The C-Terminal Domain of RNA Polymerase II Functions as a Phosphorylation-Dependent Splicing Activator in a Heterologous Protein

Scott Millhouse and James L. Manley*

Department of Biological Sciences, Columbia University, New York, New York

Received 3 September 2004/Accepted 18 October 2004

RNA polymerase II, and specifically the C-terminal domain (CTD) of its largest subunit, has been demonstrated to play important roles in capping, splicing, and 3' processing of mRNA precursors. But how the CTD functions in these reactions, especially splicing, is not well understood. To address some of the basic questions concerning CTD function in splicing, we constructed and purified two fusion proteins, a protein in which the CTD is positioned at the C terminus of the splicing factor ASF/SF2 (ASF-CTD) and an RS domain deletion mutant protein (ASFΔRS-CTD). Significantly, compared to ASF/SF2, ASF-CTD increased the reaction rate during the early stages of splicing, detected as a 20- to 60-min decrease in splicing lag time depending on the pre-mRNA substrate. The increased splicing rate correlated with enhanced production of prespliceosomal complex B and the early spliceosomal complex C but, interestingly, not the very early ATP-independent complex E. Additional assays indicate that the RS domain and CTD perform distinct functions, as exemplified by our identification of an activity that cooperates only with the CTD. Dephosphorylated ASFΔRS-CTD and a glutathione S-transferase–CTD fusion protein were both inactive, suggesting that an RNA-targeting domain and CTD phosphorylation were necessary. Our results provide new insights into the mechanism by which the CTD functions in splicing.

Synthesis of mature mRNAs in the nuclei of eukaryotic cells involves a series of pre-mRNA processing events, including the addition of a cap structure to the 5' end, removal of intronic sequences (splicing), and 3' cleavage and polyadenylation. Recently, it has become increasingly evident that all three mRNA-processing steps are integrated with transcription in vivo, presumably to facilitate accurate and efficient processing of RNA polymerase II (RNAPII) transcripts, and that the RNAPII C-terminal domain (CTD) is important for all three processing reactions (2, 24, 36, 48). The CTD is conserved among eukaryotes and consists of multiple heptapeptide repeats with the consensus sequence YSPTSTPS (9). The number of repeats varies from 26 or 27 in the yeast Saccharomyces cerevisiae to 52 in mammals. The heptapeptide consensus contains five potential phosphoacceptor amino acids, but evidence suggests that the serines (or threonines) at positions 2 and 5 are the predominant sites of phosphorylation (63). Two forms of RNAPII can be distinguished based on the phosphorylation status of CTD, a hypophosphorylated II A form, which preferentially enters the preinitiation complex at the promoter, and a hyperphosphorylated II O form, which is extensively phosphorylated in vivo. The CTD contains five potential phosphoacceptor amino acids, but experimental evidence suggests that the serines (or threonines) at positions 2 and 5 are the predominant sites of phosphorylation (63). Two forms of RNAPII can be distinguished based on the phosphorylation status of CTD, a hypophosphorylated II A form, which preferentially enters the preinitiation complex at the promoter, and a hyperphosphorylated II O form, which is associated with elongation complexes (12). More recently, different phosphorylated forms of RNAPII have been observed on genes depending on the location of the polymerase along the gene (7, 33), and specific kinases have been implicated in directing serine-specific phosphorylation of the CTD during the transcription cycle (16, 49, 56, 65).

Splicing of most mRNAs occurs in a large macromolecular complex composed of snRNPs (U1, U2, U4, U5, and U6) and non-snRNP proteins, including members of the SR protein family (23, 29, 37). In structure, SR proteins contain one or two N-terminal RNA binding domains (RBD) and a C-terminal domain enriched in arginine and serine residues (RS domain), which is extensively phosphorylated in vivo. The RBD has been demonstrated to bind RNA in a sequence-specific manner, a characteristic that allows for the interaction of SR proteins with elements such as exonic or intronic splicing enhancers. Functionally, SR proteins are believed to perform an essential role early in spliceosome assembly by stabilizing the interactions between snRNPs and splice sites, as well as by mediating protein-protein contacts through the formation of bridges formed between 5' and 3' splice sites (11, 32, 52, 59). The result is the formation of a complex composed of U1 snRNP bound to the 5' splice site and the splicing factor U2AF bound to the 3' splice site region (E complex). In the presence of ATP, splicing proceeds as E complexes are converted to A, B, and C complexes in a well-characterized kinetic pathway that requires U2, U4, U5, and U6 snRNPs and other factors (29, 35, 42, 51).

The above model of spliceosome assembly was the result of analyses of splicing in isolation from transcription, but more recent studies suggest an important role for the RNAPI O CTD in spliceosome assembly. Cytological studies have suggested that splicing in vivo can occur cotranscriptionally and that essential splicing factors are localized at sites of active transcription (1, 3). RNAPI O also colocalizes with SR proteins in large nuclear domains called “speckles” (4), and splicing of transcripts produced by RNAPII with a shortened CTD is impaired (18, 38). In biochemical studies, RNAPI O can physically interact with snRNPs and SR-like proteins (6, 31, 41, 57, 61), and we have shown that purified RNAPI O can enhance in vitro splicing of several pre-mRNAs at an early (E or
A complex step in spliceosome assembly, while IIA in fact has a repressive effect (26). RNAP II CTD-targeted antibodies and CTD peptides have also been shown to inhibit in vitro splicing (61). A recombinant CTD fusion protein (glutathione S-transferase [GST]–CTD) has been shown to enhance in vitro splicing of pre-mRNAs containing exons defined by splice sites (62).

Because RNAP IIO is the elongating form of the polymerase and functions early to enhance spliceosome assembly, we previously proposed a model where the hyperphosphorylated CTD of RNAP IIO and its associated proteins facilitate the binding of U1 and/or U2 snRNPs to the 5′ splice site and/or branch site, respectively (26). This model is supported by in vitro transcription experiments which showed that U1 and U2 snRNPs can be recruited to elongating RNAP II by interactions involving the CTD kinase P-TEFb (20). However, despite this considerable evidence linking transcription, and specifically the CTD, to splicing, little is known about how the CTD actually functions in this process.

We have continued our examination into the relationship between the CTD and the splicing machinery by constructing chimeric proteins in which the CTD is fused to the SR protein ASF/SF2 or to an RS domain-deleted derivative of ASF/SF2. These two fusion proteins, dubbed ASF-CTD and ASF/RS-CTD, were designed to test first whether recruitment of the CTD to the pre-mRNA, perhaps normally performed by the body of RNAP II, might be important for CTD function and second whether the CTD can functionally substitute for the RS domain. We show that, like ASF/SF2, ASF-CTD complements splicing in SR protein-deficient S100 extract. However, ASF-CTD markedly increased the reaction rate at or before the first catalytic step. Although ASF/Rs and ASF/Rs-CTD were generally much less active in S100, ASF-CTD and ASF/Rs-CTD, but not ASF/SF2 or ASF/Rs, increased the rate of splicing similarly in nuclear extracts. As with intact RNAP II, CTD phosphorylation was necessary for the observed splicing enhancement, and GST-CTD was inactive. We also provide evidence for the existence of different CTD cofactor requirements depending on the pre-mRNA substrate. These results are consistent with a model in which the phosphorylated CTD facilitates an early step(s) of spliceosome assembly but highlight the possibility that it plays a more sophisticated role in splicing regulation than previously thought.

**RESULTS**

**Immunoblot analysis and phosphatase treatment.** ASF/SF2-CTD was dephosphorylated with calf intestinal phosphatase (CIP; New England Biolabs) at 30°C in a buffer containing 20 mM HEPES (pH 8), 50 mM KC1, 1% glycerol, 5 mM MgCl2, 0.25 mM EDTA, and 0.1 mM EDTA. Approximately 10 U of CIP was required to fully dephosphorylate 1 μg of ASF/SF2-CTD after 3 h at 30°C. To prepare purified dephosphorylated ASF/SF2-CTD, 50 μg of ASF/SF2-CTD was treated with CIP and then denatured by dialysis in a buffer containing 8 M guanidine HCl (pH 8). Dephosphorylated ASF/SF2-CTD was purified by Ni2+-agarose chromatography and renatured as described for the phosphorylated protein except that final dialysis was in buffer D containing 2 M urea to increase solubility. Phosphorylated ASF/Rs-CTD was also dialyzed in buffer D containing 2 M urea as a control for functional studies involving the dephosphorylated protein. Immunoblot analysis was performed by separating phosphorylated and dephosphorylated proteins by SDS–8% polyacrylamide gel electrophoresis (PAGE), followed by immunoblotting with H14, H5, or 8WG16 antibodies ( Covance). Unless specifically stated otherwise in the text, recombinant ASF/SF2-CTD and ASF/Rs-CTD proteins are referred to as ASF-CTD and ASF/Rs-CTD, respectively, and the recombinant ASF/Rs-CTD protein treated with CIP and repurified is referred to as ASF-CTDp.

**In vitro splicing and spliceosome assembly.** 32P-labeled pre-mRNA substrates were prepared as described previously (54). Nuclear extract and cytoplasmic S100 were prepared by the method of Dignam et al. (15). Fractionation of HeLa nuclear extract to obtain a 20 to 40% ammonium sulfate fraction of nuclear extract (NF20-40) was performed as described previously (55). In vitro splicing reactions were performed in either 12.5- or 25-μl reaction mixtures, which contained 2 to 16% nuclear extract or 20% S100 as a control. Samples were treated without or with calf intestine phosphatase as described in the figure legends. The final concentrations of buffer components were 12 mM HEPES (pH 7.9), 40 to 60 mM KC1, 0.12 mM EDTA, 0.30 mM DTT, 12% glycerol, 3 mM MgCl2, 20 mM creatine phosphate (di-Tris), 1 mM ATP, 3% polyvinyl alcohol (PVA), and 0.5 U of RNasin (Promega), with the exception that, in experiments involving dephosphorylated ASF/SF2-CTD, urca was present at a final concentration of 0.16 mM. In addition, for ATP-dependent spliceosome assembly assays, PVA was omitted from the splicing reactions and heparin (0.8 mg/ml) was added prior to loading the samples on native agarose gels as described previously (13). For ATP-independent E complex assembly, PVA, ATP, creatine phosphate, and MgCl2 were omitted from splicing reactions and the samples were loaded on 1.5% native agarose gels (13). The adenosine major late (AdML) pre-mRNA was utilized in these studies due to better resolution of E complexes. Experiments were repeated two to four times and were highly consistent with the representative data shown.

**RESULTS**

ASF/SF2-CTD fusion proteins expressed in baculovirus-infected insect cells are phosphorylated on Ser-2 and Ser-5 of the heptapeptide repeats. To study the functional properties of the CTD relevant to splicing, the CTD was expressed as a C-terminal fusion protein with ASF/SF2 and the RS domain deletion mutant protein, ASF/SF2 (68). ASF/SF2, ASF/Rs,
ASF-CTD, and ASFΔRS-CTD proteins were purified from baculovirus-infected insect cells, and aliquots were examined by SDS-PAGE (Fig. 1A). Although the expected molecular masses of ASF-CTD and ASFΔRS-CTD are 68 and 62 kDa, respectively, the apparent molecular masses observed by SDS-PAGE were approximately 116 and 110 kDa, respectively (Fig. 1A, lanes 3 and 4). Since CTD phosphorylation imparts a dramatic decrease in gel mobility (64), it was likely that the CTDs of these proteins were highly phosphorylated during expression. Treatment of ASFΔRS-CTD with CIP shifted the protein’s gel mobility to approximately 80 kDa (Fig. 1B, compare lane 1 to 4 or 5) or resulted in intermediate-mobility forms at midrange concentrations of CIP (Fig. 1B, lanes 2 and 3).

Since the CTD is predominantly phosphorylated at serines 2 and 5 during transcription in vivo, the CTD of ASFΔRS-CTD was tested for phosphorylation at these positions by Western blotting using H5 and H14 antibodies, which are specific for heptads phosphorylated at Ser-2 or Ser-5, respectively. A Coomassie blue stain of ASFΔRS-CTD during a CIP time course is shown in Fig. 1C. The same samples were also immunoblotted with CTD-specific antibody 8WG16, which recognizes unphosphorylated epitopes in the CTD (Fig. 1C, lanes 5 to 8), H5 (Fig. 1C, lanes 9 to 12), and H14 (Fig. 1C, lanes 13 to 16). Ser-2 phosphorylation was strongly detected at the 0-min time point of CIP incubation (Fig. 1C, lane 9) but was undetectable at later times (Fig. 1C, lanes 10 to 12). Ser-5 phosphorylation, in contrast, was detected not only at the 0-min time point (Fig. 1C, lane 13) but also at the 1- and 2-h time points and was detected in the intermediate- and even high-mobility forms of ASFΔRS-CTD (Fig. 1C, lanes 14 and 15). These results indicate that H5 reactivity can be completely lost before H14 reactivity is affected. The ability of H14, but not H5, to detect intermediate levels of CTD phosphorylation has also been observed with authentic RNAP II (44) and, together with the observation of discrete bands for the untreated protein and the dephosphorylated isoform, supports the view that these forms are analogous to the IIO and IIA isoforms of RNAP II. The CTDs of our recombinant proteins are designated below either CTD0 or CTDΔ to differentiate the native phosphorylated and the CIP-dephosphorylated recombinant proteins, respectively.

Fusing the CTD to ASF/SF2 increases splicing rate in S100.

We showed previously that RNAP IIO can stimulate splicing of several pre-mRNAs in cell extracts but that, when GST-CTD was added to splicing reactions under similar conditions, no enhancement was observed (26). This result was somewhat unexpected since the CTD is sufficient both to activate 3′ cleavage in vitro (25) and to associate with a number of splicing factors (10, 31, 40). We speculated that GST-CTD may not enter spliceosomes formed in vitro, because it is unable to interact in the appropriate way with the pre-mRNA. Given that SR proteins both interact with the pre-mRNA and function, as appears to be the case with RNAP IIO, at a very early stage (21), it seemed reasonable to hypothesize that fusing the CTD with an SR protein may allow for the participation of the CTD in splicing. Alternatively, it could be that other regions in

---

FIG. 1. Characterization of histidine-tagged ASF-CTD fusion proteins purified from baculovirus-infected insect cells. (A) Silver-stained gradient (4 to 20%) SDS-polyacrylamide gel containing 40 ng of purified ASFΔRS (lane 1), ASF (lane 2), ASFΔRS-CTD (lane 3), and ASF-CTD (lane 4). (B) Coomassie-stained SDS–8% polyacrylamide gel containing purified ASFΔRS-CTD treated with 0.0 (lane 1), 0.25 (lane 2), 1.0 (lane 3), and 5.0 U (lanes 4 to 6) of CIP for 3 h at 30°C. Phosphatase activity was inhibited by addition of 50 mM β-glycerophosphate (lane 6). (C) Time course (0, 1, 2, and 3 h) of ASFΔRS-CTD dephosphorylation analyzed by Coomassie staining and immunoblotting with 8WG16, H5, and H14 antibodies.
the multisubunit RNAP II are necessary for its function in splicing.

We first examined splicing of a β-globin pre-mRNA substrate in S100 extract complemented with either ASF/SF2 or ASF-CTDO. Since S100 lacks SR proteins but contains all other essential splicing factors, splicing does not occur in S100 unless supplemented with exogenous SR proteins such as ASF/SF2. To examine the effect of the CTD on ASF/SF2 activity, a splicing time course was performed with the β-globin pre-mRNA, S100, and either ASF/SF2 or ASF-CTDO (Fig. 2A). Both ASF/SF2 and ASF-CTDO at 100 nM concentrations activated splicing. Most interestingly, the reactions with ASF-CTDO showed an increase in the rate of splicing, as shown by the formation of the first-step intermediates or final products. Quantitation (Fig. 2B) revealed that reactions with ASF-CTDO displayed a reduction in splicing lag time of approximately 20 min (from 40 to 20 min) compared to reactions with ASF/SF2. Relative to ASF-CTDO, both ASFΔRS and ASFΔRS-CTDO stimulated splicing much less efficiently (Fig. 2C, compare lanes 2 to 7 with lane 1), suggesting that the RS domains in both ASF/SF2 and ASF-CTDO were functionally important in this assay. Nevertheless, the comparison between ASF-CTDO and ASF/SF2 showed that ASFΔRS-CTDO stimulated splicing more efficiently than did ASFΔRS (Fig. 2C, compare lanes 5 to 7 with 2 to 4). These results suggested that the CTDs of both ASF-CTDO and ASFΔRS-
CTD₀ are stimulatory in the splicing of β-globin pre-mRNA substrate in S100 extract and that the CTD₀ can play a functional role in splicing that is distinct from that of the RS domain.

To extend these results, we next compared the splicing kinetics of reactions with ASF-CTD₀ (Fig. 2D, lanes 2 to 4) with those of reactions with either the same amount (lanes 5 to 7), 2.5-fold more (lanes 8 to 10), or 6-fold more (lanes 11 to 13) ASF/SF2. We observed that, even with sixfold more ASF/SF2, the first appearance of splicing intermediates (approximately 30 min) lagged behind that seen in reactions with ASF-CTD₀ (Fig. 2D, compare lanes 5, 8, and 11 with 1). These data indicate that higher ASF/SF2 levels did not substantially increase the reaction rate prior to the first catalytic step. In addition, at later times in the reaction (30 to 60 min), the ratio of final spliced product to lariat/exon 2 intermediate was noticeably higher in the reactions with ASF-CTD₀ than in all of the reactions with ASF/SF2 and this ratio did not change significantly in the ASF/SF2 reactions regardless of the amount of ASF/SF2 included (for example, compare lanes 7, 10, and 13 in Fig. 2D). We conclude from these data that the CTD alters qualitatively the effect of ASF/SF2 on splicing.

Splicing activity of ASF-CTD₀ is not specific to β-globin pre-mRNA and S100 extracts. To test both the generality of the above findings and the response of a pre-mRNA with a more specific requirement for ASF/SF2, we examined the splicing of the human immunodeficiency virus type 1 (HIV-1) tat pre-mRNA (34). In contrast to that of β-globin, the tat pre-mRNA can be specifically committed to splicing by ASF/SF2, but not by other SR proteins (21). A time course of splicing in S100 extract (Fig. 3A) shows that ASF-CTD₀ (lanes 8 to 14) again increased the splicing rate compared to ASF/SF2 at an identical concentration (lanes 1 to 7). Quantitation revealed that the splicing lag time was decreased by approximately 30 min when either splicing intermediates or final products were quantitated (Fig. 3B). The splicing of tat pre-mRNA in nuclear extracts, which unlike S100 extracts contain all of the essential splicing factors, is unusual because it is inhibited by endogenous hnRNP A1 (5, 66). However, this inhibition can be overcome, at least in part, by addition of exogenous ASF/SF2 (34) or RNAP IIO (26). Addition of ASF-CTD₀ to nuclear extract not only enabled tat pre-mRNA splicing but again increased the splicing rate compared to that for ASF/SF2 (Fig. 3C, com-
pare lanes 8 to 14 with 1 to 7). Remarkably, ASF-CTD was decreased the long lag time of tat pre-mRNA splicing in nuclear extract by nearly 60 min (from 120 to 60 min) compared to ASF/SF2 (Fig. 3D).

CTD-dependent splicing enhancement correlates with ATP-dependent spliceosome assembly, but not ATP-independent prespliceosome assembly. We next wished to investigate at which step the CTD functioned to accelerate splicing. To this end, we utilized an AdML pre-mRNA, which has frequently been used to study spliceosomal complex formation in reaction mixtures containing nuclear extracts (13). As with other pre-mRNAs tested, the AdML pre-mRNA was also spliced with increased kinetics when nuclear extracts were supplemented with ASF-CTD (Fig. 4A and B). When spliceosome assembly was examined by native gel electrophoresis, the formation of prespliceosomal A complexes as well as early spliceosomal B complexes was more efficient in reactions with ASF-CTD than in reactions with ASF/SF2 or nuclear extract alone (Fig. 4C and D). This was most clearly seen with the A complexes at the earliest (2- and 5-min) time points (Fig. 4C compare lanes 14 and 15 to 8 and 9) and D), and with B complexes at later (15- and 30-min) time points (Fig. 4C compare lanes 16 and 17 to 10 and 11) and D).

The data with ASF-CTD were very similar to results obtained with purified RNAP IIO (26) and are consistent with the possibility that the CTD may facilitate U1 and/or U2 binding to the pre-mRNA. However, the first discrete splicing complex detectable in HeLa extracts is the E complex (39, 50), which forms in the absence of ATP and which is believed to be a functional precursor of the A complex. Therefore, we next examined whether ASF-CTD could specifically enhance E complex assembly by examining prespliceosome formation in nuclear extract in the absence of ATP and creatine phosphate. Unlike nonspecific H complexes, E complexes require incubation at 30°C and can be separated from H complexes by native agarose gel electrophoresis as long as heparin is not added to the reactions (13). Surprisingly, we observed no differences in the rate of E complex assembly between ASF-CTD and ASF/SF2 (Fig. 4E, compare lanes 14 and 15 to 8 and 9) and D), and with B complexes at later (15- and 30-min) time points (Fig. 4C compare lanes 16 and 17 to 10 and 11) and D).

Addition of a nuclear fraction to S100 stimulates IgM-A3 splicing in a CTD-dependent manner. An important question is whether CTD function involves cofactors distinct from components of the general splicing machinery. To begin to address this, we took advantage of the fact that, while high concentrations of ASF/SF2 can activate splicing of the IgM-A3 substrate (an immunoglobulin M pre-mRNA-based substrate containing consensus ASF/SF2 binding sites) in S100 (54, 58), at lower concentrations, neither ASF/SF2 nor ASF-CTD had any effect on splicing (Fig. 5A, compare lanes 5 to 9). However, when we added increasing amounts of NF20-40 (53), ASF-CTD, but not ASF/SF2, activated IgM-A3 splicing (Fig. 5A, compare lanes 10 to 12 to 6 to 8). The effect of NF20-40 on the splicing of other pre-mRNAs was also examined, but the stimulation was significantly less than that observed with the IgM-A3 pre-mRNA (data not shown). These findings suggest that a splicing factor or cofactor that cooperates specifically with the CTD and not with ASF/SF2 is present in NF20-40. To confirm this, we performed an ASF/SF2 titration in the absence or presence of NF20-40. Increasing amounts of ASF/SF2 resulted in a dose-dependent increase in splicing of IgM-A3 pre-mRNA (Fig. 5B, lanes 4 to 6), but addition of a constant amount of NF20-40 had no significant effect at any ASF/SF2 concentration (Fig. 5B, compare lanes 8 to 10 with 4 to 6). This is in sharp contrast to the significant effect that NF20-40 had on ASF-CTD-activated splicing (Fig. 5B, compare lanes 1 and 2). These data provide strong evidence that a splicing cofactor(s) can specifically cooperate with the CTD to activate splicing.

The RS domain and CTD have distinct functions in splicing. The above data suggest that the CTD and RS domain perform qualitatively distinct functions in splicing. However, because both the ASF/SF2 RS domain and the RNAP IIO CTD consist of highly phosphorylated repetitive motifs, we wished to further examine differences and similarities between them. In nuclear extract, both ASF-CTD and ASF ΔRS-CTD enhanced splicing of IgM-A3, whereas ASF/SF2 and ASF ΔARS could not (Fig. 6A). This suggests that the RS domains of these proteins did not contribute to splicing under these conditions, which was not surprising since SR proteins exist naturally in nuclear extracts to perform this function. In stark contrast to what was found for IgM-A3 splicing in nuclear extracts, in reaction mixtures containing S100 and NF20-40 we observed a strong dependency on both the RS domain and the CTD of ASF-CTD (Fig. 6B). Under conditions where neither ASF-CTD nor ASF ΔARS-CTD activated splicing in S100 alone (Fig. 6B, lanes 1 to 3), the addition of an equal amount of NF20-40 to these reactions stimulated only those with ASF-CTD (Fig. 6B, compare lane 5 with 4 and 6). Therefore, it seems that the very limited amount of SR proteins in NF20-40, which from Western blots we estimate to be similar to the amount detected in S100 (53) (data not shown), requires the addition of exogenous SR proteins with intact RS domains in order to activate splicing of the IgM-A3 substrate and that the CTD is unable to perform this function. These data also suggest that the factor(s) in NF20-40 that cooperates with the CTD is not SR proteins.

CTD phosphorylation is necessary for CTD splicing activity. The importance of CTD phosphorylation in pre-mRNA splicing is supported by the observation that RNAP IIO but not RNAP IIA enhances in vitro splicing (26). Since our ASF-CTD fusion proteins were purified in a highly phosphorylated state, we wished to determine if CTD phosphorylation is necessary for ASF-CTD splicing activity. However, since ASF/SF2 itself is highly phosphorylated in its RS domain, we chose to use ASF ΔARS-CTD in these studies, so that dephosphorylation would affect only the CTD. To examine this, we dephosphorylated ASF ΔARS-CTD in vitro with CIP and repurified the resultant protein. A silver-stained SDS gel of the dephosphorylated protein, ASF ΔARS-CTD, shows the large mobility shift compared to that for the phosphorylated protein, ASF ΔARS-CTD (Fig. 7A).

The splicing activity of ASF ΔARS-CTD was first tested with the IgM-A3 substrate by addition of increasing amounts of ASF ΔARS-CTD (Fig. 7B, lanes 4 and 5) or ASF ΔARS-CTD (Fig. 7B, lanes 2 and 3) to nuclear extract. ASF ΔARS-CTD was unable to activate splicing, whereas ASF ΔARS-CTD again produced a dose-dependent increase in splicing activity. We also
tested splicing of the β-globin pre-mRNA in nuclear extract supplemented with either ASFΔRS-CTD₀ or ASFΔRS-CTD₁, and again only ASFΔRS-CTD₀ produced increased splicing activity (Fig. 7C, lanes 2 to 4 compared to 5 to 7). Importantly, GST-CTD, either phosphorylated or dephosphorylated, had no detectable effect on the splicing of IgM-A3 at any concentration tested (Fig. 7D). These data together show that CTD phosphorylation is necessary for ASF/SF2-CTD-enhanced splicing in vitro and further strengthens the view that CTD fused to ASF/SF2 and CTD in its natural context function similarly.
DISCUSSION

Considerable evidence has established that RNAP II, and specifically the CTD, can play a positive role in the splicing of mRNA precursors. However, there has been very little progress in understanding how it actually functions in this process. This is due to the complexity both of the splicing reaction and of RNAP II itself. As we showed previously and confirmed here, and unlike the situation with capping and 3' end formation, the isolated CTD does not function in standard splicing assays. We hypothesized that this might reflect the inability of the isolated CTD to be targeted to the substrate and/or splicing machinery, and the data presented here support that idea. Our findings are significant because they not only indicate that the CTD is sufficient to enhance splicing when targeted to the pre-mRNA but also provide a means for more readily analyzing CTD function in splicing. Below we discuss the evidence that the ASF-CTD fusion protein indeed recapitulates the behavior of RNAP II and the new insights it provides into how the CTD enhances splicing.

It is essential that the ASF-CTD fusion proteins function similarly to RNAP II if they are to provide insights into how the CTD functions in splicing. All the data presented here strongly support this. For example, our results showing that the fusion proteins function early in splicing are consistent with previous work in our laboratory which showed that RNAP II enhanced the splicing of several pre-mRNA substrates at an early step of spliceosome assembly. We also showed previously that CTD function could not be performed by excess SR proteins (26), which is consistent with our finding here that the CTD and RS domain have distinct functions. Additionally, in agreement with previous studies involving RNAP II and RNAP IIA (26), the CTD required phosphorylation in order to enhance splicing.

To understand the mechanism by which the CTD enhances splicing, it was important to determine where within the spliceosome assembly pathway CTDO performs this function. Previous evidence with RNAP II suggested a CTD-sensitive step during or prior to A complex formation (26). The results presented here are in agreement with this conclusion and extend our understanding by deemphasizing a CTDO role in ATP-independent E complex formation and more specifically implicating ATP-dependent A complex formation as a CTDO-sensitive step. At the snRNP level, E complex is characterized by the base pairing of the U1 snRNP to the 5' splice site whereas the U2 snRNP is not yet base paired and is loosely associated with the complex (14). We conclude that the CTDO is most likely not involved in stabilizing the U1-5' splice site interaction that occurs without ATP and is instead involved in the later events that occur in the presence of ATP, such as the base pairing of U2 to the branch site.

Our data have provided evidence that the phosphorylated CTD performs a function distinct from that of the RS domain. This is in fact consistent with our finding that the CTDO had no effect on E complex formation since SR proteins are important for the commitment of pre-mRNAs to splicing by stabilizing U1 binding to the 5' splice site through the combined interactions between the RBD and pre-mRNA and RS domain and protein components of U1 snRNP (28, 32). This is additional evidence to support the theory that the CTDO is involved in ATP-dependent transition to the A complex by facilitating U2 base pairing to the branch site. Although we never observed any effect of CTDO on E complex formation in our in vitro assays, it remains unclear if, in the presence of ATP, the CTDO might play a role in the initial interactions

FIG. 5. Addition of NF20-40 to S100 strongly stimulates in vitro splicing of IgM-A3 in the presence of ASF-CTD but not ASF/SF2. (A) Titration of NF20-40 into splicing reactions does not complement S100 alone (lanes 1 to 4) or S100 supplemented with 200 nM ASF/SF2 (lanes 5 to 8), but NF20-40 stimulates splicing of IgM-A3 in the presence of 200 nM ASF-CTD (lanes 9 to 12). The amounts of NF20-40 used in the titrations were 0, 1, 3, and 6 μl. (B) Three microliters of NF20-40 stimulates splicing in S100 in the presence of 100 nM ASF-CTD (lane 2). Larger amounts of ASF/SF2 stimulate splicing of IgM-A3 equally well in either the absence (lanes 3 to 6) or presence (lanes 7 to 10) of 3 μl of NF20-40. Concentrations of ASF/SF2 used in titrations were 0, 200, 400, and 600 nM.

540 MILLHOUSE AND MANLEY MOL. CELL. BIOL.
between the components of E complex and the pre-mRNA. This seems plausible since it has been demonstrated that ATP alters the composition of large snRNP/RNAP IIO complexes that form on short 5′ splice site RNAs (30). Notably, in these studies, ATP allowed for the incorporation of U2 into U1-containing complexes. In any event, our results were surprising because they suggest that the CTD, while functioning early in splicing, does not affect the earliest known step, i.e., E complex formation.

It remains an important question what CTD-binding proteins are necessary for CTD-enhanced splicing. However, our data suggest that different CTD-binding proteins have unique splicing functions, as illustrated by the strong nuclear protein requirement for CTD-activated splicing of IgM-A3, but not of 5′-globin or tat pre-mRNAs. Since CTDo did not show any effect on IgM-A3 splicing in S100 unless NF20-40 was added, CTDo-enhanced splicing of a subset of pre-mRNAs may be more strongly enhanced by cofactors that are absent from S100, whereas other pre-mRNAs may be less dependent on these factors due to specific strengths and weaknesses of a particular pre-mRNA. Therefore, the CTD may coordinate the actions of two or more CTDo-associated splicing regulators that have discrete functions. Accordingly, it is interesting that the CTD has also been shown to bind to the splicing factor PSF (17). In addition to linking CTDo to early spliceosomal complex assembly (43), PSF has been shown to be required much later in the splicing reaction, during the second catalytic step (22). It is tempting to speculate that CTDo can, in a substrate-specific manner, remain associated with spliceosomes from prespliceosome formation through catalysis. In line with this possibility, RNAP IIO has been shown to form stable complexes with the late spliceosomal component U5 snRNP and the [U4/U6.U5] tri-snRNP (6, 57).

Our data confirm our previous suggestion that the isolated CTD cannot activate splicing in vitro. This contrasts with the ability of the CTD alone to function in vitro capping (8, 27) and 3′ processing (25) and indicates that another property of RNAP IIO is necessary for CTDo-enhanced splicing. Fusion of the CTD to ASF/SF2, which normally enters spliceosomes, should localize the CTD closely with the pre-mRNA, perhaps in a manner analogous to that for a nascent transcript and the CTD of elongating RNAP IIO. Because we observed splicing enhancement by ASF-CTD, with or without the RS domain, we believe that it is most likely that the RNA binding activity of ASF/SF2 is in fact important for CTDo splicing activity, by providing an RNA targeting function. In possible contrast to our results, it has been reported previously that GST-CTD can enhance in vitro splicing, through the exon definition mode of splicing (62). We observed no requirement for having the exons in our substrate pre-mRNAs defined at both ends by splice sites. Perhaps the suggested exon definition role of the isolated CTD is bypassed by having the CTD fused to ASF/SF2, which itself is involved in identifying splice sites through interactions with U1 and U2AF (28, 32, 55, 59). Another difference is that our ASF-CTD was a highly phosphorylated protein preparation compared to the mostly unphosphorylated GST-CTD used by Zeng and Berget (62). Additionally, in our experiments, we never tested the high GST-CTD concentrations employed by these authors.

The splicing-stimulatory activity of ASFΔRS-CTDo, like
that of RNAP II, requires CTD phosphorylation, and therefore phosphorylated CTD heptads are likely to interact specifically with a component of the splicing machinery. Studies using yeast two-hybrid assays have revealed phospho-CTD-binding proteins that contain an RS domain, an RNA recognition motif, and a distinct CTD-binding domain (45, 61); however, the functional relevance of these proteins in splicing remains to be determined. The phosphorylated CTD has been shown to interact directly with splicing factors such as PSF (17) and the yeast splicing factor Prp40 (40). Nevertheless, the importance of these protein-protein interactions for CTD-enhanced splicing in our in vitro system requires further investigation.

CTD phosphorylation undergoes dynamic changes during transcription, and it is likely that this is critical for transcription-coupled RNA processing. For example, it has been shown that CTD Ser-5 phosphorylation is observable near the promoter and Ser-2 phosphorylation is seen throughout the gene (7). TFIIH kinase (16, 56) is involved in phosphorylation of Ser-5 during promoter escape, yielding a CTD that is competent for 5’ capping of the nascent transcript, and the transcription elongation factor P-TEFb (49, 65) is responsible for at least a fraction of the Ser-2 phosphorylation during the transcript elongation phase. The correlation between the timing of Ser-2 phosphorylation and the entry of elongating RNAP II

FIG. 7. Phosphorylation of ASFΔRS-CTD is necessary for enhanced splicing activity in nuclear extract. (A) Silver-stained SDS-polyacrylamide gel of 50 ng of purified phosphorylated ASFΔRS-CTDα (lane 2) and dephosphorylated and repurified ASFΔRS-CTDα (lane 3). Protein molecular mass markers are shown in lane 1 (New England Biolabs; broad range, 7702S). (B) Addition of 15 (lane 2) or 30 nM (lane 3) ASFΔRS-CTDα stimulates splicing of IgM-A3 in nuclear extract (NE). Addition of 15 (lane 4) or 30 nM (lane 5) ASFΔRS-CTDα has no effect on splicing of IgM-A3. (C) Addition of ASFΔRS-CTDα (lanes 2 to 4) stimulates splicing of β-globin, but ASFΔRS-CTDα (lanes 5 to 7) has no effect on splicing. (D) Neither GST-CTDα (lanes 2 to 4) nor GST-CTDα (lanes 5 to 7) stimulates splicing of IgM-A3 under conditions where ASFΔRS-CTDα (lanes 8 to 10) enhances splicing. Reaction mixtures for panel D contained a titration of 5, 20, or 80 nM respective recombinant proteins and were incubated for 75 min at 30°C.
into coding regions suggests the possibility that Ser-2 phosphorylation may be important for CTD-enhanced splicing. P-TEFb interacts with RNAP II and the HIV-1 Tat cofactor and general elongation factor, Tat-SF1 (19). Intriguingly, P-TEFb has also been shown to interact with snRNPs through a Tat-SF1–snRNPs complex (20). It was demonstrated that this complex promotes transcription elongation and supports splicing in nuclear extracts depleted of snRNA. In yeast, the CUS2 protein was identified as a suppressor of a U2 snRNA mutation and, intriguingly, is homologous to Tat-SF1 (46, 60). The relationship between the Ser-2 kinase P-TEFb, RNAP II, Tat-SF1, and snRNPs (U2 in particular) is generally consistent with our suggestion that the CTD domain may function during prespliceosomal A complex formation. The detection of P-TEFb, RNAP II, and snRNPs in large complexes formed on short 5′-splice sites of mammalian A complex formation. The detection of P-TEFb, RNAP II, and snRNPs in large complexes formed on short 5′-splice-site within large transcription/splicing complexes. EMBO J. 13:283–3367.

The phosphorylated CTD is a functionally independent domain that, when targeted to the pre-mRNA, interacts with the splicing machinery to facilitate the formation of the prespliceosomal A complex. This is reflected in a significant decrease in the lag time typically seen in vitro splicing reactions. The CTD domain function is distinct from that of the RS domain, consistent with evidence showing that CTD domain must interact with protein domains different from those that interact with RS domains. For the future, this experimental system provides a means to examine not only splicing factors that functionally associate with the CTD domain but also the role of specific heptads in splicing. Our results extend previous studies involving purified RNAP II and support the idea that the phosphorylated CTD can function as a general activator of splicing, provided it is properly recruited to the RNA substrate.

ACKNOWLEDGMENTS

We thank N. Rao and C. Shin for assistance in the production and characterization of HeLa extracts and R. Reed for plasmids and technical advice.

This work was supported by NIH R37 GM48259 to J.L.M. S.M. was the recipient of NRSF F32 GM063322.

REFERENCES
