Nucleic Acids Research, 1993, Vol. 21, No. 12 2815–2822

Identification of an snRNP-associated kinase activity that phosphorylates arginine/serine rich domains typical of splicing factors

Andreas Woppmann+, Cindy L.Will, Ute Kornstädt, Ping Zuo[1], James L.Manley[1] and Reinhard Lührmann*

Institut für Molekularbiologie und Tumorforschung, Emil Mannkopff Straße 2, D-3550 Marburg, Germany and [1]Department of Biological Sciences, Columbia University, New York, NY 10027, USA

Received April 6, 1993; Revised and Accepted May 19, 1993

ABSTRACT

The U1 snRNP-specific 70K protein is one of the few snRNP proteins from higher eukaryotic cells that is phosphorylated *in vivo* (1,2). Immunoaffinity purified spliceosomal snRNPs (U1, U2, U5, and U4/U6) were tested for their ability to phosphorylate *in vitro* the U1-specific 70K protein. An snRNP-associated kinase activity which phosphorylates all U1-70K isoelectric variants was identified. Like its *in vivo* counterpart, this snRNP-associated enzyme phosphorylates solely serine residues of the 70K protein, preferentially utilizing ATP as a phosphodonor. Tryptic phosphopeptide analysis revealed an overlapping set of at least four radiolabeled peptides in the *in vivo* and *in vitro* phosphorylated protein, suggesting that the snRNP-associated serine kinase is responsible, at least in part, for the 70K protein phosphorylation observed *in vivo*. Chymotryptic digestion of *in vitro*, ^{32}P-labeled 70K protein and *in vitro* phosphorylation studies with a synthetic peptide, indicated that the multiple 70K phosphorylation sites are limited to a highly charged, C-terminal domain of the protein. *In vitro* phosphorylation studies with the splicing factor ASF/SF2 and several deletion mutants demonstrated that, similar to the U1-70K protein, the snRNP-associated serine kinase phosphorylates the carboxy terminal RS-rich domain of ASF/SF2. A potential general role for this enzyme in the phosphorylation of splicing factors and its consequences for pre-mRNA splicing regulation are discussed.

INTRODUCTION

Splicing of nuclear pre-mRNAs requires the presence of a number of trans-acting factors which associate with a given intron during spliceosome formation (see ref. 3 for recent review). Two classes of splicing factors may be distinguished. One class is comprised of the UsnRNPs, U1, U2, U5, and U4/U6. They consist of either one (U1, U2, U5) or two (U4/U6) snRNA molecules which are associated with 10 to 15 proteins (4). While a subset of these proteins is common to each of the U-snRNPs, several proteins associate specifically with one snRNP species; U1 snRNPs, for example, contain three particle specific proteins designated 70K, A, and C. The second class of trans-acting splicing factors consists of proteins which are not associated with the snRNPs. While the factors in both classes are essential for constitutive pre-mRNA splicing, some of them, such as ASF/SF2 or U1 snRNP, have also been shown to be involved in alternative splicing events (5–7). The exact mechanism whereby these factors influence splice site selection is poorly understood.

A number of splicing factors, including ASF/SF2 (8,9), U2AF(10), SC35 (11) and regulators of Drosophila RNA splicing (reviewed in 12,13), and the U1-70K protein (14,15,16) share structural features which include an RNA binding domain and a region rich in arginine and serine residues. Recent studies have demonstrated an essential role for the latter domain in the activity of ASF/SF2 and U2AF (10,17). ASF/SF2 has recently been shown to belong to a large family of conserved, serine/arginine-rich, nuclear phosphoproteins (18). This family has been proposed to consist largely of functionally related proteins. Several members have been identified as splicing factors, suggesting that a number of proteins essential for pre-mRNA splicing exist as phosphoproteins *in vivo*. Studies with a monoclonal antibody which is known to specifically recognize a phosphorylated epitope and which was also shown to recognize ASF/SF2, provided preliminary evidence that ASF/SF2 may be phosphorylated *in vivo* (18–20). While direct evidence for the phosphorylation of ASF/SF2 is still lacking, the U1-70K protein, on the other hand, has clearly been shown to exist as a phosphoprotein *in vivo* (1,2). Phosphorylation of each of the thirteen 70K protein isoelectric variants has been demonstrated, with all of the variants sharing at least four major phosphorylation sites (2).

A central role for reversible protein phosphorylation in the regulation of both constitutive and alternative splicing has recently been postulated (21,22). Multiple protein phosphorylation events

* To whom correspondence should be addressed

+ Present address: Neurex Co., 3760 Haven Avenue, Menlo Park, CA 94025, USA

could partially explain the ATP requirement for both spliceosome formation and the subsequent cleavage/ligation steps of pre-mRNA splicing. Several recent studies indicate that the dephosphorylation of splicing factors is required for both steps of the splicing reaction. Tazi *et al.* have demonstrated through the use of ATP analogs and protein phosphatase inhibitors that the second step of *in vitro* splicing is blocked in the absence of phosphatase activity (21). Similarly, Mermoud *et al.* have demonstrated, through the use of highly specific protein phosphatase inhibitors and purified protein phosphatases, that at least two different serine/threonine phosphatases are essential for pre-mRNA splicing, but not for spliceosome assembly (22). One target of the serine/threonine phosphatases may be the U1-70K protein. Indeed, splicing complementation studies with U1 snRNP particles containing either phosphorylated or thiophosphorylated 70K protein suggest that dephosphorylation of this U1-specific protein may be critical for its function as a splicing factor *in vitro* (23).

Here, we describe the identification and initial characterization of an snRNP-associated kinase activity which phosphorylates the U1-specific 70K protein *in vitro*. Through phosphoamino acid and phosphopeptide analyses, we provide evidence that the snRNP-associated kinase is responsible, at least in part, for 70K phosphorylation *in vivo*. Protease digestion studies and the use of a synthetic peptide allowed the mapping of 70K protein phosphorylation sites to the arginine-serine rich C-terminal domain of the protein. The results of *in vitro* phosphorylation studies with purified, recombinant ASF/SF2 and a form arising from alternative splicing and lacking the RS domain, ASF-3 (8), as well as various ASF/SF2 deletion mutants, suggest that this enzyme may generally recognize the RS-rich domains typical of a number of splicing factors.

MATERIALS AND METHODS

Isolation of snRNPs

The cultivation of HeLa S3 cells and their labeling with ^{32}P-orthophosphate were carried out as described by Bringmann *et al.* (24). Nuclear extracts were prepared either as described by Bringmann *et al.* (24) (designated S1R extracts) or according to Dignam *et al.* (25) (designated splicing extracts). S1R extracts lack the high molecular weight U5-specific proteins and contain an, as yet, unidentified protein which we have designated D′ (see Figure 1 and reference 26). U1, U2, U5, and U4/U6 snRNPs were isolated from S1R nuclear extract by immunoaffinity chromatography with an anti-m_3G antibody as described by Bach *et al.* (27). U1–U6 snRNPs were isolated in an identical manner from splicing extracts except anti-m_3G antibody bound snRNPs were washed and eluted with buffer containing 250 mM, rather than 420 mM, NaCl. Kinase active U1 snRNPs were isolated from the latter snRNP preparation by Mono Q column chromatography as previously described (27). SnRNPs were dialysed against buffer containing 5 mM Hepes-KOH, pH 7.9, 5% (w/v) glycerol, 25 mM KCl, 1 mM $MgCl_2$, 0.05 mM EDTA, 0.5 mM DTE, and 0.5 mM PMSF before use in all phosphorylation experiments.

Preparation of ASF/SF2 and synthetic peptide

ASF/SF2 deletion mutants were constructed as described by Zuo and Manley (17). Deletion mutants A, B, and C (see Figure 6) were generated by digestion of the ASF/SF2 RS-rich domain with HinfI, ApaI, and PF/MI, respectively. Mutant D was generated by removing an ApaI-StyI fragment from the RS-rich domain. ASF/SF2, ASF-3, and the deletion mutants were expressed and purified from *E.coli* as previously described (8). The 32 amino acid long polypeptide (peptide-32) with the sequence DRDRER-RRSHRSERERRRDRDRDRDRDREHKR (corresponding to positions 349–380 of the human U1-70K protein) was synthesized on an Applied Biosystems 430A peptide synthesizer. The peptide was purified before use by Sephadex G25 gel filtration.

Kinase assays

In vitro phosphorylation assays were performed by incubating a mixture of anti-m_3G purified snRNPs (final concentration of 100–10 μg protein/ml) or Mono Q purified U1 snRNP (final concentration 50 μg protein/ml) at 37°C for 30 min in buffer containing 10 mM Hepes-KOH, pH 7.9, 10% (w/v) glycerol, 50 mM KCl, 3 mM $MgCl_2$, 0.1 mM EDTA, 0.5 mM DTE, 0.5 mM PMSF and 0.2 mCi/ml (66 nM) gamma-[^{32}P]-ATP (3000 Ci/mmol). ASF/SF2 variants and peptide-32 were added to the aforementioned mixture to a final concentration of 25–50 μg/ml and 150 μg/ml, respectively. For the negative control, duplicate samples were frozen in liquid nitrogen immediately after gamma-[^{32}P]-ATP addition or, alternatively, diluted 20-fold with PK buffer containing 100 mM Tris–HCl (pH 7.5), 12.5 mM EDTA, 150 mM NaCl and 0.1% (w/v) SDS, and immediately placed on ice.

PAGE analysis of phosphorylated proteins

Protein was isolated from the *in vitro* kinase assays or *in vivo* labeled, immunoaffinity purified snRNPs, by extraction with phenol/chloroform essentially as described by Lehmeier *et al.* (26). Proteins were fractionated on 7% or 15% discontinuous SDS-polyacrylamide gels as described by Laemmli *et al.* (28) or by two-dimensional PAGE as previously described (2).

Phosphoamino acid analysis and tryptic phosphopeptide mapping

Tryptic phosphopeptide mapping was performed as previously described (2). Phosphoamino acid analysis was performed generally as described by Bister *et al.* (29). In particular, radiolabeled 70K protein, phosphorylated either *in vivo* or *in vitro*, was fractionated by SDS-PAGE and isolated as described for tryptic proteolysis by Woppmann *et al.* (2). Lyophilized, radiolabeled 70K protein was then redissolved in 50 μl of 5.7 M HCl and incubated for 2 h at 110°C. After evaporation of the acid, hydrolyzed products were separated by two dimensional electrophoresis on cellulose-coated thin-layer plates. Electrophoresis in the first dimension was carried out for 20 min at 2300 V in pH 1.9 buffer [2.2% (v/v) formic acid, 7.8% (v/v) acetic acid] and in the second dimension, for 16 min at 1500 V in pH 3.5 buffer [0.5% (v/v) pyridine, 5% (v/v) acetic acid]. Phosphoamino acid markers were visualized by staining with 0.2% ninhydrin dissolved in acetone.

Chymotryptic digestion

Digestions were performed with radiolabeled 70K protein that had been eluted from an SDS-PAGE gel as described (2). Typically, 70 μl of 50 mM NH_4HCO_3 were added to a 20 μl sample containing ~0.2 μg 70K protein in gel elution buffer (10 mM NH_4HCO_3, 0.5% SDS and 0.5% β-mercaptoethanol), followed by 10 μg of freshly prepared chymotrypsin (1 mg/ml in 50 mM NH_4HCO_3), and incubated at 37°C for 15 to 240 min. The samples were lyophilized prior to SDS-PAGE analysis.

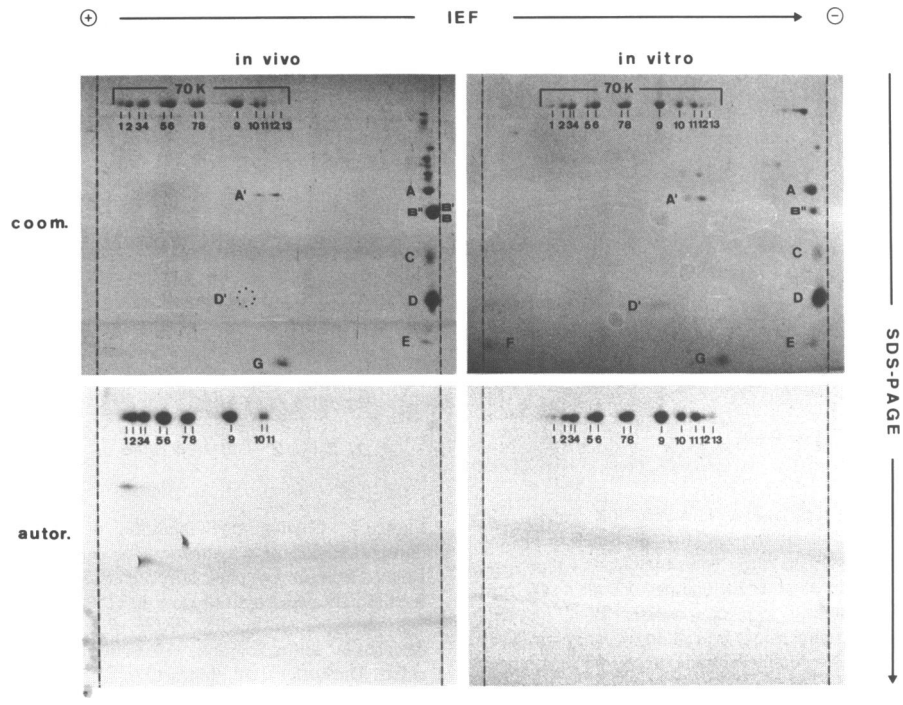

Figure 1. Phosphorylation of the U1-specific 70K protein. U1, U2, U4/U6 and U5 snRNPs were isolated from S1R nuclear extract, prepared from HeLa cells grown in either the absence or presence of ^{32}P-orthophosphate, by immunoaffinity chromatography (see Methods). Unlabeled, immunoaffinity purified U1–U6 snRNPs were phosphorylated *in vitro* as described in Materials and Methods. Protein isolated from 50 μg of *in vivo* (left panel) or *in vitro* (right panel) ^{32}P-phosphorylated snRNPs was subjected to two-dimensional PAGE (isoelectric focusing in a pH 3–11 gradient followed by 15% SDS-PAGE). The coomassie-stained 2-D gels are displayed in the upper panels, while the corresponding autoradiograms are shown in the lower. The positions of snRNP proteins and of the 70K isoelectric variants are indicated. An identical pattern is observed with snRNPs isolated from splicing extracts.

RESULTS

The U1 snRNP 70K protein kinase is associated with purified snRNPs

We have previously demonstrated through the *in vivo* labeling of HeLa cells with ^{32}P-orthophosphate that the U1 snRNP-associated 70K protein is phosphorylated *in vivo* (2). Two-dimensional PAGE analysis further revealed that all of the previously identified 70K isoelectric variants are phosphorylated. To test whether the phosphotransferase activity responsible for 70K protein phosphorylation is associated with snRNPs, U1–U6 snRNPs which had been purified from HeLa nuclear extract by anti-m_3G immunoaffinity chromatography were incubated with gamma-[^{32}P]-ATP and the phosphorylation of snRNP-associated proteins analysed by two-dimensional PAGE. For comparison, *in vivo* ^{32}P-labeled, U1–U6 snRNPs, isolated in an identical manner, were also analysed. As shown in Figure 1, a kinase activity which selectively phosphorylates all of the 70K protein isoelectric variants is present in affinity purified snRNPs. Identical results are obtained with snRNPs isolated either from S1R or splicing extracts. As previously observed with *in vivo* labeled 70K protein, the extent of *in vitro* ^{32}P-incorporation for each variant directly correlates with its mass (compare coomassie stained gel with autoradiograph), suggesting that each 70K variant is phosphorylated to a similar extent.

Additional *in vitro* phosphorylation studies with Mono Q purified snRNPs indicate that the snRNP-associated U1-70K kinase is predominantly, although not exclusively, associated with U1 snRNPs (data not shown). Kinase activity does not require an intact RNP particle, since 70K phosphorylation is observed

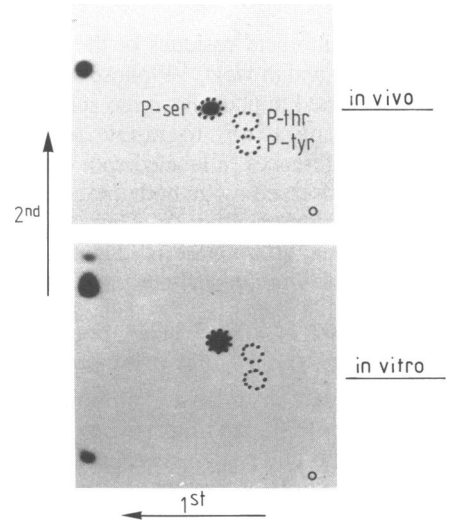

Figure 2. Phosphoamino acid analysis of phosphorylated 70K protein. ^{32}P-labeled 70K protein (600 cpm) was isolated from *in vivo* (upper panel) or *in vitro* (lower panel) phosphorylated snRNPs, hydrolyzed, and the resultant amino acids separated by two-dimensional thin-layer electrophoresis as described in the Methods section. ^{32}P-labeled amino acids were detected by autoradiography and identified by comigration with internal standards of phosphoserine (P-ser), phosphothreonine (P-thr), and phosphotyrosine (P-tyr) which were visualized by staining with ninhydrin. The positions of these standards are indicated by dotted circles. Small circles indicate the site of sample application. The left, upper most spot in both panels is free phosphate (P_i) which is generated by the hydrolysis of amino acid-phosphate bonds during the acid hydrolysis step.

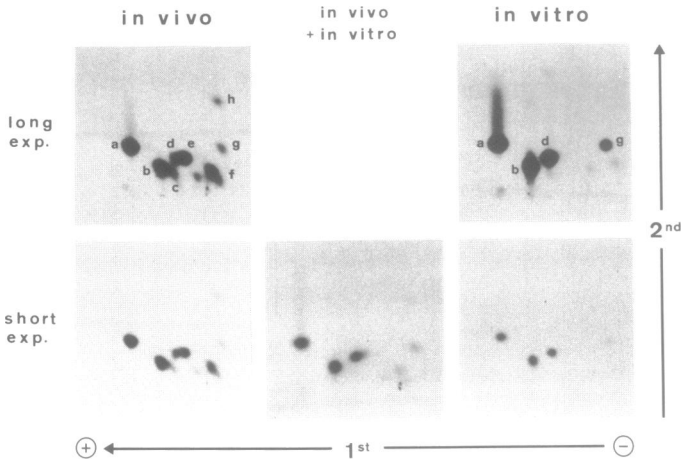

Figure 3. Tryptic phosphopeptide analysis of phosphorylated 70K protein. ^{32}P-labeled 70K protein (4000 cpm) was isolated from *in vivo* (left panel) or *in vitro* (right panel) phosphorylated S1R snRNPs and digested to completion with trypsin (see Methods). Each sample, as well as a 50:50 mixture of both (2000 cpm each)(lower middle panel), was fractionated by electrophoresis (first dimension) followed by chromatography (second dimension) on cellulose-coated thin-layer plates. Phosphopeptides (a−h) were detected by autoradiography for long (upper panels) or short (lower panels) periods of time.

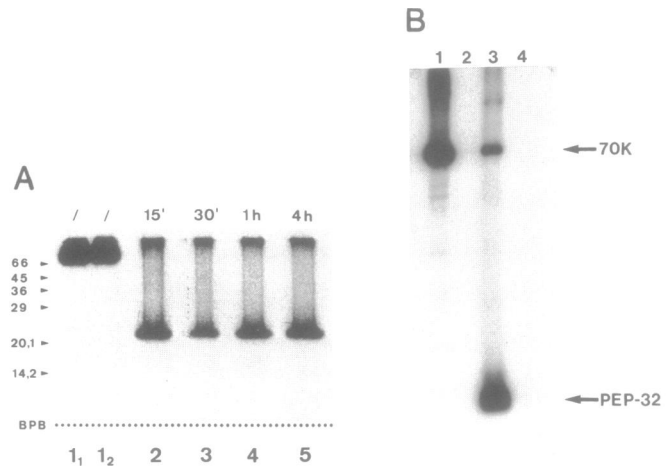

Figure 4. Chymotryptic digestion of phosphorylated 70K and *in vitro* phosphorylation of a synthetic 70K peptide. **A**. ^{32}P-labeled 70K protein was isolated from *in vivo* phosphorylated S1R snRNPs, digested with chymotrypsin for the indicated length of time at 37°C (lanes 2−5) or at O°C (lanes 1_1 and 1_2), and separated by SDS-PAGE on a 15% gel. ^{32}P-labeled fragments were detected by autoradiography. Molecular weight markers (in kDa) are indicated at left. The buffer front, indicated by a dotted line, is still visible and, thus, even the shortest possible chymotryptic fragments should be detectable. **B**. Kinase active U1 snRNPs and gamma-[^{32}P]-ATP were incubated in the absence (lanes 1−2) or presence (lanes 3−4) of a synthetic 70K peptide (pep-32) with the following sequence: DRDRERRRSHRSERERRRDRDRDRDRDREHKR. Samples were incubated at 37°C (lanes 1 and 3) or diluted with PK buffer and incubated on ice (lanes 2 and 4) for 30 min. Protein was isolated as described in Methods and separated by SDS-PAGE on a 12% gel. The radiolabeled proteins were detected by autoradiography. The position of the 70K protein and synthetic peptide are indicated at right.

with RNase-digested snRNPs or with purified snRNP proteins which have been dissociated from snRNA (data not shown). Additionally, *in vitro* phosphorylation studies demonstrate that the snRNP-associated kinase preferentially utilizes ATP as a phosphodonor; no significant 70K phosphorylation is observed with GTP or any other nucleotide, aside from ATP (not shown).

Serine residues of the 70K protein are phosphorylated *in vivo* as well as *in vitro*

To identify which amino acid residues of the 70K protein are phosphorylated *in vivo* and *in vitro*, ^{32}P-phosphorylated 70K was gel purified and subjected to phosphoamino acid analysis. After two-dimensional separation together with unlabeled phosphoamino acid references, a labeled spot corresponding to phosphoserine was observed with both *in vivo* and *in vitro* phosphorylated 70K protein (Fig.2). Thus, the amino acid specificity of the *in vivo* 70K kinase is identical to that of the enzyme associated with immunoaffinity purified snRNPs.

An overlapping subset of radiolabeled phosphopeptides is observed with *in vitro* versus *in vivo* labeled 70K protein

^{32}P-labeled 70K protein was isolated from both *in vivo* and *in vitro* phosphorylated snRNPs and digested to completion with trypsin. The resultant tryptic peptides were fractionated by two dimensional thin layer chromatography. As shown in the upper left panel of Figure 3, eight radiolabeled peptides (designated a−h) are observed in the *in vivo* labeled protein (the unlabeled spot to the left of f is not reproducibly observed), indicating at least eight different phosphorylation sites. Since seven identical SR dipeptides, one or more of which could be phosphorylated, are theoretically expected upon digestion with trypsin, there are potentially as many as fifteen phosphorylation sites. By comparison, proteolysis of the *in vitro* labeled 70K yields four phosphopeptides which appear to be identical to those of the *in vivo* labeled sample designated a, b, d, and g. To test whether these peptides are indeed identical, a mixture of the *in vivo* and *in vitro* phosphopeptides was analysed (Fig.3, lower middle panel). The exact comigration of peptides from each sample indicate that peptides a, b, d, and g of the *in vitro* phosphorylated sample possess amino acid sequences identical to those of the *in vivo* labeled sample. Thus, identical, albeit fewer, phosphorylation sites are present after *in vitro* phosphorylation. These results strongly suggest that the snRNP-associated serine kinase is responsible, at least in part, for the observed *in vivo* phosphorylation of the U1-70K protein.

The multiple U1-70K phosphorylation sites are clustered in a carboxy-terminal domain

Phosphopeptide analysis demonstrated multiple (at least four) *in vitro* phosphorylation sites within the U1-70K protein. In order to map these sites, *in vitro* ^{32}P-labeled 70K protein was digested with chymotrypsin for various lengths of time and the products separated by SDS-PAGE. As shown in Figure 4A, chymotryptic digestion yields only one labeled fragment with an apparent molecular weight of 23 kd (identical results are obtained with *in vivo* labeled protein). The fact that only one radiolabeled fragment is generated, even after lengthy digestion times, indicates that proteolysis was complete. The sequence of the U1-70K protein, with the predicted sites of chymotryptic cleavage indicated, is presented in Figure 5B. A compilation of all of the expected fragments and their molecular weights is also depicted in Table 1. Consistent with our results, chymotryptic cleavage should yield a 23 kd fragment which maps to the C-terminal half of the 70K protein. To confirm that phosphorylation occurs within this carboxy-terminal region, *in vitro* phosphorylation was carried

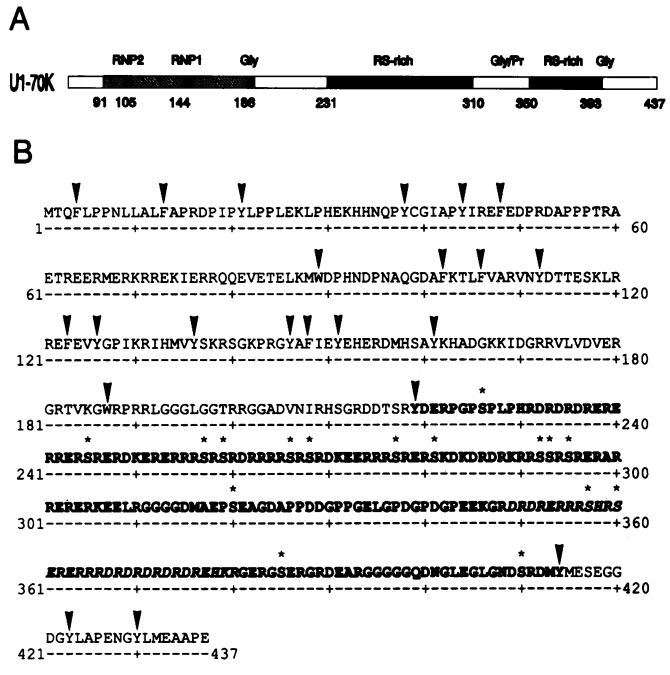

Figure 5. Domain structure and chymotrypsin cleavage sites of the 70K protein. **A.** The structure of the human 70K protein is depicted. The RNA binding domain, with the consensus RNP2 and RNP1 sequences, and glycine- and argine/serine-rich regions are indicated. The amino acid positions of each region are indicated below. **B.** The amino acid sequence of the human 70K protein with all possible chymotryptic cleavage sites (arrowheads) is presented. The 23kd chymotryptic fragment discussed in the text is indicated by bold-faced letters and potential phosphorylation sites are marked with an asterisk. The sequence of peptide-32 is italized.

out with a 32 amino acid long synthetic peptide corresponding to residues 349−380 (see Fig.5B). As shown in Figure 4B, this peptide was efficiently phosphorylated by kinase active U1 snRNPs (at the expense of 70K phosphorylation), providing further evidence that phosphorylation of the 70K protein is limited to the 23 kd region defined by chymotryptic proteolysis. Upon close inspection, it is apparent that this region of the 70K protein contains several arginine/serine rich clusters (Fig.5A and B). The predominance of potential phosphorylation sites consisting of an RS motif prompted us to investigate whether other proteins containing similar motifs could be phosphorylated by the snRNP-associated, serine kinase.

The snRNP-associated kinase phosphorylates the RS domain of ASF/SF2

Sequence analysis of alternative splicing factor (ASF/SF2) revealed a carboxy terminal domain consisting of multiple, repetitive RS dimers (8,9). The suitability of this domain as a substrate for the snRNP-associated serine kinase was tested by incubating bacterially expressed, recombinant ASF/SF2, several deletion mutants, and ASF-3 with kinase-active, purified U1 snRNPs and gamma-[^{32}P]-ATP. As shown in Figure 6A, lanes 1 and 2, ASF/SF2 is efficiently phosphorylated, whereas ASF-3, an alternative splicing variant which lacks an RS-rich domain, is not (lanes 3 and 4). The RS-domain as the site of phosphorylation is further confirmed by the results obtained with various ASF/SF2 deletion mutants (lanes 5−12). Removal of 24 amino acids from the C-terminus (mutant A) has little effect

Table 1. U1-70K chymotryptic peptides

Position	From	To	Mol. Wt.
15	147	148	236.3
12	124	126	409.4
16	149	151	423.5
9	103	106	507.6
1	1	4	525.6
6	45	48	563.6
5	39	44	622.7
10	107	112	720.8
23	431	437	759.9
22	424	430	762.8
3	14	21	928.0
21	415	423	943.9
2	5	13	997.2
14	137	146	1135.3
13	127	136	1213.5
17	152	161	1274.3
11	113	123	1381.5
8	90	102	1397.4
4	22	38	2077.4
18	162	187	2976.4
19	188	219	3466.8
7	49	89	5193.8
20	220	414	23434.7

Collection of all theoretical 70K peptides generated by chymotryptic digestion and their expected molecular weight.

on the efficiency of *in vitro* phosphorylation. Deletion of an additional 24 amino acids (mutant B) completely inhibits ^{32}P-incorporation, suggesting that the majority of phosphorylation sites are located within residues 200−224. A low level of phosphorylation is observed with mutant D which contains only the C-terminal most 20 amino acids of the RS domain, suggesting that additional phosphorylation sites are located in this region as well.

DISCUSSION

In this paper we have demonstrated the presence of an enzymatic activity in UsnRNPs (U1, U2, U5, and U4/U6) which had been immunoaffinity purified from HeLa cell nuclear extract. This snRNP-associated enzyme was shown to phosphorylate the U1-specific 70K protein *in vitro*. *In vivo* and *in vitro* phosphorylated 70K share several properties, providing evidence that the snRNP-associated kinase is responsible for the *in vivo* phosphorylation of this protein. In both instances, all thirteen 70K isoelectric variants are phosphorylated and serine residues are the sole recipients of a phosphate group (see Fig.1 and 2). More importantly, four identical phosphopeptides are generated upon tryptic digestion of the *in vivo* and *in vitro* labeled protein, indicating both are phosphorylated at identical sites (see Fig.3). The presence of four additional *in vivo* radiolabeled peptides (i.e., a total of eight) suggests that one or more additional serine kinases, distinct from the snRNP-associated one, may be involved in 70K phosphorylation. Alternatively, several sites normally phosphorylated by the snRNP-associated kinase *in vivo*, could be unaccessible to the enzyme *in vitro*, or, since at least partial dephosphorylation is required for the *in vitro* incorporation of gamma-[^{32}P]-ATP, they could be extensively phosphorylated in the isolated U1 snRNP. Nonetheless, our results suggest that the phosphorylation of the 70K protein *in vivo* is attributed, at least in part, to the activity of the snRNP-associated enzyme. We

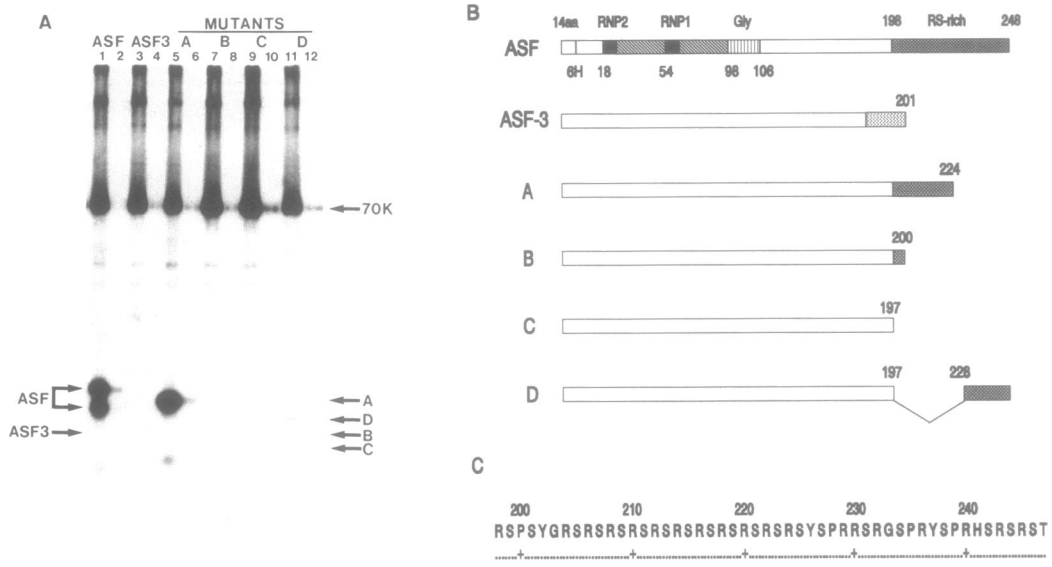

Figure 6. *In vitro* phosphorylation of recombinant ASF/SF2 and deletion mutants. **A.** Equal molar amounts of bacterially expressed ASF/SF2 (lanes 1,2), ASF-3 (lanes 3,4), and deletion mutants A (lanes 5,6), B (lanes 7,8), C (lanes 9,10) and D (lanes 11,12) were incubated with kinase active U1 snRNPs and gamma-[^{32}P]-ATP at 37°C for 30 min (odd numbered lanes) or on ice for 30 min after the addition of PK buffer (even numbered lanes). Protein was isolated as described in Methods and separated by SDS-PAGE on a 12% gel. Radiolabeled proteins were detected by autoradiography. The positions of the ASF/SF2 variants and 70K protein, as detected by coomassie stain (not shown), are indicated by arrows. Bacterially expressed ASF, like its mammalian counterpart, often migrates as a doublet (see refs. 8 and 9 for further discussion). **B.** The structure of ASF/SF2, ASF-3, and the deletion mutants is presented. The RNA binding domain, with the consensus RNP2 and RNP1 sequences, and glycine-rich region, common to all variants (open box), as well as the arginine/serine (RS)-rich region (stippled box) are indicated. The amino acid positions of the various regions are indicated below. The position of the C-terminal most amino acid or cleavage site (mutant D) is indicated. **C.** Amino acid sequence of the RS-rich domain of human ASF-1.

cannot, however, exclude the possibility that additional serine kinases with an overlapping substrate specificity also contribute to the *in vivo* phosphorylation of one or more sites recognized by the snRNP-associated kinase.

The serine kinase co-purifing with snRNPs appears to be relatively stably-associated with at least one of the snRNP particles, since U1 snRNPs fractionated by Mono Q column chromatography retain significant kinase activity. However, despite its apparent preferential association with U1, there is evidence against a particle-specific association, since other snRNP species appear to possess the U1-70K kinase activity as well (not shown). In addition, although associated with snRNP particles, kinase activity is not dependent upon an intact RNP structure. Purified snRNP proteins, prepared free of snRNA as described by Sumpter *et al.*(30), or particles treated with RNase A, still retain the ability to phosphorylate the U1-70K protein. Preliminary fractionation studies indicate that the U1-70K kinase is a minor protein distinct from the thus far identified snRNP proteins (Fetzer *et al.*, unpublished results). Due to its apparent low abundance (relative to previously characterized snRNP proteins), we cannot rule out that this enzyme represents a contaminant rather than a *bona fide* snRNP component. However, retention of the U1-70K kinase even after several purification steps, including immunoaffinity chromatography, glycerol gradient centrifugation and Mono Q chromatography, suggests a physical association of the kinase with, at least, the U1 snRNP particle. This association may arise by its interaction with the 70K protein during phosphorylation; a stable association between a kinase and its preferred substrate has been reported for nucleolar protein kinase NII and nucleolin (31).

Several lines of evidence indicate that the multiple 70K phosphorylation sites are confined to serine residues in the C-terminal half of the protein. When either *in vivo* or *in vitro* ^{32}P-labeled 70K protein was subjected to chymotryptic proteolysis (see Fig.4A) radioactivity was confined to a single digestion product with an apparent molecular weight of 23 kd. A comparison with the theoretically expected chymotryptic fragments (Fig.5B and Table 1), suggested that it corresponded to a 23 kd fragment encompassing much of the C-terminal half of 70K. Since potential chymotryptic cleavage sites could be inactivated by amino acid modification, the observed 23 kd fragment could consist of two or more uncleaved, smaller fragments. This is, however, relatively unlikely, since all other theoretical cleavage products are very small and, thus, the generation of a 23 kd fragment would require the inactivation of multiple, consecutive cleavage sites. The presumed identity of this phosphorylated fragment was confirmed by the observed efficient *in vitro* phosphorylation of a synthetic peptide which comprises 32 residues (349 to 380) of the proposed phosphorylated 70K region. The absence of phosphorylation sites in the most N-terminal third of the protein was demonstrated by *in vitro* phosphorylation studies with an N-terminal fragment of the 70K protein (not shown).

The unique character of the phosphorylated region of the 70K protein led to an initial hypothesis regarding the amino acid sequences preferentially recognized by the snRNP-associated serine kinase. In particular, this domain is highly charged and contains a high percentage of arginine and, to a lesser extent, serine residues. As mentioned above, similar RS-rich domains have been documented in a number of splicing factors, including ASF/SF2, both subunits of U2AF, SC35, and the Drosophila splicing regulators *tra*, *tra-2*, and *su(w^a)*. The likelihood that RS motifs are recognized and phosphorylated by the snRNP-associated enzyme is underscored when the potential

phosphorylation sites within the 23 kd region of the 70K protein are compared (see Figure 5B); the majority of serine residues are flanked by an arginine residue. Tryptic phosphopeptide analysis of *in vitro* phosphorylated 70K demonstrated that four unique amino acid sequences serve as substrates for the serine kinase. Although sixteen serine residues are present in this region, only nine unique peptides are expected upon tryptic proteolysis, since seven identical SR dipeptides should be generated. While the majority of unique tryptic peptides contain serines flanked by arginine, four of these fragments, and thus four potential phosphorylation sites, do not. The results of *in vitro* phosphorylation with the synthetic 32mer provided the first direct evidence that the U1-70K protein kinase may preferentially phosphorylate serines flanked by arginine; both serine residues within this peptide fall into this catagory. Preliminary phosphorylation studies with additional peptides of varying sequence supported this hypothesis as well (not shown). Significant support is further provided by the results of the *in vitro* phosphorylation studies utilizing the splicing factor ASF/SF2 (see below).

While additional studies are clearly needed before a consensus phosphorylation site can be assigned to the snRNP-associated kinase, its apparent affinity for serine residues preceeded or flanked by arginine residues suggests that it may be a novel member of the serine/threonine class of protein kinases. Few of the well-studied serine/threonine kinases phosphorylate serine residues either preceeded or flanked by arginine (32). Two well-studied kinases which fit into this catagory are cyclic nucleotide-dependent enzymes which phosphorylate histone H2B (32). The absence of cyclic nucleotides in the *in vitro* kinase reaction demonstrates the cyclic nucleotide independence of the snRNP-associated kinase; by analogy it must also not require calcium or diacylglycerol, two additional well-documented kinase activators. The inability of this kinase to phosphorylate casein or histone H1 (data not shown), further distinguishes it from casein kinase I and protein kinase C, respectively. Although it is tempting to speculate that the snRNP-associated kinase belongs to a previously unidentified group of serine kinases, such conclusions are premature in the absence of additional information. A more detailed characterization of this snRNP-associated kinase is clearly warranted and may require its purification to homogeneity.

Important evidence for the substrate specificity and potential general role of the snRNP-associated kinase in splicing regulation was provided by *in vitro* phosphorylation studies with recombinant ASF/SF2, ASF-3, and the deletion mutants. Efficient phosphorylation of ASF/SF2, but not ASF-3 was observed (see Fig.6). Since ASF-3 differs from ASF/SF2 primarily at its carboxy terminus (i.e., the RS-rich region is absent), these results already indicated that phosphorylation by the snRNP-associated kinase is confined to the RS-rich domain. The availability of ASF/SF2 mutants lacking various regions of the RS domain allowed us to map more precisely the sites of ASF phosphorylation. The significant loss of ^{32}P-incorporation upon deletion of residues 200 to 224, demonstrated that the majority of phosphorylation sites are located in this region. Interestingly, this segment is comprised almost exclusively of ten RS repeats; of the eleven serine residues present, only one is not flanked by arginine (see Fig.6C). Additional phosphorylation sites were also detected in the most C-terminal 20 amino acids. Not surprisingly, this region also contains several RS dipeptides.

By analogy with the U1-70K protein, our results suggest that the snRNP-associated kinase may also be responsible for the apparent *in vivo* phosphorylation of ASF/SF2. They additionally raise the question as to whether other splicing factors which contain RS-rich domains may be phosphorylated by this kinase as well. The functional importance of the RS domain of U2AF and ASF/SF2 has recently been demonstrated (10,17). In both instances, deletion of this domain abolishes the factor's function in constitutive pre-mRNA splicing, although its RNA binding capacity is maintained. In the case of ASF/SF2, which carries out dual functions, acting in both constitutive and alternative pre-mRNA splicing, deletion of the RS domain does not affect its capability to influence splice site selection (17). Interestingly, residues 200−224, corresponding to the heavily phosphorylated region described here, are necessary and sufficient to provide RS function in the constitutive splicing assay (17). In light of these results, the potential functional significance of modifying this region through phosphorylation is readily apparent. Due to the highly basic nature of such RS-rich domains, they have been postulated to mediate binding to RNA through electrostatic interactions (3), or conceivably interactions with other proteins. Phosphorylation (and subsequent dephosphorylation) could modulate an electrostatic-based interaction of this region with other molecules, for example, during spliceosome formation or splicing proper. The phosphorylation of ASF/SF2 (or other factors such as SC35 and U2AF) could be temporally regulated by means of the 70K protein kinase's association with snRNPs. That is, the U1 snRNP, for example, could act as a chaperone, bringing the enzyme into close proximity to the above mentioned factors during pre-spliceosome assembly.

The potential significance of protein phosphorylation in the regulation of splicing is underscored by its well documented role in the regulation of countless other biochemical processes. A number of recent studies have provided evidence that protein phosphorylation/dephosphorylation cycles play a significant role in pre-mRNA splicing. In particular, serine/threonine phosphatase activity has been shown to be essential for the first and second steps of splicing (21,22). Tazi *et al.* have provided evidence that one potential target of these phosphatases may be the U1-70K protein (23). *In vitro* phosphorylated U1 snRNPs were added to splicing extracts which had been depleted of their endogeneous U1 and their ability to complement the splicing-inactive extracts tested. U1 particles containing normally phosphorylated 70K restored splicing, whereas those particles phosphorylated with S-ATP did not. Since ^{35}S-phosphorylated 70K was not dephosphorylated in splicing extracts (in contrast to ^{32}P-phosphorylated protein), these results suggested that the dephosphorylation of the 70K protein was crucial for its *in vitro* splicing activity. Our results suggest that an additional target of the serine/threonine phosphatases may be ASF/SF2. We are currently utilizing *in vitro* splicing complementation assays to test whether the phosphorylation state of the ASF/SF2 RS domain affects its activity in constitutive pre-mRNA splicing.

ACKNOWLEDGEMENTS

We thank Thilo Patschinsky for help in performing phospho-tryptic peptide mapping and Berthold Kastner for providing kinase-active, Mono Q purified U1 snRNPs. This research was supported by grants from the Fonds der Chemischen Industrie and SFB 272 (A3). A.W. was supported in part by the Friedrich Ebert Stiftung.

REFERENCES

1. Wooley,J.C., Cone,R.D., Tartof,D. and Chung,S.Y. (1982) Proc. Natl. Acad. Sci. USA, **79**, 6762–6766.
2. Woppmann,A., Patschinsky,T., Bringmann,P., Godt,F. and Lührmann,R. (1990) Nucleic Acids Res., **18**, 4427–4438.
3. Moore,M.J., Query,C.C. and Sharp,P.A. (1993) In Cold Spring Harbor Symposium on Quantitative Biology, Cold Spring Harbor Press, Cold Spring Harbor, in press.
4. Lührmann,R., Kastner,B. and Bach,M. (1990) Biochim. Biophys. Acta Gene Struct. Expression, **1087**, 265–292.
5. Ge,H. and Manley,J.L. (1990) Cell, **62**, 25–34.
6. Krainer,A.R., Conway,G.C. and Kozak,D. (1990) Genes Dev., **4**, 1158–1171.
7. Kuo,H., Nasim,F.H. and Grabowski,P.J. (1991) Science, **251**, 1045–1050.
8. Ge,H., Zuo,P. and Manley,J.L. (1991) Cell, **66**, 373–382.
9. Krainer,A.R., Mayeda,A., Kozak,D. and Binns,G. (1991) Cell, **66**, 383–394.
10. Zamore,P.D., Patton,J.G. and Green,M.R. (1992) Nature, **355**, 609–614.
11. Fu,X.-D. and Maniatis,T. (1992) Science, **256**, 535–538.
12. Bingham,P.M., Chou,T.-B., Mims,I. and Zachar,Z. (1988) Trends in Genetics, **4**, 134–138.
13. Mattox,W., Ryner,L. and Baker,B.S. (1992) J. Biol. Chem., **267**, 19023–19026.
14. Theissen,H., Etzerodt,M., Reuter,R., Schneider,C., Lottspeich,F., Argos,P., Lührmann,R. and Philipson,L. (1986) EMBO J., **5**, 3209–3217.
15. Spritz,R.A., Strunk,K., Surowy,C.S., Hoch,S.O., Barton,D.E. and Francke,U. (1987) Nucleic Acids Res., **15**, 10373–10391.
16. Query,C.C. and Keene,J.D. (1987) Cell, **51**, 211–220.
17. Zuo,P. and Manley,J.L., in press.
18. Mayeda,A., Zahler,A.M., Krainer,A.R. and Roth,M.B. (1992) Proc. Natl. Acad. Sci. USA, **89**, 1301–1304.
19. Roth,M.B., Murphy,C. and Gall,J.G. (1990) J. Cell Biol. **111**, 2217–2223.
20. Roth,M.B., Zahler,A.M. and Stolk,J.A. (1991) J. Cell Biol. **115**, 587–596.
21. Tazi,J., Daugeron,M.-C., Cathala,G., Brunel,C. and Jeanteur,P. (1992) J. Biol. Chem., **267**, 4322–4326.
22. Mermoud,J.E., Cohen,P. and Lamond,A.I. (1992) Nucleic Acids Res., **20**, 5263–5269.
23. Tazi,J., Kornstädt,U., Rossi,F., Jeanteur,P., Cathala,G., Brunel,C. and Lührmann,R., Nature, in press.
24. Bringmann,P., Rinke,J., Appel,B., Reuter,R. and Lührmann,R. (1983) EMBO J., **2**, 1129–1135.
25. Dignam,J.D., Lebovitz,R.M. and Roeder,R.G. (1983) Nucleic Acids Res., **11**, 1475–1489.
26. Lehmeier,T., Foulaki,K. and Lührmann,R. (1990) Nucleic Acids Res., **18**, 6475–6484.
27. Bach,M., Bringmann,P. and Lührmann,R. (1990) Methods Enzymol., **181**, 232–257.
28. Laemmli,U.K. (1970) Nature, **227**, 680–685.
29. Bister,K., Trachmann,C., Jansen,H.W., Schroeer,B. and Patschinski,T. (1987) Oncogene, **1**, 97–109.
30. Sumpter,V., Kahrs,A., Fischer,U., Kornstädt,U. and Lührmann,R. (1992) Mol. Biol. Rep., **16**, 229–240.
31. Caizergues-Ferrer,M., Belenguer,P., Lapeyre,B., Amalric,F., Wallace, M.O. and Olsen,M.O.L. (1987) Biochemistry, **26**, 7876–7883.
32. Pearson,R.B. and Kemp,B. (1992) Methods Enzymol., **200**, 62–85.