Members of the Heterogeneous Nuclear Ribonucleoprotein H Family Activate Splicing of an HIV-1 Splicing Substrate by Promoting Formation of ATP-dependent Spliceosomal Complexes

Received for publication, January 26, 2007, and in revised form, March 1, 2007 Published, JBC Papers in Press, March 2, 2007, DOI 10.1074/jbc.M700774200

Michael C. Schaub, Suzette R. Lopez, and Massimo Caputi1
From the Department of Biomedical Science, Florida Atlantic University, Boca Raton, Florida 33431

In this study we analyzed members of the heterogeneous nuclear ribonucleoprotein (hnRNP) H protein family to determine their RNA binding specificities and roles in splicing regulation. Our data indicate that hnRNPs H, H*, F, 2H9, and GRSF-1 bind the consensus motif DGGGD (where D is U, G, or A) and aggregate in a multimeric complex. We analyzed the role of these proteins in the splicing of a substrate derived from the HIV-1 tat gene and have shown that hnRNP H family members are required for efficient splicing of this substrate. The hnRNP H protein family members activated splicing of the viral substrate by promoting the formation of ATP-dependent spliceosomal complexes. Mutational analysis of six consensus motifs present within the intron of the substrate indicated that only one of these motifs acts as an intronic splicing enhancer.

Alternative splicing is the process by which exons from a primary transcript (pre-mRNA) can be spliced in different arrangements to yield mRNAs that will produce functionally different protein variants. The sequencing of the human genome has revealed that up to 75% of multie exon genes are alternatively spliced (1). A clear understanding of the mechanisms regulating splicing is key to our comprehension of the complex regulation of the eukaryotic genome and can provide us with new diagnostic and therapeutic tools.

The spliceosome is a ribonucleoprotein complex that removes non-coding intervening sequences called introns from precursor mRNA. Within the intron, a 3’ splice site, 5’ splice site, and branch site are required for splicing. These are short, loosely conserved sequences that alone are not sufficient for the proper recognition of exonic and intronic sequences and regulation of splicing. Additional regulatory elements are classified according to their location and function, because either exonic and intronic splicing enhancers (ESEs and ISEs)2 or exonic and intronic splicing silencers (ESSs and ISSs). These regulatory sequences can interact with factors that promote proper recognition of the splice sites and regulate splicing in response to physiological stimuli. Among the best characterized ESEs are purine-rich sequences that recruit members of the serine/arginine-rich (SR) family of splicing activators (2). The best known ESSs are dependent on interaction with members of the heterogeneous ribonucleoprotein A/B family (hnRNPs A/B). Dissection of the mechanisms regulating alternative splicing in several genes has shown that positive and negative cis-acting sequences are organized in multipartite control elements where SR proteins and hnRNPs often play counteracting roles (3–5).

hnRNPs H, H*, F, 2H9, and GRSF-1 constitute a family of highly homologous, ubiquitously expressed proteins that have been implicated in splicing, polyadenylation, capping, export, and translation of cellular and viral mRNAs. hnRNPs H, H*, 2H9, and F have been shown to interact with both splicing enhancers and silencers. The binding of these proteins to intronic sequences activates neural-specific splicing of the c-src mRNA (6,7), elicits use of the TR62-specific 5’ splice site of the thyroid hormone receptor (8) and the intron 2.3’ splice site of the THPO gene (9), and promotes inclusion of the CI exon of the N-methyl-D-aspartic acid-type glutamate R1 receptor (GRIN1) (5) and the HIV-1 exon 6D (10). Binding of hnRNPs H and F to a regulatory element within the alternatively spliced coding exon 1 of the bcl-xs gene promotes splicing to the bcl-xxspecific 5’ splice site (11). hnRNP H has also been shown to inhibit both splicing of the rat β-tropomyosin gene, by binding an exonic silencer (12), and Rous sarcoma virus splicing by binding an intronic element (13).

Little is known of the mechanisms by which hnRNP H family members act in splicing silencing or activation, and it is not understood whether the basal splicing machinery requires such proteins to function (14). In this work, we identified the minimal RNA motif recognized by hnRNPs H, H*, F, 2H9, and GRSF-1 as the consensus sequence DGGGD. To evaluate whether the consensus DGGGD binding motif affects splicing independently from its location in respect to the splice sites and other splicing control elements, we systematically mutated the consensus motifs of a splicing substrate derived from the HIV-1 tat gene. Surprisingly, only two out of eight motifs appeared to

nitrilotriacetic acid; RAC, RNA-affinity chromatography; snRNP, small nuclear ribonucleoprotein.

1 To whom correspondence should be addressed: Dept. of Biomedical Science, Florida Atlantic University, 777 Glades Rd., Boca Raton, FL 33431. Tel.: 561-297-0627; Fax: 561-297-2221; E-mail: mcaputi@fau.edu.
2 The abbreviations used are: ESE, exonic splicing enhancer; ISE, intronic splicing enhancer; ESS, exonic splicing silencer; ISS, intronic splicing silencer; SR family, serine/arginine-rich family; hnRNP, heterogeneous nuclear ribonucleoprotein; HIV-1, human immunodeficiency virus, type 1; Ni-NTA, nickel nitrilotriacetic acid; RAC, RNA-affinity chromatography; snRNP, small nuclear ribonucleoprotein.
Splicing Regulation by hnRNP H Family Members

be regulating splicing, suggesting that interaction with other regulatory elements and/or proper positioning within a higher order RNA structure may be required for splicing control by hnRNPs of the H family. Furthermore, in vitro spaying assays and analysis of spliceosome complex formation, following depletion add-back of individual hnRNP H proteins, indicated that members of this family play a key role in ATP-dependent spliceosome formation.

EXPERIMENTAL PROCEDURES

Plasmid Construction and Protein Preparation—Splicing reporter plasmids were derived from the plasmids pTat12 and pTat12ESSMut previously described as pH51-X and pΔESS10, respectively (15). Constructs carrying the mutations of the G-runs were obtained by PCR site-directed mutagenesis. Construct pTat12ESSMut-5’M was obtained by mutating the conserved GT dinucleotide of the pTat12ESSMut 5’ splice site by site-directed mutagenesis. The hnRNP H, hnRNP F, and hnRNP 2H9 cDNAs were cloned into the expression vector PrsetB (Invitrogen). The recombinant N-terminal histidine-tagged proteins were expressed in Escherichia coli BL21(DE3). The bacterial cells were grown to an optical density of 0.6 at 600 nm before induction with 1 mM isopropyl-β-D-thiogalactopyranoside for 3 h. Induced cells were pelleted by centrifugation and washed with 50 mM Tris-HCl (pH 8.0) containing 0.1M dithiothreitol, 4 mM ATP, 4 mM MgCl₂, 5 mM creatine phosphate) were incubated with the RNA-bound beads to deplete >80% of hnRNP H, H’, F, 2H9, and GRSF-1.

Pull-down Assays—100 μl of Ni-NTA-agarose beads were incubate with variable amounts of histidine-tagged hnRNP F and H for 30 min at 4 °C in 250 μl of incubation buffer (20 mM HEPES-KOH, pH 7.9, 5% glycerol, 0.1 M KCl, 0.2 mM EDTA, 0.5 mM dithiothreitol, 4 mM ATP, 4 mM MgCl₂). The Ni-NTA beads were washed of unbound hnRNPs and incubated for 2 h at 4 °C with 250 μl of HeLa nuclear extracts in incubation buffer. Ni-NTA-agarose beads were then washed four times, and the proteins were eluted in 2% SDS, separated on polyacrylamide SDS gels, electroblotted, and probed with specific antibodies.

Spliceosome Assembly Reactions—Spliceosome assembly reactions were carried out as previously described (18). For assembly of the E complex, pre-mRNAs were incubated in 20-μl splicing reactions lacking ATP and MgCl₂. For assembly of the ATP-dependent complexes pre-mRNAs were incubated in nuclear extracts under splicing conditions. Where indicated, 4 μl of 4 mg/ml heparin was added prior to loading on gels. Reactions were loaded on 1.5% low melting point agarose (Invitrogen) gels. The running buffer was 50 mM Tris and 50 mM glycine. Gels were dried down under vacuum at room temperature.

RESULTS

Sequence analysis showed that hnRNP H, H’, F, 2H9, and GRSF-1 are closely related and that the identity between any two members of the family is between 44 and 96% (Fig. 1). We have previously shown that hnRNPs H, H’, F, and 2H9 assemble on cellular and viral substrates containing the sequence GGGA (19). This result was achieved by incubating short substrate RNAs with HeLa nuclear extracts, and proteins specifically bound to the bait RNAs were then eluted and analyzed. This RAC allows for the detection of RNA-protein interactions within functional nuclear extracts that mimic the complex cellular environment. Utilizing this technique we extended our previous observations to define a precise RNA sequence recognized by the members of the hnRNP H family. All the members of the hnRNP H family appear to share similar specificities for the consensus DGGGD sequence (where D is A, G, or U) (Table 1). C residues preceding or following the stretch of three Gs decreased the recruitment of hnRNPs to the substrate. Interestingly, a poly-G stretch constituted by 8 Gs recruits D. L. Black, University of California), and GRSF-1 (provided by Dr. J. Wilusz, Colorado State University).

hnRNPs of the H family were depleted from nuclear extracts by RAC utilizing the high affinity RNA oligonucleotide (AGGGAG), and the control C(AGUC)₄AGUCU(CAGU)₃ RNA oligonucleotide. Splicing reaction mixtures containing 250 μl of nuclear extract in a final volume of 400 μl (20 mM HEPES-KOH, pH 7.9, 5% glycerol, 0.1 M KCl, 0.2 mM EDTA, 0.5 mM dithiothreitol, 4 mM ATP, 4 mM MgCl₂) were incubated with the RNA-bound beads to deplete >80% of hnRNP H, H’, F, 2H9, and GRSF-1.

Substrate RNAs for RNA-affinity chromatography (RAC) were synthesized utilizing T7 RNA polymerase. RNAs were separated on a 6% polyacrylamide gel and visualized with a Kodak 200R Image Station. Splicing ratios were calibrated adjusting for the relative length of the spliced and unspliced RNA products.

RNA-affinity Chromatography Assays and Nuclear Extract Depletion—Substrate RNAs for RNA-affinity chromatography (RAC) were synthesized utilizing T7 RNA polymerase. RNAs were covalently linked to adipic acid dihydrazide-agarose beads as previously described (17). The beads containing immobilized RNA were incubated in a reaction mixture containing 250 μl of HeLa cell nuclear extract in a final volume of 400 μl (20 mM HEPES-KOH, pH 7.9, 5% glycerol, 0.1 M KCl, 0.2 mM EDTA, 0.5 mM dithiothreitol, 4 mM ATP, 4 mM MgCl₂). Beads were then washed, and the proteins specifically bound to the immobilized RNA were eluted in 2% SDS, separated on polyacrylamide SDS gels, electroblotted, and probed with antibodies against the following proteins: SR proteins, hnRNP A1, hnRNP K/J, and hnRNP L (provided by Dr. G. Dreyfuss, University of Pennsylvania), SF2/ASF (provided by Dr. S. L. Black, University of California), and GRSF-1 (provided by Dr. L. Wilusz, Colorado State University).
all the members of the hnRNP H family with the highest efficiency.

Because the RAC assays are performed utilizing nuclear extracts it was unclear whether each hnRNP binds the consensus DGGGD sequence directly or as part of a multiprotein complex. We performed an RAC assay utilizing the single recombinant hnRNP H, F, and 2H9 proteins. The recombinant proteins were recruited to the DGGGD-containing RNA but not to the control sequence (Fig. 2A, lanes 1 and 2) indicating that each hnRNP of the H family has the ability to independently bind the consensus sequence.

Next we examined the possibility that in native conditions the members of the hnRNP H family interact through protein-protein interaction. To this end we incubated nuclear extracts with the recombinant histidine-tagged hnRNP H and hnRNP F coupled with Ni-NTA-agarose (Fig. 2B). hnRNPs H, H', F, 2H9, and GRFS-1 were pulled down by the resin coupled with hnRNP H (lanes 3 and 4) and hnRNP F (lanes 5 and 6) but not by the resin alone (lane 7). To eliminate the possibility that these proteins were interacting with an intervening RNA, the nuclear extracts have been treated with RNase A. These data are in agreement with previous work showing that hnRNPs H and F have the ability to form heterodimers (7) and suggest that the members of the hnRNP H family form a multiprotein complex.

Because recent work has suggested a possible interaction between hnRNPs H and F and other splicing factors (5, 20), the fraction eluted from the his-tagged hnRNP F and H bound to the Ni-NTA beads was analyzed for the presence of several hnRNPs (A1, K/J, L, and PTB), SR proteins, and spliceosomal components. No proteins other than the members of the
hnRNP H family appeared to be associated with the hnRNP H family complex (data not shown). Nevertheless, it is conceivable that interactions between members of the hnRNP H family and other components of the spliceosomal machinery may occur in vivo within the dynamic pre-RNA-spliceosomal complex.

To determine the extent to which protein-protein interactions contribute to recruitment of H family members to a substrate RNA incubated with nuclear extracts, we compared the relative amounts of hnRNPs eluted from the his-tagged hnRNP F and H pull-down assays with the ones eluted from the RAC assay. Because the same incubation and washing conditions were utilized for both assays and the amount of his-tagged hnRNP F and H loaded onto the beads was roughly equivalent to the amount of those proteins recovered by the DGGGD-containing RNA (see Fig. 2B blots for hnRNP H and F, lanes 1, 3, and 5) the RNA substrate was ~6 times more efficient in recruiting the members of the hnRNP H family than the hnRNP H or F Ni-NTA-bound proteins (Fig. 2B, lanes 1, 3, and 5). Thus, the recruitment of the hnRNP H family members to the agarose (21)-bound RNA is due mainly to direct RNA-protein interactions.

To investigate the role of the DGGGD consensus motif in splicing regulation, we utilized the pTat12 substrate, which

| RNA bait (35-36 nt) | H/H' | F | 2H9 | GRSF-1 |
|---------------------|------|---|-----|--------|
| C (AGUC)₄ AGUC CU (CAGU)₃ | O | O | O | O |
| C (AGUC)₄ AAGAA CU (CAGU)₃ | x | O | O | O |
| C (AGUC)₄ AGGA CU (CAGU)₃ | x | O | O | O |
| C (AGUC)₄ AGGU CU (CAGU)₃ | O | O | O | x |
| C (AGUC)₄ AGGC CU (CAGU)₃ | O | O | O | O |
| C (AGUC)₄ AGGA CU (CAGU)₃ | x | O | O | O |
| C (AGUC)₄ AGGAG CU (CAGU)₃ | x | O | O | O |
| C (AGUC)₄ UGGG CU (CAGU)₃ | x | x | x | x |
| C (AGUC)₄ CGGG CU (CAGU)₃ | xx | x | x | x |
| C (AGUC)₄ GGGA CU (CAGU)₃ | xxx | x | xxx | xx |
| C (AGUC)₄ AGGGC CU (CAGU)₃ | xx | xx | xxx | xx |
| C (AGUC)₄ AGGGCU (CAGU)₃ | xxx | xxx | xxx | xxx |
| C (AGUC)₄ AGGGCU (CAGU)₃ | xxx | xxx | xxx | xxx |
| C (AGUC)₃ CGGGGCGGGGCG (CAGU)₃ | xxx | xxx | xxx | xx |
| C (AGUC)₃ (AGGGG)₂ (CAGU)₃ | xxxxx | xxx | xxx | xxx |
| C (AGUC)₃ (AGGGGA)₂ (CAGU)₃ | xxxxx | xxx | xxx | xxx |
| C (AGUC)₂ (AGGGGA)₄ (CAGU)₂ | xxxxx | xxxxx | xxxxx | xxxxx |
| (AGGGG)₇ | xxxxx | xxxxx | xxxxx | xxxxx |
| **Consensus DGGGD** | | | | |
Splicing Regulation by hnRNP H Family Members

contains two exons and one intron, and is derived from the HIV-1 tat gene (Fig. 3, A and B). Splicing of the tat gene is regulated by several ESEs and ESSs that have been mapped and characterized (4, 15, 22, 23) (Fig. 3B). The substrate pTat12 is weakly spliced unless the ESS2, located within tat exon 2, is mutated (pTat12-ESS2M, Fig. 3, A and B). Mutation of the ESS activates splicing by disrupting the binding of hnRNP A1, which, in turn, antagonizes the activity of SR proteins binding to nearby ESE sequences (4, 15, 24–27).

We mutated the eight DGGGD motifs that are present within the pTat12 and pTat12-ESS2M splicing substrates (Fig. 3B, mutations S1–S8). In vitro splicing reactions were carried out in HeLa nuclear extracts; splicing efficiency was measured as the ratio of spliced versus unspliced RNA. In general, mutations of the G-runs did not increase splicing in the wild-type substrates, with the exception being the mutation of the S8 G-run, which increased splicing efficiency from <0.05 to 0.15 (Fig. 3C, lane 9). The S8 G-run is located within a previously characterized splicing silencer (ESS2p), and it appears to regulate splicing by binding hnRNP A1 and hnRNP H, which decreases accessibility to the upstream 3′ splice site (23).

Mutation of the G-runs in the pTat12-ESS2M substrates revealed that one of the putative hnRNP H family binding sites, the S3 G-run, is required for efficient splicing. The S3 mutation decreases splicing efficiency from 0.8 to 0.2. Mutations of the other intronic G-runs did not appear to alter splicing efficiency (Fig. 3C, lanes 11, 12, and 17–19). The location of the S3 G-run seems crucial for its activity, because its repositioning just up- or downstream of its original location (Fig. 3B, S3-U and S3-D) did not restore splicing (pTatE12-S3-U and pTatE12-S3-D, Fig. 3C, lanes 15 and 16).

Given its location within the HIV-1 genome the S3 G-run could play a role in the regulation of not only the tat gene but other viral genes as well. All HIV-1-spliced mRNAs include the first exon and the 5′ splice site present in the pTat12 construct, referred to as the major 5′ splice site. In the viral genome the S3 G-run is located in close proximity of the major HIV-1 5′ splice site. Thus, the mechanism regulating splicing via the S3 G-run can potentially control the expression of most viral gene products.

All the G-runs present within the pTat12-ESS2M substrate have the potential to recruit members of the hnRNPs of the H family. Because splicing is down-regulated by mutation of the S3 G-run, it is conceivable that hnRNPs of the H family can activate splicing of the substrate by binding this specific G-run via imposing spatial and functional constraints in respect to other cis- or trans-acting regulatory elements. Alternatively, it is possible that the S3 mutation impairs the assembly of other factors or disrupts an RNA secondary structure required for proper splicing of the substrate.

To further investigate these possibilities, we first confirmed by RAC analysis that recruitment of hnRNP H family members to a short RNA region derived from the splicing substrate is inhibited by the S3 G-run mutation. Fig. 4A shows that hnRNPs H, H′, F, 2H9, and GRSF-1 are recruited to the wild type but not to the mutated substrate (lanes 1 and 2). Analysis of the proteins eluted from the substrates indicates that, aside from members of the hnRNP H family, no other factors show altered recruitment in response to the S3 mutation (Fig. 4A and data not shown).

Next we sought to determine whether members of the hnRNP H family were required for efficient splicing of the pTat12-ESS2M substrate. HeLa nuclear extracts were depleted from the members of the hnRNP H family by RAC utilizing a high affinity substrate (AGGGG), (Fig. 4B). More than 80% of hnRNPs H, H′, F, 2H9, and GRSF-1 were depleted from the extracts, whereas other hnRNPs or essential splicing factors, such as SR proteins, were retained (hnRNP A1, SF2/ASF, and SC35; Fig. 4B, lanes 1 and 2). Mock deletions were done by utilizing a control RNA substrate and did not result in the substantial depletion of hnRNPs from the splicing extract (Fig. 4B, lane 3).
FIGURE 3. Splicing of the HIV-1 p Tat12 substrate is altered by mutation of an intronic consensus hnRNP H family binding site. A, the pTat12 and pTat12-ESS2M substrates are derived from the HIV-1 genome as represented. The long intervening intron between the first and second tat exon has been shortened as indicated. ESS2 is mutated in the pTat12-ESS2M substrate.

B, annotated sequence of the pTat12 substrate. Green and blue lines below the nucleotide sequence depict the exon-intron organization of the pTat12 substrate with thick lines indicating exons and thin lines introns. Red boxes indicate the location of ESSs, whereas ESEs are indicated by light blue boxes. Yellow boxes indicate location of the G-runs (S1 to S8). Open boxes indicate mutation and repositioning of the S3 G-run either upstream (S3-U) or downstream (S3-D) of its original position. Mutation are indicated upstream of the boxes. C, in vitro splicing of the pTat12 and pTat12-ESS2M substrates. The G-runs were mutated as indicated in the upper side of the panel. In vitro splicing reactions were performed in HeLa nuclear extracts. Splicing efficiency was measured as the ratio of spliced versus unspliced substrate (S/US).
Splicing Regulation by hnRNP H Family Members

The pTat12-ESS2M substrate was spliced utilizing the mock depleted (Fig. 4C, lane 1) and hnRNP H family-depleted nuclear extract (Fig. 4C, lanes 2–10). The substrate RNA was spliced efficiently in the mock depleted extract, whereas a splicing product was barely detectable in the depleted extract. Addition of the single recombinants hnRNPs H, F, and 2H9 increased splicing efficiency in the depleted extract, although not to the same level of the mock depleted extract. Higher splicing efficiency was achieved by adding a mixture of hnRNPs H, F, and 2H9 to the depleted extract; nevertheless, splicing activity was not completely restored. This could have been due to the lack of GRSF-1 in the complemented extracts; unfortunately, we were unable to obtain a recombinant plasmid containing the full GRSF-1 sequence. It is also plausible that protein isoforms, other than the ones utilized in the recombination assay, are required to fully restore splicing efficiency. These data indicate that hnRNP H family members at the S3 G-run is required for efficient splicing of the pTat12-ESS2M substrate and that the functions of the members of this protein family in splicing activation are not redundant.

Spliceosomal complexes assemble in a stepwise manner in vitro; the nonspecific complex H, which consists of several hnRNPs family members, assembles first, and its formation is independent of the presence of functional splice sites. This is followed by formation of the ATP-independent pre-spliceosomal early (E) complex, which requires functional 5′ and 3′ splice sites. The A complex is the first ATP-dependent complex to enter the assembly, followed by the B complex. The two catalytic steps of splicing take place after the final C complex is assembled. To investigate what step was affected by the S3 G-run mutation, we compared spliceosomal complex formation in the substrates pTatE12-ESS2M, pTat12-S3-ESS2M, and pTat12-ESS2M-5′M. pTat12-ESS2M-5′M carries a mutation in the 5′ splice site, which prevents the formation of functional pre-spliceosomal complexes, and therefore it is not spliced efficiently (Fig. 5A). The intact 5′ splice site is required for binding of the essential U1 snRNP component of the ATP-independent pre-spliceosomal complex.

Analysis of spliceosomal complex formation utilized native agarose gel electrophoresis (18). The ATP-independent E complex can be identified due to its sensitivity to heparin; as shown in Fig. 5B the E complex is efficiently assembled by pTatE12-ESS2M and pTat12-S3-ESS2M but not by pTatE12-ESS2M-5′M carrying the 5′ splice site mutation, which prevents snRNP1 binding. In Fig. 5C, we analyzed the formation of the
Splicing Regulation by hnRNP H Family Members

FIGURE 5. hnRNP H family members regulate formation of ATP-dependent spliceosomal complexes. A, the substrates pTat12-ESS2M, pTat12-S3-ESS2Mut (ESS2 and S3 G-run mutations combined), and pTat12-ESS2M-5’M (Ess2 and S5’s mutations combined) were spliced in vitro utilizing HeLa nuclear extracts. B, native agarose gel analysis of ATP-independent spliceosomal complexes. Indicated substrates were incubated in nuclear extracts in the absence of ATP for the time indicated and then loaded onto a 1.5% agarose gel with the addition (heparin +) or omission (heparin −) of heparin before loading. E and H complexes are indicated. C, native agarose gel analysis of ATP-dependent spliceosomal complexes. Indicated substrates were incubated in nuclear extracts in the presence of ATP for the time indicated and then loaded onto a 1.5% agarose gel. Heparin was added to the reaction prior to loading. H, A, B, and C complexes are indicated. D, assembly of ATP-independent spliceosomal complexes in depleted extracts. The pTat12-ESS2M substrate was incubated with hnRNP H protein depleted (NE Dep) and mock depleted (NE) extracts in the absence of ATP for the times indicated and then loaded onto a 1.5% agarose gel. E, assembly of ATP-dependent spliceosomal complexes in hnRNP H protein-depleted nuclear extracts. pTat12-ESS2M splicing substrate was incubated in nuclear extracts in the presence of ATP for the times indicated. Nuclear extracts were either mock depleted (NE −) or depleted of the hnRNP H protein family members (NE Dep +). A mixture of hnRNPs H, F, and 2H9 was added to the preparation for a total amount of 150 ng where indicated.

ATP-dependent A, B, and C complexes. A time course shows that the pTat12-ESS2M substrate can assemble the three spliceosomal complexes efficiently. The pTat12-S3-ESS2M substrate shows less A and B complex formation, whereas no C complex can be clearly detected after a longer exposure. The decrease in complex formation correlates with the decrease in splicing efficiency seen in Fig. 3C.

Because binding of the members of the hnRNP H family to the S3 G-run was essential for splicing, we sought to determine whether these proteins are required for efficient spliceosomal formation as well. The pTat12-ESS2M substrate was incubated with the hnRNPs H family-depleted and mock depleted nuclear extracts. E complex formation was not inhibited by depletion of the hnRNPs H family (Fig. 5D), whereas the A, B, and C complexes assembled in the mock depleted but not in the hnRNPs H family-depleted extracts (Fig. 5E, lanes 1–8). After a 40-min incubation, only complex A was visible, and its presence might have been due to residual amounts of hnRNPs H proteins left in the nuclear extracts after depletion. Reconstitution of the depleted extracts with recombinant hnRNP H proteins showed that a mixture of hnRNPs H, F, and 2H9 could restore formation of spliceosomal complexes A and B, but not C, to a level close to the one obtained with the mock depleted extracts (Fig. 5E, lanes 9–12). These results indicate that members of the hnRNP H family are required for formation of ATP-dependent spliceosomal complexes but not for the ATP-independent E complex.

DISCUSSION

In this work, we studied the roles of the five members of the hnRNP H family in splicing regulation. Previous research has implicated the involvement of hnRNPs H, H9, 2H9, and F in the utilization of both splicing enhancers and silencers, and different models for their functions in splicing regulation have been proposed. Binding of hnRNP H to the 5′/3′ splice sites of several constitutively spliced exons. Thus, it is unlikely that the interplay between hnRNP H and the U1 snRNP at the 5′ splice site is the sole mechanism regulating splicing of these proteins.

Genomic studies have shown that G-runs are found preferentially in intronic sequences immediately flanking the splice sites and appear to facilitate splicing of the intron (29, 30). Accordingly, a recent report proposed that splicing of longer introns could be regulated by a looping-out mechanism in which hnRNP A1 and hnRNPs H/F may function by binding sequences proximal to the 5′/3′ splice sites and then interacting with one another to loop out most of the intron and bring the splice sites into proximity with one another (20). It has also been proposed that the interplay between exonic splicing silencers binding hnRNP A1 and intronic splicing enhancers binding hnRNP H regulate the brain-specific GRIN1 CI exon (5). Interaction between H and A1 appears to be key in both models, although hnRNP A1 function may vary depending on its association with either exonic silencers or intronic enhancers. Our data do not indicate that hnRNP A1 is directly recruited to the multimeric complex formed by the members of the hnRNP H family. Nevertheless, it is possible that weak or transient interactions between members of the hnRNP H family and hnRNP A1 do exist within the functional spliceosome but cannot be efficiently detected utilizing a pull-down assay.

These models suggest that hnRNP H and possibly F can exert different effects on splicing depending upon their relative positions in respect to the splice sites and other regulatory factors. However, these models have been developed utilizing splicing
substrates whose sequences have not been widely analyzed for the presence of splicing regulatory elements, thus not taking into consideration the likely presence and potential effects of other types of silencers or enhancers (e.g. SR protein-dependent ESEs). To better understand the interaction between different cis- and trans-regulatory elements and hnRNPs of the H family, we utilized a substrate derived from the HIV-1 tat gene, which has been characterized for the presence of several SR protein-dependent splicing enhancers and hnRNPs A1-dependent silencers (4, 15, 22, 23). Mutation of two of the eight DGGGD consensus motifs (S3 and S8) present within the substrate affected splicing. The S8 motif is located within tat exon 2 in close proximity to the 3′ splice. Previous work has shown that its mutation disrupts a splicing silencer with activity dependent on both hnRNPs A1 and hnRNPs H and may function by directly interfering with the binding of spliceosomal components to the 3′ splice site (23). S3 is located within the intron of the pTat12 substrate. The pTat12 substrate is efficiently spliced only when an hnRNPs A1-dependent ESS contained within the second tat exon is disrupted (ESS2M). Mutation of the S3 motif lowered the splicing efficiency of the pTat12-ESS2M substrate. We have demonstrated that all the members of the hnRNPs H family interact with the S3 sequence and are essential for efficient splicing of the viral substrate. Because previous studies indicate the absence of ESSs within tat exon 1 (15), binding of hnRNPs H to the S3 motif does not appear to directly counteract the effect of an hnRNPs A1-dependent ESS within the upstream exon. Thus, the proposed combinatorial theory involving hnRNPs A1-dependent exonic silencers and hnRNPs H-dependent intronic enhancers does not seem to apply to this system (5). The looping-out model can also hereby explain why mutation of only one of six intronic DGGGD motifs can decrease splicing efficiency of the pTat12-ESSM substrate. If in agreement with this model, mutation of at least two of the hnRNPs H intronic binding sites should decrease splicing efficiency (20). Because RNA substrates are highly structured molecules, it is conceivable that cis-acting sequences and splicing factors can come in contact with the hnRNPs H proteins only if the DGGGD binding site is properly positioned within a higher order structure. Indeed, repositioning the S3 site just upstream or downstream did not recover splicing efficiency (Fig. 4B, S3-U and S3-D mutations). Similar results have been achieved in a recent study that showed that only 1 of 10 potential intronic binding sites for the members of the hnRNPs H family regulate splicing of the second intron of the THPO gene (9).

The data we have presented suggest that the proper location of the S3 site within the intron is key for efficient splicing of the pTat12-ESS2M substrate. Furthermore, introduction of the S3 site and major portions of flanking intronic sequences within a heterologous splicing substrate, derived from the bcl2 gene, did not reproduce the regulatory effects seen with the pTat12-ESS2M substrate. This suggests that interaction of the hnRNPs H family members with other regulatory factors binding elsewhere on the substrate or the precise location of the S3 binding site within a complex secondary structure may be key in regulating splicing of this substrate. A complex set of ESE and ESS elements, which bind hnRNPs A1 and SR proteins, is present within the downstream exon (4, 22, 31), and thus it is conceivable that an interplay between this regulatory region and the S3 hnRNPs H binding site within a complex RNA structure may regulate splicing of the intron.

We have also demonstrated that formation of an ATP-independent spliceosomal complex, which is dependent on the presence of a functional 5′ splice site, is not inhibited by mutation of the S3 sequence or by depletion of the hnRNPs H family members. However, formation of ATP-dependent splicing complexes was inhibited by both mutation of the S3 sequence and depletion of hnRNPs H proteins. This result is in agreement with previous work showing that regulation of splicing, via hnRNPs H interaction with sequences different from the 5′ splice site, does not alter binding of the U1 snRNP (20). Previous studies on the mechanisms of splicing have also suggested that regulation occurs by modulation of U2AF and U1snRNP assembly, which triggers the formation of an ATP-independent pre-spliceosomal complex (26, 32, 33). It was also recently reported that an ESS sequence can regulate splicing of a substrate derived from the CD45 gene at a U1- or U2-ATP-dependent complex (34). Given the highly complex and dynamic nature of the spliceosome it is possible that distinct regulatory mechanisms affect different stages of its assembly pathway.

Although other authors have shown that hnRNPs F interacts with hDim, the human ortholog of Dib1p, which is a component of the Saccharomyces cerevisiae U4/U5/U6 tri snRNP (35), we were unable to detect spliceosomal snRNP components assembling with the hnRNPs H family members. It is conceivable that such interactions could occur only within the dynamic complex formed by the pre-mRNA and the spliceosome. We propose that members of the hnRNPs H family bind to intronic splicing enhancers characterized by the DGGGD sequence and form a complex that interacts with spliceosomal components. Alternatively, hnRNPs H, H’, F, 2H9, and GRSF-1 could activate spliceosome formation by interfering with the activity of exonic splicing silencers and/or intronic splicing enhancers that bind hnRNPs A1. In both models, activity of the hnRNPs H proteins could be mediated by structural motifs present in the RNA, which might bring the multimeric complex in contact with other factors binding the RNA substrate.

A growing number of reports suggests close coordination and mechanistic coupling of the events required to produce and export functional mRNAs. Members of the hnRNPs H family have been implicated in multiple aspects of mRNA biogenesis. hnRNPs F has been shown to interact with TATA BP (36) and associate with polymerase II through binding of poly-dG sequences (37). hnRNPs F has also been shown to interact with the CBP20 and CBP80 components of the Cap binding complex (38) and together with hnRNPs H and H’ to regulate mRNA 3′ end formation (39, 40). hnRNPs H has also been shown to regulate nuclear retention (41), while a role in translation regulation has been proposed for GRSF-1 (42). As more data is collected, these proteins are emerging as key regulators of mRNA biogenesis. Further evidence of their central role in gene expression can be seen in the correlation between cancer development and deregulation of expression of hnRNPs H and F (43). We are currently investigating possible roles for the members of the hnRNPs H family in coordinating the transcription-RNA processing machinery in vivo. We hope to determine whether these
proteins can regulate viral mRNA transcription, stability, export, and translation as well as splicing.

Acknowledgments—We thank Dr. Alan Zahler (University of California, Santa Cruz) and Dr. Ramaswamy Narayanan for their helpful comments.

REFERENCES

1. Johnson, J. M., Castle, J., Garrett-Engele, P., Kan, Z., Loerch, P. M., Armour, C. D., Santos, R., Schildt, E. E., Stoughton, R., and Shoemaker, D. D. (2003) Science 302, 2141–2144
2. Blencowe, B. J. (2000) Trends Biochem. Sci. 25, 106–110
3. Caceres, J. F., Stamm, S., Helfman, D. M., and Krainer, A. R. (1994) Science 265, 1706–1709
4. Min, H., Chan, R. C., and Black, D. L. (1995) Mol. Cell. Biol. 15, 2659–2671
5. Zhang, Y., Lindblom, T., Chang, A., Sudol, M., Sluder, A. E., and Golemis, E. A. (2000) Gene 257, 33–43
6. Min, H., Chan, R. C., and Black, D. L. (1995) Genes Dev. 9, 2659–2671
7. Chou, M. Y., Rooke, N., Turck, C. W., and Black, D. L. (1999) RNA (N. Y.) 7, 859–874
8. Chou, M. Y., Rooke, N., Turck, C. W., and Black, D. L. (1999) Mol. Cell. Biol. 19, 69–77
9. Chou, M. Y., Rooke, N., Turck, C. W., and Black, D. L. (1999) Mol. Cell. Biol. 19, 69–77
10. Caputi, M., and Zahler, A. M. (2002) EMBO J. 21, 845–855
11. Caputi, M., and Zahler, A. M. (2002) EMBO J. 21, 845–855
12. Chen, C. D., Kobayashi, R., and Helfman, D. M. (1999) Genes Dev. 13, 593–606
13. Fogel, B. L., and McNally, M. T. (2000) J. Biol. Chem. 275, 32371–32378
14. Krecic, A. M., and Swanson, M. S. (2003) Science 302, 2141–2144
15. Krecic, A. M., and Swanson, M. S. (2003) Science 302, 2141–2144
16. Marshand, M., Mayeda, A., Kjem, J., and Caputi, M. (2004) J. Biol. Chem. 279, 10077–10084
17. Caputi, M., and Zahler, A. M. (2001) Mol. Cell. Biol. 21, 2807–2817
18. Das, R., and Reed, R. (1999) RNA (N. Y.) 5, 1504–1508
19. Caputi, M., and Zahler, A. M. (2001) J. Biol. Chem. 276, 43850–43859
20. Martinez-Contreras, R., Fisette, J. F., Nasim, F. U., Madden, R., Cordeau, M., and Chabot, B. (2006) PLoS Biol. 4, e21
21. Brosi, R., Groning, K., Behrens, S. E., Luhrmann, R., and Kramer, A. (1993) Science 262, 102–104
22. Caputi, M., Freund, M., Kammcr, S., Asang, C., and Schaal, H. (2004) J. Virol. 78, 6517–6526
23. Jacqueten, S., Mereau, A., Bilodeau, P. S., Damier, L., Stoltzfus, C. M., and Branlant, C. (2001) J. Biol. Chem. 276, 40464–40475
24. Marchand, M., Mayeda, A., Jacqueten, S., Thomas, D., Mougin, A., Gironne, R., Stevenin, J., and Branlant, C. (2002) J. Biol. Chem. 323, 629–632
25. Tange, T. O., Damgaard, C. K., Guth, S., Valcarcel, J., and Kjem, T. O. (2001) EMBO J. 20, 5748–5758
26. Zhu, J., Mayeda, A., and Kjem, T. O. (2001) Mol. Cell 8, 1351–1361
27. Bilodeau, P. S., Domisic, J. K., and Stoltzfus, C. M. (1999) J. Virol. 73, 9764–9772
28. Buratti, E., Baralle, M., De Conti, L., Baralle, D., Romano, M., Ayala, Y. M., and Baralle, F. E. (2004) Nucleic Acids Res. 32, 4224–4236
29. McCullough, A. J., and Berget, S. M. (1997) Mol. Cell. Biol. 17, 4562–4571
30. Yeo, G., Hoon, S., Venkatesh, B., and Burge, C. B. (2004) Proc. Natl. Acad. Sci. U. S. A. 101, 15700–15705
31. Hallay, H., Locker, N., Ayadi, L., Ropers, D., Guittet, E., and Branlant, C. (2006) J. Biol. Chem. 281, 37159–37174
32. Sharma, S., Falick, A. M., and Black, D. L. (2005) Mol. Cell. Biol. 19, 485–496
33. Staknis, D., and Reed, R. (1994) Mol. Cell. Biol. 14, 7670–7682
34. House, A. E., and Lynch, K. W. (2006) Nat. Struct. Mol. Biol. 13, 937–944
35. Zhang, Y., Lindblom, T., Chang, A., Sudol, M., Shader, A. E., and Golermis, E. A. (2000) Gene (Amst.) 257, 33–43
36. Yoshida, T., Kokura, K., Makino, Y., Ossipow, V., and Tamura, T. (1999) Genes Cells 4, 707–719
37. Gamberi, C., Izaahakal, E., Beisel, C., and Mattaj, I. W. (1997) Mol. Cell. Biol. 17, 2587–2597
38. Veraldi, K. L., Arhin, G. K., Martincic, K., Chung-Ganster, L. H., Wilusz, J., and Milcarek, C. (2001) Mol. Cell. Biol. 21, 1228–1238
39. Alkan, S. A., Martincic, K., and Milcarek, C. (2006) Biochem. J. 393, 361–371
40. Kim, D. H., Langlois, M. A., Lee, K. B., Riggs, A. D., Puymirat, J., and Rossi, J. J. (2000) Proc. Natl. Acad. Sci. U. S. A. 96, 6694–6699
41. Honore, B., Baandrup, U., and Vorum, H. (2004) Exp. Cell Res. 294, 199–209
42. Bagga, P. S., Arhin, G. K., and Wilusz, J. (1998) Nucleic Acids Res. 26, 5343–5350