hnRNP A1 controls HIV-1 mRNA splicing through cooperative binding to intron and exon splicing silencers in the context of a conserved secondary structure

CHRISTIAN KROUN DAMGAARD,¹ THOMAS ØSTERGAARD TANGE,¹,² and JØRGEN KJEMS¹

¹Department of Molecular and Structural Biology, University of Aarhus, DK-8000 Aarhus C, Denmark

ABSTRACT

The removal of the second intron in the HIV-1 rev/tat pre-mRNAs, which involves the joining of splice site SD4 to SA7, is inhibited by hnRNP A1 by a mechanism that requires the intronic splicing silencer (ISS) and the exon splicing silencer (ESS3). In this study, we have determined the RNA secondary structure and the hnRNP A1 binding sites within the 3′ splice site region by phylogenetic comparison and chemical/enzymatic probing. A biochemical characterization of the RNA/protein complexes demonstrates that hnRNP A1 binds specifically to primarily three sites, the ISS, a novel UAG motif in the exon splicing enhancer (ESE) and the ESS3 element, which are all situated in experimentally supported stem loop structures. A mutational analysis of the ISS region revealed that the core hnRNP A1 binding site directly overlaps with a major branchpoint used in splicing to SA7, thereby providing a direct explanation for the inhibition of U2 snRNP association with the pre-mRNA by hnRNP A1. Binding of hnRNP A1 to the ISS core site is inhibited by at least two silencer elements in the context of a conserved RNA structure but strongly stimulated by the exonic silencer, ESS3. Moreover, the ISS also stimulate binding of hnRNP A1 to the exonic splicing regulators ESS3 and the ESE. Our results suggest a model where a network is formed between hnRNP A1 molecules situated at discrete sites in the intron and exon and that these interactions preclude the recognition of essential splicing signals including the branch point.

Keywords: exon splicing silencer; hnRNP A1; intron splicing silencer; RNA footprinting; RNA splicing; RNA structure

INTRODUCTION

The precise excision of introns from precursor-messenger RNAs (pre-mRNAs) by splicing is an essential process in the expression of most metazoan genes. Retroviruses are dependent on alternative splicing of a single primary transcript to form multiple mRNA species. In human immunodeficiency virus type-1 (HIV-1), inefficient splicing of the 9.2-kb primary transcript produces more than 30 different mRNAs (Purcell & Martin, 1993) that can be divided into three main classes. The multiply spliced 2-kb class that encodes Rev, Tat, and Nef proteins, the singly spliced 4-kb class that encodes the Env, Vpu, Vpr, and Vif proteins, and the unspliced transcript, which both encode the structural proteins Gag/Gag-Pol and serves as the RNA genome of progeny virions. An important feature for this regulation is the presence of non-consensus 3′ splice sites, where the presence of purines in the pyrimidine tracts and nonconsensus branchpoint sequences contribute to inefficient spliceosome assembly (Staffa & Cochrane, 1994; Amendt et al., 1995; Dyhr-Mikkelsen & Kjems, 1995; O’Reilly et al., 1995; Si et al., 1997, 1998). In addition, splicing is also regulated by cellular splicing factors that function via splicing enhancers and silencers in the HIV-1 transcript. The most thoroughly investigated splicing process in HIV-1 is the splicing of the two tat introns. The first tat intron is negatively regulated by at least two exon splicing silencer (ESS) elements in the second tat exon, named ESS2 and ESS2p (Amendt et al., 1994, 1995; Staffa & Cochrane, 1995; Jacquenet et al., 2001a; Fig. 1). Splicing of the second tat intron is stimulated by two exonic splicing enhancers, ESE2 and ESE3, situated in second and third tat exon, respectively (Amendt et al., 1995; Staffa & Cochrane, 1995; Kammler et al., 2001) and inhibited by at least two silencer elements, the ESS3 and the intron splicing silencer (ISS) positioned in the...
third tat exon and second tat intron, respectively (Amendt et al., 1995; Staffa & Cochrane, 1995; Tange et al., 2001; Fig. 1). The trans-acting splicing factors that recognize the positive and negative splicing signals have in some instances been characterized. The SR protein, SF2/ASF, has been implicated as a positive splicing factor acting via degenerate enhancer elements in ESE2 and ESE3 (Mayeda et al., 1999; Kammler et al., 2001; Tange & Kjems, 2001). The negative factors that have been characterized all belong to the group of heterogeneous nuclear ribonucleoproteins (hnRNPs). The ESS2, ESS3, and ISS function at an early step in spliceosome assembly by recruiting members of hnRNP A/B protein family (Amendt et al., 1995; Si et al., 1998; Caputi et al., 1999; Del Gatto-Konczak et al., 1999; Tange et al., 2001), whereas the inhibition of splicing by the ESS2p appears to be mediated by hnRNP H (Jacquenet et al., 2001a). The mechanism by which hnRNP A1 inhibits splicing of the second tat intron via the ISS and ESS3 has recently been investigated (Tange et al., 2001; Zhu et al., 2001). Zhu et al. presented evidence for hnRNP A1 binding to the ESS3, leading to a nucleation of upstream exon sequences, which interferes with the binding of splicing factors and that this effect was counteracted by SF2/ASF (Zhu et al., 2001). We found that binding of hnRNP A1 to both the ISS and ESS3 contributes to the silencing of tat intron splicing by blocking the association of the U2 snRNP with the branchpoints (Tange et al., 2001).

The hnRNP A/B proteins have in numerous cases been implicated in the regulation of alternative splice site selection, the regulation of telomere biogenesis, and mRNA transport/trafficking (Mayeda & Krainer, 1992; Eperon et al., 1993, 2000; Mayeda et al., 1994, 1998; Yang et al., 1994; Visa et al., 1996; Chabot et al., 1997; Izaurralde et al., 1997; Carson et al., 1998; Hoek et al., 1998; LaBranche et al., 1998; Blanchette & Chabot, 1999; Munro et al., 1999), suggesting that these proteins possess multifunctional roles in the cell. The binding properties of hnRNP A1 to different nucleic acid substrates have been thoroughly investigated (Cobianchi et al., 1988; Swanson & Dreyfuss, 1988; Ishikawa et al., 1993; Mayeda et al., 1994, 1998; Shamoo et al., 1995; Abdul-Manan et al., 1996; Abdul-Manan & Williams, 1996; LaBranche et al., 1998; Blanchette & Chabot, 1999; Dallaire et al., 2000; Eperon et al., 2000; Shan et al., 2000; Fiset & Chabot, 2001) and a consensus sequence (UAGGGA/U) for high affinity bind-
hnRNP A1 binding sites in HIV-1 RNA

The recognition of RNA by cellular factors is presumably influenced by RNA structure. To address the question of whether RNA structure may play a role in regulating the splicing of the second HIV-1 tat intron (SD4 to SA7), we first analyzed an RNA fragment spanning positions 8314–8486 of HIV-1 strain HXB3 (Tat3’s ss; Fig. 1) using phylogenetic comparison and enzymatic/chemical probing techniques. This RNA encompasses all known regulatory elements reported to influence the 3’ splice site recognition, including the ISS, ESE3, and ESS3. Using the free energy minimization computer program, MFOLD (version 3.1), this region is predicted to fold into a stable secondary structure that, except for minor modifications, resembles the experimentally supported structure shown in Fig. 2C. The predicted secondary structure was insensitive to addition of 300 nt of authentic HIV-1 sequence at the termini, supporting the existence of such a structure in HIV-1 genomic RNA.

The suggested RNA structure is composed of three major stem loops (SL1–3) that contain the ISS, ESE3, and ESS3, respectively (Fig. 2C). SL1 and SL3 can be formed among the HIV-1 clades belonging to the major (M) group (Fig. 3), whereas SL2 was not readily formed in all clades (data not shown). Sequences of isolates belonging to the outliner- (O) type viruses did not support the folding of any of the proposed stem loops. However, the presence of sequence elements containing UAG motifs at conserved positions in the majority of isolates belonging to group O suggests that these viruses might also be subject to a similar regulation by members of the hnRNP A/B family.

The proposed structure was generally consistent with our enzymatic and chemical probing experiments, and the minor discrepancies could generally be accounted for by sliding of bulges and internal loops (Fig. 2A, B; the results are summarized in Fig. 2C). The apical loops of SL1, SL2, SL3, and the loop, containing a GAA-repeat in SL2 (8420–8429), were highly reactive towards single-strand-specific RNases and chemicals, whereas the double-strand-specific enzyme, RNase V1, generally cleaved within double-stranded segments. Some regions were inert to all probing agents. In particular, the 3’ side of SL1 (8341–8352) that contains two of the three hnRNP A1 consensus binding sites, S2 and S3, within the ISS (Fig. 1; Tange et al., 2001) was much more inaccessible to RNase V1 than the complementary region (8322–8335; Fig. 2C), suggesting that it may be buried in tertiary interactions. In contrast, the SL3 that encompasses the ESS3 element was exposed to double-strand-specific cleavages on both sides of the stem region and hypersensitive to RNases T1 and T2 in the apical loop, clearly supporting the validity of the SL3 structure (Fig. 2A, C).

**hnRNP A1 binds to discrete sites in the ISS, ESE3, and the ESS3 elements**

We have recently shown that hnRNP A1 binds directly to an intron RNA fragment containing the ISS element and to an exon RNA fragment containing the ESS3 (Tange et al., 2001). A sequence analysis of the fragments revealed that the ISS fragment contains three putative UAG-containing hnRNP A1 binding motifs and the exon fragment 2 sites (Fig. 2C). To further delineate the binding sites of hnRNP A1 within these regions, we performed RNA footprinting experiments on both 5’- and 3’-end-labeled Tat3’s ss RNA in the presence of increasing concentrations of recombinant hnRNP A1 protein (Fig. 4A, B; the results are summarized in Fig. 4C). The most pronounced protections were seen in the apical loops of SL1 and SL3 using RNase T1 and T2 (Fig. 4A, lanes 3–5 and 9–11; Fig. 4B, lanes 4–9), using RNase T2 in the GAA repeat (Fig. 4A, lanes 9–11), and using RNase V1 at the S3 sequence within the ISS and immediately downstream of this site (Fig. 4A; lanes 15–17), suggesting that these elements are the major binding sites for hnRNP A1 proteins (denoted by boxes in Fig. 4C). Weaker protections were observed using RNase V1 and T2 at nt 8325–8335 that com-
FIGURE 2. Enzymatic and chemical probing of Tat3'sss RNA. A: 5' or 3' end-labeled Tat3'sss RNA (left and right panels, respectively) was subjected to partial RNase cleavage using the indicated RNases followed by separation of RNA fragments in denaturing polyacrylamide gels. RNases T1 cleaves after single-stranded G-residues, RNase T2 cleaves nonspecifically in single-stranded regions with some preference for A residues, and RNase V1 cleaves double-stranded regions. An alkaline hydrolysis ladder is included (Alk) together with a partial RNase U2 (A-specific; A) and T1 (G-specific; G) digestion under denaturing conditions. B: Chemical probing of Tat3'sss. Primer extension, using primers complementary to position 8402–8440 (left panel) and 8476–8495 (right panel), was used to detect DMS modifications at A and C residues (lane 1) and DEP modifications at A residues (lane 2). K (lane 3) denotes a control where no chemical was added. A dideoxynucleotide sequencing ladder is included (lane 4–7). A, B: The positions corresponding to the loops of SL1 and SL3, the GAA-rich internal loop of SL2, and the 3' splice site are indicated to the left. C: A secondary structure model of Tat3'sss. The structure is based on computer folding that was modified to conform to the enzymatic and chemical probing data. The three major stem loop structures (SL1–3) are indicated. The data from at least three independent data sets are summarized using the symbols shown in the inserted legend. Nucleotide numbers (HXB2), the RNA termini and 3' splice site (3'sss) are marked. Outlined fonts denote sequences with homology to the hnRNP A1 binding consensus sequences.
prises the complementary strand to the S2 and S3 elements (Fig. 4A, lanes 9–11 and 15–17), against RNaseV1 at nt 8450–8451 at the three-way junction of the structure (Fig. 4A, lanes 15–17, and Fig. 4B, lanes 11–13). In general, we found a similar hnRNP A1 protection pattern using both 5'- and 3'-end-labeled Tat3'ss RNA, confirming that it is primary cleavage events that are protected by hnRNP A1 (compare Fig. 4A and B). We conclude from the footprinting data that hnRNP A1 binds most strongly to discrete sites in the RNA that overlap with the functionally characterized splicing silencers, ISS and ESS3. In addition, hnRNP A1 also interacts strongly with the GAA repeat sequence, which previously has been found to have both enhancing and silencing properties depending on the structural context (Tange & Kjems, 2001; see discussion below).

Defining a core recognition site for hnRNP A1

To further investigate the individual roles of the hnRNP A1 binding sites, we performed mobility shift assays using wild type and various point and deletion mutants of the Tat3'ss RNA (Fig. 1). Binding of hnRNP A1 to wild-type Tat3'ss yielded a distinct first-order complex and several slower migrating complexes (Fig. 5, lanes 1–3). Deletion of the ESS3 element (Tat3'ss-D) significantly lowered the formation of all complexes, suggesting that this region is involved in formation of the first complex (Fig. 5, lanes 7–9). Mutating S3 (Tat3'ss-S3) had an intermediate effect on complex formation (Fig. 5, lanes 16–18), whereas mutations in the S1 or the GAA region (Tat3'ss-S1 and Tat3'ss-GAA, respectively) had only minor effects on...
FIGURE 4. Footprinting of hnRNP A1 on Tat3'ss and Tat3'ss-ΔESS3 RNA (ΔESS3) constructs. A: 5'-end labeled Tat3'ss or Tat3'ss-ΔESS3 RNA (see Fig. 1) was subjected to partial digestion with the indicated RNases in the presence of 0, 250, and 500 ng hnRNP A1 (lanes 3–20); A hydrolysis ladder (lane 1) and A-specific sequence ladder (lane 2) were included as markers. Control lanes for both constructs where no enzyme was added are included (lanes 21–22). The positions of the branchpoint 26 nucleotides upstream of the 3'splice site (BP-26); SA7, and the loop sequences are indicated to the left. The regions that are specifically protected by hnRNP A1 are indicated with bars to the right. B: Same experiment as in A, but using 3'-end labeled Tat3'ss. C: The secondary structure model of Tat3'ss summarizing hnRNP A1-specific protections. Bold and thin circles indicate strong and weak protection, respectively. The regions in the RNA exhibiting major protections from hnRNP A1 are shown by large circles and the cooperativity in binding is indicated by arrows. Thick and stippled arrows indicate that a particular site has a strong or a weak influence by presence of another hnRNP A1 binding site, respectively.
the first order complex, but showed some redistribution of the intensities of the bands corresponding to the higher order complexes. Interestingly, the S3 mutation affected the overall formation of higher order complexes most severely, suggesting that the progression from first to higher order complexes involves this element. Mutating the S2 element (Tat3’ss-S2) increased complex formation (Fig. 5A, lanes 13–15), indicating that the S2 sequence is inhibitory to hnRNP A1 binding. These results correlate with the footprinting result and confirm that S3, the GAA-repeat, and the ESS3 constitute the primary binding sites for hnRNP A1.

The accessibility of the hnRNP A1 binding site at S3 is regulated by secondary structure

Several lines of evidence suggest that hnRNP A1 bind single-stranded RNA (see Discussion). One explanation for increased binding of hnRNP A1 to the S2 mutant may therefore be that the SL1 stem becomes partially disrupted in the S2 mutant exposing the adjacent hnRNP A1 binding site at S3. To investigate this possibility, we constructed another mutant that potentially restored the base-pairing ability of SL1 by mutating the sequence complementary to the S2 site (Fig. 1; Tat3’ss-cS2). Binding of hnRNP A1 to this mutant was significantly lower than to Tat3’ss-S2 and approached that of wild type (Fig. 5, lanes 19–21). To study this phenomenon in further detail, we performed additional binding experiments using an RNA spanning the SL1 alone (Tat-SL1; Fig. 1). The result of the gel shift assay is shown in Figure 6A. In this context, the S3 mutation (Tat-SL1-S3) had the most severe phenotype, almost completely abolishing hnRNP A1 binding (Fig. 6A, lanes 17–20) whereas the S1 mutation (Tat-SL1-S1) only had a minor effect (Fig. 6A, lanes 9–12). The S2 mutation (Tat-SL1-S2) dramatically increased the hnRNP A1 binding (Fig. 6A, lanes 13–16) and restoration of the base pairing in the SL1 stem (Tat-SL1-cS2) reduced the hnRNP A1 binding to wild-type levels. To confirm that the mutations in S2 alone and its complementary strand were accompanied with an opening and closure of the SL1 stem, respectively, we employed enzymatic probing to investigate the secondary structure of wild-type and mutant SL1 RNAs (Fig. 6B; data is summarized in Fig. 6C). Mutating the S2 element led to a strong induction of RNase T1 and T2.
FIGURE 6. hnRNP A1 binding analysis of SL1 alone. 

A: Gel shift analysis showing the binding of hnRNP A1 to various Tat-SL1 RNA mutants (see Fig. 1 for details). The binding of hnRNP A1 to Tat-SL1 forms two discrete complexes.

B: Enzymatic probing of 5' end labeled wild-type or mutated Tat-SL1 RNA as indicated. T1, T2, and V1 indicate the enzymes used (see legend to Fig. 2). Alkaline digestion and A- and G-specific lanes are included as markers (lanes 1–3 and 8–10).

C: Summary of the experimental probing data (legend is shown below). Shaded regions indicate the three hnRNP A1 binding motifs (S1–3) in the ISS.

See legend to Figure 2 for details.
cleavages in the S3 region and a more moderate induction of T1 cleavage at the opposite side of SL1, which is consistent with an opening of the S3 region (Fig. 6B, lanes 19–21; indicated by an asterisk in Fig. 6C). Some degree of structural rearrangement appears to have occurred in the terminal loop in this mutant, which is reflected by the disappearance of the RNase T1 cleavage at G-3339 and the appearance of RNase V1 cleavages in the same region. As predicted, the compensating mutations in the 5′ strand of SL1 reduced the accessibility of the S3 region back to the wild-type level (Fig. 6B, lanes 27–29). The probing of the RNA mutated in the S1 region showed only minor changes compared to the wild-type RNA (Fig. 6B, lanes 15–17), whereas the S3 mutated RNA exhibited some changes in RNaseV1 cleavage activity on the 5′ side of SL1 (Fig. 6B, lanes 23–25). Taken together, these results suggest that binding of hnRNP A1 to the S3 motif is inhibited by local RNA structure, which is disrupted in the S2 mutant.

Cooperative binding of hnRNP A1 to the ESS3, the GAA loop, and ISS

The ability of hnRNP A1 to form dimers raises the possibility that hnRNP A1 binding to the multiple binding sites for hnRNP A1 in the 3′ splice site region occurs in a cooperative fashion. To investigate a potential cooperativity, we footprinted the binding of hnRNP A1 using Tat 3′ss substrates that contained mutations in the ISS, GAA loop, or ESS3 regions. Interestingly, the protections of the S1, S3, and the GAA loop induced by hnRNP A1 were clearly dependent on the presence of the ESS3 element. In the absence of the ESS3 element, the protections at S3 and the GAA repeat loop was abolished and it was significantly reduced at S1 (marked by arrows in Fig. 4A, lanes 6–8, 12–14, and 18–20). To perform the reverse experiment, we mutated all three UAG sequences in the ISS (Fig. 1, Tat3′ss-S123) and repeated the footprinting experiment. As expected, this mutant did not bind any hnRNP A1 protein in the ISS, but, more interestingly, the binding of hnRNP A1 to the ESS3 and GAA-loop was significantly reduced (Fig. 7). Finally, we found that mutating the UAG sequence in the GAA loop had an intermediate effect on the hnRNP A1-induced protection at S3 and S1 but no effect on the ESS3 binding (data not shown). We conclude that the binding of hnRNP A1 to the ESS3 plays a major role in recruiting hnRNP A1 to other sites that are distantly positioned in the primary sequence and that the ISS has a more moderate effect on hnRNP A1 binding to the ESS3 and GAA loop regions. Moreover, hnRNPA1 binding to the GAA loop also stimulates hnRNP A1 binding to the ISS. The suggested communication scheme between hnRNP A1 binding sites is summarized by arrows in Fig. 4C.

The S3 sequence in the ISS constitutes an hnRNP A1 responsive core in splicing

In agreement with our hnRNP A1 binding data, we have previously found that the ISS and the ESS3 elements constitute the major hnRNP A1 responsive silencing elements in an in vitro splicing assay (Tange et al., 2001). To investigate the functional role of the individual S1–S3 sequences elements of the ISS in splicing, we introduced the S1, S2, and S3 mutations, together and individually, into a previously described chimeric splicing substrate PiPtat (Tange et al., 2001), producing the constructs PiPtat-S123, PiPtat-S1, PiPtat-S2, and PiPtat-S3 (see Fig. 1 for details). In these constructs, the 5′ exon and part of the intron originates from PiP7A, and the 3′ part of the intron and the 3′ exon derive from Tat3′ss. We have previously shown that splicing of PiP7A is insensitive to hnRNP A1, whereas the splicing of PiPtat is efficient in hnRNP A1-depleted nuclear extracts but strongly compromised in the presence of recombinant hnRNP A1 protein.
(Tange et al., 2001). Using this construct allowed us to
directly assess the inhibitory effect of hnRNP A1 on the
usage of the HIV-1 SA7. In agreement with our earlier
results (Tange et al., 2001), the splicing of wild-type
PiP tat pre-mRNA was clearly inhibited by hnRNP A1
(Fig. 8, lane 2–5), and this effect was almost fully re-
lieved upon scrambling all three hnRNP A1 consensus
binding sites in the intron (PiP tat-S123; Fig. 8, lanes,
7–10). Introducing the S1 mutation in PiP tat (PiP tat-
S1) had only a very little effect (Fig. 8, compare
lanes 21–24 with 25–28), whereas the S2 mutation
(PiP tat-S2) resulted in a slight increase in splicing,
along with a stronger response to hnRNP A1 than wild
type (Fig. 8, compare lanes 2–5 with 12–15). Mutating
the S3 element (PiP tat-S3) almost eliminated the respon-
siveness to hnRNP A1, resembling the effect of scram-
bling all S1–S3 sequences (Fig. 8, compare lanes 7–10
and 17–20). We conclude that S3 constitutes the func-
tional hnRNP A1 responsive core of the ISS.

**DISCUSSION**

Inefficient splicing of the HIV-1 genome is required to
preserve a pool of unspliced and singly spliced tran-
scripts, to be transported across the nuclear mem-
brane via the Rev/RRE dependent pathway. Besides
having suboptimal splice sites, exonic and intronic se-
quences in the tat pre-mRNA function as binding sites
for cellular factors of the hnRNP A/B and hnRNP H
classes, which leads to further inhibition of splicing
(Amendt et al., 1994, 1995; Staffa & Cochrane, 1994,
1995; Si et al., 1997, 1998; Caputi et al., 1999; Del
Gatto-Konczak et al., 1999; Jacquenet et al., 2001a;
Tange et al., 2001; Zhu et al., 2001). Here, we have
studied the binding of hnRNP A1 in the context of the native RNA structure.

**Determination of a conserved secondary structure in the 3’ splice site region**

Our computer predictions and probing results are consistent with the existence of three major stem loop structures, SL1, SL2, and SL3, of which SL1 and SL3 are conserved in M-type HIV-1. The ISS and the ESS3 are positioned within SL1 and SL3, respectively, whereas the SL2 harbors a GAA repeat at positions 8420–8430, which is part of the degenerate SF2/ASF-dependent enhancer element ESE3 (Mayeda et al., 1999; Tange & Kjems, 2001). This GAA repeat is highly accessible to chemical modifications by DMS/DEP and single-strand specific RNase cleavage (Fig. 2A, B), consistent with these sequences acting as binding sites for cellular splicing proteins including hnRNP A1 and presumably also SF2/ASF. In contrast, we found that S3, which is the functionally most hnRNP A1 responsive site of the ISS, was less accessible to enzymatic probing reagents and situated in a helix structure adjacent to an internal bulge. Intriguingly, this structure is very similar to the recently reported secondary structural context of the ESS2 (Jacquenet et al., 2001b). The functional significance of this structure is supported by the observation that this base-pairing scheme can form in all investigated M-group viruses (Fig. 3). Another similarity between the ISS and the ESS2 is the appearance of a base-paired UAG element immediately upstream, which corresponds to the S2 sequence of the ISS (Jacquenet et al., 2001b; this report).

**Modulation of hnRNP A1 binding by RNA structure**

hnRNP A1 has originally been reported to have a general nonspecific RNA-binding activity, although several reports have specified RNA sequences that are bound by hnRNP A1 with 10–1,000-fold increased affinity (Swanson & Dreyfuss, 1988; Buvoli et al., 1990; Matunis et al., 1993; Burd & Dreyfuss, 1994; Mayeda et al., 1994, 1998; Abdul-Manan et al., 1996; Abdul-Manan & Williams, 1996; Chabot et al., 1997; Blanchette & Chabot, 1999; Caputi et al., 1999; Del Gatto-Konczak et al., 1999; Tange et al., 2001; Zhu et al., 2001). A high-affinity binding substrate containing the UAGGGA/U sequence has been defined by SELEX (Burd & Dreyfuss, 1994), which is consistent with the observation that all known natural targets, including those characterized in this report, contain a conserved UAG motif (see Table 1).

The reported binding affinity \( K_d \) of hnRNP A1 for various nucleic acid substrates is highly variable, ranging from 1 nM for the strongest binding SELEX RNA substrate UAUGAUAGGACUUGGUGU (Burd & Dreyfuss, 1994) to 1 \( \mu \)M for nonspecific RNA (Mayeda et al., 1994, 1998; Abdul-Manan & Williams, 1996; Shan et al., 2000). Based on the presented gel shift analysis and filter binding assays (C.K. Damgaard & J. Kjems, unpubl. observations) we have estimated an apparent dissociation constant for hnRNP A1 binding to the HIV-1 ISS and the ESS3 to be approximately 100 nM. A possible explanation for the apparent lower affinity of hnRNP A1 to larger naturally derived substrate RNAs may be that binding occurs in competition with RNA structure. The inhibitory effect that secondary structure may pose on hnRNP A1 binding could explain the observation that the functionality of the ESS3 element is highly dependent on the context in which it is inserted (Staffa & Cochrane, 1995). In accordance with this, we found that a close match to the extended consensus sequence is not always sufficient for hnRNP A1 binding. Although the S1, S2, and S3 sequences and the ESS3 all contain the UAG motif and exhibit 5 out of 6-nt matches to the in vitro-selected sequence, we find that binding of S3 in the RNA most likely is inhibited by RNA structure and dependent on other elements in the RNA. The preference for hnRNP A1 binding to single-stranded regions is consistent with the crystal structure of the complex between the UP1 domain of hnRNP A1 and a TTAGGGTTAGGG (TR2) oligonucleotide (Ding et al., 1999). Here the dTAGG core sequence is recognized in a single-stranded configuration by the beta-sheet platform of the RRMs. Thus, it seems plausible that the

### Table 1. Splicing silencer elements that bind hnRNP A1.a

<table>
<thead>
<tr>
<th>Origin</th>
<th>Sequence</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>FGF receptor 2, K-SAM</td>
<td>UAGGCCAGGC</td>
<td>Del Gatto-Konczak et al. (1999)</td>
</tr>
<tr>
<td>hnRNP A1 pre-mRNA (CE1)</td>
<td>UACCUGAAGAUAGG</td>
<td>Chabot et al. (1997)</td>
</tr>
<tr>
<td>hnRNP A1 pre-mRNA (CE4)</td>
<td>UAGAUAUGCCCAGUAGG</td>
<td>Blanchette and Chabot (1999)</td>
</tr>
<tr>
<td>HIV-1 (ESS2 element)</td>
<td>CAUGACUAGA</td>
<td>Amendt et al. (1994, 1995); Caputi et al. (1999)</td>
</tr>
<tr>
<td>HIV-1 (ESS3 element)</td>
<td>AGAUCAAUUCGUAAGUGA</td>
<td>Amendt et al. (1995); Staffa and Cochrane (1995); this report</td>
</tr>
<tr>
<td>HIV-1 (ESSv element)</td>
<td>UAGAGCUAGUALUGCUAGGUG</td>
<td>Bilodeau et al. (2001)</td>
</tr>
<tr>
<td>HIV-1 (ISS element)</td>
<td>UAGUGAUAUGAGUAGGCA</td>
<td>Tange et al., 2001; this report</td>
</tr>
<tr>
<td>HIV-1 (GAA loop element)</td>
<td>UAGAGAGAGAG</td>
<td>This report</td>
</tr>
</tbody>
</table>

aThe UAG sequence motifs that represent the core recognition motif for hnRNP A1 (Burd & Dreyfuss, 1994; Ding et al., 1999) are underlined.
S3 sequence also is bound in a single-stranded configuration and that a structural rearrangement in the S3 region is required prior to binding. In support of such a model, hnRNP A1 has been reported to exhibit unwinding/annealing activity (Kumar & Wilson, 1990; Pontius & Berg, 1990; Munroe & Dong, 1992; Mayeda et al., 1994; Portman & Dreyfuss, 1994) and it is possible that hnRNP A1 proteins associated with the putative single-stranded UAG motifs in the ESS3 and/or S1 are responsible for unwinding or rearranging the RNA structure prior to specific recognition of S3. The requirement for the ESS3 element for S3 binding in the Tat3’ss RNA and the increased binding of hnRNP A1 to the Tat-SL1 RNA upon destabilization of the helix is clearly consistent with this model. The cooperativity between the hnRNP A1 binding sites may also involve multimerization of the protein (see discussion below).

**hnRNP A1 multimerization**

Our RNA footprinting analysis suggests that the UAGUGA sequence in the ESS3 constitutes a primary recognition site for hnRNPA1 and that binding of hnRNPA1 to this site leads to cooperative binding of additional hnRNPA1 molecules to the GAA repeat loop and the ISS. Similarly, we also found that the ISS also stabilizes binding of hnRNPA1 to the ESS3 to some extent. Interestingly, the ISS and ESS3 are positioned relatively close to each other in the secondary structure model, where the SL1 and SL3 can form a coaxial stacked helix (Fig. 2C). It is possible that this structure promotes cooperative binding of hnRNPA1 to the ESS3 and the ISS. A model for cooperative binding was recently suggested by Zhu et al. (2001), which involves binding of hnRNPA1 to the ESS3 followed by a linear oligomerization along the upstream sequence. This may lead to nucleation of the ESE3 and possibly also regions around SA7 resulting in a decrease in splicing. Our results do support such a cooperative model, although the linear oligomerization of hnRNPA1 is not strictly compatible with our footprinting result, as we observe discrete ribonuclease protection sites at an hnRNPA1 concentration that is sufficient to inhibit splicing completely. Our observations are more consistent with the mechanism that has been suggested for the regulation of the alternative exon 7B in the hnRNPA1 mRNA, where hnRNPA1 binds to high-affinity sites flanking the exon and dimerizes, and that this bridging may promote exon skipping (Blanchette & Chabot, 1999). The ISS, GAA-loop, and ESS3 elements might function in a similar manner, precluding the recognition of essential splicing signals.

In a previous study, we found that the ISS, part of the ESE (corresponding to the GAA repeat), and ESS3 elements inhibited splicing synergistically and in an hnRNPA1-dependent fashion (Tange & Kjems, 2001; Tange et al., 2001). In the latter study, the ISS was absolutely required for the inhibition of splicing, and its function was further potentiated by the presence of the ESS3. Our present data provide a mechanistic explanation for this observation. Binding of hnRNPA1 to the ISS element blocks U2 snRNP and abrogates splicing, and this process is aided by hnRNPA1 binding to sites in the exon.

In this report, we identified the S3 sequence as the major determinant for ISS function. This conclusion is consistent with an early in vivo study by Staffa and Cochrane (1994) where a UAGG-to-GAGC mutation in S3 was made to generate an optimized branchpoint. Accordingly, a marked increase in splicing in COS-7 cells was observed (Staffa & Cochrane, 1994). At that time, this effect was attributed solely as an effect of increased branching, but our results indicate that this mutation would significantly lower hnRNPA1 binding and thereby contribute to the observed increase in splicing efficiency.

We have recently shown that hnRNPA1 inhibits the formation of the spliceosomal A complex at a step subsequent to U2AF65 binding to the polypyrimidine tract, and we speculated that hnRNPA1 binding sterically blocks the association of SF1/mBBP and/or U2 snRNP (Tange et al., 2001). The primary role of the S3 for splicing inhibition provides further support for this model, as the S3 element coincides precisely with one of the major branchpoints at position −26 (T.O. Tange & J. Kjems, submitted). Concurrently, binding of hnRNPA1 to the S3 region probably also inhibits binding of U2 snRNP to the branchpoint at position −16 based on the finding that U2 snRNP association involves nonspecific binding of an anchoring region 6–15 nt upstream of the branchpoint (Newnham & Query, 2001). Binding of hnRNPA1 to other important splicing signals may also contribute to its inhibitory effect. For instance, hnRNPA1 may interfere with binding of positive splicing factors to the GAA repeats in the ESE3 as suggested by Zhu et al. (2001), whereas the unchanged accessibility of regions near the 3′ splice site upon hnRNPA1 binding implies that the recognition of the splice site itself by the spliceosome may not be influenced by hnRNPA1.

One limitation of our study is that only a subfragment of the HIV-1 transcript has been investigated, and it remains to be investigated whether hnRNPA1 molecules associated with other sites in the HIV-1 premRNA may contribute to the binding pattern of hnRNPA1 in the SA7 splice site region.

**MATERIALS AND METHODS**

**Construction of plasmids**

The GST-hnRNPA1 plasmid has been described previously (Blanchette & Chabot, 1999). All constructs were made using standard PCR and cloning techniques (Sambrook et al., 1989).
hnRNP A1 binding sites in HIV-1 RNA

The pBS-tat, which covers the HIV-1 isolate HXB3 (GenBank M141100) nt 5811–6158 and 8141–8486 (GenBank K03455: HXB2 numbering) has also been described elsewhere (Tange et al., 2001). This construct was used as parental plasmid for constructing wild-type and point-mutation/deletion mutants in pBS-Tat3’ss covering the last 65 nt of the tat/rev intron and the first 108 nt of the third tat exon (8314–8486; HXB2 numbering). pBS-Tat3’ss, pBS-Tat3’ss-S123, pBS-Tat3’ss-S1, pBS-Tat3’ss-S2, pBS-Tat3’ss-S3, and pBS-Tat3’ss-cS2 were generated by standard PCR techniques and site-directed mutagenesis, producing PCR fragments with tags containing EcoRI and HindIII sites that were inserted between the EcoRI and HindIII sites of pBSI(+). See Figure 1 for sequences of mutagenized regions. These plasmids were linearized at either Aval or HindIII to generate RNAs spanning regions 8314–8394 or 8314–8486 (HXB2), pBS-PIP tat-S123, pBS-PIP tat-S1, pBS-PIP tat-S2, and pBS-PIP tat-S3 were constructed by standard site-directed mutagenesis, producing PCR fragments with tags containing SalI and HindIII sites that were inserted into either SalI or HindIII digested pBS-PIP tat (Tange et al., 2001). All constructs were verified by sequencing.

Protein expression

GST-hnRNP A1 was expressed and purified essentially as described in Blanchette and Chabot (1999) and kept in GST-hnRNP A1 buffer (20 mM HEPES/KOH, pH 7.6, 150 mM KCl, 20% glycerol, 1 mM DTT, 0.2 mM EDTA). Protein concentrations were measured by the method of Bradford (BioRad) and/or estimated from Coomassie Blue stained SDS-polyacrylamide gels using serial dilutions of bovine serum albumin as standard.

In vitro transcription

Internally labeled Tat3’ss-derived RNAs were prepared by standard in vitro run-off transcription with [α-32P]UTP (20 μCi/mL, 800 Ci/mmol; Pharmacia) using T7 RNA polymerase (USB). Final nucleotide concentrations in reactions were: GpppG, 1 mM; ATP, 0.4 mM; CTP, 0.4 mM; GTP, 0.1 mM; UTP, 0.04 mM; and [α-32P]-UTP, 0.002 mM. All transcripts were gel purified and concentrations determined by scintillation counting. pBS-Tat3’ss or pBS-PIP tat-derived templates were generated by linearization with Aval to produce SL1 constructs or HindIII for full-length exon constructs. Preparative amounts of nonlabeled Tat3’ss RNA was generated using Megascript T7 transcription kit (Ambion) according to the manufacturer’s protocol and RNA was subsequently gel purified. 5’ and 3’ end labeling was done as described in Damgaard et al. (1998).

Electrophoretic mobility shift assays

RNA-hnRNP A1 complexes were prepared by mixing 1 μL (0–300 ng) GST-hnRNP A1 in GST-hnRNP A1 buffer, 5 μL 2× binding buffer (20 mM HEPES/KOH, pH 7.6, 200 mM KCl, 4 mM MgCl₂, 1 mM EDTA, 2 mM DTT, 20% glycerol, 240 ng/μL rRNA; 0.1 μL RNasin (40 U/μL), 2.9 μL HzO, and 1 μL labeled RNA (5 fmol) that was renatured by heating to 75°C for 5 min and subsequently cooled slowly to room temperature in renaturation buffer (20 mM HEPES/KOH, pH 7.6, 100 mM KCl) and kept on ice. The reactions were mixed and incubated for 15 min at room temperature and run on a native 6% polyacrylamide gel containing 100 mM Tris/borate, pH 8.3, and 1 mM EDTA (1× TBE).

Enzymatic and chemical RNA probing

Probing was essentially done as described previously (Damgaard et al., 1998). Five hundred nanograms of either 5’ or 3’ end-labeled RNA was renatured as described above. A mixture of 48 μL 5× probing buffer (50 mM HEPES/KOH, pH 7.6; 500 mM KCl, 10 mM MgCl₂, 2.5 mM EDTA, 5 mM DTT, 50% glycerol (v/v)), 50 μg RNA, and ddH₂O to a final volume of 230 μL was added to 10 μL of renatured RNA. This mixture was split into twelve 20-μL aliquots, each of which was incubated with either 1 μL ddH₂O (negative control) or with 1 μL of RNase T₁ (1–10 U/μL) or 1 μL RNase T₂ (2–10 U/μL) or 1 μL RNase V₁ (10–50 U/μL). After incubation for 13 min on ice, reactions were terminated with 180 μL 300 mM NaAc, pH 6.0 and 2 mM EDTA and reactions were phenol extracted twice and the RNA was precipitated. The RNA pellets were redissolved in 10 μL 95% formamide loading buffer. Samples were heated to 95°C for 3 min and resolved on a 6–8% denaturing polyacrylamide gel containing 8 M urea, 100 mM Tris/borate, pH 8.3, and 1 mM EDTA. Alkaline hydrolysis ladder was prepared by incubating end-labeled RNA (50,000 cpm) in 2× alkaline hydrolysis buffer (100 mM NaHCO₃/Na₂CO₃, pH 9.0, 2 mM EDTA, 0.5 μg/μL RNA) at 96°C for 5–10 min. G and A sequences were generated by incubating end-labeled lyophilized RNA (50,000 cpm) with 1 μL 0.05 U/μL RNase T₁ (G specific) or 1 μL 0.005 U/μL RNase U₂ (A specific) in either 10 μL G buffer (20 mM NaCitrate, pH 5.0, 1 mM EDTA, 7 M urea, 0.025% xylene cyanol, 0.025% bromphenol blue) or 10 μL A buffer (20 mM NaCitrate, pH 3.5, 1 mM EDTA, 7 M urea, 0.025% xylene cyanol, 0.025% bromphenol blue), respectively, at 50°C for 15 min. Chemical probing was done by incubating 200 μL modification buffer (70 mM HEPES KOH, pH 7.8, 10 mM MgCl₂, 270 mM KCl), 10 pmol renatured RNA, and either 1 μL DMS (freshly diluted 1:2 in 96% EtOH (v/v)) or 5 μL DEP at 0°C for 30–60 min. DMS reaction was terminated by adding 50 μL DMS termination buffer (1 M Tris acetate, pH 7.5, 1 M β-mercaptoethanol, 1.5 M NaAc, 0.1 mM EDTA) and RNA was precipitated. DEP reaction was terminated by adding 20 μL 3 M NaAc, pH 6.0, and RNA precipitated with EtOH. Both pellets are redissolved in 0.3 M NaAc, pH 6.0, phenol extracted, and precipitated again. The RNA was redissolved in 5 μL and subjected to primer extension using a 5’ end-labeled primer. Each of the 2.4 μL RNA reactions were annealed to 0.2 pmol (3 μL of a 5’ end-labeled primer by adding 0.6 μL 10× anneal buffer (100 μM Tris-HCl, pH 6.9, 400 mM KCl, 5 mM EDTA) and incubating at 95°C for 1 min, then transferred to a 50°C water bath for another 10 min, after which the reactions were put on ice. To each reaction, 1 μL of a 10× RT buffer (500 mM Tris-HCl, pH 8.4, 100 mM MgCl₂, 20 mM DTT) was added plus 2.9 μL DNTP mix (2.5 mM of each nucleotide) and 0.1 μL of AMV reverse transcriptase (Pharmacia Biotech, 2 U/reaction), and incubated at 46°C for 30 min. The reactions were terminated by addition of 40 μL of 300 mM NaAc, pH 6.0, and samples were precipitated. A DNA marker sequence was obtained by adding deoxy nucleotides to reverse transcription reactions using untreated RNA as a template. The cDNA
samples were resolved in a 6–8% denaturing polyacrylamide gel containing 8 M urea, 100 mM Tris/borate, pH 8.3, and 1 mM EDTA.

RNA footprinting

Five hundred nanograms of either 5’ or 3’ end-labeled RNA were renatured as described above. A mixture of 48 μL 5× probing buffer, 50 μg rRNA, and ddH2O to a final volume of 210 μL was added to the renatured RNA. This mixture was split into three 70-μL aliquots, each of which was incubated with 10 μL of either hnRNP A1 storage buffer or 0.75 to 1.5 μg of hnRNP A1. After incubation for 15 min at room temperature, each reaction was further split into four 15-μL aliquots to which either 1 μL ddH2O (negative control) or 1 μL RNase T1 (1 U/mL) or 1 μL RNase T2 (2 U/mL) or 1 μL RNase V1 (30 U/mL) was added on ice. After incubation for 13 min on ice, reactions were terminated with 185 μL 300 mM NaAc, pH 6.0, and reactions were phenol extracted twice and the RNA was precipitated. The RNA pellets were redissolved in 10 μL 95% formamide loading buffer and 1 mL RNasin (Promega)) and 0.5% bromophenol blue, 2 mM EDTA, and 10 μL RNase T1 (1 U/mL) or storage buffer (20 mM HEPES/KOH, pH 7.5, 1 mM MgCl2, 25 mM MgCl2, 20 mM HEPES/KOH, pH 7.9, 0.2 mM EDTA, 1 mM DTT, 1 μL U2AF or 1 μL buffer D, 1 μL 10× splice buffer (25 mM MgCl2, 5 mM ATP, 200 mM creatine phosphate, 3.5 mM DTT, 3.2 U/μL RNasin (Promega), 2 μL 15% polyvinyl alcohol (PVA), 1 μL GST-hnRNP A1 (100–500 ng) or storage buffer (20 mM HEPES/KOH, pH 7.9, 150 mM KCl, 20% glycerol, 1 mL DTT, 0.2 mM EDTA), and 1 μL labeled PIPtat-derived pre-mRNA (5–10 fmol). The reactions were incubated at 30°C for 2 h and stopped by addition of 190 μL splice stop buffer (300 mM NaCH3COO, pH 6.0, 0.5% SDS, 100 mM NaCl, 20 mM Tris/ClHCl, pH 7.5, 5 mM EDTA, 0.03 μg/μL bulk Escherichia coli rRNA). After 1× phenol extraction, 1× chloroform extraction, and ethanol precipitation, the samples were resuspended in RNA loading buffer (90% formamide, 1 mM EDTA, 0.05% bromophenol blue, and 0.05% xylenol cyanol), heated at 95°C for 3 min, and applied to a denaturing 6% polyacrylamide gel containing 8 M urea, 100 mM Tris/borate, pH 8.3, and 1 mM EDTA.

In vitro splicing assays

Splicing reactions with oligo(dT)ΔNE (HeLa nuclear extract; ~10 μg/mL; 4C, Belgium; Tangye et al., 2001) were prepared by mixing 3 μL oligo(dT)ΔNE, 1 μL buffer D (100 mM KCl, 20% glycerol, 20 mM HEPES KOH, pH 7.9, 0.2 mM EDTA, 1 mM DTT), 1 μL U2AF or 1 μL buffer D, 1 μL 10× splice buffer (25 mM MgCl2, 5 mM ATP, 200 mM creatine phosphate, 3.5 mM DTT, 3.2 U/μL RNasin (Promega), 2 μL 15% polyvinyl alcohol (PVA), 1 μL GST-hnRNP A1 (100–500 ng) or storage buffer (20 mM HEPES/KOH, pH 7.9, 150 mM KCl, 20% glycerol, 1 mL DTT, 0.2 mM EDTA), and 1 μL labeled PIPtat-derived pre-mRNA (5–10 fmol). The reactions were incubated at 30°C for 2 h and stopped by addition of 190 μL splice stop buffer (300 mM NaCH3COO, pH 6.0, 0.5% SDS, 100 mM NaCl, 20 mM Tris/ClHCl, pH 7.5, 5 mM EDTA, 0.03 μg/μL bulk Escherichia coli rRNA). After 1× phenol extraction, 1× chloroform extraction, and ethanol precipitation, the samples were resuspended in RNA loading buffer (90% formamide, 1 mM EDTA, 0.05% bromophenol blue, and 0.05% xylenol cyanol), heated at 95°C for 3 min, and applied to a denaturing 6% polyacrylamide gel containing 0.75 × TBE.

ACKNOWLEDGMENTS

We are grateful to Jim Manley for SF2/ASF plasmid, Benoit Chabot for hnRNPA1 plasmid, and Rita Rosendahl for excellent technical assistance. The work was supported in parts by grants from the Danish Medical and Natural Research Councils, the Karen Elise Jensen Foundation, the Danish AIDS Foundation, and the Carlsberg Foundation. C.K.D. and T.O.T. were supported by University of Aarhus.

Received January 22, 2002; revised for return February 14, 2002; revised accepted August 30, 2002

REFERENCES

hnRNP A1 binding sites in HIV-1 RNA


Newnham CM, Query CC. 2001. The ATP requirement for U2 snRNP addition is linked to the pre-mRNA region 5’ to the branch site. RNA 7:1298–1309.


hnRNP A1 controls HIV-1 mRNA splicing through cooperative binding to intron and exon splicing silencers in the context of a conserved secondary structure.

Christian Kroun Damgaard, Thomas Ostergaard Tange and Jørgen Kjems

RNA 2002 8: 1401-1415

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here.