Nucleic Acids Research, 1994, Vol. 22, No. 24 5223–5228

Splicing factor SF3a60 is the mammalian homologue of PRP9 of *S.cerevisiae*: the conserved zinc finger-like motif is functionally exchangeable *in vivo*

Angela Krämer*, Pierre Legrain[1], Frank Mulhauser, Karsten Gröning, Reto Brosi and Graeme Bilbe[2]

Département de Biologie Cellulaire, Université de Genève, 30 quai Ernest-Ansermet, CH-1211 Genève 4, Switzerland, [1]Département de Biologie Moléculaire, Institut Pasteur, 25 rue du Dr Roux, F-75724 Paris Cedex 15, France and [2]Pharma Research, Department of Molecular and Cell Biology, Ciba-Geigy Ltd, CH-4002 Basel, Switzerland

Received September 23, 1994; Revised and Accepted October 28, 1994 EMBL accession no. X81789

ABSTRACT

A cDNA encoding the 60 kDa subunit of mammalian splicing factor SF3a has been isolated. The deduced protein sequence reveals a 30% identity to the PRP9 splicing protein of the yeast *S.cerevisiae*. The highest homology is present in a zinc finger-like region in the C-terminal domain of both proteins. The PRP9 zinc finger-like motif has been replaced by the equivalent region of mammalian SF3a60. The chimeric protein rescues the temperature-sensitive phenotype of the *prp9-1* mutant strain demonstrating that not only the structure but also the function of this domain has been conserved during evolution.

INTRODUCTION

Introns are removed from nuclear pre-mRNA in two transesterification reactions which are likely to be RNA catalyzed (for review and recent discussions see refs 1–3). Splicing takes place in a large ribonucleoprotein particle, the spliceosome, that is assembled in multiple well-defined steps involving dynamic interactions between pre-mRNA, small nuclear ribonucleoprotein particles (snRNPs) and non-snRNP protein factors.

Mammalian snRNPs involved in pre-mRNA splicing consist of relatively short RNA molecules (snRNAs) and several proteins that can be grouped into three categories: (i) the Sm proteins that are common to all snRNPs, (ii) proteins that are characteristic of each particle and relatively tightly bound and (iii) unique proteins that dissociate from the particles at salt concentrations exceeding 150–200 mM (4–6). The roles of the snRNPs in splicing and their interactions with the pre-mRNA have been extensively studied (1). U1 snRNP binds to the 5′ splice site forming the first specific intermediate in spliceosome assembly, complex E. The interaction of U2 snRNP with the pre-mRNA branch site, which is located in the intron close to the 3′ splice site, gives rise to pre-splicing complex A. In the following step the U4/U6.U5 triple snRNP binds and complex B is generated. A structural reorganization results in the transition to the active spliceosomal complex C and spliced RNA is generated by two transesterification reactions.

With a few exceptions, all of the mammalian splicing proteins identified participate in the early steps of spliceosome assembly. Members of the SR family of splicing factors act in concert with U1 snRNP to commit the pre-mRNA to the splicing pathway (7–9). U2AF, required for the U2 snRNP–branch site interaction, binds to the polypyrimidine tract at the 3′ end of the intron and is present in complex E (10–12). At least four splicing factors participate in the formation of pre-splicing complex A. The exact roles of an 88 kDa protein (13) and splicing factor SF1 (14) at this stage are unknown. SF3a and SF3b function in the assembly of the 17S U2 snRNP and are constituents of this particle (15,16; K.Gröning and A.Krämer, unpublished). Thus, although SF3a and SF3b have been identified initially as non-snRNP splicing factors, they can be viewed as loosely associated specific snRNP proteins. SF3b binds to the 12S U2 snRNP (16) which contains the Sm proteins as well as the U2 snRNP-specific proteins A′ and B″ (4). The 15S particle thus formed is converted into the active 17S U2 snRNP after interaction with SF3a; by itself, SF3a does not associate with the 12S particle. Recent work has shown that two of the three SF3a subunits (SF3a60 and SF3a66) are structurally related to the *S.cerevisiae* splicing factors PRP9 and PRP11 (16–19). PRP9 and PRP11 function during complex A assembly in yeast (20–22) and interact genetically and physically with a third protein, PRP21 (initially termed SPP91), that acts at the same step of splicing (22–25). By analogy to the yeast system the third subunit of SF3a (SF3a120) may represent PRP21.

As a first step to analyse the structural relationship between the mammalian and yeast proteins and to determine the interactions between the SF3a subunits and their particular function during splicing, we have isolated a cDNA clone encoding

*To whom correspondence should be addressed

SF3a60. We show here that SF3a60 and PRP9 are homologous and that the *in vivo* function of a zinc finger-like motif present in both proteins has been conserved from yeast to man.

MATERIALS AND METHODS

Protein purification and microsequencing

SF3a was purified from HeLa cell nuclear extracts as described (15). Purified protein (~40 μg/SF3a subunit) obtained after Superose 12 gel filtration was separated by SDS−PAGE and transferred to an Immobilon P membrane (Millipore) by electroblotting (26). The appropriate band was located after staining with Coomassie blue in 40% (v/v) methanol/1% (v/v) acetic acid and destaining in 40% (v/v) methanol, excised from the membrane and washed in H_2O. SF3a60 was digested *in situ* with 1.5 μg trypsin in 50 mM ammonium bicarbonate, 0.5 mM $CaCl_2$ and 10% (v/v) acetonitrile in a total volume of 200 μl. After incubation at 37°C overnight the supernatant was removed, the membrane was washed with 50 μl H_2O and briefly sonicated in 150 μl 80% (v/v) formic acid. Peptides in the combined supernatants were applied to a C4 reverse phase column (VYDAC 218TP52, 2.1×250 mm) and eluted with a 90 min gradient of 0−60% (v/v) acetonitrile in 0.05% (w/v) trifluoroacetic acid. The eluate was monitored at 214 and 280 nm. Amino acid sequencing of five peptides, two of which gave partially overlapping sequences, was performed by Paul Jenö (Protein Chemistry Laboratory, Biozentrum, University of Basel) on an Applied Biosystems sequencer model 477A connected on-line to an Applied Biosystems model 120A PTH-amino acid analyser.

cDNA cloning

Degenerate oligonucleotides were designed from two tryptic peptides of SF3a60: oligonucleotide 1 was a 30mer of sequence 5′-AARCAYCCIAATGARATHTGYGTICCCATG-3′ (R = A or G, Y = C or T, I = inosine, H = A or C or T) representing amino acids KHPNEICVPM; oligonucleotide 2 was a 44mer of sequence 5′-GAGGARGCYCARAAYCTIGTGGARTTYACIG-ATGARGAGGGCTA-3′ representing amino acids EEAQNLVEFTDEEGY (27,28). Recombinant phage (5×10^5) from an amplified HeLa cell cDNA library in λgt11 (a generous gift from G.McMaster) were screened on duplicate filters with ^{32}P-labelled oligonucleotide 2. Hybridization was performed at 42°C in 2×SSC, 5×Denhardt's solution, 50 mM sodium phosphate, pH 7.0, 10 mM EDTA, 1% SDS, 200 μg/ml tRNA, 100 μg/ml poly(A) and 50 μg/ml single-stranded salmon sperm DNA. The final wash was performed in 1×SSC and 1% SDS at 50°C. Of fourteen double-positive clones identified eleven were picked and purified by two further rounds of screening. DNA of four phage that were positive with both oligonucleotide probes was isolated and restriction digests were subjected to Southern blot hybridization. The *Eco*RI insert of the phage containing the largest insert (3a60.1) was isolated and subcloned into pBluescript (Stratagene). *Exo*III deletions were generated with the Erase-A-Base system (Promega) and both strands of the insert were sequenced with Sequenase (USB or Pharmacia) by dideoxy chain termination (29).

A cDNA containing the complete coding sequence for SF3a60 was isolated with standard methods (28) from a 5′ stretch HeLa cDNA library in λgt11 (Clontech) by screening with the 5′ terminal *Eco*RI−*Hin*dIII restriction fragment (~900 bp) of clone 3a60.1 that was random prime-labelled with [α-^{32}P]CTP using the Megaprime kit (Amersham). The inserts of four phage were isolated and cloned into pBluescript. The insert of clone 3a60.2 was sequenced on both strands.

Northern blotting

A commercial human Multiple Tissue Northern blot (Clontech) was probed with the 3′ terminal *Hin*dIII−*Eco*RI fragment of clone 3a60.1 containing ~300 bp of coding sequence and the entire 3′ untranslated region of ~1200 bp as described (30).

Expression of recombinant SF3a60 in *E.coli* and antibody preparation

A *Bgl*II restriction fragment of plasmid 3a60.1 (encoding amino acids 102−253 and 336−501 of SF3a60; Figure 1) was ligated into the *Bam*HI site of vector pQE11 (Qiagen). The recombinant protein contains an N-terminal His_6 tag and has an estimated molecular mass of 39 kDa. The plasmid was transformed into the *E.coli* strain JM101. A 1 l culture was induced with 2 mM isopropyl β-D-thiogalactoside for 4 h. A bacterial extract was prepared by resuspending the cells in buffer A (15% glycerol, 20 mM HEPES−KOH, pH 7.9, 100 mM KCl, 3 mM $MgCl_2$, 0.5 mM dithiothreitol, 0.5 mM phenylmethylsulfonyl fluoride) followed by sonification and centrifugation. Proteins were extracted from the pellet in 6 M guanidinium hydrochloride, 0.1 M sodium phosphate, 0.01 M Tris−HCl, pH 8.0, for 1 h at 4°C with gentle shaking and insoluble material was removed by centrifugation. The supernatant was loaded onto a Ni^{2+}−nitrilotriacetic acid affinity column (Qiagen) and proteins were eluted with a pH step gradient in 8 M urea, 0.1 M sodium phosphate and 0.01 M Tris−HCl following the instructions of the supplier. Fractions containing recombinant SF3a60 (rSF3a60; eluted at pH 4.5) were dialysed against buffer A containing 20% glycerol. Most of the recombinant protein precipitated during dialysis; this material was resuspended in SDS sample solution (31), separated by SDS−PAGE and rSF3a60 was electro-eluted from the gel in 25 mM Tris, 192 mM glycine and 0.025% SDS. Following dialysis against buffer A containing 20% glycerol, the precipitate of rSF3a60 was resuspended in 8 M urea. Antibodies were raised in rabbits by several injections with 150−600 mg of recombinant SF3a60 with Specol (Central Veterinary Institute, Lelystad, The Netherlands) used as adjuvant.

For Western blot analysis proteins were separated in a 10% SDS−polyacrylamide gel and transferred to nitrocellulose (26). Antibody detection was performed with the ECL kit (Amersham) following the manufacturer's instructions.

Expression of a chimeric PRP9−SF3a60 protein in yeast

Two oligonucleotides (5′-ATTAGTCGACTACTGGC-TGTATAAGCTTCATGG-3′ and 5′-GTGTTGTCGACGCCC-AAACACCTCATGCC-3′; *Sal*I restriction sites introduced for cloning are underlined) were used as primers for the PCR amplification of the SF3a60 zinc finger region (amino acids 394−439; Figure 4). The amplified DNA fragment was cut with *Sal*I and cloned in place of the corresponding sequences of the PRP9 gene (amino acids 409−454) in plasmid pPL55. This construct is a derivative of the multicopy vector pEMBLYe31 (2m, *LEU2*$^+$) containing a PRP9 coding sequence that has been modified by introduction of two *Sal*I restriction sites flanking the zinc finger motif to allow cloning of PCR fragments; the *Sal*I sequence GTCGAC was inserted after the codon for amino acid 408 of PRP9 and substitutes for the codons of amino acids 455 and 456. Accurate cloning was verified by sequencing of the relevant region in the final construct pPL214, which encodes the

```
  1 M E T I L E Q Q R R Y H E E K E R L M D V M A K E M L T K K S T L R D Q I N S D H R T R A M Q D R Y
 51 M E V S G N L R D L Y D D K D G L R K E E L N A I S G P N E F A E F Y N R L K Q I K E F H R K H P N
101 E I C V P M S V E F E E L L K A R E N P S E E A Q N L V E F T D E E G Y G R Y L D L H D C Y L K Y I
151 N L K A S E K L D Y I T Y L S I F D Q L F D I P K G R K N A E Y K R Y L E M L L E Y L Q D Y T D R V
201 K P L Q D Q N E L F G K I Q A E F E K K W E N G T F P G W P K E T S S A L T H A G A H L D L S A F S
251 S W E E L A S L G L D R L K S A L L A L G L K C G G T L E E R A Q R L F S T K G K S L E S L D T S L
301 E A K N P K S K G T K R D T E R N K D I A F L E A Q I Y E Y V E I L G E Q R H L T H E N V Q R K Q A
351 R T G E E R E E E E E E Q I S E S E S E D E E N E I I Y N P K N L P L G W D G K P I P Y W L Y K L H
401 G L N I N Y N C E I C G N Y T Y R G P K A F Q R H F A E W R H A H G M R C L G I P N T A H F A N V T
451 Q I E D A V S L W A K L K L Q K A S E R W Q P D T E E E Y E D S S G N V V N K K T Y E D L K R Q G L
501 L*
```

Figure 1. Predicted amino acid sequence of SF3a60 cDNA. The amino acid sequence is shown in the one letter code and positions are indicated on the left. Tryptic peptides sequenced are underlined, the cysteine and histidine residues of the zinc finger motif are double underlined.

chimeric PRP9−SF3a60 protein. A *prp9-1* ts strain (Cy105; 23) was transformed with pPL214, transformants were selected at 25°C and replica-plated at 25, 31 and 35°C.

Sequence analysis

Sequence alignments and structural analyses were performed with programs of the GCG software package of the University of Wisconsin (32). Nucleotide and protein sequences were compared with entries in the GENEMBL (release 40) and SWISSPROT (release 29) databases.

RESULTS

A HeLa cDNA library in λgt11 was screened with degenerate oligonucleotide primers that were designed to the sequences of two peptides derived from purified SF3a60. This procedure yielded cDNA 3a60.1 (2206 bp) that extends from an internal *Eco*RI site (at amino acid 93; Figure 1) to a stretch of 17 adenosines preceded by a potential polyadenylation signal (not shown). Both peptides to which oligonucleotides had been generated are located in the immediate vicinity of the 5′ end of this cDNA and no methionine is present upstream. To obtain a cDNA encoding the complete SF3a60 protein, a second library was screened with the 5′ terminal *Eco*RI−*Hin*dIII fragment of clone 3a60.1. cDNA 3a60.2 (2850 bp in length), which contains 842 bp at the 5′ end in addition to those found in clone 3a60.1 and lacks 427 bp of 3′ untranslated sequences, was obtained. Clone 3a60.2, as well as several other independent isolates, contains a 246 bp sequence (encoding amino acids 255−336; Figure 1) that is not present in clone 3a60.1. Within this region the sequence of a fourth SF3a60 peptide was found that is missing in clone 3a60.1. No splice site consensus sequences (33) are apparent at either end of this region, making it unlikely that different SF3a60 mRNAs are produced by use of alternative 5′ or 3′ splice sites, as for example is observed in the case of splicing factors ASF/SF2 or PSF (34,35). Although the additional internal sequence could represent an exon that is alternatively spliced by exon inclusion or skipping we assume that cDNA 3a60.1 arose from a recombination event during the preparation of the cDNA library because no mRNA was detected by Northern blotting that could correspond to a cDNA lacking the 246 bp sequence (see below). Furthermore, PCR amplification of oligo(dT)-primed HeLa cDNA with primers flanking the 246 bp region resulted in the synthesis of only one fragment of the size expected from

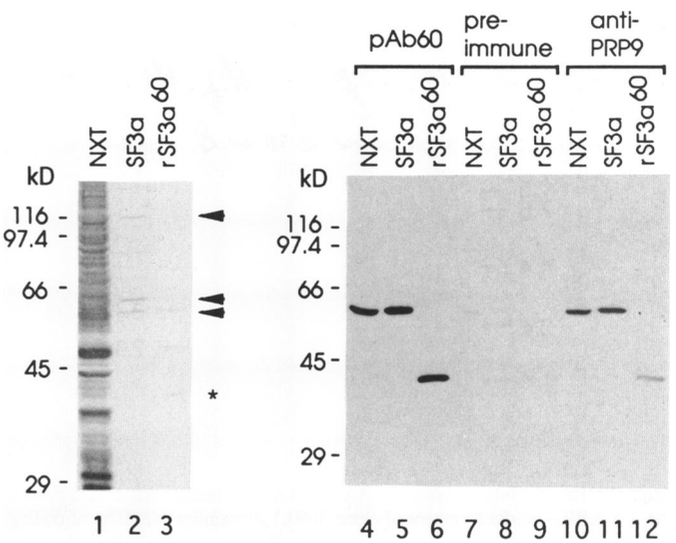

Figure 2. Western blot analysis of SF3a. HeLa cell nuclear extract, purified SF3a and recombinant SF3a60 were separated by SDS−PAGE and proteins were transferred to nitrocellulose. Filters were incubated with anti-SF3a60 antibodies (lanes 4−6), pre-immune serum (lanes 7−9) and anti-PRP9 antibodies (lanes 10−12). A Coomassie blue stain of the fractions used is shown in lanes 1−3. Fractions and antibodies are indicated on top of the lanes, molecular masses (in kDa) of marker proteins are shown on the left. Arrows indicate the SF3a subunits of 60, 66 and 120 kDa, the star indicates the recombinant protein.

clone 3a60.2 (D.Nesic and A.Krämer, unpublished observation).

The open reading frame of clone 3a60.2 encodes a protein of 501 amino acids with a calculated molecular mass of 58.848 kDa, which agrees well with the size of SF3a60 determined by SDS−PAGE (15). All peptide sequences obtained from the tryptic digest of SF3a60 are present in the deduced amino acid sequence (Figure 1). The initiator methionine is located 566 base pairs from the 5′ end of the cDNA and the surrounding sequence conforms to the consensus sequence for eukaryotic translation initiation (36). An in-frame stop codon is present nine base pairs upstream of the putative initiator methionine. The *in vitro* translated product of a derivative of clone 3a60.2 yields a protein of the same size as the protein purified from HeLa cells (D.Nesic and A.Krämer, unpublished). Thus, we assume that clone 3a60.2 encodes the complete SF3a60 protein.

To verify that the cDNA clones isolated encode SF3a60, a partial protein (see Materials and Methods) was expressed in *E.coli*, the recombinant polypeptide was purified and antibodies were raised in rabbits. Figure 2 shows that these antibodies react with the protein used as antigen (rSF3a60) and with SF3a60 in HeLa cell nuclear extracts and purified SF3a, whereas the pre-immune serum is negative. Furthermore, anti-PRP9 antiserum which reacts with SF3a60 (16) also recognizes rSF3a60. Taken together, these results indicate that the cDNAs isolated encode SF3a60.

The 3′ untranslated region is 1205 bp in length and contains two potential polyadenylation signals 16 and 989 bp upstream of the site of poly(A) addition. Northern blotting of poly(A^+) RNA of different human tissues revealed two mRNAs of ~2.1 and 3.1 kb in length (Figure 3); hence both potential polyadenylation signals appear to be utilized. It is also apparent that the ratio of the two mRNAs varies between the tissues analysed. For example, the longer form is predominant in poly(A^+) RNA of lung, whereas the shorter form is more abundant in RNA isolated from liver, suggesting that the two mRNAs are alternatively expressed in different tissues.

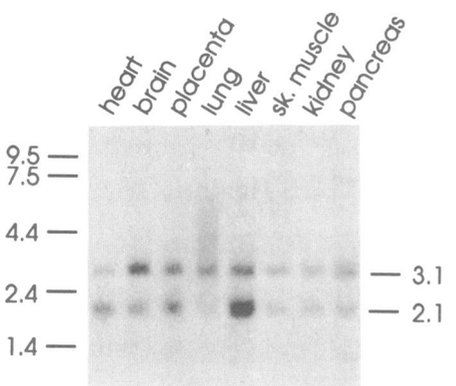

Figure 3. Northern blot analysis. Poly(A^+) RNA from human tissues was probed with a *Hin*dIII−*Eco*RI fragment of clone 3a60.1 containing ~300 bp of coding sequence and the entire 3′ untranslated region. The tissues are indicated on top of the lanes, size markers (in kb) are indicated on the left, the sizes of the SF3a60 mRNAs are shown on the right.

It has previously been demonstrated that SF3a60 is structurally related to the yeast splicing protein PRP9 (16−19). An alignment of the deduced amino acid sequences of SF3a60 and PRP9 (37) reveals an identity of about 30% and a 56% similarity (Figure 4). Both proteins are relatively rich in charged amino acids (36%) and sequences in the C-terminal thirds of both proteins are highly homologous. This region contains a zinc finger motif of the C_2H_2 type (Figure 4; see 38 for review) which is essential for viability (37). A leucine-rich region and a second zinc finger-like motif present in PRP9 are not conserved in SF3a60. Also, no other motifs that are present in splicing proteins have been detected in the SF3a60 protein sequence.

To assess the functional significance of the homology in the zinc finger region, a protein was constructed in which amino acids 409−454 of the PRP9 protein (Figure 4) were exchanged for the corresponding sequences of SF3a60 (amino acids 394−439). The chimeric protein was then tested for its ability to rescue the temperature-sensitive phenotype of the *prp9-1* mutant strain (Table 1; 23). Growth at 31 and 35°C can be rescued by transformation with PRP9 wild-type or pPL55 (PRP9/*Sal*I; see Materials and Methods), but not with the vector alone. Consistent with previous results (24) demonstrating that deletion of the second zinc finger motif of the PRP9 protein (PRP9ΔCH2) leads to a dominant negative phenotype, the *prp9-1* ts strain expressing this mutant gene is not viable at 25°C. In contrast, cells transformed with the plasmid expressing the chimeric PRP9−SF3a60 protein were viable at 25°C and, more importantly, these transformants also grew at 31 and 35°C. The rescue is specific for the SF3a60 zinc finger-like motif, since mutant PRP9 genes containing the equivalent region of yeast PRP11 (39) or the mammalian U1 snRNP C protein (40) encode proteins that have a dominant negative effect on viability (data not shown). In summary, the chimeric PRP9−SF3a60 gene complements the temperature-sensitive phenotype of the *prp9-1*

```
3A60  METILEQQRRYHEEKERLMD.....................VMAKEMLTKKSTLRDQ.......INSDHRTRAMQDRYMEVSGNLRDLYDDKDGLRKEE 71
PRP9  MN.LLETRRSLLEEMEIIENAIAERIQRNPELYYHYIQESSKVFPDTKLPRSSLIAENKIYKFKKVKRKRKQIILQQH..EINIFLRDYQEKQQTFNKIN 97

3a60  ..LNAISGPNEFAEFYNRLKQI.KEFHRKHPN.EICVPMSVEFEELLKARENPSEEAQNLVE.....FTDEEGYGRYLDLHDCYLKYINLKASEK 157
PRP9  RPEETQEDDKDLPNFERKLQQLEKELKNEDENFELDINSKKDKYALFSSSSDPSRRTNILSDRARDLDLNEIFTRDEQYGEYMELEQFHSLWLNVIKRGD 197

3a60  LDYITYLSIFDQLFD.......IPKERKNAEYKRYLEMLLEYLQDYTDRVKPLQDQNELFGKIQAEFEKKWENGTFPGWPKETSSALTHAGAHLDLSAFS 250
PRP9  CSLLQFLDILELFLDDEKYLLTPPMDRKNDRYMAFLLKLSKYVETFFFKSYALLDAAAVENLIKSDFEHSYCRGSLRSEAKGI.....YCPFC 285

3a60  SWEELASLGLDRLKSALLALGLKCGGTLEERAQRLFSTKGKSLESLDTSLFAKNPKSKGTKRDTERNKDIAFLEAQIYEYVEILGEQRHLTHENVQRKQA 350
PRP9  SR.....WFKTS.....SVFESHLVGKIHKKNESKRRNFVYSEYKLHRYLKYLNDEFSRTRSFVERKLA 344

3a60  RTGEER.....EEEEEEQISESESEDEENEIIYNPK..NLPLGWDGKPIPYWLYKLHGLNINYNCEICGNYTYRGPKAFQRHFAEW 429
PRP9  FTANERMAEMDILTQKYEAPAYDSTEKEGAEQVDGEQRDGQLQEEHLSGKSFDMPLGPDGLPMPYWLYKLHGLDREYRCEICSNKVYNGRRTFERHFNEE 444

3a60  RHAHGMRCLGIPNTAHFANVTQIEDAVSLWAKLKLQ.....KASERWQP.DTEEEYEDSSGNVVNKKTYEDLKRQGLL 501
PRP9  RHIYHLRCLGIEPSSVFKGITKIKEAQELWKNMQGQSQLTSIAAVPPKPNPSQLKVPTELELEEEDEEGNVMSKKVYDELKKQGLV 530
```

Figure 4. Amino acid sequence alignment of SF3a60 and PRP9. SF3a60 and PRP9 protein sequences were aligned with the BESTFIT program. Identical amino acids are indicated by a solid line between the sequences, similar amino acids (D = E, K = R, N = Q, I = L = V, F = W = Y, S = T) are shown by a colon. The zinc finger region is underlined; the cysteine and histidine residues within the motif are indicated by stars. A putative leucine zipper and a second zinc finger present in PRP9 but not in SF3a60 are shown by dotted lines. The region exchanged between PRP9 and SF3a60 for construction of the chimeric protein is delimited by arrowheads.

mutation, indicating that the essential function of the zinc finger-like region is evolutionarily conserved.

DISCUSSION

We have isolated a cDNA encoding the 60 kDa subunit of mammalian splicing factor SF3a. The deduced protein sequence is identical to that of the spliceosome-associated protein SAP61 (19), confirming the identity of the two polypeptides as previously suggested on the basis of co-migration in 2D polyacrylamide gels (M.Bennett, R.Brosi, A.Krämer and R.Reed, unpublished results).

Immunological studies have shown that SF3a60 (or SAP61) and the PRP9 protein of *S.cerevisiae* are structurally related (16–18) and both proteins function in splicing during the formation of pre-splicing complex A (15,20). Extending these observations, direct comparison of the amino acid sequences of SF3a60 and PRP9 demonstrates a 30% overall structural identity between the proteins. The region of highest homology is located in the C-terminal third of the proteins and includes a zinc finger-like motif of the C_2H_2 type (Figure 4). Mutational analyses of the cysteine and histidine residues of this motif in PRP9 have demonstrated that it is essential for PRP9 function *in vivo* (37). The second zinc finger-like motif of PRP9 is not conserved in SF3a60 and only the mutation of one histidine in this motif is detrimental to PRP9 function, whereas the other histidine and both cysteine residues can be mutated without obvious growth defects (37). In addition, a leucine zipper-like region detected in PRP9 is absent from SF3a60. The function of this motif has not been analysed in detail. Mutation of amino acids 78 and 177 in the N-terminal portion of PRP9 leads to a temperature-sensitive phenotype and interferes with the interaction between PRP9 and PRP21 when tested in the two-hybrid system (24). These two positions are conserved in SF3a60.

As a first attempt to understand whether structurally related domains in mammalian and yeast splicing factors provide equivalent functions, the zinc finger-like and flanking sequences of SF3a60 and PRP9 have been exchanged. Our results clearly demonstrate that a functional PRP9 protein can include the region of either origin, yeast or human. Thus, despite a number of non-conservative amino acid replacements within the exchanged region, the essential function *in vivo* of this domain has been conserved from yeast to man. An alignment of the zinc finger-like motifs of several proteins involved in splicing has been proposed previously (37; Figure 5) and these domains may sustain similar or related functions, such as RNA binding or protein–protein interactions (18,19,24). Several amino acids that are identical in the yeast PRP11 and human U1 snRNP C proteins are also found in the mammalian PRP11 homologue SAP62 (18,39,40). These residues are not conserved in zinc finger-like motifs present in other splicing proteins. Consistent with these differences in amino acid sequence, PRP9 chimeric proteins formed with zinc finger-like domains of PRP11 or the U1 snRNP C protein are non-functional, demonstrating that these domains are not strictly exchangeable.

Apart from a conservation of the zinc finger-like region of SF3a60 and PRP9, the sequences flanking this motif also display a high content of identical amino acids. Although the homology upstream of a zinc finger-like domain in mammalian SAP62 and yeast PRP11 (18) is not as striking as in SF3a60 and PRP9, several amino acids are shared by all four proteins (Figure 5). In contrast, amino acids located between the cysteine and histidine residues are less well conserved between the two groups of splicing factors. A particular function of this upstream sequence has not been investigated; however, the conservation may hint to a similar role, either in combination with the zinc finger or a function on its own. Future experiments in which sequences are exchanged between mammalian and yeast homologues as well as detailed mutational analyses of conserved regions will clarify the *in vivo* function as well as the role of these domains in spliceosome formation *in vitro*.

Table 1. Expression of a chimeric PRP9–SF3a60 protein in a *prp9-1* ts strain

Plasmid	Temperature[a] 25° C	31° C	35° C
Vector	+[b]	–	–
PRP9 wild–type	+	+	+
PRP9/*SalI*	+	+	+
PRP9ΔCH2	0[c]		
PRP9–SF3a60	+	+	+

[a]Transformed cells were grown at 25°C for 3 days and replica-plated at the indicated temperatures.
[b]Over 300 colonies were obtained.
[c]No transformants were obtained.

```
U1C                        MPKFY C DY C DTYLTHDSPSVRKT H CSGRK H KENVKDYYQK
                                 |  |   |  |   | | |     |   | | |   ||
Prp11    EKDDQVRSNPY.IYKNHSG...KLV C KL C NTM..HMSWSSVER H LGGKK H GLNVLRRGIS
                  ||   ||| |       | || |  |   |    |    |   ||| |   |  ||
Sap62    ......NKDPYFM.KNHLG...SYE C KL C LTL..HNNEGSYLA H TQGKK H QTNLARRAAK
                ** ** * *    *  *    *                * *       *        **
Prp9#2   PLGPDGLPMPYWLYKLH.GLDREYR C EI C SNK.VYNGRRTFER H FNEER H IYHLRCLGIE
         ||| || | |||||||| ||   |  | || |  |   | |   | | | | | | |     |||||
SF3a60   PLGWDGKPIPYWLYKLH.GLNINYN C EI C GNY.TYRGPKAFQR H FAEWR H AHGMRCLGIP

Prp9#1   ENLIKSDFEHSYCRGSLRSEAKGIY C PF C SRW..FKTSSVFES H LVGKI H KKNESKRRNF

Prp6     GYIASARLEEKARKFSVAKKIIENG C QE C PRSSDIWLENIRL. H ESDV. H YCKTLVATAI
```

Figure 5. Alignment of zinc finger-like motifs and flanking regions of proteins involved in pre-mRNA splicing. Amino acid sequences of the zinc finger regions of the human U1 snRNP C protein (40), PRP11 (39), SAP62 (18), PRP6 (37), PRP9 (37) and SF3a60 are shown. The conserved cysteine and histidine residues of the zinc finger motif are shown in bold. Amino acids that are conserved between PRP11, SAP62 and the U1 snRNP C protein and between the second zinc finger of PRP9 (PRP9#2) and SF3a60 are indicated by vertical lines between the sequences. Amino acids that are conserved in at least three positions between PRP11, SAP62, PRP9 and SF3a60 are indicated by a star.

ACKNOWLEDGEMENTS

We thank G.McMaster for the λgt11 HeLa cDNA library, S.Backes for helpful suggestions, M.Quentin for technical assistance and members of the Krämer laboratory for comments on the manuscript. The help of D.Blank (Biozentrum, University of Basel) during the early stages of this work is gratefully acknowledged. This work was supported by grants from the Schweizerischer Nationalfonds and the Kanton of Geneva to A.K., from the CNRS (URA 1140) to P.L and by the Boehringer Ingelheim Fonds to K.G.

REFERENCES

1. Moore, M.J., Query, C.C. and Sharp, P.A. (1993) In: Gesteland, R.F. and Atkins, J.F. (eds), The RNA World. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 303–357.
2. Madhani, H.D. and Guthrie, C. (1994) *Annu. Rev. Genet.*, **28**, 1–26.
3. Nilsen, T.W. (1994) *Cell*, **78**, 1–4.

4. Lührmann, R., Kastner, B. and Bach, M. (1990) *Biochim. Biophys. Acta*, **1087**, 265–292.
5. Behrens, S.E. and Lührmann, R. (1991) *Genes Dev.*, **5**, 1439–1452.
6. Behrens, S.E., Tyc, K., Kastner, B., Reichelt, J. and Lührmann, R. (1993) *Mol. Cell. Biol.*, **13**, 307–319.
7. Fu, X.-D. (1993) *Nature*, **365**, 82–85.
8. Wu, J.Y. and Maniatis, T. (1993) *Cell*, **75**, 1061–1070.
9. Kohtz, J.D., Jamison, S.F., Will, C.L., Zuo, P., Lührmann, R., Garcia-Blanco, M.A. and Manley, J.L. (1994) *Nature*, **368**, 119–124.
10. Ruskin, B., Zamore, P.D. and Green, M.R. (1988) *Cell*, **52**, 207–219.
11. Zamore, P.D. and Green, M.R. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 9243–9247.
12. Gozani, O., Patton, J.G. and Reed, R. (1994) *EMBO J.*, **13**, 3356–3367.
13. Ast, G., Goldblatt, D., Offen, D., Sperling, J. and Sperling, R. (1991) *EMBO J.*, **10**, 425–432.
14. Krämer, A. (1992) *Mol. Cell. Biol.*, **12**, 4545–4552.
15. Brosi, R., Hauri, H.P. and Krämer, A. (1993) *J. Biol. Chem.*, **268**, 17640–17646.
16. Brosi, R., Gröning, K., Behrens, S.E., Lührmann, R. and Krämer, A. (1993) *Science*, **262**, 102–105.
17. Behrens, S.E., Galisson, F., Legrain, P. and Lührmann, R. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 8229–8233.
18. Bennett, M. and Reed, R. (1993) *Science*, **262**, 105–108.
19. Chiara, M.D., Champion-Arnaud, P., Buvoli, M., Nadal-Ginard, B. and Reed, R. (1994) *Proc. Natl. Acad. Sci. USA*, **91**, 6403–6407.
20. Abovich, N., Legrain, P. and Rosbash, M. (1990) *Mol. Cell. Biol.*, **10**, 6417–6425.
21. Arenas, J.E. and Abelson, J.N. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 6771–6775.
22. Ruby, S.W., Chang, T.-H. and Abelson, J. (1993) *Genes Dev.*, **7**, 1909–1925.
23. Chapon, C. and Legrain, P. (1992) *EMBO J.*, **11**, 3279–3288.
24. Legrain, P., Chapon, C. and Galisson, F. (1993) *Genes Dev.*, **7**, 1390–1399.
25. Legrain, P. and Chapon, C. (1993) *Science*, **262**, 108–110.
26. Kyhse-Anderson, J. (1984) *J. Biochem. Biophys. Methods*, **10**, 203–209.
27. Lathe, R. (1985) *J. Mol. Biol.*, **183**, 1–12.
28. Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
29. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
30. van Hille, B., Richener, H., Evans, D.B., Green, J.R. and Bilbe, G. (1993) *J. Biol. Chem.*, **268**, 7075–7080.
31. Laemmli, U.K. (1970) *Nature*, **227**, 680–685.
32. Devereux, J., Haeberli, P. and Smithies, O. (1984) *Nucleic Acids Res.*, **12**, 387–395.
33. Ohshima, Y. and Gotoh, Y. (1987) *J. Mol. Biol.*, **195**, 247–259.
34. Ge, H., Zuo, P. and Manley, J.L. (1991) *Cell*, **66**, 373–382.
35. Patton, J.G., Porro, E.B., Galceran, J., Tempst, P. and Nadal-Ginard, B. (1993) *Genes Dev.*, **7**, 393–406.
36. Kozak, M. (1989) *J. Cell Biol.*, **108**, 229–241.
37. Legrain, P. and Choulika, A. (1990) *EMBO J.*, **9**, 2775–2781.
38. Burd, C.G. and Dreyfuss, G. (1994) *EMBO J.*, **13**, 1197–1204.
39. Chang, T.H., Clark, M.W., Lustig, A.J., Cusick, M.E. and Abelson, J. (1988) *Mol. Cell. Biol.*, **8**, 2379–2393.
40. Sillekens, P.T.G., Beijer, R.P., Habets, W.J. and van Venrooij, W.J. (1988) *Nucleic Acids Res.*, **16**, 8307–8328.