Regulation of Alternative Polyadenylation by U1 snRNPs and SRp20

HUA LOU,^{1,2*} KARLA M. NEUGEBAUER,³ ROBERT F. GAGEL,² AND SUSAN M. BERGET¹

Verna and Marrs McLean Department of Biochemistry, Baylor College of Medicine,¹ and Section of Endocrine Neoplasia and Hormonal Disorders, University of Texas M.D. Anderson Cancer Center,² Houston, Texas 77030, and Division of Basic Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington 98109³

Received 20 April 1998/Returned for modification 20 May 1998/Accepted 2 June 1998

Although considerable information is currently available about the factors involved in constitutive vertebrate polyadenylation, the factors and mechanisms involved in facilitating communication between polyadenylation and splicing are largely unknown. Even less is known about the regulation of polyadenylation in genes in which 3'-terminal exons are alternatively recognized. Here we demonstrate that an SR protein, SRp20, affects recognition of an alternative 3'-terminal exon via an effect on the efficiency of binding of a polyadenylation factor to an alternative polyadenylation site. The gene under study codes for the peptides calcitonin and calcitonin gene-related peptide. Its pre-mRNA is alternatively processed by the tissue-specific inclusion or exclusion of an embedded 3'-terminal exon, exon 4, via factors binding to an intronic enhancer element that contains both 3' and 5' splice site consensus sequence elements. In cell types that preferentially exclude exon 4, addition of wild-type SRp20 enhances exon 4 inclusion via recognition of the intronic enhancer. In contrast, in cell types that preferentially include exon 4, addition of a mutant form of SRp20 containing the RNA-binding domain but missing the SR domain inhibits exon 4 inclusion. Inhibition is likely at the level of polyadenylation, because the mutant SRp20 inhibits binding of CstF to the exon 4 poly(A) site. This is the first demonstration that an SR protein can influence alternative polyadenylation and suggests that this family of proteins may play a role in recognition of 3'-terminal exons and perhaps in the communication between polyadenylation and splicing.

Polyadenylation is a ubiquitous step during posttranscriptional pre-mRNA processing. The biochemistry and basal machinery of polyadenylation have been well characterized. Polyadenylation is a two-step process: cleavage and poly(A) addition (10, 17, 55). Both *cis*-acting elements and *trans*-acting factors are required for this process. The *cis* elements include an AAUAAA hexanucleotide upstream of and a G/U-rich sequence downstream of the cleavage site. A multiprotein complex including four components, cleavage-polyadenylation specificity factor, cleavage stimulation factor (CstF), cleavage factor, and poly(A) polymerase, assembles on a constitutive poly(A) site during polyadenylation. Many of these polyadenylation factors have been purified and cloned.

Expression of a number of genes is regulated at the level of polyadenylation. Little, however, is known about the molecular mechanisms affording regulation. It has been suggested that differential binding of a basal polyadenylation factor, the CstF 64-kDa protein, contributes to developmentally regulated usage of an immunoglobulin alternative poly(A) site (16, 52). Auxiliary factors that interact with the basal polyadenylation machinery have also been implicated in regulated polyadenylation. In the latter class are two U1 snRNP-associated proteins, U1A and the U1 70K protein (22–24, 34, 35). To date, no non-snRNP proteins have been reported to be important for regulation of polyadenylation.

SR proteins are a family of serine- and arginine-rich RNA-binding proteins (reviewed in references 18, 37, 42, 47, and 53). They were initially isolated as essential splicing factors and

regulators of 5' splice site selection (20, 29, 30, 59, 60). The classical members of this family include SRp20, SC35, ASF/SF2, SRp30c, SRp40, SRp55, and SRp75. Subsequently, 9G8 and SRp54 were added to the family (8, 61). Classical SR proteins have been shown to bind to exon enhancer sequences and/or splice sites to enhance exon inclusion and to regulate splice site recognition. During splicing, SR proteins bound to RNA interact through their SR domains, both with other SR proteins and with constitutive splicing factors that contain SR domains, such as U2AF and the U1 snRNP 70K protein (47, 53, 57). These interactions have been postulated to play a role in bridging splice sites in exons and introns during early assembly of the spliceosome. Therefore, classical SR proteins are normally considered to be associated with splicing.

Recently, one subunit of a constitutive polyadenylation factor, cleavage factor I (CFI), has been observed to possess a domain containing SR dipeptides (49). The presence of this domain suggests that polyadenylation could communicate to splicing via SR proteins and that SR proteins could participate in regulation of polyadenylation during alternative 3' exon recognition. One system of alternative polyadenylation that has the potential to reveal the role of SR proteins in polyadenylation is that occurring during alternative processing of pre-mRNA from the human calcitonin/calcitonin gene-related peptide (CT/CGRP) gene.

This processing choice involves tissue-specific recognition of one of two alternate 3'-terminal exons, exon 4 or 6 (1, 48). In thyroid C cells, exon 4 is included to generate an mRNA molecule containing exons 1 to 4, with the usage of the exon 4 polyadenylation site. The CT peptide is produced from this mRNA. In neuronal cells, exon 4 is excluded to generate an mRNA molecule containing exons 1 to 3 and 5 to 6, with usage of the exon 6 polyadenylation site (Fig. 1A). The CGRP is

* Corresponding author. Mailing address: Verna and Marrs McLean Department of Biochemistry, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030. Phone: (713) 798-4622. Fax: (713) 795-5487. E-mail: hlou@bcm.tmc.edu.

FIG. 2. In vivo inclusion of CT/CGRP exon 4 requires hybridization of U1 snRNA to the enhancer core 5' splice site sequence. (A) Diagram of RT-PCR analysis of CT/CGRP RNA produced in cells cotransfected with a CT/CGRP minigene containing a mutated enhancer 5' splice site (ss) sequence and a U1 snRNA containing a compensating point mutation at its 5' terminus. The utilized minigene and the oligonucleotides used for PCR amplification (arrows) are diagrammed. Inclusion or exclusion of exon 4 is revealed by 282- or 263-nt amplification products, respectively. The wild-type and mutant enhancer sequences are shown (each is depicted as a 5' splice site sequence). The wild-type and mutant U1 RNA sequences are also shown. (B) CHO cells were cotransfected with the wild-type CT/CGRP minigene and no U1 (lane 1), wild-type U1 (lane 2), or mutant U1 (lane 3) or with the mutant CT/CGRP minigene and no U1 (lane 4), wild-type U1 (lane 5), or mutant U1 (lane 6). The percentage of exon 4 inclusion is indicated below each lane.

5' splice site sequence within the enhancer core. Blocking the 5' end of U1 snRNA with an antisense ribo-oligonucleotide abolished in vitro exon 4 polyadenylation, a result that could reflect competitive inhibition of the binding of U1 snRNPs or other factors that bind 5' splice sites.

We sought to more directly test the role of U1 snRNA in enhancer function. To accomplish this goal, we performed an in vivo rescue experiment similar to experiments originally used to establish a requirement for U1 snRNA hybridization in 5' splice site recognition (62). In this assay, a CT/CGRP reporter gene carrying an enhancer core mutated at the +1 position within the 5' splice site sequence was transfected along with a U1 snRNA gene harboring a compensatory mutation to restore base pairing with the mutated enhancer 5' splice site sequence (Fig. 2A). The transfections were performed with Chinese hamster ovary (CHO) cells, which normally include exon 4. Exon 4 inclusion was increased from 8 to 17% by coexpression of the compensatory mutant U1 snRNA but not by wild-type U1 snRNA (Fig. 2B, lanes 4 to 6). This observation suggests that U1 snRNA does hybridize to the core 5' splice site sequence. Increasing the amount of U1 snRNA plasmid used for transfection, however, did not increase exon 4 inclusion to more than 17% (data not shown), indicating only partial rescue of inclusion (Fig. 2B, compare lanes 1 and 6). This experimental limitation suggested that the compensatory U1 snRNA did not completely relieve a requirement for a functional 5' splice site sequence within the enhancer core and indicated a need for additional factors for maximal recognition of the enhancer 5' splice site sequence. The latter conclusion is consistent with studies of other genes showing assembly of complexes including both non-snRNP proteins and U1 snRNPs on real and pseudo-5' splice sites (28, 51).

SRp20 binds to the enhancer core. To identify additional protein factors involved in enhancer recognition, we performed UV cross-linking with site-specifically labeled short ribo-oligonucleotides containing only the core sequence of the intron enhancer element, including the pyrimidine tract and 5' splice site sequence (Fig. 3A). We observed three major cross-linked proteins of approximately 75, 45, and 25 kDa (Fig. 3A). Cross-linking of all three proteins was inhibited by competitor ribo-oligonucleotides containing the pyrimidine tract. Only the 25-kDa protein was competed by a competitor ribo-oligonucleotide containing the 5' splice site sequence.

We suspected that the 25-kDa protein might be SRp20, because of its molecular mass and because of a previous observation that an SR protein bound to the enhancer (33). Immunoprecipitation of UV-cross-linked proteins with an SRp20-specific antibody confirmed that the 25-kDa species was indeed SRp20. For this experiment we used a longer precursor RNA containing both the poly(A) site and the intron enhancer (Fig. 3B) because of the availability of versions of this precursor in which the enhancer core elements had been mutated (33). An anti-SRp20 antibody selectively immunoprecipitated the 25-kDa UV-cross-linked protein bound to the wild-type substrate (Fig. 3B, lane 1). Mutation of either the 5' splice site or pyrimidine tract sequence greatly inhibited cross-linking (Fig. 3B, lanes 2 and 3). SRp20 was also immunoprecipitated with the wild-type short, site-specifically labeled substrate shown in Fig. 3A (data not shown). These results suggest that SRp20 binds to the enhancer core.

We also analyzed the binding of SRp20 to the enhancer by using a gel shift assay. Purified recombinant SRp20 was incubated with wild-type or mutant RNA substrates containing the complete 127-nucleotide (nt) intron 4 enhancer sequence (Fig. 1). SRp20 bound to the wild-type RNA, as demonstrated by the appearance of two complexes (Fig. 3C, lanes 1 to 4). The origin of multiple complexes is unknown. When the 5' splice site sequence in the enhancer core was mutated, complex formation was partially reduced at low protein concentrations but not affected at high protein concentrations (Fig. 3C, lanes 5 to 8). When the pyrimidine tract sequence was mutated, complex production was almost abolished at all protein concentrations (Fig. 3C, lanes 9 to 12). This experiment indicates the direct interaction of SRp20 with the enhancer core sequence and suggests a greater importance for the pyrimidine tract than for the 5' splice site sequence for binding.

Binding of SRp20 to the enhancer core pyrimidine tract sequence is very interesting in light of recent studies by Jumma et al. (26, 27). Those authors suggest that splicing of SRp20 undergoes autoregulation involving binding of the SRp20 protein to the 3' splice site of an alternatively included exon within the SRp20 pre-mRNA. They further identified a region immediately upstream of the polypyrimidine tract that was also required for regulation. To investigate whether flanking upstream sequences are important for SRp20 binding to the CT/CGRP enhancer, we examined the sequence immediately upstream of the enhancer core (herein termed the upstream sequence). We noticed that this sequence is extremely conserved in the human, mouse, and rat CT/CGRP genes (Fig. 1B). Disruption of this sequence by deletion or point mutations decreased exon 4 inclusion in vivo and exon 4 polyadenylation cleavage in vitro (Fig. 4B and C), suggesting an important role for this sequence in the regulation of CT/CGRP alternative RNA processing. However, deletion of this sequence did not affect the binding of SRp20 in the UV-cross-linking-immunoprecipitation assay (Fig. 3B, lane 4). These results suggest that the upstream sequence does not directly bind SRp20; rather, this sequence functions by binding factors

FIG. 3. SRp20 binds to the CT/CGRP enhancer core. (A) UV cross-linking of HeLa cell nuclear extract proteins to a site-specifically labeled ribo-oligonucleotide containing the enhancer core sequence. The sequence of the utilized oligonucleotide (oligo) containing the enhancer core is indicated, with the pyrimidine tract (Py) and 5' splice site sequence (5' ss) (the 5' splice site is underlined) marked. The position of the single introduced labeled phosphate is indicated with an asterisk. UV cross-linking was performed in the presence of no competitor RNA (lane 1), a competitor RNA consisting of the pyrimidine tract of the enhancer core (lane 2), a competitor RNA consisting of the 5' splice site sequence of the enhancer core (lane 3), or a competitor RNA consisting of U3 RNA sequences (lane 4). The molecular masses of cross-linked species are indicated. An arrow marks the position of SRp20. The identities of the high-molecular-mass bands are unknown; the 45-kDa band is probably hnRNP C. (B) Immunoprecipitation of SRp20 UV cross-linked to RNA containing the CT exon 4 poly(A) site with wild-type (wt) or mutated enhancer. The utilized precursor RNA is diagrammed and consisted of all known exon 4 polyadenylation signals and the region downstream of the cleavage site including the enhancer. Substrates contained a wild-type enhancer (lane 1), an enhancer in which the 5' splice site sequence had been altered to CAG/CAAGAC (lane 2), an enhancer in which the pyrimidine tract had been altered to CUACGCGCAUCGUC (lane 3), or an enhancer in which the sequence upstream of the pyrimidine tract was deleted (lane 4) (Fig. 1B and 4A). (C) Gel shift analysis with increasing amounts of recombinant glutathione S-transferase-SRp20 and in vitro-transcribed RNA substrates containing a 127-nt wild-type intron enhancer (lanes 1 to 4), an enhancer in which the 5' splice site sequence had been altered to CAG/CAAGAC (lanes 5 to 8), or an enhancer in which the pyrimidine tract had been altered to CUACGCGCAUCGUC (lanes 9 to 12).

FIG. 4. A sequence upstream of the enhancer core is required for exon 4 inclusion in vivo and exon 4 polyadenylation in vitro. (A) Sequence of the region upstream of the enhancer core. Point mutations introduced into the minigene construct are indicated. The utilized deletion removed the upstream sequence shown by the bracket. (B) RT-PCR analysis of CT/CGRP RNA produced in cells transfected with a CT/CGRP minigene containing the wild-type (lane 1), deleted (lane 2), or point-mutated (lane 3) upstream sequence. The utilized minigene and the oligonucleotides used for PCR amplification (arrows) are diagrammed. Inclusion or exclusion of exon 4 is indicated. (C) In vitro cleavage assay of RNAs containing the CT exon 4 poly(A) site with wild-type or mutated enhancer. The RNA substrate contains the wild-type (lanes 1 to 4), deleted (lanes 5 to 8), or point-mutated (lanes 9 to 12) upstream sequence. Samples were taken for analysis at 1, 20, 40, and 60 min. The precursor and cleaved products are indicated.

yet to be identified. Furthermore, it limits the region of SRp20 binding to the core pyrimidine tract and 5' splice site sequence.

Overexpression of SRp20 in T98G cells increases exon 4 inclusion. To address the functionality of SRp20 binding in enhancer function, we performed in vivo transfections with wild-type or mutant SRp20 cDNA and monitored the effect of increased SRp20 concentration on the processing of a cotransfected CT/CGRP minigene (Fig. 5 and 6). In the first experiment, we transfected T98G cells (a human glioblastoma cell line that preferentially excludes exon 4 [31, 32]) in an attempt to increase exon 4 inclusion. Transfection of wild-type SRp20 increased exon 4 inclusion in T98G cells from 10 to 46% (Fig. 5B, lanes 1 and 2). The increase was dependent on the core 5' splice site sequence, as evidenced by a lesser effect when the core 5' splice site was mutated (Fig. 5B, lanes 3 and 4). Transfection of other SR proteins, including SRp75, SRp55, SRp54, U2AF, and SC35, had no effect on inclusion (Fig. 5C, lane 3, and data not shown), suggesting that the positive effect on enhancer-dependent inclusion was specific to SRp20. Transfection of ASF/SF2, however, also increased exon 4 inclusion, but the increase was independent of the enhancer core 5' splice site sequence (Fig. 5C, lane 2, and data not shown), suggesting that ASF/SF2 can affect exon 4 inclusion by mechanisms other than binding to the enhancer core 5' splice site sequence.

Expression of mutant SRp20 decreases exon 4 inclusion in CHO cells. In the second experiment, we transfected CHO cells that normally include exon 4 with a mutant form of SRp20 in an attempt to depress exon 4 inclusion (Fig. 6A). A mutant SRp20 was generated by truncating the SRp20 cDNA to produce a protein containing the N-terminal RRM but lacking the C-terminal SR domain. Expression of wild-type or mutant protein in CHO cells after transfection is shown in Fig. 6B. The shortened protein was still able to bind to RNA (data not shown). By analogy with the behavior of other SR proteins, the

FIG. 5. In vivo inclusion of CT/CGRP exon 4 is stimulated with wild-type SRp20 in T98G cells. (A) Diagram of the CT/CGRP minigene, RT-PCR oligonucleotides (arrows), and experimental strategy. T98G cells that preferentially exclude exon 4 were cotransfected with a CT/CGRP minigene and wild-type SRp20 to test the ability of the protein to stimulate inclusion. (B) RT-PCR assay of total RNA from transfections of T98G cells with the diagrammed CT/CGRP minigene with a wild-type enhancer (lanes 1 and 2) or an enhancer in which the core 5' splice site (ss) had been mutated (lanes 3 and 4) and wild-type SRp20. The presence or absence of a cotransfected SRp20 expression plasmid is indicated by + or -, respectively. Amplification bands resulting from inclusion (319 nt) or exclusion (280 nt) are indicated. The percentage of inclusion of exon 4 is indicated below each lane. Higher-molecular-weight amplification products in the + lanes result from activation of cryptic splicing within the intron downstream of exon 4 (32). (C) RT-PCR assay of total RNA from transfections of T98G cells with the CT/CGRP minigene with a wild-type enhancer and a vector plasmid control (lane 1) or an expression plasmid for ASF/SF2 (lane 2) or dSRp55 (B52) (lane 3). The SRp20 cDNA used for these experiments was tagged at its N terminus; expression of recombinant SRp20 was monitored by Western blotting with tag-specific antibodies.

truncated protein should not be able to interact with other SR proteins because of its missing SR domain (6, 28, 63). Expression of truncated SRp20, but not wild-type SRp20, depressed exon 4 inclusion in CHO cells from 70 to 38% (Fig. 6C). The observed effect was specific to SRp20, because a similarly truncated SC35 had no effect (data not shown). Depression of inclusion required that the reporter CT/CGRP minigene contain a wild-type enhancer (data not shown).

Addition of a mutant SR protein to cells could potentially alter CT/CGRP processing through a nonspecific effect on general or SR-dependent processing. It should be noted that processing of CT/CGRP via exon 4 exclusion increased following addition of the mutant SRp20, indicating that generic splicing was not affected. The splicing of a control constitutively spliced RNA also was not altered in parallel experiments (data not shown). Most importantly, expression of the mutant SRp20 did not change the splicing phenotype of a control construct derived from the adenovirus late sequence containing two competing 5' splice sites and which will respond to wild-type SR proteins (Fig. 6D, lanes 1 and 3). The 5' splice site selection

FIG. 6. In vivo inclusion of CT/CGRP exon 4 in CHO cells is repressed with a truncated SRp20 lacking its SR domain. (A) Diagram of wild-type SRp20 and a truncated form of SRp20 (SRp20Δ) that contained the RRM but lacked the SR domain. CHO cells that preferentially include exon 4 were cotransfected with a CT/CGRP minigene and SRp20 or SRp20Δ. (B) Expression of proteins from cells transfected with an expression vector coding for SRp20 or SRp20Δ. A Western blot with tag-specific antibodies is shown to document production of the desired SRp20 forms following transfection. (C) RT-PCR assay of total RNA from transfections of CHO cells with the diagrammed CT/CGRP minigene with a wild-type enhancer and a vector plasmid control (lane 1), wild-type SRp20 (lane 2), or truncated SRp20 missing its SR domain (SRp20Δ) (lane 3). Products and percentages of inclusion are indicated as in Fig. 5B. (D) RT-PCR assay of total RNA from transfections of HeLa cells with the diagrammed control construct containing duplicated 5' splice sites (ss) and vector plasmid (lane 1), SC35 (lane 2), or SRp20Δ (lane 3). The products for using either 5' splice site are indicated. The construct contains two identical 5' splice sites derived from the adenovirus major late transcription unit exon 2. The SRp20 cDNA used for these experiments was tagged at its N terminus; expression of recombinant SRp20 was monitored by Western blotting with tag-specific antibodies.

of this substrate can be regulated by varying SR protein levels. For example, coexpression of the SR protein SC35 increased usage of the proximal 5' splice site (Fig. 6D, lane 2). These results indicate that the truncated SRp20 disrupted neither general splicing nor an SR-sensitive splicing phenotype. When other SR proteins have been tested for their ability to regulate 5' splice site usage after truncation to remove the SR domain, ASF/SF2 has been observed to still retain switching activity, but SC35 did not (7, 56). We do not yet know why different SR proteins behave differently in this assay, but we note that ASF/SF2 has two RRM domains, whereas SRp20 and SC35 have only one. Regardless, the inability of the truncated SRp20 to alter alternative splicing but still affect CGRP processing suggests that the effect of the mutant SRp20 on pre-mRNA processing of the CT/CGRP minigene is specific to CT/CGRP enhancer function and not a general inhibition of constitutive splicing factors or general SR proteins.

Mutant SRp20 disrupts binding of the CstF 64-kDa subunit to the CT exon 4 poly(A) site. Although the above-described experiments indicated that SRp20 functionally binds to the enhancer core and facilitates exon 4 inclusion, they did not indicate whether the bound SRp20 affected exon 4 splicing or polyadenylation. To address this question, we turned to the in

FIG. 7. Association of a polyadenylation factor with the CT/CGRP exon 4 poly(A) site in vitro is depressed by a mutant form of SRp20 lacking the SR domain. In vitro polyadenylation extracts containing either wild-type or mutant SRp20 were created to test the ability of SRp20 to modulate association of polyadenylation factors with the exon 4 poly(A) site (A). The factor monitored was the 64-kDa subunit of CstF, which was detected by UV cross-linking and immunoprecipitation of cross-linked protein (B). Association of this factor with the exon 4 poly(A) site has previously been shown to be enhancer dependent (33). (A) Detection of CstF 64-kDa protein and SRp20 in extracts by Western blotting. To make extracts containing wild-type or mutant SRp20, HeLa cells were transfected with an expression plasmid for wild-type or truncated SRp20 (SRp20Δ) that had been His tagged to create protein products with unique molecular masses. Standard polyadenylation extracts were prepared from these transfected cells (15). Both extracts were monitored for levels of the 64-kDa subunit of CstF (lanes 1 and 2) and endogenous and exogenous SRp20 (lanes 3 and 4) by Western blotting. Exogenous forms of SRp20 can be distinguished from endogenous forms by their molecular masses. (B) UV cross-linking of CstF 64-kDa protein to the CT exon 4 poly(A) site. The polyadenylation substrate shown in Fig. 3B was used as a substrate for UV cross-linking and immunoprecipitation of the 64-kDa protein in extracts from cells expressing wild-type SRp20, truncated SRp20 (SRp20Δ), or a control LacZ protein (lanes 1 to 4). A control poly(A) site derived from the CHO adenine phosphoribosyltransferase gene (32) was also used for cross-linking to monitor potential nonspecific effects of expression of the truncated SRp20 (lanes 5 and 6).

in vitro polyadenylation system and an examination of the association of polyadenylation factors with the exon 4 poly(A) site. This assay was based on a previous observation that a wild-type enhancer was required for maximal binding of CstF to the exon 4 poly(A) site (33). HeLa cells were transfected with either wild-type or truncated SRp20. Nuclear extract was prepared from these transfected cells. Levels of SRp20 (both endogenous and recombinant) were detected by Western blotting with the anti-SRp20 antibody. Transfections resulted in levels of recombinant SRp20 in the extract that were in excess of the levels of endogenous SRp20 (Fig. 7A).

These nuclear extracts were used to assay the binding of polyadenylation factors to the exon 4 poly(A) site. We monitored binding of the 64-kDa subunit of CstF, a general indicator for polyadenylation activity (36, 52), by UV cross-linking. Equal amounts of the 64-kDa protein were detected in the nuclear extracts isolated from HeLa cells transfected with either the wild-type or mutant form of SRp20 (Fig. 7A). Compared to extract containing wild-type SRp20, extract containing truncated SRp20 supported less UV cross-linking of the 64-kDa protein to the exon 4 poly(A) site (Fig. 7B, lanes 1 and 2), indicating that increased levels of a mutant form of SRp20 prevented normal binding of the 64-kDa polyadenylation factor to the exon 4 poly(A) site. No effect was observed with a control, enhancer-independent poly(A) site, the adenine phos-

phoribosyltransferase poly(A) site (Fig. 7B, lanes 5 and 6), suggesting that the inhibitory effect of the mutant form of SRp20 was specific to the exon 4 poly(A) site.

The observed effect reflected a depression of standard levels of binding of CstF, because extract from cells overexpressing a control protein, LacZ, and extract from cells overexpressing wild-type SRp20 gave indistinguishable levels of 64-kDa protein cross-linking (Fig. 7B, lanes 3 and 4). Therefore, increasing the level of SRp20 in a HeLa extract does not boost CstF binding. This result agrees with our observed inability to enhance exon 4 inclusion in HeLa cells by transfection with wild-type SRp20 and suggests that SRp20 levels are not rate limiting for either exon 4 polyadenylation or exon 4 inclusion in HeLa cells or CHO cells. Although it would be extremely informative to do a similar experiment using T98G cells to examine whether overexpression of wild-type SRp20 would result in enhanced CstF 64-kDa subunit binding, we have not been able to transfect T98G cells with high efficiency. It is also not presently possible to prepare active nuclear extract for an in vitro polyadenylation cleavage reaction from these cells. Inhibition of the binding of polyadenylation factors in the presence of a truncated form of SRp20, however, strongly indicates that the binding of this mutant form of the protein to the enhancer negates enhancer-mediated polyadenylation. Therefore, we conclude that SRp20 facilitates exon 4 polyadenylation by a direct or indirect effect on the initial binding of polyadenylation factors to the poly(A) site.

DISCUSSION

Our results demonstrate that the SR protein SRp20 binds to splicing signals within the CT/CGRP polyadenylation enhancer and increases exon 4 inclusion in vivo in cell lines that preferentially skip the exon. Furthermore, we show that expression of a mutant form of SRp20 lacking its SR domain but still able to bind to RNA causes reduced binding of a polyadenylation factor to the exon 4 poly(A) site, suggesting that SRp20 stimulates exon 4 inclusion by an effect on polyadenylation. This is the first indication that a canonical SR protein can influence polyadenylation and suggests that other SR proteins could play a general role in recognition of 3'-terminal exons. The idea that SR proteins provide a link between the splicing and polyadenylation machineries is very attractive in light of the recently discovered presence of an arginine-rich domain containing multiple SR dipeptides in one of the subunits of the polyadenylation factor CFI (49).

We do not know if the binding of SRp20 to the enhancer has effects on splicing of exon 4 as well as effects on polyadenylation cleavage. Transfection of cells with SRp20 cDNAs can alter 5' splice site usage (50), although natural target genes for SRp20-mediated regulation of splicing have not been reported. Therefore, it seems possible that SRp20 affects both the splicing and polyadenylation of CT/CGRP exon 4.

One possible model for SRp20 function during polyadenylation of exon 4 (Fig. 8A) invokes interactions between enhancer-bound SRp20 and poly(A) site-bound CFI and resembles previous models for SR protein-mediated enhancement of binding of U2AF to the 3' splice site during splicing. Like U2AF, CFI is one of the earliest factors to associate with the site to be processed, suggesting that SRp20 could activate early events in the recognition of the CT/CGRP exon 4 poly(A) site. Indeed, we have observed that both the wild-type enhancer and SRp20 are needed for maximal association of CstF with the poly(A) site.

Alternatively, SRp20 may not directly interact with poly(A) factors (Fig. 8B). Instead, binding of SRp20 might stabilize the

FIG. 8. Models for how the CT intron enhancer facilitates polyadenylation of exon 4. (A) In model 1, SRp20 is directly involved in interacting with poly(A) factors. (B) In model 2, SRp20 is indirectly involved in enhancing polyadenylation of exon 4 by stabilizing U1 snRNP interaction with the enhancer core 5' splice site sequence.

interaction of U1 snRNPs to the enhancer 5' splice site sequence. The direct link to polyadenylation would be provided by U1 snRNP proteins. U1 snRNPs or snRNP proteins, either free or in snRNPs, have been implicated in both constitutive and regulated polyadenylation (2, 3, 19, 22–24, 34, 35). U1A protein has been reported to interact positively with the cleavage-polyadenylation specificity factor 160-kDa protein and negatively with poly(A) polymerase (22, 23, 35). Inhibition of polyadenylation by interaction between poly(A) polymerase and the U1 70K protein of a U1 snRNP bound to a 5' splice site upstream of a papillomavirus poly(A) site has also been reported (19, 24). Given the observation of both positive and negative control of polyadenylation by U1 snRNP proteins, interactions between SR proteins and U1 snRNPs could have a variety of outcomes for the polyadenylation of target sites.

Our previous results suggest that PTB also binds to the enhancer core pyrimidine sequence and enhances exon 4 polyadenylation (33). In this study, we demonstrate binding of an additional protein, SRp20, to the same core sequence. If both PTB and SRp20 are capable of binding to the core pyrimidine sequence, they could bind simultaneously or sequentially. It is also possible that they antagonize each other for binding. We previously demonstrated involvement of PTB and ASF/SF2 in CT enhancer recognition by using binding studies (33). The functional relevance of this binding remained unclear. In this report, we provide evidence that at least two of the identified factors, U1 snRNA and SRp20, are required for CT exon 4 inclusion by enhancing exon 4 polyadenylation. Further studies are necessary to understand the individual role of each of the identified factors.

An important question for CT/CGRP alternative RNA processing is how exon 4 exclusion is produced in CGRP-producing cells. Two possibilities exist for this regulation: lack of a positive factor(s) or presence of a negative factor. To date, we have been unable to observe any differences in the factors known to be important for CT/CGRP alternative processing between the two model cell lines used to recapitulate this processing decision. Western blot analysis detected similar levels of SRp20 protein in HeLa and T98G cells (data not shown). This result does not rule out a role for SRp20 in neuronal exclusion of exon 4. Our observation of increased exon 4 inclusion in T98G cells following overexpression of SRp20 suggests several possible ways in which SRp20 could participate in

exclusion. First, SRp20 protein modification, for example, phosphorylation, could be different in neuronal cells through tissue-specific regulation of either the phosphorylation event or the required kinase. Second, SRp20 may be inactivated by binding to a cell-specific factor, thereby losing its ability to bind to either its RNA target or other required protein factors. Third, a cell-specific factor that binds the enhancer sequence and inhibits the positive functions of the enhancer could exist. In transgenic mice expressing the rat CT/CGRP gene, exon 4 inclusion occurs in multiple tissues, with the exon 4 exclusion limited to a few tissues, including neurons and heart (13). This result suggests the presence of a tissue-specific factor(s) as regulator of CT/CGRP alternative RNA processing.

The last possibility is the presence of an alternative form(s) of SRp20 in cells that exclude exon 4. Recently, an alternate form of SRp20 has been identified in mouse cells (26, 27). This protein is a naturally occurring truncated SRp20 that contains the RNA-binding domain but lacks the SR domain, almost identical to the mutant protein created for this study. The truncated SRp20 protein is produced by inclusion in the final RNA of an alternatively spliced exon containing an in-frame translational stop codon. In vivo, the truncated form accumulates in tissue culture cells in a resting state (G_0 cells) (26). Interestingly, equal amounts of the two forms of SRp20 were detected in whole brain (4), the natural site for CGRP production. This observation, coupled with results reported in this study, suggests that the truncated form of SRp20 may exert a natural negative effect on CT/CGRP pre-mRNA processing in selective cell types in neuronal or cardiac cells that produce CGRP. Our results with T98G cells, however, indicate that mRNA lacking exon 4 can be produced in a cell type not naturally producing the truncated form of SRp20 (data not shown), suggesting the involvement of additional factors in exon 4 exclusion in these cells.

The postulated ability of SRp20 to regulate recognition of a 3' splice site within its own mRNA also raises the possibility that SRp20 participates in CT/CGRP regulation by effects on splicing as well as polyadenylation. The regulated 3' splice site within the SRp20 gene bears little sequence resemblance to that bordering CT/CGRP exon 4, suggesting a different binding site should SRp20 affect splicing of exon 4. Given the number of regulatory sequence elements that have been observed in the CT/CGRP gene, however, there may be additional roles for SRp20 in CT/CGRP alternative RNA processing other than those reported here.

Recognition of 3'-terminal exons has been postulated to involve an interaction of splicing factors and polyadenylation factors (11, 21–23, 45, 46). The CT/CGRP gene provides a unique model system to study the nature of the relationship between polyadenylation and splicing. In this report, we provide evidence that an SR protein bound to an intron enhancer element containing splice sites can facilitate inclusion of an alternatively included 3'-terminal exon. The link between splicing and polyadenylation by an SR protein is also reminiscent of the emerging concept of RNA processing at transcription units (reviewed in reference 44). Several RNA-processing factors, including cap-binding complex and polyadenylation factors have been shown to associate with RNA polymerase II through its carboxy-terminal domain after transcription initiation (5, 9, 14, 38, 39, 41). The carboxy-terminal domain also binds a unique set of proteins containing SR domains (5, 41, 58), leading to suggestions that SR proteins play a role in the communication between transcription and RNA processing. Thus, SR proteins may be involved in multiple interactions to coordinate individual recognition events occurring during the early steps of RNA processing.

ACKNOWLEDGMENTS

We thank Yun Yang for her assistance in DNA preparation and sequencing. We thank Alan Weiner, Clinton MacDonald, James Manley, and Mark Roth for providing the U1 snRNA clone, the CstF 64-kDa subunit antibody, the ASF expression clone, and the dSRp55 expression clone, respectively. We acknowledge the helpful advice of members of the Gagel and Berget laboratories, specifically Andrew McCullough and Leslie Elrick for help with nuclear extract preparation.

This work was supported by an ACS grant to S.M.B. and USPHS grants (RO1-DK38146 to R.F.G. and 2P30-CA16672) to the M.D. Anderson Cancer Center. K.M.N. was supported by NIH grant GM488435 to Mark Roth.

REFERENCES

- Amara, S. G., V. Jonas, and M. G. Rosenfeld. 1982. Alternative RNA processing in calcitonin gene expression generates mRNAs encoding different polypeptide products. *Nature* **298**:240–244.
- Ashe, M. P., P. Griffin, W. James, and N. J. Proudfoot. 1995. Poly(A) site selection in the HIV-1 provirus: inhibition of promoter-proximal polyadenylation by the downstream major splice donor site. *Genes Dev.* **9**:3008–3025.
- Ashe, M. P., L. H. Pearson, and N. J. Proudfoot. 1997. The HIV-1 5' LTR poly(A) site is inactivated by U1 snRNP interaction with the downstream major splice donor site. *EMBO J.* **16**:5752–5763.
- Ayane, M., U. Preuss, G. Kohler, and P. J. Nielsen. 1991. A differentially expressed murine RNA encoding a protein with similarities to two types of nucleic acid binding motifs. *Nucleic Acids Res.* **19**:1273–1278.
- Blencowe, B. J., R. Issner, J. Kim, P. McCaw, and P. A. Sharp. 1995. New proteins related to the Ser-Arg family of splicing factors. *RNA* **1**:852–865.
- Caceres, J. F., and A. R. Krainer. 1993. Functional analysis of pre-mRNA splicing factor SF2/ASF structural domains. *EMBO J.* **12**:4715–4726.
- Caceres, J. F., T. Misteli, G. R. Sreaton, D. L. Spector, and A. R. Krainer. 1997. Role of the modular domains of SR proteins in subnuclear localization and alternative splicing specificity. *J. Cell Biol.* **138**:225–238.
- Cavaloc, Y., M. Popielarz, J. P. Fuchs, R. Gattoni, and J. Stevenin. 1994. Characterization and cloning of the human splicing factor 9G8: a novel 35 kDa factor of the serine/arginine protein family. *EMBO J.* **13**:2639–2649.
- Cho, E.-J., T. Takagi, C. R. Moore, and S. Buratowski. 1997. mRNA capping enzyme is recruited to the transcriptional complex by phosphorylation of the RNA polymerase II carboxy-terminal domain. *Genes Dev.* **11**:3319–3326.
- Colgan, D. F., and J. L. Manley. 1997. Mechanism and regulation of mRNA polyadenylation. *Genes Dev.* **11**:2755–2766.
- Cooke, C., and J. Alwine. 1996. The cap and 3' splice site similarly affect polyadenylation efficiency. *Mol. Cell. Biol.* **16**:2579–2584.
- Cote, G. J., D. T. Stelow, S. Peleg, S. M. Berget, and R. F. Gagel. 1992. Identification of exon sequences and an exon binding protein involved in alternative RNA splicing of calcitonin/CGRP. *Nucleic Acids Res.* **20**:2361–2366.
- Crenshaw, E. B., A. F. Russo, L. W. Swanson, and M. G. Rosenfeld. 1987. Neuron-specific alternative RNA processing in transgenic mice expressing a metallothionein-calcitonin fusion gene. *Cell* **49**:389–398.
- Dantonel, J.-C., K. G. K. Murthy, J. L. Manley, and L. Tora. 1997. Transcription factor TFIID recruits factor CPSF for formation of 3' ends of mRNA. *Nature* **389**:399–402.
- Dignam, J. D., R. M. Lebovitz, and R. G. Roeder. 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res.* **11**:1475–1489.
- Edwards-Gilbert, G., and C. Milcarek. 1995. Regulation of poly(A) site use during mouse B-cell development involves a change in the binding of a general polyadenylation factor in a B-cell stage-specific manner. *Mol. Cell. Biol.* **15**:6420–6429.
- Edwards-Gilbert, G., K. L. Veraldi, and C. Milcarek. 1997. Alternative poly(A) site selection in complex transcription units: means to an end? *Nucleic Acids Res.* **25**:2547–2561.
- Fu, X. D. 1995. The superfamily of arginine/serine-rich splicing factors. *RNA* **1**:663–680.
- Furth, P., W. Choe, J. Rex, J. C. Byrne, and C. Baker. 1994. Sequences homologous to 5' splice sites are required for the inhibitory activity of papillomavirus late 3' untranslated regions. *Mol. Cell. Biol.* **14**:5278–5289.
- Ge, H., and J. L. Manley. 1990. A protein factor, ASF, controls cell-specific alternative splicing of SV40 early pre-mRNA in vitro. *Cell* **62**:25–34.
- Gilmartin, G. M., M. A. McDevitt, and J. R. Nevins. 1988. Multiple factors are required for specific RNA cleavage at a poly(A) addition site. *Genes Dev.* **2**:578–587.
- Gunderson, S. I., K. Beyer, G. Martin, Keller, W. C. Boelens, and I. W. Mattaj. 1994. The human U1A snRNP protein regulates polyadenylation via a direct interaction with poly(A) polymerase. *Cell* **11**:531–541.
- Gunderson, S. I., S. Vagner, M. Polycarpou-Schwarz, and I. W. Mattaj. 1997. Involvement of the carboxyl terminus of vertebrate poly(A) polymerase in U1A autoregulation and in the coupling of splicing and polyadenylation. *Genes Dev.* **11**:761–773.
- Gunderson, S. I., M. Polycarpou-Schwarz, and I. W. Mattaj. 1998. U1 snRNP inhibits pre-mRNA polyadenylation through a direct interaction between U1 70K and poly(A) polymerase. *Mol. Cell* **1**:255–264.
- Hedjran, F., J. M. Yeakley, G. S. Huh, R. O. Hynes, and M. G. Rosenfeld. 1997. Control of alternative pre-mRNA splicing by distributed pentameric repeats. *Proc. Natl. Acad. Sci. USA* **94**:12343–12347.
- Jumaa, H., J.-L. Guenet, and P. J. Nielsen. 1997. Regulated expression and RNA processing of transcripts from the SRp20 splicing factor gene during the cell cycle. *Mol. Cell. Biol.* **17**:3116–3124.
- Jumaa, H., and P. J. Nielsen. 1997. The splicing factor SRp20 modifies splicing of its own mRNA and ASF/SF2 antagonizes this regulation. *EMBO J.* **16**:5077–5085.
- Kohtz, J. D., S. F. Jamison, C. L. Will, P. Zuo, R. Luhrmann, M. A. Garcia-Blanco, and J. L. Manley. 1994. Protein-protein interactions and 5' splice site recognition in mammalian mRNA precursors. *Nature* **368**:119–124.
- Krainer, A. R., G. C. Conway, and D. Kozak. 1990. Purification and characterization of pre-mRNA splicing factor SF2 from HeLa cells. *Genes Dev.* **4**:1158–1171.
- Krainer, A. R., G. C. Conway, and D. Kozak. 1990. The essential pre-mRNA splicing factor SF2 influences 5' splice site selection by activating proximal sites. *Cell* **62**:35–42.
- Lou, H., G. J. Cote, and R. F. Gagel. 1994. The calcitonin exon and its flanking intronic sequences are sufficient for the regulation of human calcitonin/calcitonin gene-related peptide alternative splicing. *Mol. Endocrinol.* **8**:1618–1626.
- Lou, H., Y. Yang, G. J. Cote, S. M. Berget, and R. F. Gagel. 1995. An intron splicing enhancer containing a 5' splice site sequence in the human calcitonin/calcitonin gene-related peptide gene. *Mol. Cell. Biol.* **15**:3979–3988.
- Lou, H., R. F. Gagel, and S. M. Berget. 1996. An intron enhancer recognized by splicing factors activates polyadenylation. *Genes Dev.* **1**:208–219.
- Lutz, C. S., and J. C. Alwine. 1994. Direct interaction of the U1 snRNP-A protein with the upstream efficiency element of the SV40 late polyadenylation signal. *Genes Dev.* **8**:576–586.
- Lutz, C. S., K. G. K. Murthy, N. Schek, J. P. O'Connor, J. L. Manley, and J. C. Alwine. 1996. Interaction between the U1 snRNP-A protein and the 160-kD subunit of cleavage-polyadenylation specificity factor increases polyadenylation efficiency in vitro. *Genes Dev.* **10**:325–337.
- MacDonald, C. C., J. Wilusz, and T. Shenk. 1994. The 64-kilodalton subunit of the CstF polyadenylation factor binds to pre-mRNAs downstream of the cleavage site and influences cleavage site location. *Mol. Cell. Biol.* **14**:6647–6654.
- Manley, J. L., and R. Tacke. 1996. SR proteins and splicing control. *Genes Dev.* **10**:1569–1579.
- McCracken, S., N. Fong, K. Yankulov, S. Ballantyne, G. Pan, J. Greenblatt, S. D. Patterson, M. Wickens, and D. L. Bentley. 1997. The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. *Nature* **385**:357–361.
- McCracken, S., N. Fong, E. Rosonia, K. Yankulov, G. Borthers, D. Siderovski, A. Hessel, S. Foster, Amgen EST Program, S. Shuman, and D. L. Bentley. 1997. 5'-capping enzymes are targeted to pre-mRNA by binding to the phosphorylated carboxy-terminal domain of RNA polymerase II. *Genes Dev.* **11**:3306–3318.
- Moore, M. J., and P. A. Sharp. 1993. Evidence for two active sites in the spliceosome provided by stereochemistry of pre-mRNA splicing. *Nature* **365**:364–368.
- Mortillaro, M. J., B. J. Blencowe, X. Wei, H. Nakayasu, L. Du, S. L. Warren, P. A. Sharp, and R. Berezney. 1996. A hyperphosphorylated form of the large subunit of RNA polymerase II is associated with splicing complexes and the nuclear matrix. *Proc. Natl. Acad. Sci. USA* **93**:8253–8257.
- Neugebauer, K. M., J. A. Stolk, and M. B. Roth. 1995. A conserved epitope on a subset of SR proteins defines a larger family of pre-mRNA splicing factors. *J. Cell Biol.* **129**:899–908.
- Neugebauer, K. M., and M. B. Roth. 1997. Distribution of pre-mRNA splicing factors at sites of RNA polymerase II transcription. *Genes Dev.* **11**:1148–1159.
- Neugebauer, K. M., and M. B. Roth. 1997. Transcription units as RNA processing units. *Genes Dev.* **11**:3279–3285.
- Niwa, M., S. D. Rose, and S. M. Berget. 1990. Polyadenylation is stimulated by the presence of an upstream intron. *Genes Dev.* **4**:1552–1559.
- Niwa, M., and S. M. Berget. 1991. Mutation of the AAUAAA polyadenylation signal depresses in vitro splicing of proximal but not distal introns. *Genes Dev.* **5**:2086–2095.
- Reed, R. 1996. Initial splice-site recognition and pairing during pre-mRNA splicing. *Curr. Opin. Genet. Dev.* **6**:215–220.
- Rosenfeld, M. G., S. G. Amara, and R. M. Evans. 1984. Alternative RNA processing: determining neuronal phenotype. *Science* **225**:1315–1320.
- Rüeggesser, U., D. Blank, and W. Keller. 1998. Human pre-mRNA cleavage factor Im is related to spliceosomal SR proteins and can be reconstituted in vitro from recombinant subunits. *Mol. Cell* **1**:243–253.
- Sreaton, G. R., J. F. Caceres, A. Mayeda, M. V. Bell, M. Plebanski, D. G.

- Jackson, J. I. Bell, and A. R. Krainer. 1995. Identification and characterization of three members of the human SR family of pre-mRNA splicing factors. *EMBO J.* **14**:4336–4349.
51. Siebel, C. W., L. D. Fresco, and D. C. Rio. 1992. The mechanism of somatic inhibition: multiprotein complexes at an exon pseudo-5' splice site control U1 snRNP binding. *Genes Dev.* **6**:1386–401.
52. Takagaki, Y., R. L. Seipelt, M. L. Peterson, and J. L. Manley. 1996. The polyadenylation factor CstF-64 regulates alternative processing of IgM heavy chain pre-mRNA during B cell differentiation. *Cell* **87**:941–952.
53. Valcrel, J., and M. R. Green. 1996. The SR protein family: pleiotropic functions in pre-mRNA splicing. *Trends Biochem. Sci.* **21**:296–301.
54. van Oers, C. C. M., G. J. Adema, H. Zandberg, Y. C. Moen, and P. D. Baas. 1994. Two different sequence elements within exon 4 are necessary for calcitonin-specific splicing of the human calcitonin/calcitonin gene-related peptide I pre-mRNA. *Mol. Cell. Biol.* **14**:951–960.
55. Wahle, E., and W. Keller. 1996. The biochemistry of polyadenylation. *Trends Biochem. Sci.* **21**:247–250.
56. Wang, J., and J. L. Manley. 1995. Over expression of the SR proteins ASF/SF2 and SC35 influences alternative splicing in vivo in diverse ways. *RNA* **1**:335–346.
57. Wu, J. Y., and T. Maniatis. 1993. Specific interactions between proteins implicated in splice site selection and regulated alternative splicing. *Cell* **75**:1061–1070.
58. Yuryev, A., M. Patturajan, Y. Litingtung, R. V. Joshi, C. Gnetile, M. Gebara, and J. L. Cordon. 1996. The C-terminal domain of the largest subunit of RNA polymerase II interacts with a novel set of serine/arginine proteins. *Proc. Natl. Acad. Sci. USA* **93**:6975–6980.
59. Zahler, A. M., W. S. Lane, J. A. Stolk, and M. B. Roth. 1992. SR proteins: a conserved family of pre-mRNA splicing factors. *Genes Dev.* **6**:837–847.
60. Zahler, A. M., K. M. Neugebauer, W. S. Lane, and M. B. Roth. 1993. Distinct functions of SR proteins in alternative pre-mRNA splicing. *Science* **260**:219–222.
61. Zhang, W. J., and J. Y. Wu. 1996. Functional properties of p54, a novel SR protein active in constitutive and alternative splicing. *Mol. Cell. Biol.* **16**:5400–5408.
62. Zhuang, Y., and A. Weiner. 1986. A compensatory base change in U1 snRNA suppresses a 5' splice site mutation. *Cell* **46**:827–835.
63. Zuo, P., and J. L. Manley. 1994. The human splicing factor ASF/SF2 can specifically recognize pre-mRNA 5' splice sites. *Proc. Natl. Acad. Sci. USA* **91**:3363–3367.