

## TAP (NXF1) Belongs to a Multigene Family of Putative RNA Export Factors with a Conserved Modular Architecture

ANDREA HEROLD,<sup>1</sup> MIKITA SUYAMA,<sup>1</sup> JOÃO P. RODRIGUES,<sup>2</sup> ISABELLE C. BRAUN,<sup>1</sup> ULRIKE KUTAY,<sup>3</sup>  
MARIA CARMO-FONSECA,<sup>2</sup> PEER BORK,<sup>1</sup> AND ELISA IZAURRALDE<sup>1\*</sup>

*European Molecular Biology Laboratory, D-69117 Heidelberg, Germany<sup>1</sup>; Institute of Histology and Embryology, Faculty of Medicine, University of Lisbon, 1699 Lisbon Codex, Portugal<sup>2</sup>; and Institute of Biochemistry, Swiss Federal Institute of Technology, CH-8092 Zürich, Switzerland<sup>3</sup>*

Received 12 July 2000/Returned for modification 15 August 2000/Accepted 6 September 2000

**Vertebrate TAP (also called NXF1) and its yeast orthologue, Mex67p, have been implicated in the export of mRNAs from the nucleus. The TAP protein includes a noncanonical RNP-type RNA binding domain, four leucine-rich repeats, an NTF2-like domain that allows heterodimerization with p15 (also called NXT1), and a ubiquitin-associated domain that mediates the interaction with nucleoporins. Here we show that TAP belongs to an evolutionarily conserved family of proteins that has more than one member in higher eukaryotes. Not only the overall domain organization but also residues important for p15 and nucleoporin interaction are conserved in most family members. We characterize two of four human TAP homologues and show that one of them, NXF2, binds RNA, localizes to the nuclear envelope, and exhibits RNA export activity. NXF3, which does not bind RNA or localize to the nuclear rim, has no RNA export activity. Database searches revealed that although only one *p15 (nxt)* gene is present in the *Drosophila melanogaster* and *Caenorhabditis elegans* genomes, there is at least one additional p15 homologue (p15-2 [also called NXT2]) encoded by the human genome. Both human p15 homologues bind TAP, NXF2, and NXF3. Together, our results indicate that the TAP-p15 mRNA export pathway has diversified in higher eukaryotes compared to yeast, perhaps reflecting a greater substrate complexity.**

mRNAs are exported from the nucleus as large ribonucleoprotein complexes (mRNPs). To date, proteins directly implicated in this process include several nucleoporins and RNA binding proteins (hnRNPs), an RNA helicase of the DEAD-box family (Dbp5), and the nuclear pore complex (NPC)-associated proteins Gle1p, TAP and Mex67p, and RAE1 (also called Gle2p) (reviewed in references 22, 28, and 32). Mex67p is essential for mRNA export in *Saccharomyces cerevisiae*, while RAE1 is essential for mRNA export in *Schizosaccharomyces pombe* (9, 27, 36). Their vertebrate homologues, TAP and RAE1, have also been implicated in the export of cellular mRNAs (6, 8, 12, 15, 20, 31).

We identified TAP as the cellular factor which is recruited by the constitutive transport element (CTE) of simian type D retroviruses to promote nuclear export of their genomic RNAs (12). In *Xenopus* oocytes, titration of TAP with an excess of CTE RNA prevents cellular mRNAs from exiting the nucleus (12, 30, 33). This suggests a role for this protein in the export of cellular mRNA.

Considerable progress has been made in defining TAP structural and functional domains (see Fig. 1) and in identifying its binding partners. TAP partners include various nucleoporins (4, 17); p15 (also called NXT1), a protein related to nuclear transport factor 2 (NTF2) (7, 17); transportin, which mediates TAP nuclear import (4); and several mRNP-associated proteins, such as E1B-AP5, RAE1 (4), and members of the Yra1p/REF family of proteins (37, 39). Binding of TAP to these mRNP-associated proteins is mediated by its N-terminal domain (residues 1 to 372) (4, 39). This domain includes a non-canonical RNP-type RNA binding domain (RBD) and four

leucine-rich repeats (LRRs) and exhibits general RNA binding affinity and specific binding to the CTE RNA (8, 12, 24).

TAP heterodimerizes with p15 via its NTF2-like domain (residues 371 to 551) (4, 40). Nucleoporin binding by TAP in vitro and in vivo is mediated by a domain located at the very C-terminal end of the protein (residues 508 to 619) (4, 5). The similarities of this domain to the ubiquitin-associated (UBA) domain allowed the prediction that residues located in a conserved loop were implicated in TAP-nucleoporin interaction (40).

In this study, we identify TAP homologues in *Homo sapiens*, *Caenorhabditis elegans*, and *Drosophila melanogaster*. The overall domain organization of these proteins, including residues important for p15 and nucleoporin interaction, is conserved. Two of four putative human TAP homologues, NXF2 and NXF3, have been characterized in detail. These proteins show 57 and 44% sequence identity to TAP. Neither binds specifically to the CTE RNA, indicating that they may exhibit different substrate specificities. Both proteins are localized within the nucleoplasm, but only NXF2 associates with the nuclear envelope and exhibits detectable RNA export activity. In addition, a human homologue of p15 (p15-2 [also called NXT2]) was characterized. p15-2 binds to TAP, NXF2, and NXF3 with affinities similar to those of p15. Like p15-1, a fraction of p15-2 localizes to the nuclear envelope. Both human p15 homologues participate in RNA nuclear export. Together, our results indicate that the TAP-p15-mediated export pathway has diversified in higher eukaryotes and in humans includes at least two p15 proteins and multiple TAP-like putative mRNA export factors.

### MATERIALS AND METHODS

**Homology searches and sequence analysis.** Homologues of TAP and p15 were retrieved in expressed sequence tag (EST) databases, human genomic DNA, and the *Drosophila* and *C. elegans* genomes, as well as in various protein sequence databases, using the BLAST suite of programs (3). Multiple sequence alignments were constructed by using CLUSTAL W (42) and manually refined on the SEAVIEW alignment editor (11). To predict the genomic structure of TAP and

\* Corresponding author. Mailing address: EMBL, Meyerhofstrasse 1, D-69117 Heidelberg, Germany. Phone: 0049 6221 387 389. Fax: 0049 6221 387 518. E-mail: izaurrealde@embl-heidelberg.de.

p15 homologues in various organisms, the Genewise program (<http://www.sanger.ac.uk/Software/Wise2/>) was employed. If cDNA and/or EST sequences were available, the BLASTN program for genomic sequences was also used to determine genomic structures.

**Plasmids.** Full-length human NXF2, NXF3, p15-2a, and p15-2b cDNAs were amplified by PCR using the human testis Marathon-Ready cDNA library (Clontech) as a template and primers introducing the appropriate restriction sites. The 3' primers were designed according to the predicted cDNAs and included stop codons. The 5' primers were designed using sequence information obtained after performing 5' rapid amplification of cDNA ends reactions with the same human testis cDNA library and oligonucleotides lying in the p15-2a, p15-2b, NXF2, and NXF3 coding region that is represented in the corresponding human genomic sequences or I.M.A.G.E. Consortium cDNA clones. The NXF2 sequence is represented in the IMAGp998P054498Q2 and IMAGp998C234110Q2 clones. NXF3 is represented in the IMAGp998M08574Q2 clone ([see http://www.rzpd.de](http://www.rzpd.de) for further information). The p15-2a and p15-2b cDNAs are represented in the genomic sequence AL031387. The complete NXF2 and NXF3 cDNA was cloned into pGEMT-easy (Promega) and sequenced. p15-2a cDNA was directly cloned in pGEXCS (29) and sequenced. The cDNAs present in these plasmids were used for all further subcloning steps.

To generate a plasmid for *in vitro* translation, NXF2 cDNA was excised from pGEM T-easy with *Bam*HI and *Hind*III and inserted into the same restriction sites present in pBspALTER, a derivative of pALTER-Ex1 (Promega). NXF3 was excised with the enzymes *Nco*I and *Bam*HI and cloned into the *Nco*I/*Bam*HI sites of pBSSK-HA, a derivative of the pBSSK(+) vector (Stratagene) with the  $\beta$ -globin 5' untranslated region inserted between the *Hind*III and *Eco*RI sites. For the coexpression assay (see Fig. 5F), p15-1 and p15-2a cDNAs were cloned in the *Nco*I/*Not*I or *Nco*I/*Bam*HI sites of vector pET28c (Novagen), respectively.

Plasmids allowing the expression of glutathione *S*-transferase (GST) fusions of full-length NXF2 and NXF3 in *Escherichia coli* were generated by inserting the coding sequences into the *Bam*HI and *Not*I sites of pGEX4T-1 (NXF2) (Pharmacia) or the *Nco*I and *Bam*HI sites present in pGEXCS (NXF3). The NXF2 point mutants E598R, W599A, and N600A and the triple alanine substitution 3xAla598 were generated by using an oligonucleotide-directed *in vitro* mutagenesis system from Stratagene (Quick-change site-directed mutagenesis) in the context of pGEX4T-1-NXF2. The mutation 3xAla598 introduces a *Not*I site into the coding sequence of NXF2. Plasmid pGEX4T-1-NXF2 $\Delta$ 598-626 was made by digesting pGEX4T-1-NXF2 3xAla598 with *Ppu*MI and *Not*I and inserting the released fragment into pGEX4T-1-NXF2 cut with the same enzymes. Constructs encoding GST fusions of the NXF2 amino acids 1 to 377 and 102 to 203 (RBD) were generated by inserting the corresponding PCR products into pGEX4T-1 digested with *Bam*HI and *Not*I. A plasmid encoding a GST fusion of NXF3 amino acids 90 to 192 (RBD) was generated by inserting the corresponding PCR product into pGEXCS digested with *Nco*I-*Bam*HI. All PCRs were performed with the Expand high-fidelity PCR system (Roche). The integrity of the PCR products was confirmed by sequencing.

For expression of GFP fusions in mammalian cells, NXF2 and NXF3 cDNAs were excised from the pGEMT-easy clones with the restriction enzymes *Bam*HI/*Hind*III (NXF2) or *Eco*RI/*Bam*HI (NXF3) and inserted into the vector pEGFP-C1 (Clontech) digested with *Bgl*II/*Hind*III or *Eco*RI/*Bam*HI, respectively. pEGFP-N3 (Clontech) derivatives encoding *Staphylococcus aureus* protein A (zz) fusions of p15-1 and p15-2a were constructed in two steps: first, p15-1 and p15-2a cDNAs were cloned into the vector pRN3zz using the *Nco*I/*Bam*HI restriction sites. Then, fragments encoding the zz tag fusions were released by digesting the corresponding plasmids with *Hind*III/*Not*I and inserted into pEGFP-N3 cut with the same enzymes. Additional plasmids used in this study were previously described (4, 8, 24, 39).

**In vitro translation and expression of recombinant proteins.** For generation of <sup>35</sup>S-labeled *in vitro*-translated proteins, the combined *in vitro* transcription-translation (TnT) kit from Promega was used following the instructions of the manufacturer. Translation was checked by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and subsequent autoradiography. *E. coli* BL21(DE3) was used when proteins were expressed with the pGEX vectors, while *E. coli* M15[pREP4] was used for expressing proteins cloned into the pQE vectors. Recombinant proteins were purified as previously described (12).

**GST pull-down and *in vitro* RNA binding assays.** Gel retardation assays and GST pull-down assays were performed as previously described (4, 8, 39). For *in vitro* synthesis of a 43-nucleotide RNA probe, pBluescribe was linearized with *Bam*HI and transcribed with T3 RNA polymerase. The amounts of unlabeled competitor RNAs used per binding reaction are indicated in the figure legends.

**Ran binding assays.** The expression and purification of NTF2 and the expression of importin fragments have been described previously (21). Wild-type Ran or zzRanQ69L was immobilized on immunoglobulin G (IgG)-Sepharose in the presence of the Rna1 or RCC1 and energy-regenerating system, respectively. *E. coli* lysates (250  $\mu$ l) supplemented with equal amounts (5  $\mu$ M) of purified NTF2, p15-1, or p15-2a, and a lysate from *E. coli* expressing importin  $\beta$  (fragment 1-452) were subjected to binding to 15  $\mu$ l of IgG-Sepharose beads coated with RanGDP or RanQ69L-GTP. Binding was performed for 4 h at 4°C in a final volume of 1.5 ml of binding buffer (50 mM HEPES [pH 7.6], 200 mM NaCl, 2 mM magnesium acetate). After being washed three times with binding buffer, the bound protein was eluted with 1.5 M MgCl<sub>2</sub> and precipitated with isopropanol.

The starting material and the bound fractions were analyzed by SDS-PAGE and Coomassie blue staining.

**Immunofluorescence.** HeLa cells were transfected with FuGENE6 (Roche). Approximately 20 h after transfection, the cells were fixed with 3.7% formaldehyde for 10 min and subsequently permeabilized with 0.5% Triton X-100 for 15 min or were extracted first with 0.5% Triton X-100 for 1 min on ice and then fixed in formaldehyde. Indirect immunofluorescence assays were performed as previously described (2). The zz fusion proteins were visualized with a rabbit polyclonal anti-protein A antibody (Sigma) (dilution, 1:1,000 in phosphate-buffered saline supplemented with 5% fetal calf serum and 0.05% Tween 20) and a secondary Cy3-coupled anti-rabbit IgG antibody (diluted 1:4,000). Coverslips were mounted in VectaShield medium (Vector Labs).

**DNA transfection and CAT assays.** Human 293 cells were transfected by the calcium phosphate method. The cells were transfected at 50% confluency in 6-cm-diameter dishes with a plasmid DNA mixture. This mixture consisted of 0.5  $\mu$ g of pDM138 (13, 14), 1  $\mu$ g of pEGFP-C1 plasmid encoding green fluorescent protein (GFP)-tagged versions of TAP or TAP homologues, and 1  $\mu$ g of pEGFP-N3 derivatives encoding zz-tagged versions of p15-1 or p15-2a. The total amount of plasmid DNA transfected in each sample was held constant by adding the appropriate amount of the corresponding parental plasmids without an insert. The transfection efficiency was determined by including 0.5  $\mu$ g of pCH110 plasmid (Pharmacia) encoding  $\beta$ -galactosidase ( $\beta$ -Gal), as  $\beta$ -Gal expression from this vector is not affected by increasing levels of TAP expression. Quail QT6 cells were transfected using Lipofectamine Plus (Life Technologies). The cells were transfected at 50% confluency in six-well dishes with a plasmid DNA mixture. This mixture consisted of 0.5  $\mu$ g of pCH110, 0.1  $\mu$ g of pDM138 or pDM138-CTE, 0.5  $\mu$ g of pEGFP-C1 plasmids encoding GFP-tagged versions of TAP or TAP homologues, 0.5  $\mu$ g of pEGFP-N3zzp15-1, and 0.8  $\mu$ g of pBSISK. The cells were harvested 40 h posttransfection, and chloramphenicol acetyltransferase (CAT) activity was measured as described previously (26). Protein expression levels were analyzed by Western blotting with anti-GFP antibodies kindly provided by Patrick Keller.

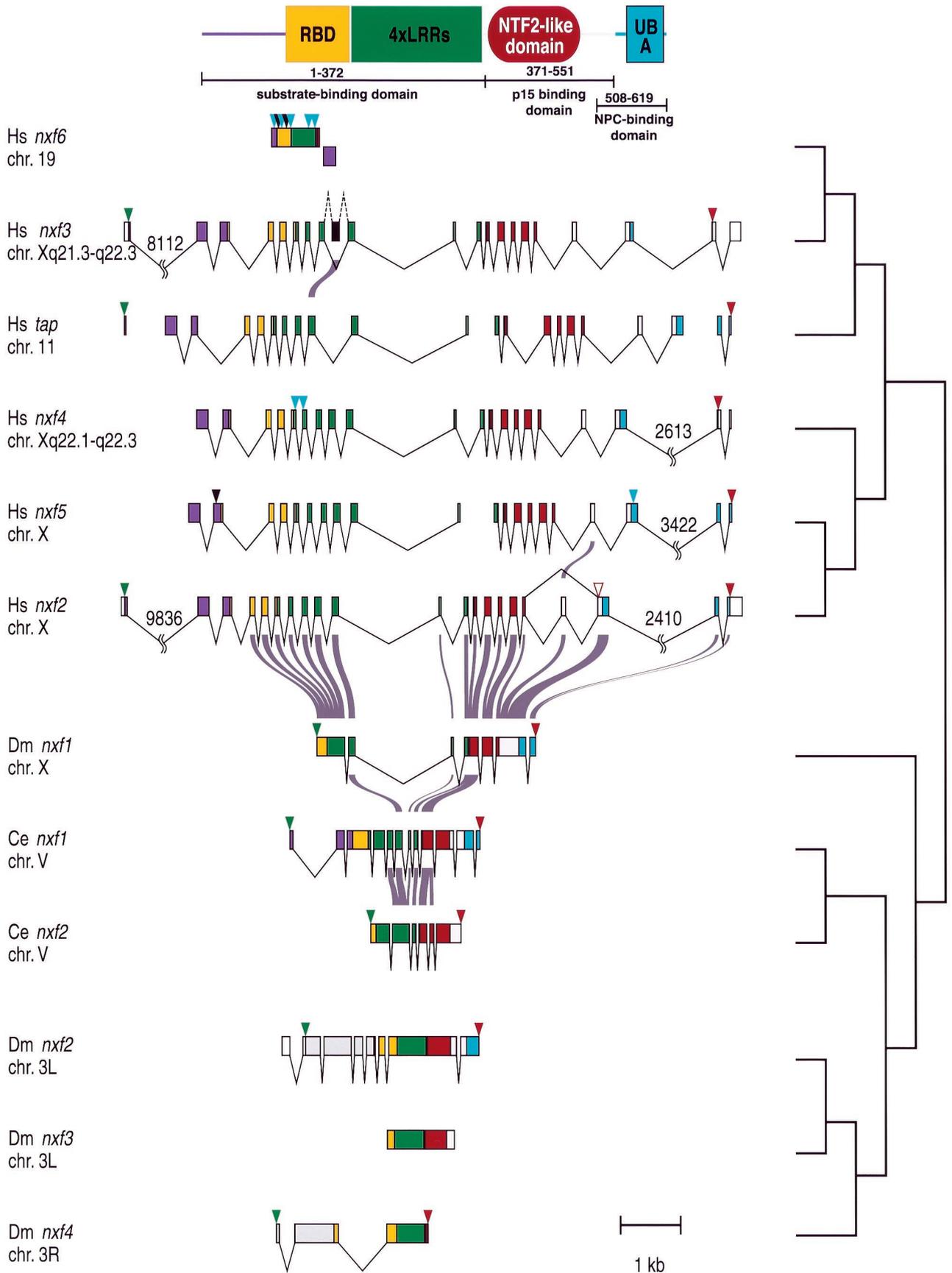
**Nucleotide sequence accession numbers.** The NXF2, NXF3, p15-2a, and p15-2b sequence data have been submitted to the EMBL database under accession numbers AJ277526, AJ277527, AJ277591, and AJ278323.

## RESULTS

**TAP belongs to a multigene family of evolutionarily conserved proteins.** We performed an extensive search for TAP homologues. Using BLAST searches (3), we identified two homologues in the *C. elegans* genome, four in the *Drosophila* genome, and, in addition to TAP itself, four putative homologues in the human genome, including the two previously identified (5) (Fig. 1). In agreement with the Human Genome Nomenclature Committee, genes encoding TAP-like proteins in higher eukaryotes were named *nxf* (for nuclear export factor). To avoid confusion, we will call human NXF1 TAP.

In *C. elegans*, the two TAP-like genes were annotated only as hypothetical proteins (10). In *Drosophila*, three sequences were predicted genes (*nxf1*, *nxf2*, and *nxf4*) that show partial similarity to TAP, while the fourth sequence (*nxf3*) was detected in a portion of the genome for which no genes were predicted (1). In the human genome, several ESTs and regions in genomic DNA showed high similarity to TAP. These could be assembled into six nonidentical cDNA fragments and were found to correspond to four distinct TAP homologues when full-length inserts of various EST clones were sequenced. TAP cDNA sequence and sequences from the fragments were used to scan human genomic DNA to identify coding exons. Using the Genewise program, the genomic structure of the four human candidate TAP homologues, in addition to those of TAP itself and of a pseudogene, were predicted (Fig. 1). TAP mapped on chromosome 11. An intronless region corresponding to *nxf6* was located on chromosome 19. Due to its fragmentary nature and the multiple frameshifts and stop codons (Fig. 1), *nxf6* is likely to be a pseudogene. All four of the other homologous candidate genes were located on the X chromosome (Fig. 1).

When the predicted gene structures are compared, all human *nxf* genes show similar intron-exon patterns, with the exception of *nxf6* (Fig. 1). In spite of the presence of frameshifts and in-frame stop codons (Fig. 1), *nxf5* and *nxf4* have



retained this genomic structure. cDNAs encoding various alternative splice forms of *nxf5* have recently been isolated (Y. Lin, S. Frints, G. Froyen, and P. Marynen, submitted for publication), and at least one EST (accession number, AI150002) that corresponds to *nxf4* was identified, indicating that some splice forms of these genes are expressed. Thus, the genes might be subjected to alternative splicing, thereby avoiding the frameshifted exons. Alternative splicing seems to be an important mechanism in this gene family, as ESTs representing multiple alternative spliced forms exist in the database. Furthermore, multiple splice forms were cloned by reverse transcription-PCR (data not shown; Lin et al., submitted; A. S. Zolotukhin and B. K. Felber, personal communication), suggesting that multiple protein variants may result from the expression of human *nxf* genes.

**Multiplication of *nxf* genes has occurred independently in different lineages.** Phylogenetic analysis of the NXF protein family indicates that separate gene duplication events have occurred in several eukaryotic lineages (Fig. 2). Thus, although higher eukaryotes have several TAP homologues, they have evolved independently, so there is no clear one-to-one relationship of any isoforms.

Comparison of the deduced amino acid sequences of the TAP-like proteins in all species, including yeast, indicates that the overall domain organization of the protein family has been evolutionarily conserved (40) (Fig. 1 and 3), although, due to alternative splicing, not all domains are always expressed. At the genomic level, the LRRs are present in all members of the family; however, ESTs representing *H. sapiens* NXF3 and the cDNA we have isolated lack exon 9 (Fig. 1) and, as a consequence, part of the LRRs (Fig. 3). With the exception of *D. melanogaster nxf4*, the NTF2-like domain is also present in all *nxf* genes, including yeast *mex67*; however, in some forms of *H. sapiens nxf2*, exon 18 is skipped, resulting in an internal deletion within this domain and the introduction of a premature stop codon downstream of the skipped exon (Fig. 1). The UBA-like domain is absent in *D. melanogaster nxf3* and *nxf4* and *C. elegans nxf2* but is present in all human *nxf* genes. Nevertheless, the presence of premature stop codons in the *H. sapiens* NXF3, NXF4, and NXF5 genes (Fig. 3) results in proteins lacking part of this domain.

**Characterization of the substrate binding domain of NXF2 and NXF3.** cDNAs encoding human NXF2 and NXF3 were cloned by reverse transcription-PCR and sequenced (Fig. 3). To investigate whether these proteins interact with E1B-AP5 or the REF proteins and exhibit RNA binding activity, we performed *in vitro* binding assays.

[<sup>35</sup>S]methionine-labeled NXF2 and NXF3 were synthesized *in vitro* in rabbit reticulocyte lysates and assayed for binding to glutathione agarose beads coated with either GST-E1B-AP5, GST-REF1-II, or GST. Binding of TAP to the recombinant proteins was tested in parallel. Figure 4A shows that TAP, NXF2, and NXF3 could be selected on glutathione agarose beads coated with E1B-AP5 (lane 3) but not on beads coated

with GST (lane 2). TAP and NXF2 bound to full-length REF1-II (fragment 1–163 [lane 6]) and to its C-terminal domain (fragment 103–163 [lane 5]) but not to its RBD (fragment 14–102 [lane 4]). In contrast, NXF3 did not interact with REF1-II. None of these interactions was affected by the presence of RNase A, indicating that they were not RNA mediated (data not shown; see reference 39 for TAP).

RNA and CTE binding activities were assayed by electrophoretic gel mobility retardation assays. To test general RNA binding affinity, a 43-nucleotide-long <sup>32</sup>P-labeled RNA probe was incubated with purified recombinant N-terminal domain of NXF2 (residues 1 to 377) fused to GST, and the resulting complexes were resolved in a native polyacrylamide gel and visualized by autoradiography (Fig. 4B). As a control, binding of GST-TAP (1 to 372) to the RNA probe was tested in parallel. The N-terminal domains of TAP and NXF2 bound to the RNA probe (Fig. 4B, lanes 2 and 6). Formation of both protein-RNA complexes was competed by the addition of increasing amounts of tRNA (Fig. 4B, lanes 3 to 5 and 7 to 9), suggesting that the proteins exhibit similar RNA binding affinities. The N-terminal domain of NXF3 could not be expressed in *E. coli*; however, full-length NXF3 coexpressed with p15-2a (see below) did not bind RNA (data not shown). The N-terminal domain of TAP contains a noncanonical RBD that exhibits general RNA binding activity (24). This domain is conserved in all vertebrate NXF proteins (24) (Fig. 1 and 3). Figure 4C shows that while the isolated RBDs of TAP and NXF2 bind RNA, the corresponding domain of NXF3 did not exhibit detectable RNA binding activity.

Binding to the CTE RNA probe was tested using *in vitro*-translated proteins. To assess the specificity of the interaction, the binding reactions were supplemented with increasing amounts of unlabeled CTE RNA or M36 RNA, a CTE derivative which does not bind TAP (12). Under conditions in which TAP bound specifically and with high affinity to the CTE RNA (Fig. 4D, lanes 3 to 6), neither NXF2 nor NXF3 interacted with the CTE probe (Fig. 4D, lanes 7 and 8). Moreover, recombinant NXF2 and NXF3, coexpressed in *E. coli* with p15-1 or p15-2a, did not show specific CTE binding (data not shown).

Next, we tested the abilities of NXF2 and NXF3 to promote CTE-dependent export of a precursor mRNA in quail cells. In this assay, TAP and TAP-like proteins were cotransfected with the reporter plasmid pDM138 (14) and its derivative pDM138-CTE. These plasmids harbor a single intron containing the CAT coding sequence, which is excised when the RNA is spliced (Fig. 4E). Cells transfected with the pDM138 plasmid express only the spliced transcripts in the cytoplasm and thus yield only trace levels of CAT enzyme activity (14). The presence of the CTE in the intron (pDM138-CTE) allows nuclear retention to be bypassed and export of the unspliced transcripts to be promoted (15). In quail cells, CTE-mediated export of this precursor mRNA requires coexpression of human TAP (15), which leads to an increase in CAT activity (Fig. 4F). This activation can be abolished by preventing TAP-nucleo-

FIG. 1. Intron-exon structure of the *nxf* genes. The domain organization of human TAP is indicated at the top. Hs, *H. sapiens*; Dm, *D. melanogaster*; Ce, *C. elegans*. When known, the chromosomal (chr.) locations are indicated below the gene names. Exons are colored according to the domains: purple, N-terminal portion found only in human homologues and *C. elegans nxf1*; yellow, RBD; green, LRR; red, NTF2-like domain; pink, linkers upstream and downstream of NTF2-like domain; cyan, UBA domain. 5' and 3' untranslated regions present in the ESTs or cDNA sequences are indicated by open boxes. Exons having no similarity to human TAP are gray. Introns are depicted as lines. The intron-exon structures are drawn to scale except for the long introns, which have breaks in the middle with the lengths indicated by numbers. An alternative splicing pathway is shown by lines above the gene (*nxf2*). A skipped exon is shown in black and connected by dotted lines (*nxf3*). Some characteristics of the sequences are indicated above the exon by filled triangles: green, initiation codon; red, termination codon; black, in-frame stop codon; cyan, frameshift. In *nxf2*, a stop codon created by alternative splicing is indicated as a red open triangle. On the complementary strand of *nxf6*, there is a region showing weak similarity to other TAP gene sequences; this region is shown in shift vertically. Some introns for the human TAP and NXF5 genes are not shown because the genes are mapped on distinct fragments of the genome sequences and the lengths of the introns are not clear. On the right, a simplified phylogenetic tree (not to scale) is shown.

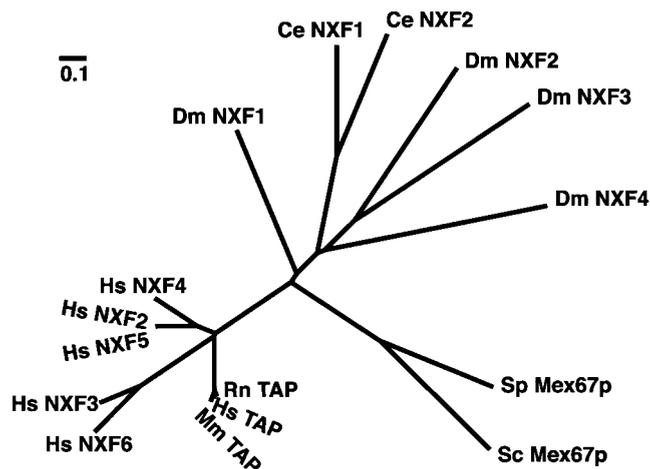


FIG. 2. Phylogenetic tree of NXF family sequences. The tree was drawn by the neighbor-joining method (34). Abbreviations (other than those in Fig. 1): Mm, *Mus musculus*; Rn, *Rattus norvegicus*; Sc, *S. cerevisiae*; Sp, *S. pombe*.

porin interaction (15) (Fig. 4F, TAP $\Delta$ NPC) or by deleting the RBD required for CTE binding (24). In contrast to TAP, neither NXF2 nor NXF3 could promote specific CTE-dependent export (Fig. 4F), although these proteins were expressed at comparable levels (data not shown).

Thus, NXF2 displays general affinity for RNA as reported previously for TAP and Mex67p (12, 17, 35, 39) and interacts with E1B-AP5 and Ref1-II, whereas NXF3 only interacts with E1B-AP5 and does not exhibit detectable RNA binding activity. Neither protein specifically interacts with the CTE RNA or mediates CTE-dependent export.

**p15-2, a human p15 homologue that interacts with TAP and localizes to the nuclear rim.** The NTF2-like domain of TAP heterodimerizes with p15 and is conserved in most members of the NXF family. Searches for p15 homologues revealed one *nxt*-like gene in *Drosophila* and *C. elegans* and two in available human genomic sequences (Fig. 5A). The additional human *nxt* gene was named *nxt2*. The proteins encoded by these human genes will be referred to as p15-1 and p15-2, respectively. *nxt2* is located on the X chromosome, and its genomic structure more closely resembles those of the homologues in other species than does the intronless *nxt1* gene on chromosome 20 (Fig. 5A). Human EST sequences include two alternative splice variants of *nxt2* (p15-2a and p15-2b) which differ in their 5' exons (Fig. 5A). The predicted p15-2 forms were confirmed by cloning and sequencing the corresponding cDNAs. Alignment of the deduced amino acid sequence of p15-2a with known homologues and phylogenetic analysis suggest that the two human *nxt* genes are the result of a recent duplication in the *nxt* lineage (Fig. 5B and C).

The subcellular localization of p15-2a fused to two IgG-binding units of protein A from *Staphylococcus aureus* (zz tag) was analyzed in transfected HeLa cells. Figure 5D shows that p15-2a was evenly distributed in the nucleoplasm and was excluded from the nucleolus. Furthermore, a fraction of the protein was detected in the cytoplasm. To investigate whether

p15-2a associates with the nuclear envelope, transfected HeLa cells were extracted with Triton X-100 prior to fixation (Fig. 5D, +Triton X-100). Under these conditions, most of the nucleoplasmic and cytoplasmic pools of the protein were solubilized. However, a fraction of p15-2a was resistant to detergent extraction and was clearly visualized in a rim at the nuclear periphery. Similar results were obtained when p15-2a was fused to GFP (data not shown). Thus, the subcellular localization of p15-2a is similar to that previously reported for p15-1 (7, 17).

p15-1 interacts with TAP and is closely related to NTF2 (Fig. 5C) (7, 17, 40). We therefore investigated whether p15-2a could interact with TAP or Ran. Lysates from *E. coli* supplemented with equimolar amounts of recombinant purified p15-1, p15-2a, or NTF2 were incubated with IgG-Sepharose beads coated with purified zzRanGDP, zzRanQ69L-GTP, or zzTAP. The RanQ69L mutant is GTPase deficient and remains in the GTP-bound form (19). In contrast to Black et al. (7), but in agreement with Katahira et al. (17), we could not observe binding of p15-1 or p15-2a to Ran (Fig. 5E, lanes 7, 8, 11, and 12). However, both proteins were selected on immobilized TAP (Fig. 5E, lanes 15 and 16), suggesting that they were properly folded. In addition, we could not detect binding of TAP-p15 heterodimers to RanGDP or to RanGTP (data not shown). Under the same conditions, NTF2 bound to RanGDP (Fig. 5E, lane 6), while the Ran binding domain of Importin  $\beta$  (fragment 1-452) bound RanQ69L-GTP (Fig. 5E, lane 13). Thus, both p15-1 and p15-2a directly interact with TAP but not with Ran.

Next, we tested whether the p15 proteins could also interact with NXF2 or NXF3. Untagged p15-1 and p15-2a were coexpressed in *E. coli* along with TAP, NXF2, or NXF3 fused to GST. The bacterial lysates were incubated with glutathione agarose beads, and after extensive washes, the bound proteins were eluted with SDS-sample buffer. Both p15-1 and p15-2a were copurified with TAP, NXF2, and NXF3 but not with GST (Fig. 5F). Moreover, coexpression of the NXF2 and NXF3 proteins with p15-1 or p15-2a significantly increased their stability, as full-length NXF3 could not be expressed in *E. coli* in the absence of p15 (data not shown).

**NXF2, but not NXF3, binds to nucleoporins and localizes to the nuclear rim.** To test nucleoporin binding of NXF2 and NXF3, we immobilized bacterially expressed GST-CAN (fragment 1690-2090) or GST on glutathione agarose beads. The beads were then incubated with in vitro-translated TAP, NXF2, or NXF3. Both TAP and NXF2 bound CAN, while no significant binding was observed for NXF3 (Fig. 6A). In order to determine whether NXF2 could interact with other FG-repeat containing nucleoporins, pull-down assays were performed with recombinant NXF2 or TAP fused to GST and in vitro-translated CAN (fragment 1690-1894), Nup153 (fragment 895-1475), or full-length Nup98 and p62. The TAP mutants W594A and D595R were used as negative controls (40). With the exception of Nup98, the nucleoporins tested bound to NXF2 with efficiencies similar to those with TAP (Fig. 6B, lane 6 versus lane 3).

Residues located at positions 593 to 595 in the TAP sequence (NWD [Fig. 3]) have been implicated in nucleoporin

FIG. 3. Multiple sequence alignment of NXF family sequences. First column, species names (Hs, *H. sapiens*; Dm, *D. melanogaster*; Ce, *C. elegans*); second column, protein names; third column, positions of the first aligned residues in each of the sequences. The positions conserved in 80% of the sequences are indicated in the consensus line: a, aromatic (FHXY); c, charged (DEHKR); h, hydrophobic (ACFGHIKLMRTVWY); l, aliphatic (LIV); o, hydroxyl (ST); p, polar (CDEHKRQST); s, small (ACDGNPSTV); t, turnlike (ACDEGHKRNQST); u, tiny (AGS). The assigned domains are indicated below the consensus line. Highly conserved residues are indicated by colored boldface characters: orange, polar; light green, tiny; dark green, hydrophobic; blue, proline; light blue, hydroxyl; purple, cysteine. Exon boundaries are indicated by red marks.



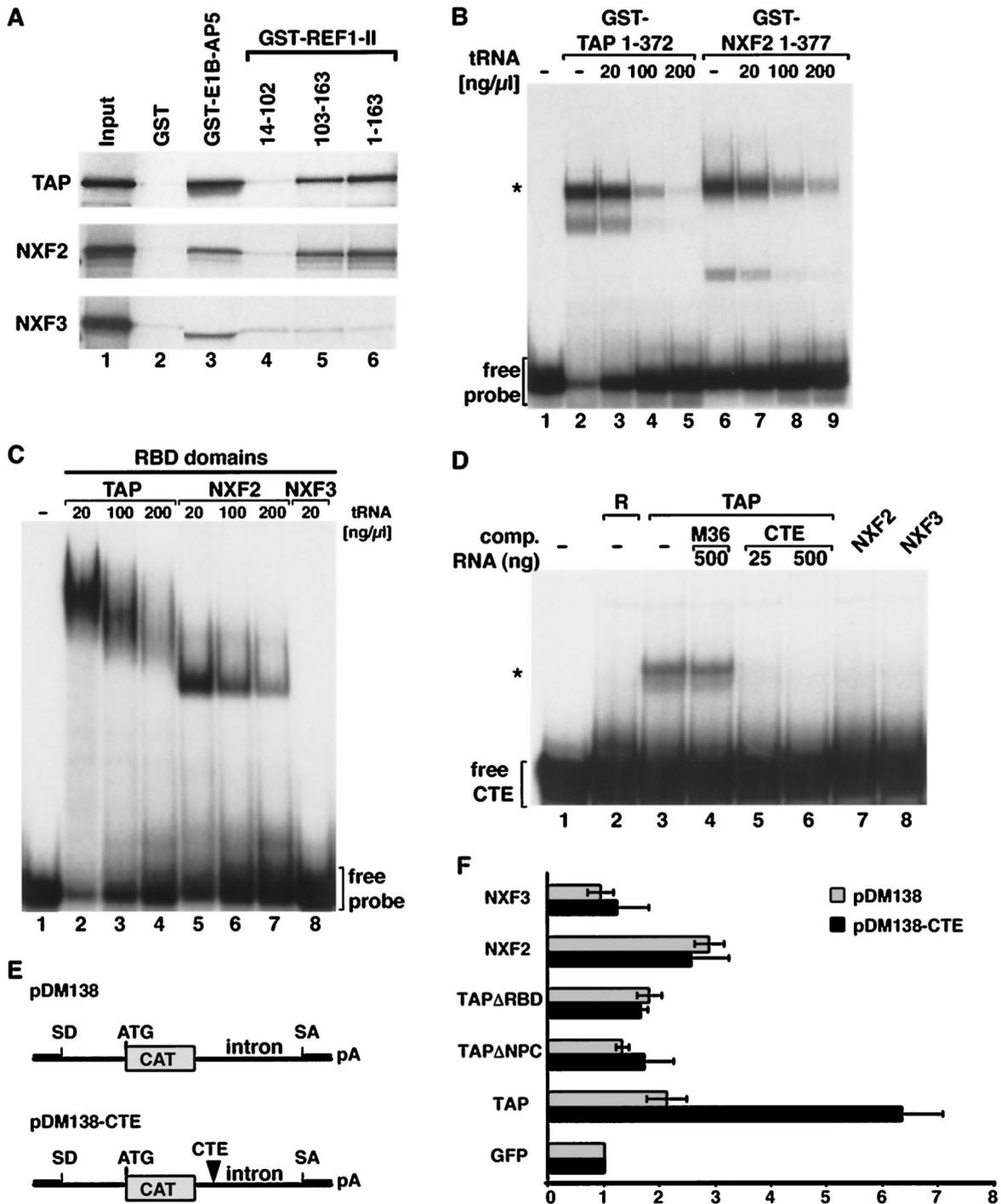


FIG. 4. Characterization of the N-terminal domains of NXF2 and NXF3. (A) GST pull-down assays were performed with [ $^{35}$ S]methionine-labeled TAP, NXF2, NXF3, and the recombinant proteins indicated above the lanes. Lanes 2, background obtained with glutathione agarose beads coated with GST; lanes 3, proteins selected on immobilized GST-E1B-AP5 (fragment 101–453); lanes 4 to 6, binding to GST-REF1-II or fragments of this protein as indicated. In all panels, 1/10 of the inputs (lanes 1) and 1/3 of the bound fractions (lanes 2 to 6) were analyzed on SDS-PAGE followed by fluorography. Supernatant fractions were analyzed in parallel in order to confirm that the absence of binding was not due to protein degradation (data not shown). (B and C) Gel mobility retardation assays were performed using a radiolabeled RNA probe derived from pBS polylinker and purified recombinant proteins fused to GST. (B) TAP or NXF2 N-terminal fragments (50 ng) were used; (C), 1.5  $\mu$ g of the corresponding RNA binding domains were used. Unlabeled competitor tRNA was added as indicated above the lanes. The position of the free RNA probe is indicated. The asterisk indicates the positions of the RNA-protein complexes. (D) A gel mobility retardation assay was performed with reticulocyte lysates unprogrammed (R) or programmed with cDNAs encoding TAP, NXF2, or NXF3. In lane 4, unlabeled M36 competitor RNA was added, while in lanes 5 and 6, CTE competitor RNA was included in the reaction mixtures. The amounts of the competitor RNAs are indicated above the lanes. The position of the free RNA probe is indicated on the left. (E) Schematic representation of pDM138 and pDM138-CTE vectors (14). (F) Quail cells were transfected with plasmids pDM138 or pDM138-CTE along with various plasmids encoding either GFP alone or fused to the N termini of TAP, NXF2, NXF3, or the TAP mutants indicated on the left. TAP $\Delta$ NPC has a deletion of residues 541 to 613. The cells were collected 40 h after transfection, and CAT activity was determined. Data from three separate experiments were expressed relative to the activities measured when GFP alone was coexpressed with pDM138 with or without CTE. The data are means  $\pm$  standard deviations.

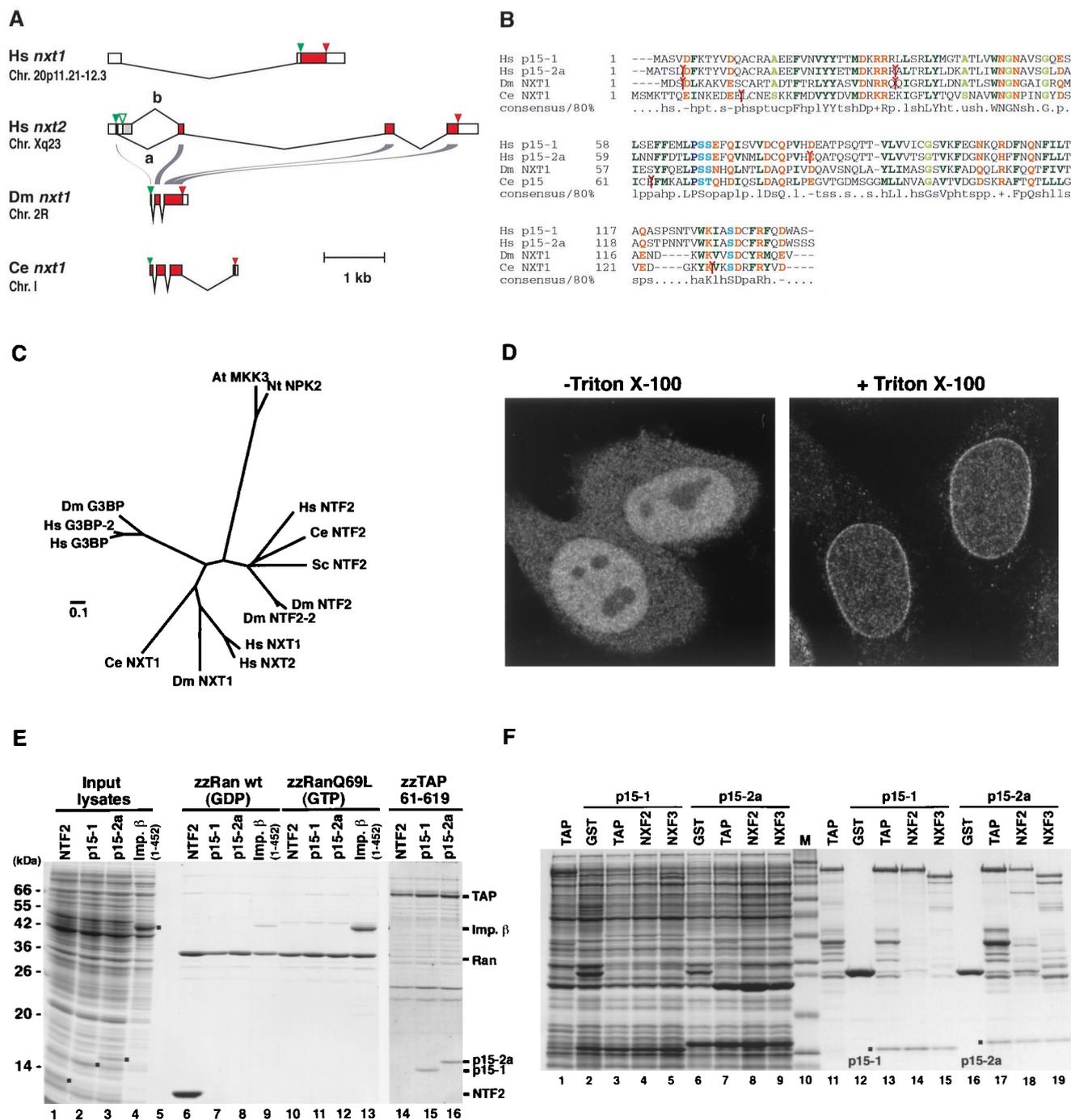


FIG. 5. p15-2a, a human p15-1 homologue, interacts with TAP and localizes to the nuclear rim. (A) Intron-exon structures of p15 family sequences. Protein coding regions and untranslated regions are colored red and white, respectively. The positions of initiation and termination codons are indicated by green and red triangles, respectively. For human p15-2a and -b, the alternative splicing pathway is shown by lines above the gene. The 5' exon of the p15-2a gene contains an open reading frame of five amino acids, while p15-2b gene 5' exon sequences contain an in-frame stop codon but no in-frame ATG. Therefore, translation of this mRNA may generate a truncated protein starting at methionine 29 in the p15-2a sequence. Alternatively, translation may start at the GTG codon (green open triangle), resulting in an open reading frame (gray), since this region does not show any similarity to the other p15 sequences. (B) Multiple sequence alignment of p15 family sequences. The symbols are as in Fig. 3. (C) Phylogenetic tree of the NTF2 family sequences. The tree was drawn by the neighbor-joining method (34). Abbreviations (other than those in Fig. 1 and 2): *At*, *Arabidopsis thaliana*; *Nt*, *Nicotiana tabacum*; G3BP, Ras-GAP SH3 domain binding protein; MKK3, MAP kinase 3; NPK2, a tobacco protein kinase. (D) Subcellular localization of p15-2a. HeLa cells were transfected with a pEGFP-N3 plasmid derivative expressing a zz fusion of p15-2a. The fusion protein was detected throughout the nucleoplasm and cytoplasm and was excluded from the nucleolus (left). On the right, HeLa cells were extracted with (+) Triton X-100 prior to fixation. A punctate labeling pattern was visible at the nuclear periphery for the p15-2a protein. (E) Lysates from *E. coli* expressing the Ran binding domain of Importin  $\beta$  (fragment 1-452) or supplemented with equimolar amounts of NTF2, p15-1, or p15-2a were incubated with IgG-Sepharose beads coated with purified zzRanGDP, zzRanQ69L-GTP, or zzTAP (fragment 61-619). After extensive washes, the bound proteins were eluted. One-hundredth of the inputs (lanes 1 to 4) and 1/10 of the bound fractions (lanes 6 to 16) were analyzed by SDS-PAGE followed by Coomassie blue staining. (F) Lysates from *E. coli* expressing GST fusions of TAP, NXF2, or NXF3 together with untagged versions of p15-1 or p15-2a were incubated with glutathione agarose beads. After extensive washes, the bound proteins were eluted with SDS sample buffer and analyzed by SDS-PAGE followed by Coomassie blue staining. Lanes 1 to 9, input lysates; lanes 11 to 19, bound fractions; lane 10, molecular mass markers (116, 97, 84, 66, 55, 45, 36, 29, 24, 20, and 14.2 kDa).

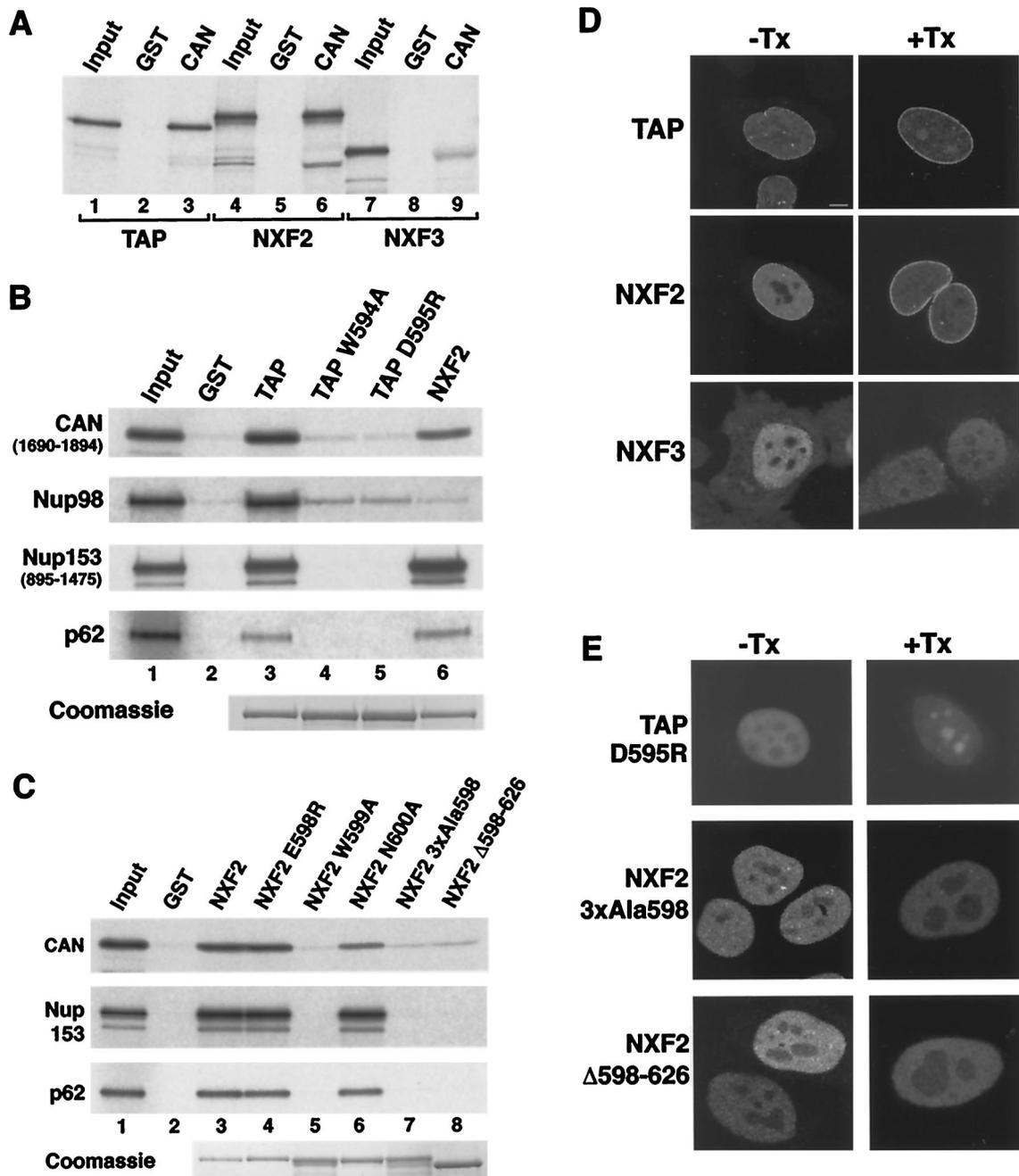


FIG. 6. A fraction of NXF2 localizes to the nuclear rim. (A) GST pull-down assays were performed with [ $^{35}$ S]methionine-labeled TAP, NXF2, or NXF3, and the recombinant proteins indicated above the lanes. One-tenth of the inputs (lanes 1, 4, and 7) and one-third of the bound fractions (lanes 2, 3, 5, 6, 8, and 9) were analyzed by SDS-PAGE followed by fluorography. (B) NXF2 interacts with multiple components of the NPC. GST pull-down assays were performed with the [ $^{35}$ S]methionine-labeled nucleoporins indicated on the left of the panels and recombinant GST or TAP, TAP W594A, TAP D595R, or NXF2 fused to GST, as indicated above the lanes. One-tenth of the input (lane 1) and one-third of the bound fractions (lanes 2 to 6) were analyzed on SDS-PAGE followed by fluorography. (C) GST pull-down assays were performed with the [ $^{35}$ S]methionine-labeled nucleoporins indicated on the left of the panels and recombinant GST, GST-NXF2, or various NXF2 mutants as indicated above the lanes. Samples were analyzed as indicated for panel A. (D and E) HeLa cells were transfected with pEGFP-C1 plasmid derivatives expressing GFP fusions of TAP, NXF2, and NXF3 or various TAP or NXF2 mutants as indicated on the left. Approximately 20 h after transfection, the cells were fixed in formaldehyde, permeabilized with Triton X-100, and directly observed with a fluorescence microscope. For all proteins, the GFP signal was detected throughout the nucleoplasm (-Tx). Cytoplasmic staining was also detected for NXF3. On the right (+Tx), HeLa cells were extracted with Triton X-100 prior to fixation. A punctate labeling pattern was visible at the nuclear periphery for TAP and NXF2, while in cells transfected with NXF3, TAP D595R, or the NXF2 mutants, no GFP signal was detected at the nuclear rim.

binding (40). We tested the effect of replacing the three loop residues of NXF2 with alanines. This substitution (3xAla598) disrupts nucleoporin binding in vitro as efficiently as the deletion of the entire NXF2 C-terminal domain ( $\Delta$ 598-626) (Fig. 6C,

lanes 7 and 8). As is the case for TAP, a single substitution of alanine for W599 was sufficient to disrupt nucleoporin binding (Fig. 6C, lane 5), while substitution of alanine for N600 or of arginine for E598 had no significant effect (Fig. 6C, lanes 4 and 6).

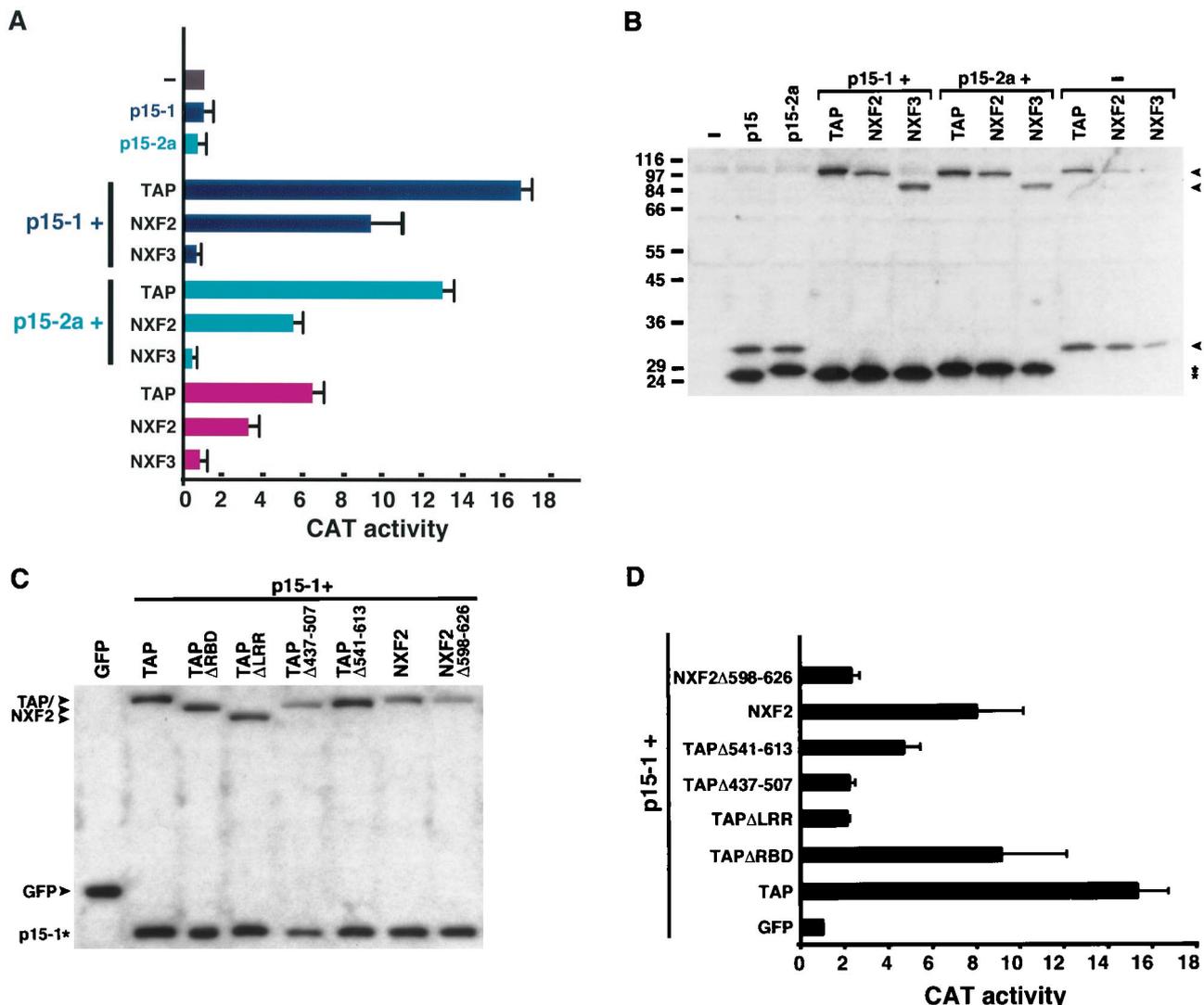


FIG. 7. NXF2 exhibits general RNA nuclear export activity. (A to D) Human 293 cells were transfected with a mixture of plasmids encoding  $\beta$ -Gal, CAT, and either GFP alone (-) or fused to the N termini of TAP, NXF2, NXF3, and various TAP and NXF2 mutants as indicated on the left. pEGFP-N3 derivatives encoding zz-tagged versions of p15-1 and p15-2a were cotransfected as indicated. The cells were collected 40 h after transfection, and  $\beta$ -Gal and CAT activities were determined. Data from three separate experiments were expressed relative to the activities measured when GFP alone was coexpressed with pDM138. The data are means  $\pm$  standard deviations. (B and C) Protein expression levels were analyzed by Western blotting with anti-GFP antibodies. (B) Arrowheads indicate the positions of NXF proteins fused to GFP or of GFP itself, while the asterisks show the positions of p15 proteins.

To analyze nucleoporin binding *in vivo*, the subcellular localization of NXF2 and NXF3 fused to GFP was investigated in transfected HeLa cells and compared to that of GFP-tagged TAP. GFP-tagged NXF2 and NXF3 were evenly detected in the nucleoplasm and were excluded from the nucleolus. Moreover, a fraction of NXF3 could be detected in the cytoplasm (Fig. 6D, -Tx). To visualize a potential nuclear envelope association, transfected cells were extracted with Triton X-100 prior to fixation (Fig. 6D, +Tx). As previously observed for TAP, a fraction of NXF2 was resistant to detergent extraction and remained associated with the nuclear envelope while NXF3 was not detected at the nuclear rim following Triton extraction. Furthermore, NXF2 and TAP mutants that do not interact with nucleoporins *in vitro* were not detected at the nuclear envelope upon Triton extraction (Fig. 6E).

**NXF2, but not NXF3, can stimulate RNA export.** We have developed an assay that tests mRNA export stimulation of TAP and its homologues. In this assay, TAP and TAP-like

proteins are cotransfected with the reporter plasmid pDM138 (14) in human 293 cells. As mentioned above, transfection with this plasmid yields only trace levels of CAT enzyme activity (13, 14). However, cotransfection with vectors expressing TAP and p15 nuclear retention to be bypassed and export of the unspliced transcripts to be promoted, increasing CAT activity (Fig. 7A). To test the effect of TAP-like proteins on *cat* gene expression, human 293 cells were cotransfected with a constant amount of pDM138 reporter plasmid, plasmids encoding TAP, NXF2, or NXF3 together with plasmids encoding p15-1 or p15-2a (Fig. 7A and D). Protein expression levels were analyzed by Western blotting (Fig. 7B and C). Overexpression of the TAP protein together with p15-1 increased *cat* expression by 15- to 17-fold compared to the expression levels obtained when only pDM138 vector was transfected, whereas coexpression of TAP with p15-2a resulted in 13-fold activation of *cat* expression (Fig. 7A). Coexpression of NXF2 with p15-1 or p15-2a activated *cat* expression up to 10- or 6-fold, respectively.

This level of induction may reflect the lower expression levels of NXF2 compared to TAP (Fig. 7B and C). In contrast, when coexpressed with p15-1 or p15-2a, NXF3 had no effect on *cat* activity, even though its expression levels in the presence of p15-1 were similar to that of TAP (Fig. 7A and B). Overexpression of p15-1 or p15-2a, in the absence of exogenous TAP or in the presence of a deletion mutant of TAP (TAP $\Delta$ 437–507) that no longer interacts with p15, had no effect on CAT activity (Fig. 7A and D). In contrast, overexpression of TAP or NXF2 in the absence of exogenous p15 resulted in a significant but modest increase in CAT activity (Fig. 7A); however, the levels of expression of TAP and NXF2 in the absence of p15 were also reduced (Fig. 7B). Therefore, p15-1 and p15-2a appear to increase the stability of TAP-like proteins. In summary, TAP and NXF2, when coexpressed with p15-1 or p15-2a, stimulate the export of the unspliced pre-mRNA, while NXF3 did not stimulate CAT gene expression. The ability of TAP and NXF2 to export pDM138 pre-mRNA is likely to reflect the genuine export activity of these proteins, as this activity could be abolished with various mutations in the TAP sequence, implying the need for a functional protein (Fig. 7D). In particular, this activity requires the LRRs and the NTF2-like domain. In contrast, the RBD and the NPC binding domain of TAP were not strictly required (Fig. 7D) (see Discussion). Thus, the lack of RNA export activity of NXF3 may not be due to the truncation of the UBA domain but to the deletion of part of the LRRs.

## DISCUSSION

In this study, we present a combined evolutionary and functional characterization of the family of TAP-like proteins in higher eukaryotes. Members of this protein family were called NXF. We have further characterized two of four putative human NXF proteins, namely, NXF2 and NXF3, and have shown that NXF2 exhibits RNA export activity.

**NXF proteins heterodimerize with p15.** Previously, we have shown that TAP heterodimerizes with p15-1 via its NTF2-like domain (40). In this study, we show that the human genome encodes at least two p15 homologues and that both interact with TAP, NXF2, and NXF3 and participate in RNA nuclear export. However, the precise role of p15 in TAP-mediated export is not well defined. p15 binding is not strictly required for TAP-mediated export of CTE-bearing substrates (4, 8, 16) but appears to play a role in the general export activity of the NXF proteins (Fig. 7A and D). Whether these effects are due to a direct role of p15 in export or to an increased stability of NXF proteins when coexpressed with p15 is unclear.

The interaction between the NTF2-like domains of TAP and p15-1 is mediated by multiple rather than a few critical residues (40). Indeed, mutation of residues located at the heterodimer interface reduced but did not completely abolish heterodimerization in vitro (40). This provides an explanation for the TAP, NXF2, and NXF3 interactions with p15-1 and p15-2a studied here, as these interactions occur even though not all interface residues are conserved. The NTF2-like domain also occurs in Mex67p, the *S. pombe* and *S. cerevisiae*, NXF homologue (40). In *S. cerevisiae*, this domain is implicated in interaction with a protein known as Mtr2p (35, 38). Mtr2p does not exhibit obvious sequence similarities with p15, but secondary-structure predictions suggest that it could be a p15 analogue (40).

**Functional conservation of the UBA-like domain.** The nucleoporin binding domain of TAP includes a UBA domain (40). The known structure of UBA domains suggests a conserved loop (NWD at positions 593 to 595 in human TAP) is

involved in the interaction with nucleoporins. Replacement of these three residues by alanines was previously shown to block the ability of TAP to promote CTE-dependent export of a precursor mRNA in quail cells (15). Moreover, single-amino-acid changes in this loop are sufficient to impair binding of TAP and *C. elegans* NXF1 to nucleoporins in vitro and in vivo (40, 41) (Fig. 6). The mutational analysis of the UBA-like domain of NXF2 presented here further supports the prediction that the conserved loop residues of the UBA-like domain have a critical role in the interaction of NXF proteins with nucleoporins. Conversely, based on these observations, it was possible to predict that NXF proteins having a truncated UBA domain are unlikely to directly interact with nucleoporins. This prediction was confirmed for NXF3.

Surprisingly, in spite of its conservation, the UBA domain of Mex67p is not sufficient for nucleoporin binding, and Mex67p lacking the UBA domain can bind to nucleoporins in the presence of Mtr2p (38). The possibility that vertebrate NXF proteins lacking the NPC binding domain can still interact with nucleoporins in the presence of p15 requires further investigation; however, NXF3 did not associate with nucleoporins either in vivo or in vitro when coexpressed with p15 (data not shown).

TAP binds Nup98 in vitro, but this interaction is not required for its localization at the nuclear rim, as in *Nup98*<sup>-/-</sup> cells, TAP remains at the nuclear envelope (43). The observation that NXF2 localizes at the nuclear rim and can promote export of the pDM138 pre-mRNA in the absence of Nup98 binding is consistent with the observation that *Nup98*<sup>-/-</sup> cells do not display an overt export inhibition phenotype (43). Thus, although TAP and NXF2 interact with multiple nucleoporins in vitro, the nucleoporins that are responsible for their localization at the nuclear rim remain to be identified. Moreover, the mechanism by which NXF proteins regulate their binding to nucleoporins is unknown.

**Role of NXF proteins in RNA nuclear export.** In yeast, Mex67p is involved in the export of bulk polyadenylated RNAs (36). Evidence for an essential role in mRNP export has been recently obtained for the *C. elegans* protein NXF1 (41). Moreover, TAP is directly implicated in the export of simian type D retroviral RNAs bearing the CTE (8, 12). Although we cannot exclude the possibility that some TAP homologues may have a function other than export, the conservation of their structural organization and the observation that at least three members of the NXF family are implicated in RNA export suggest that these proteins are likely to participate, directly or indirectly, in the export of cellular mRNAs to the cytoplasm. The diversification of NXF proteins in higher eukaryotes compared to yeast may reflect a greater substrate complexity or tissue-specific requirements. Consistent with this possibility is the observation that human NXF5 is expressed in brain (Lin et al., submitted).

NXF proteins may associate with cellular mRNPs directly, by binding to the mRNA, or indirectly, through interaction with hnRNP-like proteins. Recently, several TAP partners that could facilitate its interaction with cellular mRNPs have been identified. These include E1B-AP5, RAE1, and members of an evolutionarily conserved family of hnRNP-like proteins, the Yra1p/REF proteins (4, 37, 39). Aside from these, other RNA binding proteins may also facilitate TAP binding to the mRNPs, as Y14, an hnRNP-like protein that preferentially associates with spliced mRNAs, has recently been shown to be present in a complex containing TAP (18). In vitro, Y14 interacts with TAP and NXF2 (data not shown); therefore, it is possible that Y14, directly or indirectly, recruits NXF proteins to mRNPs in a post-splicing-dependent manner (18). Besides Y14, several other proteins form splicing-dependent inter-

actions with mRNA. These include the splicing coactivator SRm160, the acute myeloid leukemia-associated protein DEK, and several unidentified proteins (23, 25). These proteins may also link NXFs with mRNAs. NXF proteins may then be more efficiently loaded on mRNAs that have been produced by splicing. Therefore, it is possible that the pDM138 precursor mRNA cannot efficiently compete with other mRNAs for binding to NXFs, i.e., NXFs may be limiting for this substrate; hence, its export can be stimulated by overexpression of NXF-p15 heterodimers. Consistent with this is the observation that overexpression of TAP-p15 heterodimers does not stimulate export of reporter mRNAs that are efficiently expressed (I. C. Braun and E. Izaurralde, unpublished data).

TAP-mediated export of pDM138 pre-mRNA requires the LRRs and the NTF2-like domain, while deletion of the RBD or of the NPC binding domain impaired but did not abolish export (Fig. 7D). This is in sharp contrast to the situation observed when TAP-mediated export of pDM138-CTE is monitored in quail cells. In this case, both the RBD and the NPC binding domain are absolutely required (Fig. 4F). The RBD of TAP is required in *cis* to the LRRs for specific binding to the CTE RNA (24). Hence, this domain is essential for TAP-mediated export of CTE-containing substrates, whereas binding of TAP to pDM138 pre-mRNA may be mediated by protein-protein interactions. We have previously reported that the requirements for the NTF2-like and UBA domains of TAP are substrate dependent (4, 8). Indeed, in *Xenopus* oocytes, TAP-mediated export of CTE-bearing intron lariats is independent of these domains (8), while export of U6-CTE requires the UBA domain but not the NTF2-like domain (4). In quail cells, TAP-mediated export of pDM138-CTE pre-mRNA is abolished by mutations or deletions of the UBA domain (15) (Fig. 4) but is only reduced by mutations preventing p15 binding (16). Finally, the C-terminal domain of Mex67p is important but not essential for mRNA export (38). Thus, the functionality of NXF protein domains may be influenced by the specific set of proteins associated with the RNP export substrate. Discovery of the precise roles of NXF protein domains and the mechanism by which these proteins mediate directional transport of RNP substrates across the NPC remains an important goal for the future.

#### ACKNOWLEDGMENTS

The technical support of Michaela Rode is gratefully acknowledged. We thank Tom Hope and Matthias Dobbstein for the kind gift of plasmids pDM138 and pDM138-CTE. We thank Birthe Fahrenkrog, Maarten Fornerod, Iain W. Mattaj, and Christel Schmitt for critical reading of the manuscript. We are grateful to Barbara Felber, Guy Froyen, and Peter Marynen for communicating their results prior to publication.

This study was supported by the German Ministry of Research and Technology (BMBF), the Swiss National Science Foundation, the European Molecular Biology Organization (EMBO), and the Human Frontier Science Program Organization (HFSPO).

A.H. and M.S. contributed equally to this work.

#### REFERENCES

- Adams, M. D., S. E. Celniker, R. A. Holt, C. A. Evans, J. D. Gocayne, P. G. Amanatides, S. E. Scherer, P. W. Li, R. A. Hoskins, et al. 2000. The genome sequence of *Drosophila melanogaster*. *Science* **287**:2185–2195.
- Almeida, F., R. Saffrich, W. Ansorge, and M. Carmo-Fonseca. 1998. Microinjection of anti-coilin antibodies affects the structure of coiled bodies. *J. Cell Biol.* **142**:899–912.
- Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**:3389–3402.
- Bachi, A., I. C. Braun, J. P. Rodrigues, N. Panté, K. Ribbeck, C. von Kobbe, U. Kutay, M. Wilm, D. Görlich, M. Carmo-Fonseca, and E. Izaurralde. 2000. The C-terminal domain of TAP interacts with the nuclear pore complex and promotes export of specific CTE-bearing RNA substrates. *RNA* **6**:136–158.
- Bear, J., W. Tan, A. S. Zolotukhin, C. Taberero, E. A. Hudson, and B. K. Felber. 1999. Identification of novel import and export signals of human TAP, the protein that binds to the CTE element of the type D retrovirus mRNAs. *Mol. Cell. Biol.* **19**:6306–6317.
- Bharathi, A., A. Ghosh, W. A. Whalen, J. H. Yoon, R. Pu, M. Dasso, and R. Dhar. 1997. The human RAE1 gene is a functional homologue of *Schizosaccharomyces pombe* rae1 gene involved in nuclear export of Poly(A)+ RNA. *Gene* **198**:251–258.
- Black, B. E., L. Lévesque, J. M. Holaska, T. C. Wood, and B. Paschal. 1999. Identification of an NTF2-related factor that binds RanGTP and regulates nuclear protein export. *Mol. Cell. Biol.* **19**:8616–8624.
- Braun, I. C., E. Rohrbach, C. Schmitt, and E. Izaurralde. 1999. TAP binds to the constitutive transport element (CTE) through a novel RNA-binding motif that is sufficient to promote CTE-dependent RNA export from the nucleus. *EMBO J.* **18**:1953–1965.
- Brown, J. A., A. Bharathi, A. Ghosh, W. Whalen, E. Fitzgerald, and R. Dhar. 1995. A mutation in the *Schizosaccharomyces pombe* rae1 gene causes defects in poly(A)+ RNA export and in the cytoskeleton. *J. Biol. Chem.* **270**:7411–7419.
- The *C. elegans* Sequencing Consortium. 1998. Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* **282**:2012–2018.
- Galtier, N., M. Gouy, and C. Gautier. 1996. SEAVIEW and PHYLO\_WIN: two graphic tools for sequence alignment and molecular phylogeny. *Comput. Appl. Biosci.* **12**:543–548.
- Grüter, P., C. Taberero, C. von Kobbe, C. Schmitt, C. Saavedra, A. Bachi, M. Wilm, B. K. Felber, and E. Izaurralde. 1998. TAP, the human homologue of Mex67p, mediates CTE-dependent RNA export from the nucleus. *Mol. Cell* **1**:649–659.
- Hope, T. J., X. Huang, D. McDonald, and T. G. Parslow. 1990. Steroid-receptor fusion of the human immunodeficiency virus type 1 Rev transactivator: mapping cryptic functions of the arginine-rich motif. *Proc. Natl. Acad. Sci. USA* **87**:7787–7791.
- Huang, X., T. J. Hope, B. L. Bond, D. McDonald, K. Grahl, and T. G. Parslow. 1991. Minimal Rev-response element for type 1 human immunodeficiency virus. *J. Virol.* **65**:2131–2134.
- Kang, Y., and B. R. Cullen. 1999. The human TAP protein is a nuclear mRNA export factor that contains novel RNA-binding and nucleocytoplasmic transport sequences. *Genes Dev.* **13**:1126–1139.
- Kang, Y., H. P. Bogerd, and B. R. Cullen. 2000. Analysis of cellular factors that mediate nuclear export of RNAs bearing the Mason-Pfizer monkey virus constitutive transport element. *J. Virol.* **74**:5863–5871.
- Katahira, J., K. Sträßer, A. Podtelejnikov, M. Mann, J. U. Jung, and E. Hurt. 1999. The Mex67p-mediated nuclear mRNA export pathway is conserved from yeast to human. *EMBO J.* **18**:2593–2609.
- Kataoka, N., J. Yong, V. N. Kim, F. Velazquez, R. A. Perkinson, F. Wang, and G. Dreyfuss. 2000. Pre-mRNA splicing imprints mRNA in the nucleus with a novel RNA-binding protein that persists in the cytoplasm. *Mol. Cell* **6**:673–682.
- Klebe, C., F. R. Bischoff, H. Pönstingl, and A. Wittinghofer. 1995. Interaction of the nuclear GTP-binding protein Ran with its regulatory proteins RCC1 and RanGAP1. *Biochemistry* **34**:639–647.
- Kraemer, D., and G. Blobel. 1997. mRNA binding protein mrnp 41 localizes to both nucleus and cytoplasm. *Proc. Natl. Acad. Sci. USA* **94**:9119–9124.
- Kutay, U., E. Izaurralde, F. R. Bischoff, I. W. Mattaj, and D. Görlich. 1997. Dominant-negative mutants of importin- $\beta$  block multiple pathways of import and export through the nuclear pore complex. *EMBO J.* **16**:1153–1163.
- Mattaj, I. W., and L. Englmeier. 1998. Nucleocytoplasmic transport: the soluble phase. *Annu. Rev. Biochem.* **67**:256–306.
- Le Hir, H., E. L. Maquat, and J. M. Moore. 2000. Pre-mRNA splicing alters mRNP composition: evidence for stable association of proteins at exon-exon junctions. *Genes Dev.* **14**:1098–1108.
- Liker, E., E. Fernandez, E. Izaurralde, and E. Conti. The structure of the mRNA nuclear export factor TAP reveals a *cis* arrangement of a non-canonical RNP domain and a leucine-rich-repeat domain. *EMBO J.*, in press.
- McGarvey, T., E. Rosonina, S. McCracken, Q. Li, R. Arnaout, E. Mientjes, A. J. Nickerson, D. Awrey, J. Greenblatt, G. Grosveld, and B. J. Blencowe. 2000. The acute myeloid leukemia-associated protein DEK forms a splicing-dependent interaction with exon-product complexes. *J. Cell Biol.* **150**:309–320.
- Morency, C. A., J. R. Neumann, and K. O. Russian. 1987. A novel rapid assay for Chloramphenicol-acetyl transferase gene expression. *BioTechniques* **5**:444–448.
- Murphy, R., J. L. Watkins, and S. R. Wentz. 1996. GLE2, a *Saccharomyces cerevisiae* homologue of the *Schizosaccharomyces pombe* export factor RAE1, is required for nuclear pore complex structure and function. *Mol. Biol. Cell* **7**:1921–1937.
- Nakielnny, S., and G. Dreyfuss. 1999. Transport of proteins and RNAs in and out of the nucleus. *Cell* **99**:677–690.
- Parks, T. D., K. K. Leather, E. D. Howard, S. A. Johnston, and W. G.

- Dougherty. 1994. Release of proteins and peptides from fusion proteins with a recombinant plant virus proteinase. *Anal. Biochem.* **216**:413–417.
30. Pasquinelli, A. E., R. K. Ernst, E. Lund, C. Grimm, M. L. Zapp, D. Rekosh, M.-L. Hammarskjöld, and J. E. Dahlberg. 1997. The constitutive transport element (CTE) of Mason-Pfizer Monkey Virus (MPMV) accesses an RNA export pathway utilized by cellular messenger RNAs. *EMBO J.* **16**:7500–7510.
  31. Pritchard, C. E. J., M. Fornerod, L. H. Kasper, and J. M. A. van Deursen. 1999. RAE1 is a shuttling mRNA export factor that binds to a GLEBS-like NUP98 motif at the nuclear pore complex through multiple domains. *J. Cell Biol.* **145**:237–253.
  32. Ryan, K., and S. R. Wente. 2000. The nuclear pore complex: a protein machine bridging the nucleus and cytoplasm. *Curr. Opin. Cell Biol.* **12**:361–371.
  33. Saavedra, C. A., B. K. Felber, and E. Izaurralde. 1997. The simian retrovirus-1 constitutive transport element CTE, unlike HIV-1 RRE, utilises factors required for cellular RNA export. *Curr. Biol.* **7**:619–628.
  34. Saitou, N., and M. Nei. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**:406–425.
  35. Santos-Rosa, H., H. Moreno, G. Simos, A. Segref, B. Fahrenkrog, N. Panté, and E. Hurt. 1998. Nuclear mRNA export requires complex formation between Mex67p and Mtr2p at the nuclear pores. *Mol. Cell Biol.* **18**:6826–6838.
  36. Segref, A., K. Sharma, V. Doye, A. Hellwig, J. Huber, R. Lührmann, and E. Hurt. 1997. Mex67p, a novel factor for nuclear mRNA export, binds to both poly(A)<sup>+</sup> RNA and nuclear pores. *EMBO J.* **16**:3256–3271.
  37. Sträßer, K., and E. Hurt. 2000. Yra1p, a conserved nuclear RNA-binding protein, interacts directly with Mex67p and is required for mRNA export. *EMBO J.* **19**:410–420.
  38. Sträßer, K., J. Baßler, and E. Hurt. 2000. Binding of the Mex67p/Mtr2p heterodimer to FXFG, GLFG, and FG repeat nucleoporins is essential for nuclear mRNA export. *J. Cell Biol.* **150**:695–706.
  39. Stutz, F., A. Bachi, T. Doerks, I. C. Braun, B. Séraphin, M. Wilm, P. Bork, and E. Izaurralde. 2000. REF, an evolutionarily conserved family of hnRNP-like proteins, interacts with TAP/Mex67p and participates in mRNA nuclear export. *RNA* **6**:638–650.
  40. Suyama, M., T. Doerks, I. C. Braun, M. Sattler, I. Izaurralde, and P. Bork. 2000. Prediction of structural domains of TAP reveals details of its interaction with p15 and nucleoporins. *EMBO Rep.* **1**:53–58.
  41. Tan, W., A. S. Zolotukhin, J. Bear, D. J. Patenaude, and B. K. Felber. The mRNA export in *C. elegans* is mediated by Ce-TAP1, an orthologue of human TAP and *S. cerevisiae* Mex67p. *RNA*, in press.
  42. Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**:4673–4680.
  43. Wu, X., L. H. Kasper, T. Ralitsa, R. T. Mantcheva, B. M. A. Fontoura, G. T. Mantchev, G. Blobel, and J. M. A. van Deursen. NUP98 is required for gastrulation, proper nuclear pore complex assembly and efficient protein import. *J. Cell Biol.*, in press.