function of the highly conserved NES is still missing. The NES sequence was always found at the same position between KH domains 9 and 10 except in yeast, where a candidate export signal is positioned at the very amino-terminal end of the protein sequence [27]. The presence of 14–15 KH domains is a distinct structural feature of vigilins, since other KH proteins in general contain only 1–5 KH domains [1, 28]. The presence of KH domains in such numbers as seen in vigilin could explain why a variety of apparently contradictory interaction profiles with different RNA species have been reported [9, 13, 29]. Potentially, these observations account for the fact that vigilin may be capable of binding more than one type of RNA species at the same time or may change the binding pattern depending on the metabolic state of the cell.

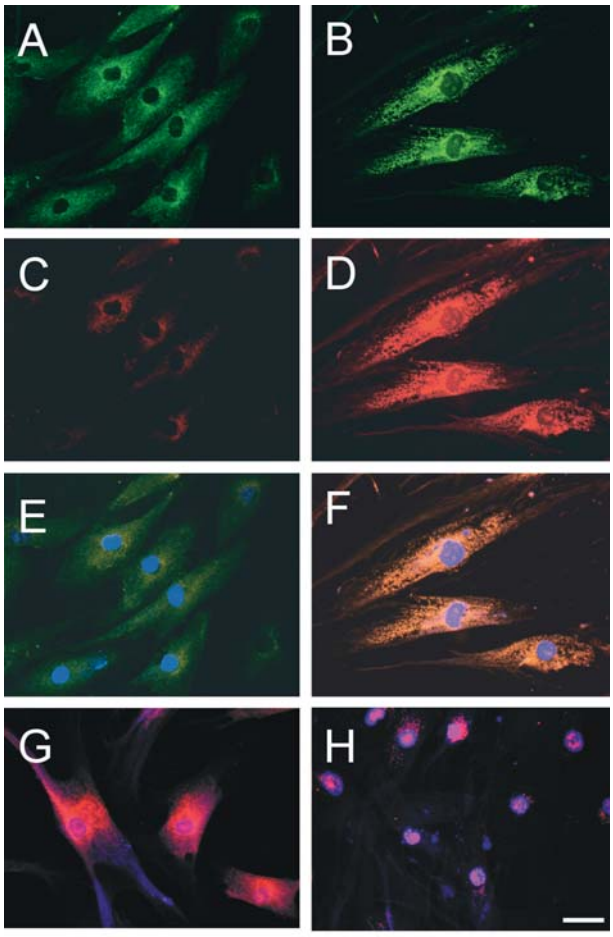

Figure 8. The translation-controling peptide encoded by the exon 1A uORF is expressed in fibroblasts at the time of reduced vigilin synthesis and localized to ribosomes. The ribosomal protein P (*A*, *B*) and the exon-1A-uORF-encoded peptide (*C*, *D*) colocalize (*E*, *F*) in primary human skin fibroblasts immunostained 2 (*A, C, E, G*) or 5 (*B, D, F, H*) days after seeding. During this cultivation time, immunostaining showed that vigilin (*G*, *H*) was downregulated, confirming the Western blot analysis in figure 6A. Bar, 15 µm.

Looking at the motif alignment in vigilin and other KH proteins (see fig. 2) the tight consecutive arrangement of the 14/15 KH domains is clearly exceptional as is the localization of both the NLS and NES at unique sites within the KH domains. Yet another unique feature of vigilin in contrast to several other KH proteins is the fact that it does not carry any other RNA binding domain such as RGG and RNP motifs beside the KH domains. Together, these features allow, vigilins to be considered as a seperate protein subfamily [9].

The mouse *vigilin* gene contains two alternatively spliced first exons

Recent investigations showed, that the human *vigilin* gene differs from the chicken gene by two additional 5′ exons, designated exon 1A and 1B, which are alternatively though concurrently spliced to exon 1C, which is homologous to exon 1 in the chicken gene [11]. Exon 1A contains three start codons [20], all of which are followed by stop codons at a relatively short distance. The nucleotides surrounding the first start codon of exon 1A match the so called Kozak sequence for strong translation initiation [30], while the second and third start codons have no similarity to this sequence [20]. As demonstrated here, the first start codon is followed by an uORF, which is translated into an immunoreactive peptide of 13 amino acids whose expression suggests a modulating impact on translational events. Exon 1B does not contain this ORF. Interestingly, the uORF encoding the peptide in exon 1A is highly conserved in the mouse and human gene, including the start and stop codons, giving further support for a possible functional role. Using RT-PCR analysis we have shown that the two splice variants 1A and 1B were endogenously transcribed throughout murine embryonic stem cell differentiation in vitro without evidence for any developmental or tissue-specific regulation (data not shown). Similarly, both these variants have been shown to be expressed in many different tissues and cell lines of human and animal origin and we never observed that a change in the relative ratio of the two vigilin splice variants correlated with vigilin synthesis [20; C. Kruse, unpublished data]. These results are in line with the notion that the splice variants probably have no tissue-specific or developmental function but rather may be distinctive regulatory targets in response to housekeeping activities, e. g. proliferation, stress response or nutritional activity [22].

Vigilin translation is regulated by the exon-1A-uORF-encoded peptide

To focus in on any functional implication of the peptide encoded by the *vigilin* exon 1A uORF, we used cell-free and cellular translation assays and found that the two transcripts are translated into similar protein amounts, but disruption of the uORF in the VIG-1A transcript results in a doubling of protein synthesis from the mutated transcript. uORFs which repress downstream translation have only been described in a limited number of non-vertebrate and vertebrate genes showing a tissue-specific expression pattern [19]. Here we demonstrate for the first time that an uORF-encoded peptide which downregulates translation is also found in the *vigilin* gene. From the observation that both *vigilin* mRNA species are translated into similar amounts of protein, one may argue that the *vigilin* exon 1B splice variant has a translational threshold formed by its high GC content. A detailed analysis of several hundred eukaryotic mRNAs has shown that numerous genes contain either uORFs or GC-burdened 5′ leader sequences [17, 31]. Both mechanisms seem to prevent excessive protein synthesis by interfering with the process of ribosomal scanning and initiation. Both uORFs and GC-rich sequences have been shown to have

an inhibitory effect on protein synthesis from these transcripts [32–37].

We found further evidence that the vigilin exon-1A-uORF-encoded peptide plays an inhibitory role in vigilin synthesis when analyzing primary cultures of fibroblasts. In contrast to other cells analyzed, we found that primary human skin fibroblasts did not produce any vigilin after 5 days in culture, whereas the mRNA level remained unchanged. Thus, the reduced translation rate had not been brought about by downregulation of the transcriptional activity or enhanced degradation of *vigilin* mRNA. At this time of downregulated vigilin synthesis, the exon-1A-uORF-encoded peptide was localized at ribosomes. From these results, we may assume that the exon-1A-uORF-encoded peptide is involved in posttranscriptional mechanisms of vigilin downregulation. The mechanism by which the uORF-encoded peptide in the *vigilin* mRNA inhibits downstream translation remains to be elucidated. In summary, our results show that the highly conserved uORF peptide in exon 1A of the *vigilin* gene fulfills an important function in the regulation of vigilin synthesis in mammals.

Acknowledegments. The skilful technical assistance of E. Klink and B. Lembrich is gratefully acknowledged. The work was supported by the Deutsche Forschungsgemeinschaft (Kr 1512/1-4).

1 Gibson T. J., Thompson J. D. and Heringa J. (1993) The KH domain occurs in a diverse set of RNA-binding proteins that include the antiterminator NusA and is probably involved in binding to nucleic acid. FEBS Lett. **324:** 361–366
2 Musco G., Stier G., Joseph C., Castiglione Morelli M. A., Nilges M., Gibson T. J. et al. (1996) Three-dimensional structure and stability of the KH domain: molecular insights into the fragile X syndrome. Cell **85:** 237–245
3 Siomi H., Matunis M. J., Michael W. M. and Dreyfuss G. (1993) The pre-mRNA binding K protein contains a novel evolutionarily conserved motif. Nucleic Acids Res. **21:** 1193–1198
4 Ashley C. T. Jr, Wilkinson K. D., Reines D. and Warren S. T. (1993) FMR1 protein: conserved RNP family domains and selective RNA binding. Science **262:** 563–566
5 Liu K. and Hanna M. M. (1995) NusA contacts nascent RNA in *Escherichia coli* transcription complexes. J. Mol. Biol. **247:** 547–558
6 Nandabalan K., Price L. and Roeder G. S. (1993) Mutations in U1 snRNA bypass the requirement for a cell type-specific RNA splicing factor. Cell **73:** 407–415
7 Siebel C. W., Admon A. and Rio D. C. (1995) Soma-specific expression and cloning of PSI, a negative regulator of P element pre-mRNA splicing. Genes Dev. **9:** 269–283
8 Schmidt C., Henkel B., Pöschl E., Zorbas H., Purschke W. G., Gloe T. R. et al. (1992) Complete cDNA sequence of chicken vigilin, a novel protein with amplified and evolutionary conserved domains. Eur. J. Biochem. **206:** 625–634
9 Lang B. D. and Fridovich-Keil J. L. (2000) Scp160p, a multiple KH-domain protein, is a component of mRNP complexes in yeast. Nucleic Acids Res. **28:** 1576–1584
10 McKnight G. L., Reasoner J., Gilbert T., Sundquist K. O., Hokland B., McKernan P. A. et al. (1992) Cloning and expression of a cellular high density lipoprotein-binding protein that is up-regulated by cholesterol loading of cells. J. Biol. Chem. **267:** 12131–12141
11 Henkel B., Schmidt C., Zorbas H., Pöschl E., Gloe T. R., Purschke W. G. et al. (1992) Chicken vigilin gene organization and expression pattern. The domain structure of the protein is reflected by the exon structure. Eur. J. Biochem. **209:** 321–328
12 Kruse C., Grünweller A., Willkomm D. K., Pfeiffer T., Hartmann R. K. and Müller P. K. (1998) tRNA is entrapped in similar, but distinct, nuclear and cytoplasmic ribonucleoprotein complexes, both of which contain vigilin and elongation factor 1 alpha. Biochem. J. **329:** 615–621
13 Kruse C., Willkomm D. K., Grünweller A., Vollbrandt T., Sommer S., Busch S. et al. (2000) Export and transport of tRNA are coupled to a multi-protein complex. Biochem. J. **346:** 107–115
14 Kruse C., Emmrich J., Rumpel E., Klinger M. H., Grünweller A., Rohwedel J. et al. (1998) Production of trypsin by cells of the exocrine pancreas is paralleled by the expression of the KH protein vigilin. Exp. Cell Res. **239:** 111–118
15 Day D. A. and Tuite M. F. (1998) Post-transcriptional gene regulatory mechanisms in eukaryotes: an overview. J. Endocrinol. **157:** 361–371
16 Klausner R. D. and Harford J. B. (1989) Cis-trans models for post-transcriptional gene regulation. Science **246:** 870-872
17 Kozak M. (1987) An analysis of 5′-noncoding sequences from 699 vertebrate messenger RNAs. Nucleic Acids Res. **15:** 8125–8148
18 Kozak M. (1991) An analysis of vertebrate mRNA sequences: intimations of translational control. J.Cell Biol. **115:** 887-903
19 Morris D. R. and Geballe A. P. (2000) Upstream open reading frames as regulators of mRNA translation. Mol. Cell Biol. **20:** 8635–8642
20 Kügler S., Plenz G. and Müller P. K. (1996) Two additional 5′ exons in the human Vigilin gene distinguish it from the chicken gene and provide the structural basis for differential routes of gene expression. Eur. J. Biochem. **238:** 410–417
21 Aumailley M., Krieg T., Razaka G., Müller P. K. and Bricaud H. (1982) Influence of cell density on collagen biosynthesis in fibroblast cultures. Biochem. J. **206:** 505–510
22 Kruse C., Grünweller A., Notbohm H., Kügler S., Purschke W. G. and Müller P. K. (1996) Evidence for a novel cytoplasmic tRNA-protein complex containing the KH-multidomain protein vigilin. Biochem. J. **320:** 247–252
23 Kügler S., Grünweller A., Probst C., Klinger M., Müller P. K. and Kruse C. (1996) Vigilin contains a functional nuclear localisation sequence and is present in both the cytoplasm and the nucleus. FEBS Lett. **382:** 330–334
24 Chomczynski P. and Sacchi N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. **162:** 156–159
25 Grünweller A., Purschke W. G., Kügler S., Kruse C. and Müller P. K. (1997) Chicken vigilin gene: a distinctive pattern of hypersensitive sites is characteristic for its transcriptional activity. Biochem. J. **326:** 601–607
26 Nigg E. A. (1997) Nucleocytoplasmic transport: signals, mechanisms and regulation. Nature **386:** 779–787
27 Weber V., Wernitznig A., Hager G., Harata M., Frank P. and Wintersberger U. (1997) Purification and nucleic-acid-binding properties of a *Saccharomyces cerevisiae* protein involved in the control of ploidy. Eur. J. Biochem. **249:** 309–317
28 Mahone M., Saffman E. E. and Lasko P. F. (1995) Localized bicaudal-C RNA encodes a protein containing a KH domain, the RNA binding motif of FMR1. EMBO J. **14:** 2043–2055
29 Dodson R. E. and Shapiro D. J. (1997) Vigilin, a ubiquitous protein with 14 K homology domains, is the estrogen-inducible vitellogenin mRNA 3′-untranslated region-binding protein. J. Biol. Chem. **272:** 12249–12252
30 Kozak M. (1986) Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. Cell **44:** 283–292

31 Kozak M. (1996) Interpreting cDNA sequences: some insights from studies on translation. Mamm. Genome **7:** 563–574
32 Kauppinen L. (1995) Regulation of the human spermidine synthase mRNA translation by its 5′-untranslated region. FEBS Lett. **365:** 61–65
33 Reynolds K., Zimmer A. M. and Zimmer A. (1996) Regulation of RAR beta 2 mRNA expression: evidence for an inhibitory peptide encoded in the 5′-untranslated region. J. Cell Biol. **134:** 827–835
34 Harigai M., Miyashita T., Hanada M. and Reed J. C. (1996) A cis-acting element in the BCL-2 gene controls expression through translational mechanisms. Oncogene **12:** 1369-1374
35 Tang W. and Tseng H. (1999) A GC-rich sequence within the 5′ untranslated region of human basonuclin mRNA inhibits its translation. Gene **237:** 35–44
36 Raney A., Baron A. C., Mize G. J., Law G. L. and Morris D. R. (2000) In vitro translation of the upstream open reading frame in the mammalian mRNA encoding S-adenosylmethionine decarboxylase. J. Biol. Chem. **275:** 24444–24450
37 Jousse C., Bruhat A., Carraro V., Urano F., Ferrara M., Ron D. et al. (2001) Inhibition of CHOP translation by a peptide encoded by an open reading frame localized in the chop 5′ UTR. Nucleic Acids Res. **29:** 4341–4351