



Two-dimensional gel of the spliceosome. Full gel showing purified spliceosome-associated factors (**a**). The area marked with the broken line is enlarged in (**b**). The numbered spots were excised and analysed by mass spectrometry (Tables 1,2).

Table 1 • Novel spliceosome-associated proteins identified by mass spectrometry as EST sequences

Spot numbers	Apparent/calculated molecular weight	EST accession number (GenBank) ^a	Protein accession number (GenBank)	Genetic mapping ^b	Protein name or homologue	Comments
4	24 /26.6	R20293	AF081788	-	SPF 27	no function assigned
9	28/26.9	N30391	U89876 ^c	-	'ALY' (mouse)	RRM domain two RGG motifs
10	27/29.9	T91668	AF083385	-	SPF 30	partial homology to survival motor neuron protein
15	28/30.8	W56224	AF083190	+	SPF 31	Dna j-like domain E/D-R/K-rich region
22	30/38.4	AA014692 (mouse)	AF083383	-	SPF 38	β transducin-like; seven WD-domains ^d
27	34/32.8	R89115	AF026029 ^e	+	Poly A tail binding protein II	evidence for Arg methylation; N-terminal acetylation; RRM2-domain
28	40/45.2	AA052680 (mouse)	AF083384	-	SPF 45	DNA damage repair protein (<i>A. thaliana</i>) N-terminal acetylation; RRM domain; E/D-R/K-rich region ^d
29	42/-	AA289699 (mouse)	50% sequence obtained	-	U23412 (<i>C. elegans</i>)	similar to Prp19 in yeast at least three WD domains ^f
30	44/57.2	N57351	AF044333 ^e	-	PRL-1; phosphatase regulatory subunit (<i>A. thaliana</i>)	seven WD domains ^g
32,33	46/58.5	R19664	AF001687 ^e	+	Hprp4 U4/U6 60-kD protein	Prp4-like 6-7 WD domains
40,46,52,55,56	50-60/69.6	W16579	AF037448 ^e	+	Gry-Rbp	possibly alternatively spliced; three RRM domains one RGG motif
54	55/-	Z42055	80% sequence obtained	-	IK factor, S74221	IK factor sequence is identical to the central 40% of the sequence of protein 54
57, 58	58/68.7	AA346837	AC004475 ^e	-	'F23858 1'	two SURP domains
59	60/77.5	AA192518	AF016370 ^e	-	hprp3p U4/U6 associated splicing factor	Prp3-like
61	75/92.3	R13357	U86753 ^e	-	Cdc 5 related protein	two Myb-kinase domains ^d
62	80/95.6	H24106	AF026402 ^e	+	U5 snRNP 100-kD protein	DEAD box RNA helicase; similar to Prp28 (<i>S. cerevisiae</i>)
64	90/100.2	AA056357	U94836 ^e	-	Erprot 213-21 mRNA	unknown function
67, 68	100/124	W16921	AF017789 ^e	+	CA150	putative transcription factor ^g
69	120/-	AA317471	15% sequence obtained	-	—	—

^aThe EST numbers refer to the database entries with which the protein was identified. ^bThe mapping information can be found bracketed by Généthon markers (<http://www.ncbi.nlm.nih.gov/UniGene/>). ^cHighly homologous protein sequence from *M. musculus*. ^dColocalization of GFP-fusion protein with snRNPs. ^eAfter identification of the EST, the protein was cloned by other researchers. ^fData from assembled EST sequences. ^gGFP-fusion proteins shown to colocalize with splicing snRNPs *in vivo* using confocal fluorescence microscopy of transiently transfected HeLa cells expressing the cognate GFP fusion proteins and double-labelled with either anti-Sm or anti-U1A antibodies to detect splicing snRNPs.

Table 2 • Spliceosome-associated proteins identified in protein sequence databases

Spot numbers	Apparent /calculated molecular weight	Genetic mapping ^a	Accession number	Protein name
<i>Previously characterized spliceosomal proteins</i>				
1, 2	23/24.6	+	P14678 SwissProt	SmB/B'
3	23/25.5	+	P08579 SwissProt	U2-snRNP B''
6, 7	25/28.6	+	P09661 SwissProt	U2-snRNP A'
16	27/31.3	-	P09012 SwissProt	U1-snRNP A
17	29/38.7	-	P09651 SwissProt	hnRNP A1
19, 20	30/27.9	-	Q01081 SwissProt	U2AF-35kD
21	30/37.5	+	P22626 SwissProt	hnRNP A2/B1
22, 23	31/33.4	+	P07910 SwissProt	hnRNP C
25, 26	35/44.4	-	L35013 trembl	SAP 49
31	45/58.8	+	U08815 trembl	SAP 61
34, 35, 36	46/53.9	-	P26368 SwissProt	U2AF-65kD
37	46/49.5	-	L21990 trembl	SAP 62
38, 39	48/70.1	+	P08621 SwissProt	U1-snRNP 70kD
45	48/68.8; 68.6	+	Y08765/ Y08766 (GenBank)	SF1 isoforms
48, 49, 50, 51	50-55/77.8	-	P52272 SwissProt	hnRNP M
63	77/88.9	+	X85237 trembl	SAP 114
65, 66	100/97.8	-	U41371 trembl	SAP 145
<i>Potentially new spliceosome-associated proteins^b</i>				
8	24/29.5	+	P12750 (SwissProt)	40S ribosomal Protein Rps4x
41, 42, 43	50/61.5	+	U51432 (trembl)	Skip ^c
43, 44	50/69.2	+	P17844 (SwissProt)	p68 RNA helicase
47	52/70.5	+	P11940 (SwissProt)	PolyA binding protein (PAB 1)
<i>Contaminants^d</i>				
5	23/49.3		P07219 (SwissProt)	Phaseolin (<i>Phaseolus vulgaris</i>)
12, 13, 14	27/24.5		P02662	
	27/26.0		P02663 (SwissProt)	alpha S1 and S2 caseins (<i>B. taurus</i>)
53	54/51.8		P55081 (SwissProt)	Microfibrillar protein
<i>Spots not identified</i>				
11, 18, 24, 60				

^aThe mapping information can be found bracketed by Généthon markers (<http://www.ncbi.nlm.nih.gov/UniGene/>). ^bThese proteins have been characterized in a different biological context, but so far have not been shown to be involved in splicing. ^cCo-localization of GFP-fusion protein with snRNPs. ^dThe contaminants were only detected as very faint spots (Fig. 1), thus their identification reflects the overall sensitivity of the method.