11-1-2019

Genome-wide analysis of heterogeneous nuclear ribonucleoprotein (hnRNP) binding to HIV-1 RNA reveals a key role for hnRNP H1 in alternative viral mRNA splicing

Sebla B Kutluay

Ann Emery

Srinivasa R Penumutchu

Dana Townsend

Kasyap Tenneti

See next page for additional authors

Authors

Sebla B Kutluay, Ann Emery, Srinivasa R Penumutchu, Dana Townsend, Kasyap Tenneti, Michaela K Madison, Amanda M Stukenbroeker, Chelsea Powell, David Jannain, Blanton S Tolbert, Ronald I Swanstrom, and Paul D Bieniasz

AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Virology®

Genome-Wide Analysis of Heterogeneous Nuclear Ribonucleoprotein (hnRNP) Binding to HIV-1 RNA Reveals a Key Role for hnRNP H1 in Alternative Viral mRNA Splicing

Sebla B. Kutluay,[a] Ann Emery,[b] Srinivasa R. Penumutchu,[c] Dana Townsend,[a] Kasyap Tenneti,[a] Michaela K. Madison,[a] Amanda M. Stukenbroeker,[a] Chelsea Powell,[f] David Jannain,[f] Blanton S. Tolbert,[c] Ronald I. Swanstrom,[b,d,e] Paul D. Bieniasz[f,g]

[a]Department of Molecular Microbiology, Washington University School of Medicine, Saint Louis, Missouri, USA
[b]Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA
[c]Department of Chemistry, Case Western Reserve University, Cleveland, Ohio, USA
[d]Department of Biochemistry and Biophysics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA
[e]UNC Center for AIDS Research, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA
[f]Laboratory of Retrovirology, The Rockefeller University, New York, New York, USA
[g]Howard Hughes Medical Institute, The Rockefeller University, New York, New York, USA

ABSTRACT Alternative splicing of HIV-1 mRNAs increases viral coding potential and controls the levels and timing of gene expression. HIV-1 splicing is regulated in part by heterogeneous nuclear ribonucleoproteins (hnRNPs) and their viral target sequences, which typically repress splicing when studied outside their native viral context. Here, we determined the location and extent of hnRNP binding to HIV-1 mRNAs and their impact on splicing in a native viral context. Notably, hnRNP A1, hnRNP A2, and hnRNP B1 bound to many dispersed sites across viral mRNAs. Conversely, hnRNP H1 bound to a few discrete purine-rich sequences, a finding that was mirrored *in vitro*. hnRNP H1 depletion and mutation of a prominent viral RNA hnRNP H1 binding site decreased the use of splice acceptor A1, causing a deficit in Vif expression and replicative fitness. This quantitative framework for determining the regulatory inputs governing alternative HIV-1 splicing revealed an unexpected splicing enhancer role for hnRNP H1 through binding to its target element.

IMPORTANCE Alternative splicing of HIV-1 mRNAs is an essential yet quite poorly understood step of virus replication that enhances the coding potential of the viral genome and allows the temporal regulation of viral gene expression. Although HIV-1 constitutes an important model system for general studies of the regulation of alternative splicing, the inputs that determine the efficiency with which splice sites are utilized remain poorly defined. Our studies provide an experimental framework to study an essential step of HIV-1 replication more comprehensively and in much greater detail than was previously possible and reveal novel *cis*-acting elements regulating HIV-1 splicing.

KEYWORDS HIV-1, hnRNP, splicing

During HIV-1 infection, a single viral transcript undergoes extensive alternative splicing, generating over 50 alternative spliced mRNAs that are classified by size (1–3). The 1.8-kb class comprises completely spliced mRNAs that encode Tat, Rev, and Nef proteins, while the incompletely spliced 4-kb-size-class mRNAs encode Vif, Vpr, Env, Vpu, and a truncated form of Tat. The unspliced primary transcript encodes the Gag and Gag-Pro-Pol proteins and serves as genomic RNA. A novel 1-kb class of transcripts was recently identified (4), but no functionality has yet been assigned to this class of transcripts. While completely spliced transcripts are exported from the nucleus by

Citation Kutluay SB, Emery A, Penumutchu SR, Townsend D, Tenneti K, Madison MK, Stukenbroeker AM, Powell C, Jannain D, Tolbert BS, Swanstrom RI, Bieniasz PD. 2019. Genome-wide analysis of heterogeneous nuclear ribonucleoprotein (hnRNP) binding to HIV-1 RNA reveals a key role for hnRNP H1 in alternative viral mRNA splicing. J Virol 93:e01048-19. https://doi.org/10.1128/JVI.01048-19.

Editor Viviana Simon, Icahn School of Medicine at Mount Sinai

Address correspondence to Sebla B. Kutluay, kutluay@wustl.edu, or Paul D. Bieniasz, pbieniasz@rockefeller.edu.

Received 23 June 2019
Accepted 23 July 2019
Accepted manuscript posted online 14 August 2019
Published 15 October 2019

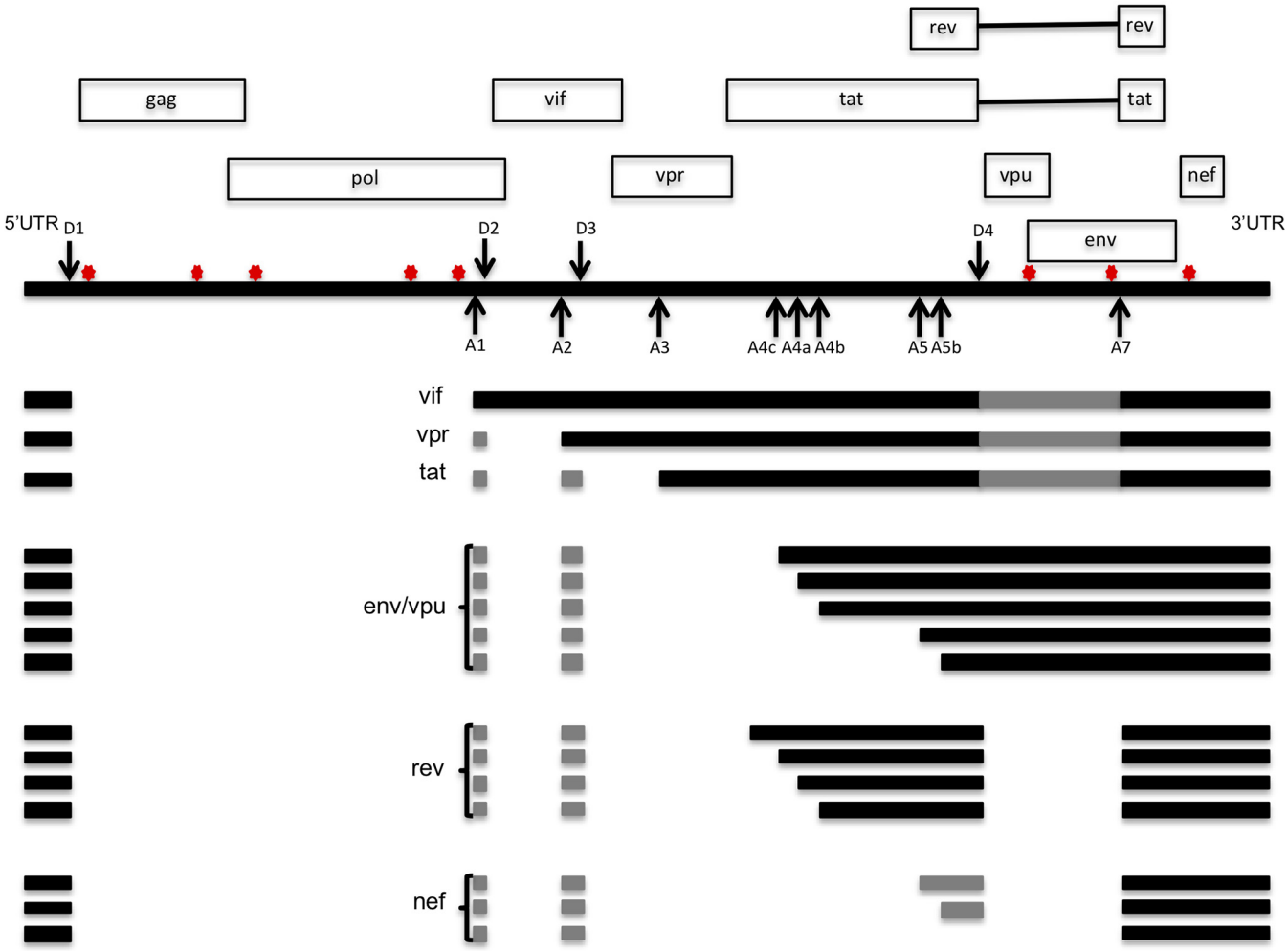

FIG 1 Schematic diagram of HIV-1 alternative splicing events. The positions of splicing donor (D1 to D4) and acceptor (A1 to A7) sites are indicated. The products of alternative splicing events are indicated at the bottom.

canonical cellular mRNA export pathways, export of incompletely spliced and unspliced transcripts is mediated by the viral Rev protein (5). Rev binds to a highly structured RNA element termed the Rev-responsive element (RRE) and mediates RNA export by recruiting the host karyopherin Crm1 (6). This mode of regulation allows for the initial appearance of regulatory proteins, followed by the production of virion proteins a few hours later (7).

The numerous spliced HIV-1 mRNAs are generated through the utilization of at least four splice donor sites (D1, D2, D3, D4) and eight acceptor sites (A1, A2, A3, A4c, A4a, A4b, A5, and A7) (1–3) (Fig. 1). All spliced RNAs use D1, located in the 5′ noncoding region, and one of the downstream acceptor sites (1, 2). The members of the fully spliced 1.8-kb class of RNAs undergo further splicing to remove the D4-A7 *env* intron (1, 2). A small fraction of RNAs contain two additional small exons in their 5′ untranslated regions (UTRs), A1-D2 and A2-D3, generated by utilization of the D2 and D3 donor sites (1, 2). It has been proposed that inclusion of these exons can impact transcript stability, splicing, and viral gene expression in heterologous expression systems (1, 2, 8–10). However, it remains to be determined whether they have a functional impact on HIV-1 replication. Additional cryptic splice donor and acceptor sites are utilized at lower frequencies (4), but it is currently unknown whether these alternative splicing events have a functional significance.

The intrinsic strengths of the donor and acceptor sites are governed in part by the

extent of their similarity to the consensus donor and acceptor sequences (2). For example, the D1 sequence is 100% identical to the consensus sequence, and all subsequent splicing events require the use of D1. Other donor and acceptor sites are generally suboptimal, and their utilization is further regulated by proximal *cis*-acting elements, including exonic splicing enhancers (ESEs), exonic splicing silencers (ESSs), and intronic splicing silencers (ISSs). These *cis*-acting elements can be bound by members of the serine-arginine-rich protein family (SR proteins) and the heterogenous nuclear ribonucleoprotein (hnRNP) family, which generally enhance and repress the use of nearby splice sites, respectively. However, splice site utilization is often context dependent and reflects the combinatorial nature of both *cis*- and *trans*-acting inputs (11).

hnRNPs have also been ascribed additional roles, including the retention of unspliced viral genomic RNA in the nucleus (12, 13), nuclear export of viral/cellular RNAs (14), and viral mRNA translation (14, 15). The majority of hnRNP proteins are localized primarily in the nucleus, but certain hnRNPs can shuttle between the nucleus and the cytosol (16). Cytosolic hnRNPs can localize to RNA/stress granules and regulate processes such as RNA translation, storage, degradation, and stress responses (17, 18). Notably, HIV-1 infection may prevent the assembly of stress granules (19, 20). Thus, hnRNPs could affect HIV-1 replication by influencing splice site utilization or later events in the virus life cycle, including nuclear export, RNA stability, trafficking, and translation.

hnRNPs are modular proteins that commonly encode 2 to 4 RNA recognition motifs (RRMs) or K homology (KH) domains. Most hnRNPs also contain RGG boxes (Arg-Gly-Gly repeats) and auxiliary domains with distinctive amino acid compositions (e.g., glycine-rich, acidic, or proline-rich domains) that primarily mediate protein-protein and nonspecific nucleic acid-protein interactions (21). The general consensus, based on *in vitro* binding studies, is that hnRNPs form specific contacts with short degenerate sequences using RRM or KH domains. For some hnRNPs, the sequence preferences have been further substantiated using global and competitive binding methods (22–24).

Based primarily on experiments employing *in vitro* splicing reporters and genetic assays, a few hnRNP proteins are thought to play particularly important roles in HIV-1 alternative splicing. Members of the hnRNP A/B family have been studied most extensively and shown to bind to the exonic splicing silencers ESS2 within *tat* exon 2 (25, 26); ESS3, located within *tat/rev* exon 3 (27, 28); ESSV, located in the *vif*-coding exon (29); and an ISS element within *tat* intron 2 (28, 30, 31). Similarly, hnRNP H1 has been suggested to bind to ESS2p (32), a cryptic exon within the Env open reading frame (33, 34), and G-rich motifs within *tat* exon 2 (35) and *tat* intron 3 (36). Other studies have proposed that hnRNP D (37), hnRNP E (38), and hnRNP K (39) can regulate HIV-1 splicing. In most studies, regulation of HIV-1 splicing has been studied using *in vitro* and *in vivo* splicing reporters, wherein viral subgenomic fragments were moved to foreign genetic environments so as to determine how they influence the distribution of spliced reporter products. In addition, the binding of hnRNPs to target elements was assessed largely *in vitro*, using defined targets. Whether the presumed target sites are also selectively bound by hnRNPs in cells and whether other binding sites on HIV-1 RNA exist are unknown.

Cross-linking immunoprecipitation (CLIP) coupled to deep sequencing (CLIP-seq) is a powerful method to identify the RNA targets of RNA-binding proteins in physiological settings (40, 41). We employed a version of CLIP-seq, referred to as photoactivatable ribonucleoside enhanced-CLIP (PAR-CLIP), to determine the binding sites, in cells, of some hnRNP proteins previously reported to regulate HIV-1 RNA splicing (42). PAR-CLIP relies on the incorporation of ribonucleoside analogs (i.e., 4-thiouridine [4-SU] in place of uridine, 6-thioguanosine [6-SG] in place of guanosine) into nascent RNA molecules. This allows enhanced protein-RNA cross-linking efficiency and provides near-nucleotide resolution of binding sites due to T-to-C (for 4-SU) or G-to-A (for 6-SG) substitutions introduced by reverse transcriptase precisely at the cross-linking site (42). We also used deep sequencing to quantify splice site utilization under conditions where hnRNPs

were depleted or their binding sites were mutated. We found that the hnRNP A/B group of proteins does not bind to discrete elements on HIV-1 RNA. Rather, they bind to many sites containing AGG motifs that are widespread in HIV-1 RNA. In stark contrast, hnRNP H1 bound to a small number of discrete sites within the viral genome, comprised of four or more consecutive guanosines, frequently embedded in A-rich elements. Complementary biochemical and NMR experiments with isolated RNA-binding domains of hnRNP H1 revealed a clear preference for G-rich sequences. Notably, depletion of hnRNPs using RNA interference (RNAi) perturbed splicing, most prominently at A1, A2, and/or A3. Mutations at discrete hnRNP H1 binding sites on viral RNA also perturbed HIV-1 splicing and, in particular, decreased the use of splice acceptor A1, causing a reduction in Vif expression and viral fitness. Overall, our studies provide an experimental framework to study an essential step of HIV-1 replication in a quantitative and comprehensive manner and in greater detail than was previously possible, and reveal a novel hnRNP H1 binding *cis*-regulatory element that controls HIV-1 splicing.

RESULTS

Expression of 3×HA-tagged hnRNP proteins for CLIP assays. We previously showed that the pattern of HIV-1 RNA splicing is strikingly similar in HEK293T cells transfected with proviral plasmids and infected T cell lines (43). Thus, for the ease of further manipulation, we first generated HEK293T-derived cell lines that stably express 3×HA-tagged versions of the hnRNP A1, A2, B1, H1, and K proteins at lower-than-endogenous levels. Thus, the total levels of each hnRNP were only marginally affected (Fig. 2A). Importantly, these manipulations did not affect viral protein (Gag) expression in cells transfected with proviral plasmids (Fig. 2B), nor did they affect particle release (Fig. 2B) or virion infectiousness (Fig. 2C).

Cells were grown in the presence of 4-SU and UV irradiated to induce protein-RNA cross-linking at 4-SU-modified nucleotides. Following RNase digestion, hnRNP-RNA adducts were immunoprecipitated, the cross-linked RNA was labeled with polynucleotide kinase, and the adducts were visualized by autoradiography following separation by SDS-PAGE and transfer to nitrocellulose membranes. All of the tagged hnRNPs cross-linked to cellular RNAs and were efficiently immunoprecipitated (Fig. 2D). To identify binding sites for hnRNPs on viral RNA, similar experiments were conducted in cells transfected with a plasmid containing a full-length DNA copy of the HIV-1_{NL4-3} (subtype B) or HIV-1_{NDK} (subtype D) viral genome. Cross-linked hnRNP-RNA species were isolated by immunoprecipitation and enzymatic digestion, to generate CLIP-seq libraries. Two independent libraries for each hnRNP and each virus strain were generated. After high-throughput sequencing, the CLIP reads were mapped to the human and HIV-1 genomes (Fig. 2E). The majority of sequencing reads mapped to the human genome, and only a small fraction of reads (~1% or less) were derived from viral RNAs (Fig. 2E). Because these values reflect the approximate relative abundance of viral and cellular mRNAs in similar settings (44), we speculate that hnRNPs may bind with similar frequencies to viral and cellular RNAs.

hnRNPs bind primarily to defined motifs within introns of cellular mRNAs. We first examined the properties of the cellular RNAs bound by hnRNPs. The majority of CLIP-seq reads for all hnRNPs were derived from the transcripts of known genes (Fig. 3A), and a large fraction of reads were from introns (Fig. 3B), as was observed in a previous study (45). Preferred sequence motifs bound by hnRNPs, revealed using the cERMIT algorithm (46), and the frequencies with which they occurred in hnRNP binding sites in cellular RNAs are depicted in Fig. 3C. For hnRNP A1, A2, and B1, the UNAGG motif resembles the well-characterized UAGG previously reported to be an hnRNP A1 binding site (47, 48). hnRNP H1 binding occurred at oligonucleotide G motifs, similar to the previously described hnRNP H1 binding GGGA motif (49). The frequent occurrence of T in the motifs bound by hnRNPs in our analysis should be interpreted cautiously and may reflect the fact that 4-SU was used to drive cross-linking and identify binding sites. Although hnRNP K has been proposed to bind short C-rich tracks, as inferred from a systematic evolution of ligands by exponential enrichment (SELEX) study (50), CLIP-seq

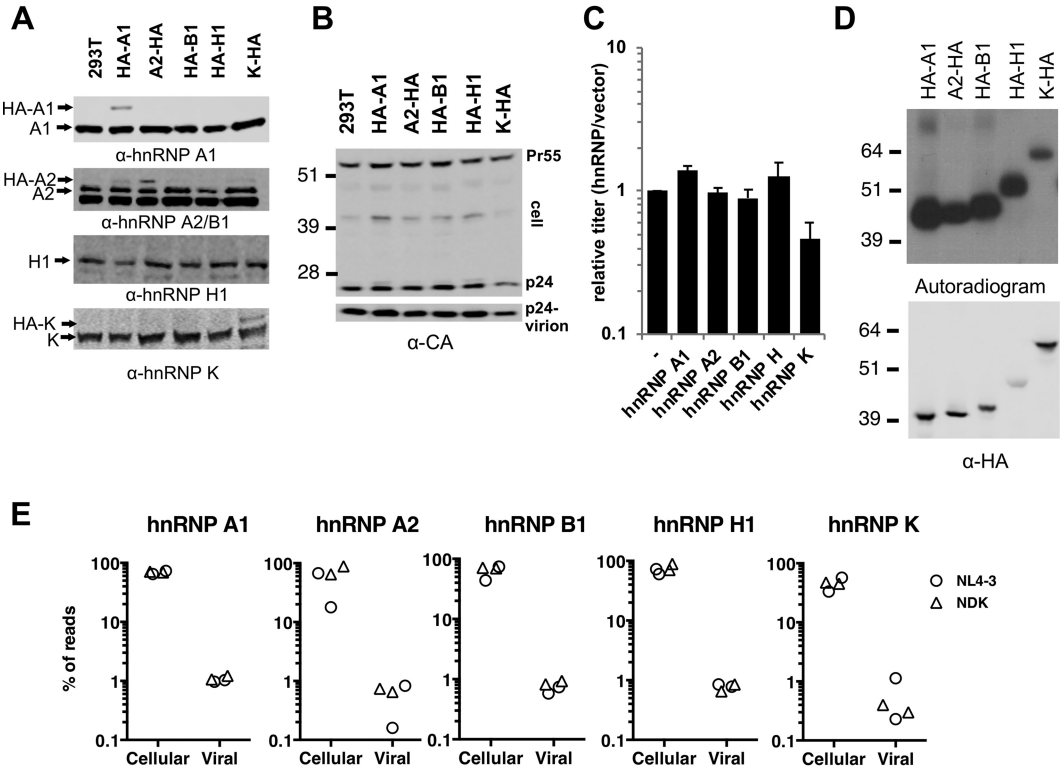

FIG 2 Generation of cell lines stably expressing HA-tagged hnRNPs and summary of CLIP experiments. (A) Western blot analysis of stably expressed HA-tagged and endogenous hnRNP proteins in the HEK293T cells used for CLIP experiments. (B, C) HEK293T cells expressing 3×HA-tagged hnRNPs were transfected with HIV-1_{NL4-3}. (B) Western blot analysis of cell lysates and virions using an anti-CA antibody. (C) Infectious virion yield was determined using MT4-green fluorescent protein reporter cells. Values are normalized to the infectious virion yield from cells transduced with the empty vector-only control and represent the average from two independent experiments. (D) Immunoprecipitated and [^{32}P]ATP- and polynucleotide kinase-labeled hnRNP-RNA adducts from 4-SU-treated cells were visualized by autoradiography (top). Western blot analysis of the same membranes (bottom). (E) Proportion of raw CLIP-seq reads that map to the human (cellular) and viral genomes, obtained from experiments conducted with HIV-1_{NL4-3} and HIV-1_{NDK}. Data represent those from four independent CLIP-seq experiments (two with HIV-1_{NL4-3} and two with HIV-1_{NDK}). Numbers to the left of gels in panels B and D are molecular masses (in kilodaltons).

revealed the presence of CGAU and CCAUAGA motifs within a subset of cellular binding sites (Fig. 3C).

Binding sites of hnRNPs on viral RNAs. Next, we determined where on viral RNAs hnRNPs bound and whether the binding specificities of hnRNPs for viral RNAs mirrored those for cellular RNAs. Although hnRNP A1, A2, and B1 proteins were previously proposed to bind to discrete sequences on HIV-1 RNAs (25–31), our CLIP analysis revealed many binding sites dispersed across the viral genome, which appeared to indicate a lack of specificity (Fig. 4 and 5). Nevertheless, regions overlapping splice acceptor A7 (at approximately nucleotide [nt] 7860) and sequences positioned 3′ to A7 (at approximately nt 8300) were frequently bound in both HIV-1_{NL4-3} and HIV-1_{NDK} (Fig. 4 and 5). Possibly, these patterns of binding may reflect the spreading of hnRNPs in the 5′ and 3′ direction through oligomerization, following an initial interaction with a high-affinity binding site (51). The binding patterns of the closely related hnRNP A1, A2, and B1 proteins were similar on both the HIV-1_{NL4-3} and HIV-1_{NDK} genomes. However, hnRNP B1 binding appeared to be more prominent at discrete locations (Fig. 4 and 5). Potentially, this property might be due to the presence of an extended N-terminal domain in hnRNP B1 that contains several basic amino acids that may alter its RNA-binding affinity and specificity.

Although the binding patterns of the hnRNP A1, A2, and B1 proteins on viral RNAs appeared to indicate promiscuous binding (Fig. 4 and 5), the HIV-1 genome is purine (especially adenosine) rich; thus, frequent binding to AGG motifs along the length of the viral genome might be expected. Indeed, the majority of the binding sites (or read

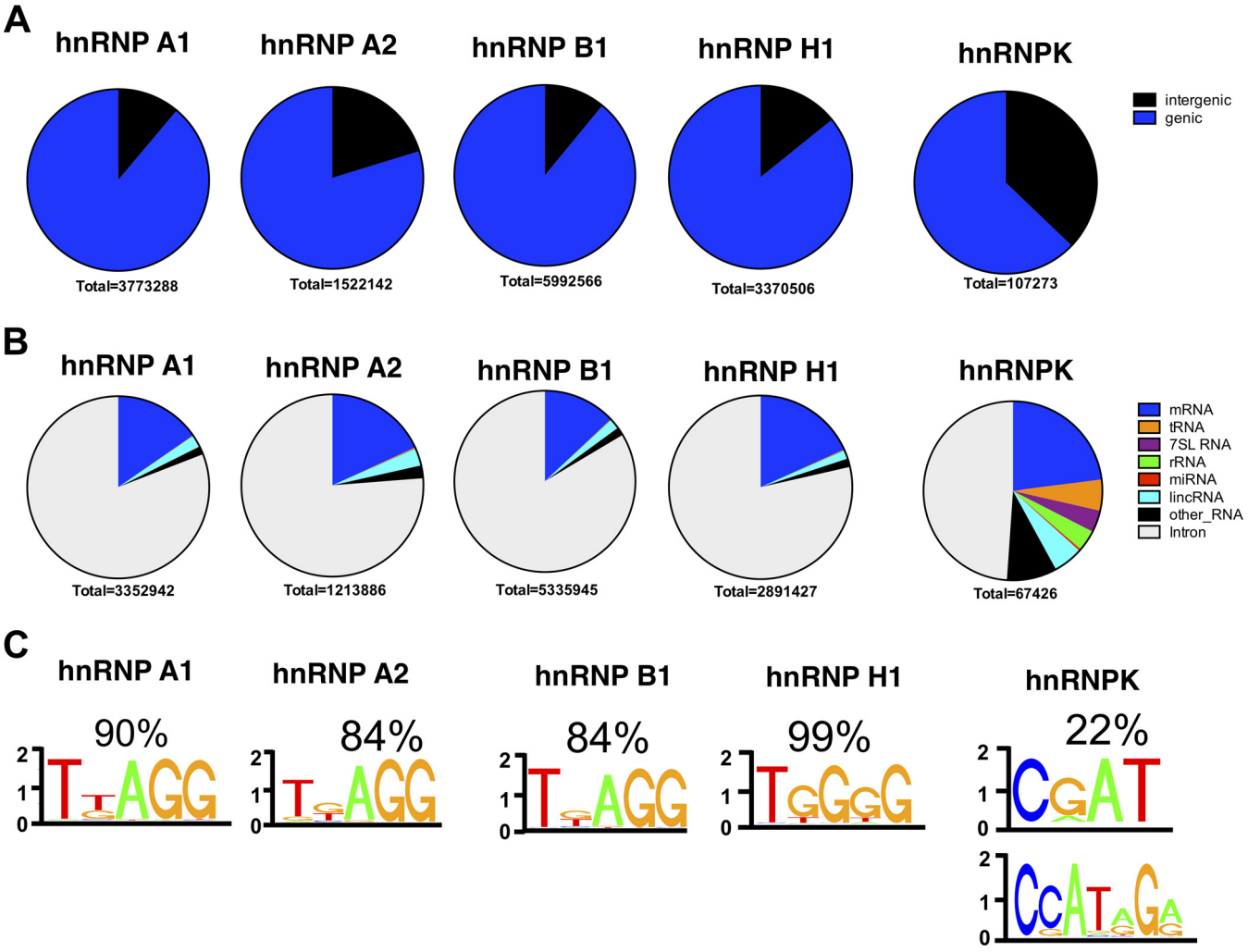

FIG 3 Classification of hnRNP-bound cellular RNAs and the binding motifs identified within them. (A) hnRNP CLIP-seq-derived sequencing reads were mapped to the human genome, clustered, and classified based on whether they map to RNA-coding genes (genic) or intergenic regions. The total number of sequencing reads in each RNA class is represented below the pie charts. (B) Further classification of genic clusters is detailed. The total number of sequencing reads in each group is represented below the pie charts. miRNA, microRNA; lincRNA, long intergenic noncoding RNAs. (C) Binding motifs within these clusters and the cumulative percentages of their occurrence within clusters.

clusters, as determined by use of the PARalyzer tool) contained one or often multiple AGG motifs (Fig. 4 and 5). Furthermore, an increased number of AGG repeats within a cluster of reads in viral RNA correlated with an increased read density (Fig. 6A to C) and an increased absolute number of reads associated with that cluster (Fig. 6D and E).

In contrast to the widespread binding pattern of hnRNP A/B proteins, hnRNP H1 bound to a small number of discrete, well-defined sequences on HIV-1 RNAs (Fig. 7A and B). While the majority of the binding sites were conserved in HIV-1_{NL4-3} and HIV-1_{NDK}, the most frequently bound site within HIV-1_{NL4-3} (at approximately nt 5800) was not present in HIV-1_{NDK} (Fig. 7A and B). Although a previous study indicated a role for hnRNP K in the regulation of HIV-1 splicing (39), hnRNP K binding to the viral genome was at background levels (Fig. 7A and B). The paucity of C nucleotides in the HIV-1 genome (52) and the even greater rarity of CG dinucleotides (53) may contribute to the relative lack of hnRNP K binding to the HIV-1 genome.

hnRNP H1 and related proteins bind to purine-rich sequences on viral RNAs. The majority of the eight most frequently hnRNP H1-bound sites consisted of one or more G-rich motifs (four or more consecutive guanosines) flanked by A-rich sequences (Fig. 8A). Binding site 4 (nucleotides 2868 to 2899) contained tandem GGG motifs and

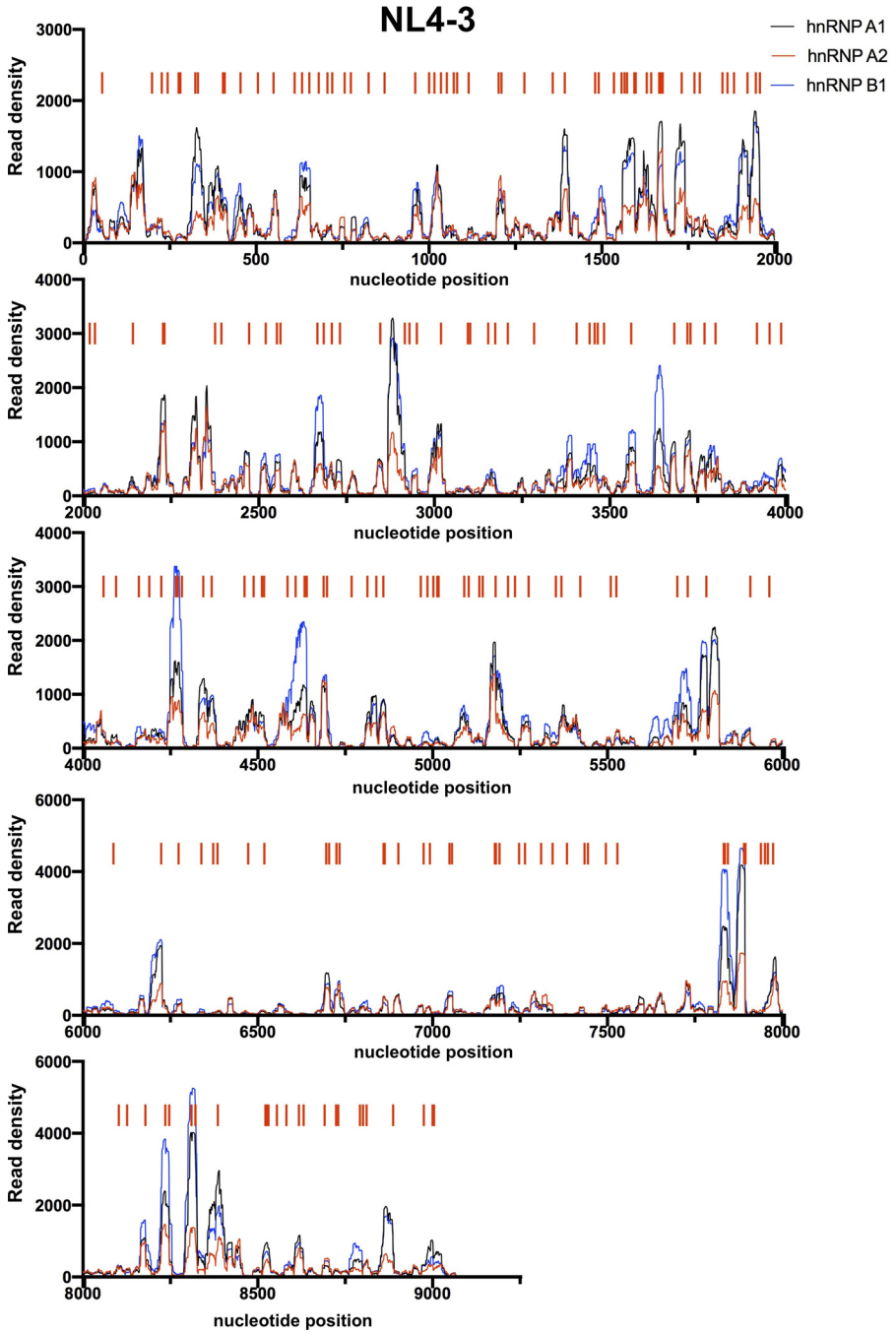

FIG 4 Binding sites of the hnRNP A1, A2, and B1 proteins on viral RNAs (NL4-3) derived from CLIP-seq. HEK293T cells stably expressing 3×HA-tagged hnRNP A1, A2, and B1 proteins were transfected with a full-length HIV-1_{NL4-3} proviral plasmid prior to CLIP-seq analysis. The frequency distribution of nucleotide occurrence (read density, *y* axis) of reads mapping to HIV-1_{NL4-3} is shown for hnRNP A1 (black), hnRNP A2 (red), and hnRNP B1 (blue). The positions of the AGG motifs in the HIV-1 genome are indicated as red lines and are overlaid on the CLIP data.

A-rich flanking sequences rather than a GGGG motif (Fig. 8A and B). Of the 41 GGGG motifs in the HIV-1_{NL4-3} genome, 39 were coincident with hnRNP H1 binding peaks, with a large range of read values (80 to 10,000 reads, compared to a background of 10 to 30 reads) being mapped to each site (Fig. 8B). Given this preference for G-rich sequences, we conducted comparative hnRNP H1 CLIP-seq experiments with cells grown in the presence of 6-SG rather than 4-SU to uncover any potential bias toward

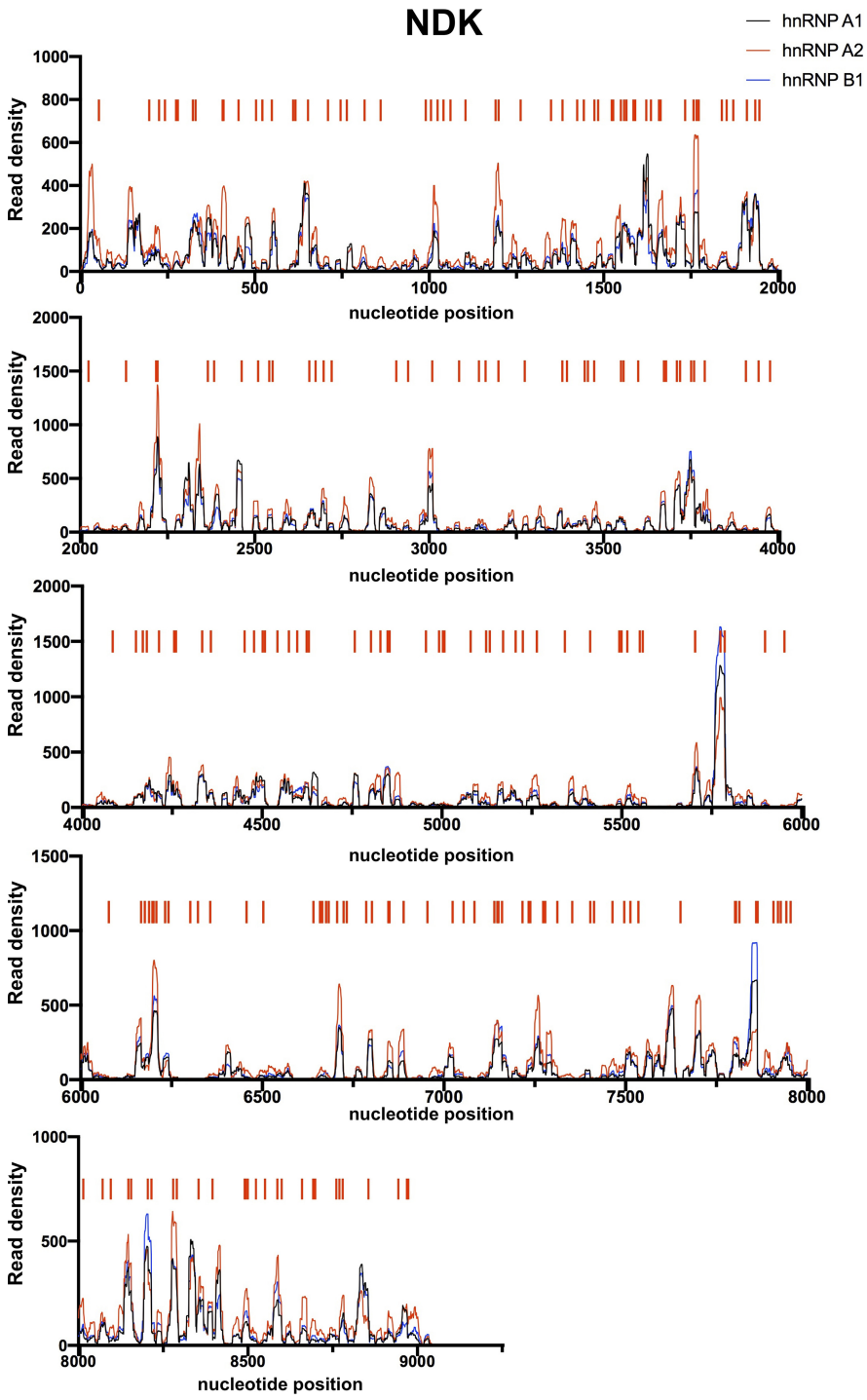

FIG 5 Binding sites of the hnRNP A1, A2, and B1 proteins on viral RNAs (NDK) derived from CLIP-seq. HEK293T cells stably expressing 3×HA-tagged hnRNP A1, A2, and B1 proteins were transfected with a full-length HIV-1_{NL4-3} proviral plasmid prior to CLIP-seq analysis. The frequency distribution of nucleotide occurrence (read density, *y* axis) of reads mapping to HIV-1_{NL4-3} is shown for hnRNP A1 (black), hnRNP A2 (red), and hnRNP B1 (blue). The positions of AGG motifs in the HIV-1 genome are indicated as red lines and are overlaid on the CLIP data.

elements with proximal uridines. The hnRNP H1 binding pattern was similar in 4-SU- and 6-SG-based CLIP experiments, but the relative frequencies of binding to certain sites varied somewhat. Binding sites near approximately nucleotide 3000 (site 4) and approximately nucleotide 5800 (site 6) were less frequently bound in 6-SG-based

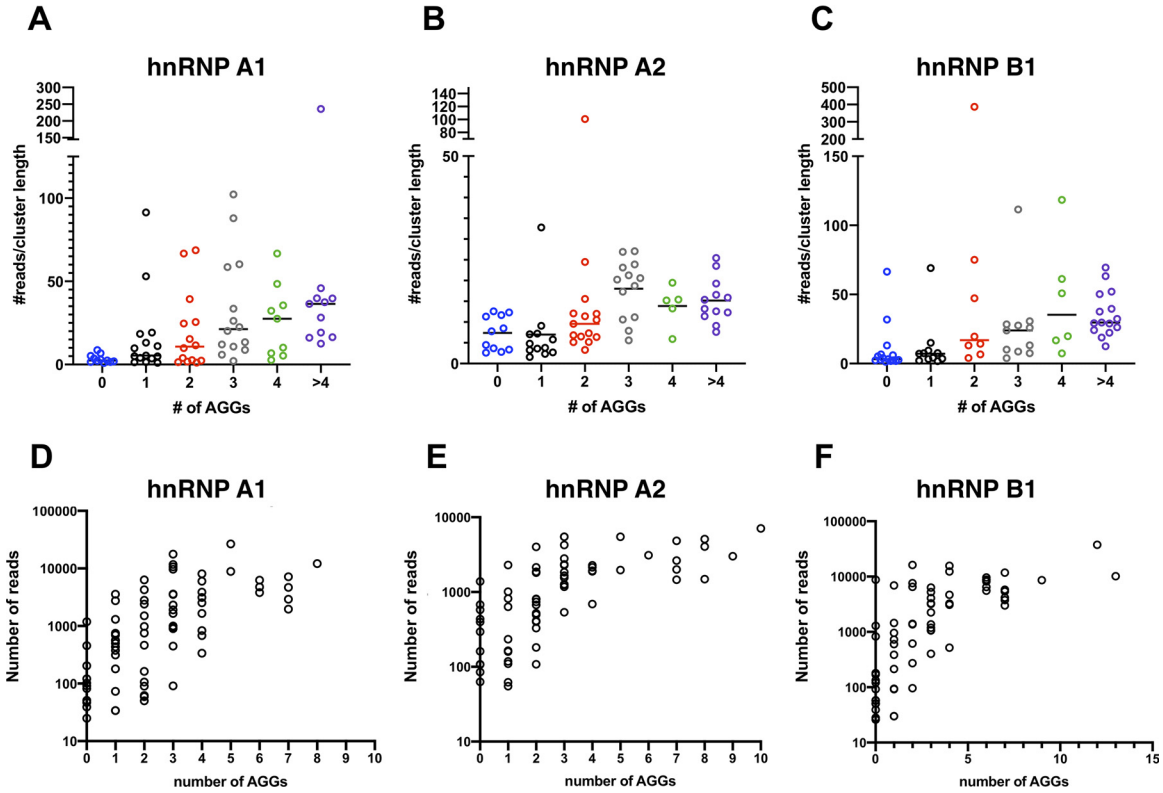

FIG 6 Increased occurrence of AGG motifs correlates with increased read density in hnRNP A1-, hnRNP A2-, and hnRNP B1-bound viral clusters. Reads that map to the viral genome (HIV-1_{NL4-3}) derived from the hnRNP A1, A2, and B1 CLIP experiments were formed into clusters (i.e., discrete binding sites) using the PARalyzer tool. (A to C) The number of AGG motifs in each cluster (*x* axis) and the corresponding normalized read density (number of reads/cluster length, *y* axis) in each cluster are shown. (D to F) Number of AGG motifs (*x* axis) and the corresponding number of reads (*y* axis) in each cluster are shown.

CLIP-seq experiments (Fig. 8C). Other sites, including site 5 (nucleotides 4325 to 4376), which includes the central polypurine tract, remained among the most frequently hnRNP H1-bound sites (Fig. 8C).

hnRNP H1 is one of a family of three closely related proteins (the others being hnRNP H2 and hnRNP F). To determine whether another member of this family exhibited similar binding properties, we mapped hnRNP F binding sites using CLIP-seq. The majority of the hnRNP F and hnRNP H1 binding sites, including site 2 (nucleotides 1384 to 1413), site 5 (nucleotides 4325 to 4376), and site 6 (nucleotides 5786 to 5836), overlapped, but some additional lower-frequency binding sites were observed (Fig. 8D). Overall, these findings corroborate the findings of the aforementioned CLIP-seq studies with hnRNP H1 (Fig. 7A and B) and indicate the existence of several specific binding sites for hnRNP H1 and closely related proteins on HIV-1 RNAs.

Calorimetric and NMR titrations validate hnRNP H1 RNA-binding preferences. Human hnRNP H1 consists of three quasi-RNA recognition motifs (qRRMs) and two glycine-rich domains. At its N terminus, qRRM1 and qRRM2 are connected through an approximately 20-amino-acid linker (referred to here as HqRRM1,2), whereas qRRM3 is embedded between the glycine-rich domains at its C terminus. While members of the hnRNP H/F family are known to bind G-rich sequences, our CLIP-seq experiments reveal an obvious A-rich preference surrounding these sequences (Fig. 8A). To quantitatively describe how hnRNP H1 binds purine-rich sequences, we carried out calorimetric and nuclear magnetic resonance (NMR) titrations. Calorimetric titrations of its HqRRM1,2 domain with an RNA oligonucleotide corresponding to site 5 showed that HqRRM1,2 binds site 5 with a moderate affinity (K_D [dissociation constant] = 0.62 μM) and with a stoichiometry of >2, likely as a result of the presence of three G tracts that are 4 to 6 nt each (Fig. 9A). The oligonucleotide A portion of the sequence did not appear to

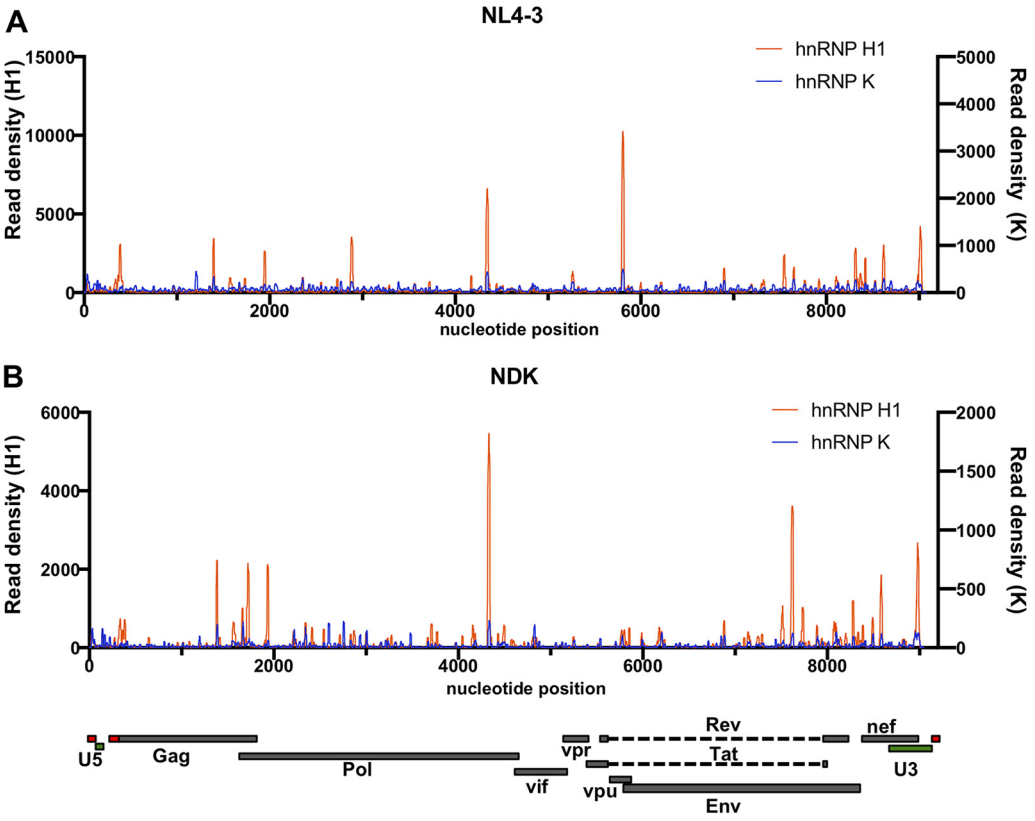

FIG 7 Binding of hnRNP H1 and K on the NL4-3 and NDK genomes. HEK293T cells stably expressing 3×HA-tagged hnRNP H1 and hnRNP K proteins were transfected with full-length HIV-1_{NL4-3} and HIV-1_{NDK} proviral plasmids prior to CLIP-seq analysis. For each chart, the position in the HIV-1 genome is represented on the *x* axis, and the schematic diagram of HIV-1 genome features shown below the charts is colinear. The frequency distribution of nucleotide occurrence (read density) of reads mapping to the HIV-1_{NL4-3} (A) and HIV-1_{NDK} (B) genomes derived from the hnRNP H1 (red) and hnRNP K (blue) CLIP experiments is shown.

contribute to binding affinity since titrations of HqRRM1,2 with a 5′-AAAAA-3′ oligonucleotide showed no binding (Fig. 9A). Furthermore, titrations with 5′-AGGAC-3′ revealed that binding requires >2 consecutive guanosines (Fig. 9A).

We next performed calorimetric titrations with model RNA oligonucleotides: 5′-AGGGG-3′, 5′-AGGGA-3′, and 5′-AGGGC-3′ (Fig. 9A). Similar experiments with a 5′-AGGGU-3′ oligonucleotide were previously reported (54). HqRRM1,2 bound all three oligonucleotides with moderate affinities (K_D [binding dissociation constant] = 0.19 to 2.4 μM). However, it bound 5′-AGGGG-3′ the tightest, followed by 5′-AGGGA-3′, 5′-AGGGU-3′ (54), and 5′-AGGGC-3′ (Fig. 9A), indicating a preference for a purine and for G in particular at position 5. The binding stoichiometry for the HqRRM1,2-(5′-AGGGG-3′) complex was <0.5, which we determined by NMR to be a result of a preexisting single-strand-to-G-quadruplex equilibrium of 5′-AGGGG-3′. The stoichiometric ratios for all other complexes are consistent with 1:1 binding, suggesting that these oligonucleotides predominantly adopt a single-strand conformation under solution conditions. Heteronuclear single quantum coherence (HSQC) titrations, wherein uniformly ^{15}N-labled HqRRM1,2 was titrated with unlabeled 5′-AGGGC-3′, revealed broadening of the NMR signals for HqRRM1,2 and indicated that binding specificity is governed by both the qRRM1 and qRRM2 domains (Fig. 9B and C), similar to our previous observations with 5′-AGGGU-3′ (54).

Effects of hnRNP depletion on HIV-1 splicing. We next tested how the RNAi-mediated knockdown of hnRNPs might alter HIV-1 alternative splicing, as well as viral protein expression and infectious virion yield, in single-cycle replication assays. Cells were infected with vesicular stomatitis virus G protein (VSV-G)-pseudotyped HIV-1_{NL4-3}

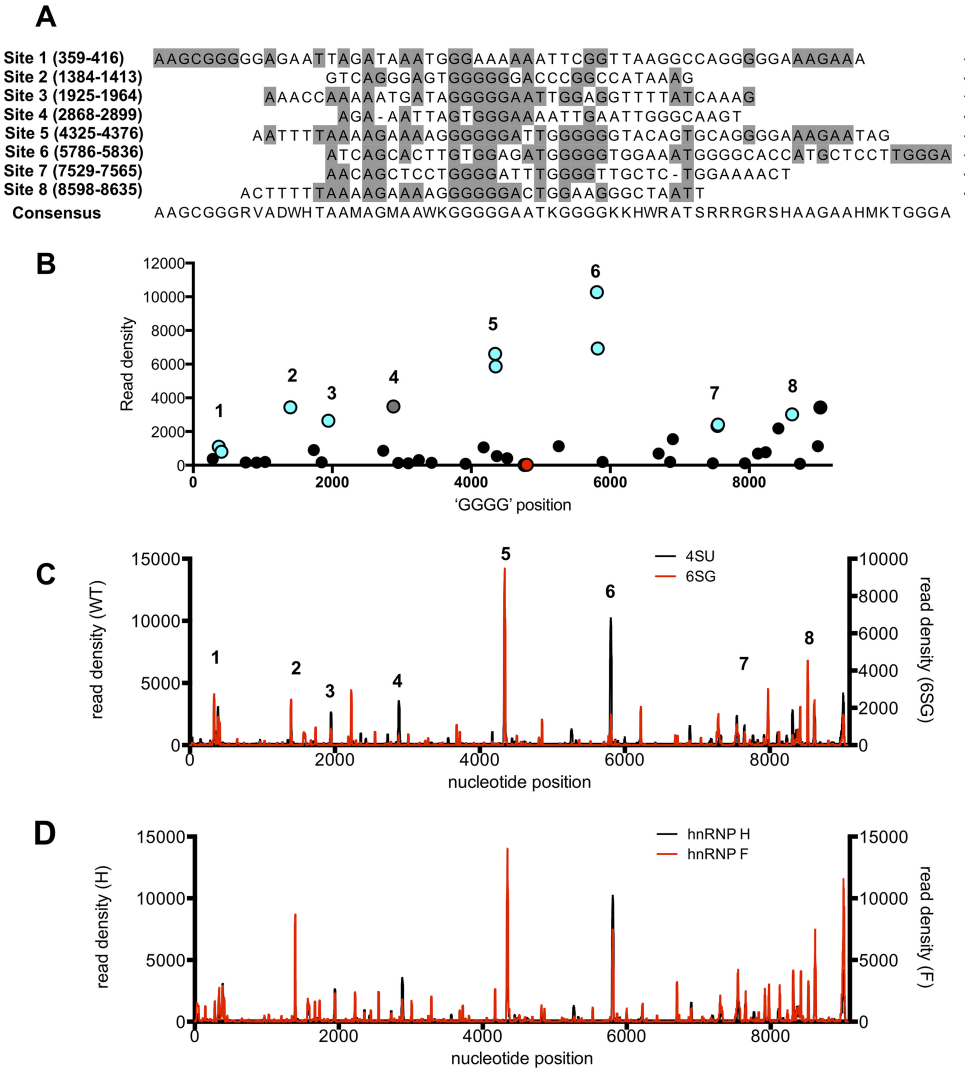

FIG 8 hnRNP H1 binds to purine-rich sequences on the viral genome. (A) Sequence alignment of the eight most frequently hnRNP H1-bound sites, as derived from CLIP-seq analysis shown in Fig. 2. (B) The locations of the 41 GGGG motifs on the HIV-1 genome (*x* axis) and the corresponding read density derived from the CLIP experiments (*y* axis) are shown. Blue circles indicate the eight most frequently bound sites, and red circles denote the GGGG motifs with the background read density. Note the presence of a high-frequency binding site that does not contain a GGGG motif, shown by a gray circle. (C and D) Comparison of the frequency distribution of nucleotide occurrence (read density) of reads mapping to the HIV-1_{NL4-3} genome derived from CLIP-seq experiments conducted using 4-SU- versus 6-SG-fed cells (C) or in hnRNP H1- versus hnRNP F-expressing cell lines (D).

(HIV-1_{NL4-3}/VSV-G) and HIV-1_{NDK} (HIV-1_{NDK}/VSV-G) following the transfection of small interfering RNA (siRNA) duplexes targeting hnRNP A1, A2/B1, and H1. As expected, the level of each hnRNP protein was reduced upon siRNA transfection, although in the case of hnRNP A1, the magnitude of the knockdown was modest (~2-fold) (Fig. 10A). We determined the effects of hnRNP knockdown on HIV-1 alternative splicing in cells using a next-generation sequencing-based assay incorporating the primer identifier (primer ID) approach (43) and targeting the 1.8-kb completely spliced class of mRNAs. Notably, each hnRNP depletion increased the use of splice acceptor A3 (Fig. 10B). Additionally, hnRNP A2/B1 and hnRNP H1 depletion reduced A1 utilization, and hnRNP H1 depletion also decreased A2 usage (Fig. 4B). hnRNP H1 depletion reduced both HIV-1_{NL4-3} (Fig. 10C) and HIV-1_{NDK} (Fig. 10D) Gag expression in cells, resulting in a commensurate decrease in the yield of infectious virus particles in cell culture supernatants (Fig. 10E). Conversely, and despite altering viral splicing, hnRNP A1 and hnRNP A2/B1 depletion had little or no effect on Gag expression and modestly decreased and increased the

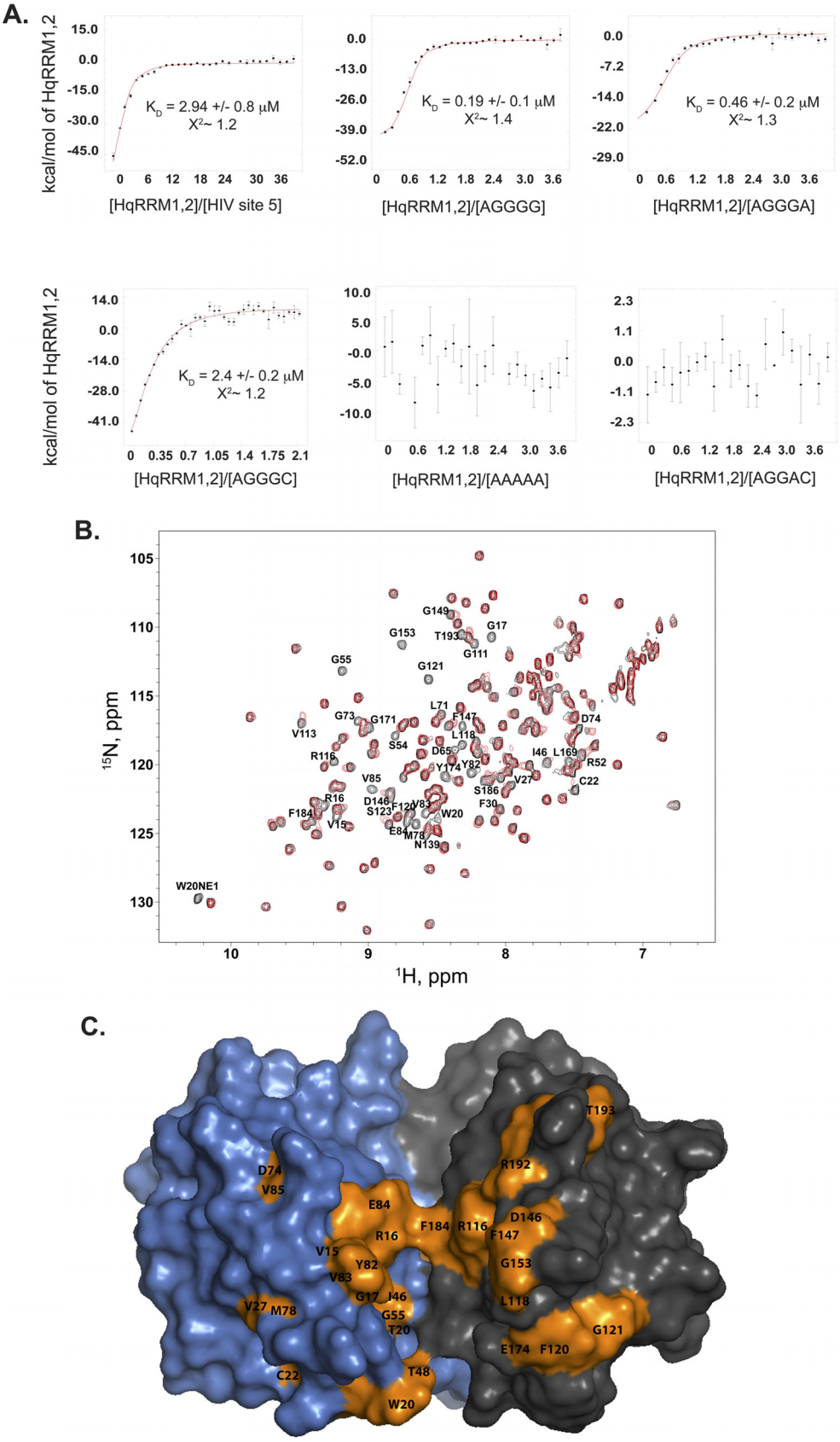

FIG 9 hnRNP H1 exhibits specificity for G-rich sequences *in vitro*. (A) Representative calorimetric titration profiles of HqRRM1,2 titrated into model oligomers. The titrations were performed at 298 K and in 20 mM sodium phosphate

(Continued on next page)

infectious virion yield, respectively (Fig. 10C to E), with the caveat that the magnitude of hnRNP A1 depletion was modest. Overall, depletion of each hnRNP had a distinguishable outcome on the overall pattern of viral splicing, particularly on the relative utilization of A1, A2, and A3, but had relatively modest effects on overall virion production. However, deleterious effects on infectious virion yield were clear and significant in the case of hnRNP H1 depletion.

Effect of mutations within hnRNP H1 binding sites on hnRNP H1 binding and HIV-1 splicing. Because hnRNP H1 bound to discrete sites on HIV-1 RNA (Fig. 7A and B) and its depletion was deleterious for the infectious virion yield (Fig. 10C to E), we next determined whether and how individual hnRNP H1 binding sites affected HIV-1 alternative mRNA splicing. Mutations within the eight most frequently bound sites were introduced, such that no more than two consecutive A or G nucleotides remained (Fig. 11A; Table 1). Although the majority of these mutations were synonymous, conservative amino acid substitutions (i.e., lysine to arginine) were introduced within hnRNP H1 binding sites when necessary (Table 1). Transfection of HEK293T cells with proviral plasmids yielded similar levels of infectious virus relative to the wild type (WT) for the majority of the mutants (Fig. 11B), with the exception of a mutation within site 8, within the 3′ polypurine tract. Mutations in site 8 are expected to reduce the infectiousness of virions from transfected cells irrespective of any potential effects on splicing, due to the requirement for the 3′ polypurine tract during reverse transcription. Finally, mutations introduced within all eight of the prominent hnRNP H1 binding sites abolished hnRNP H1 binding to those sites and did not lead to the emergence of other high-frequency binding sites (Fig. 11C).

HEK293T cells were transfected with one of eight HIV-1_{NL4-3} proviral plasmids containing mutations at an individual prominent hnRNP H1 binding site or a plasmid in which all eight hnRNP H1 binding site mutations were represented (Fig. 11A; Table 1). Thereafter, mRNA splicing was analyzed using the primer ID/next-generation sequencing approach (43). While the majority of mutations had modest effects on HIV-1 splicing, mutation of site 5 reduced the use of splice acceptor A1 to a similar degree as did hnRNP H1 knockdown but did not affect the utilization of other splice sites (Fig. 11D). This discrepancy could be due to indirect effects of hnRNP H1 knockdown and/or the complex combinatorial effects involving multiple hnRNP H1 binding sites. Notably, site 5 constituted the most frequently bound site in all hnRNP H1 CLIP-seq experiments and is situated proximal to splice acceptor A1 (Fig. 7A and 11A). Mutation of all eight prominent hnRNP H1 binding sites led to a more dramatic reduction of A1 utilization and modest perturbations of the use of other splice sites, suggesting a contribution of distal hnRNP H1 binding sites to the regulation of A1 use (Fig. 11D).

While hnRNP H1 depletion diminished A2 usage and slightly enhanced A3 usage, the hnRNP binding site mutations did not recapitulate these effects (compare Fig. 10B versus Fig. 11D). The enhancement of A3 usage following hnRNP H1 depletion was small (<2-fold), and because A3 is comparatively infrequently used, small changes in its utilization have a comparatively large fold effect. Nevertheless, the inability of the binding site mutations to completely recapitulate the effect of hnRNP H1 knockdown could be due to residual hnRNP H1 binding at sites other than sites 1 to 8. Alternatively, the effects of hnRNP H1 depletion on A2 utilization might be indirect.

FIG 9 Legend (Continued)
(pH 6.2), 20 mM NaCl, 4 mM TCEP. All titration data were processed and analyzed using Affinimeter software. The processed thermograms were fit to a 1:1 stoichiometric binding model; therefore, the stoichiometries reflect the concentrations of active binding sites. The values of the binding dissociation constants (K_D) and the corresponding standard deviations are from duplicate experiments. The goodness of fit (x^2) of the experimental data to the 1:1 binding model is reported for each titration. (B) Overlay of ^{1}H-^{15}N HSQC spectra of free HqRRM1,2 (red) and AGGGC-bound HqRRM1,2 (black) at a 4:1 molar ratio. Residues that completely disappear in the presence of saturating amounts of AGGGC are labeled. (C) Surface representations of HqRRM1,2 color coded by residues that disappear in the presence of saturating (4:1 molar ratio) amounts of AGGGC.

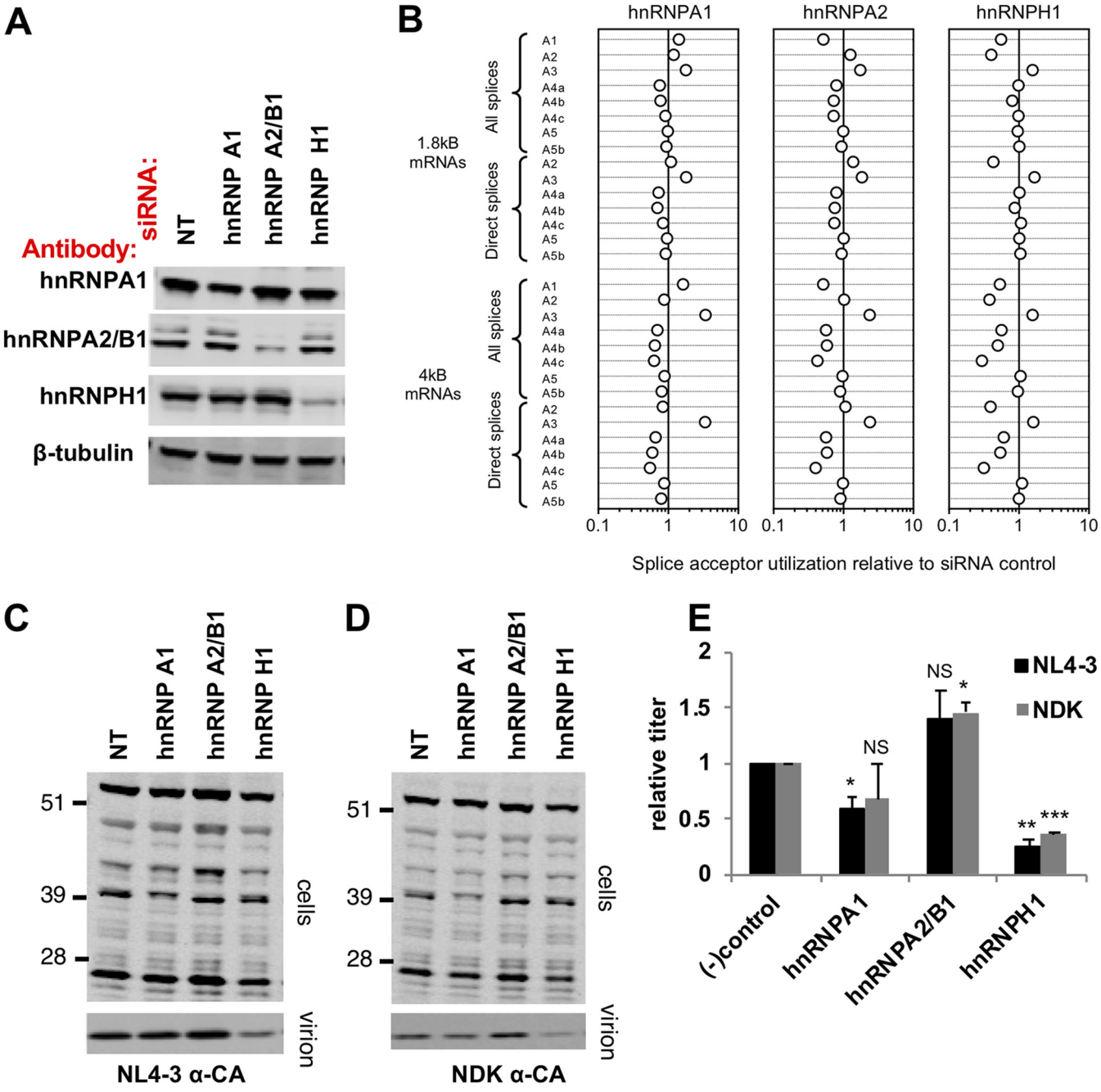

FIG 10 Effects of hnRNP depletion on HIV-1 splicing, Gag protein expression, and virion yield. HEK293T cells were transfected with siRNAs targeting the indicated hnRNPs and infected with HIV_{NL4-3}/VSV-G or HIV_{NDK}/VSV-G at 1 day posttransfection, and then the cells and viruses were harvested at 2 days postinfection for further analysis. (A) Expression of targeted hnRNPs was determined by immunoblotting against the indicated hnRNPs. (B) Analysis of the direct and all splicing events that occur at acceptors A1 to A5b in the 1.8-kb (completely spliced) class of viral RNAs. The fold change for unmanipulated controls in independent experiments and the fold change in hnRNP-depleted cells versus unmanipulated cells in a representative experiment are shown. (C, D) Western blot analysis of cell lysates and virions using an anti-CA antibody for HIV_{NL4-3} (C)- or HIV_{NDK} (D)-infected cells. Numbers to the left of the gels are molecular masses (in kilodaltons). (E) Virus titers in cell culture supernatants was determined on MT4-GFP indicator cells. Data are normalized to these for nontargeting siRNA-transfected cells (NT) and show the average ± SEM from three independent experiments. A paired *t* test was performed to assign statistical significance. NS, not significant; *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

Finally, we determined whether changes in splicing induced by perturbation of hnRNP H1 binding sites would affect viral gene expression. MT4 cells were infected with VSV-G-pseudotyped WT virus and viruses with a mutation within site 1 (mut1) through site 8 (mut8). The number of infected cells was held constant by adjusting the inoculum to a multiplicity of infection (MOI) of 0.5 IU per cell. Then, Gag, Vif, Nef, and Env protein

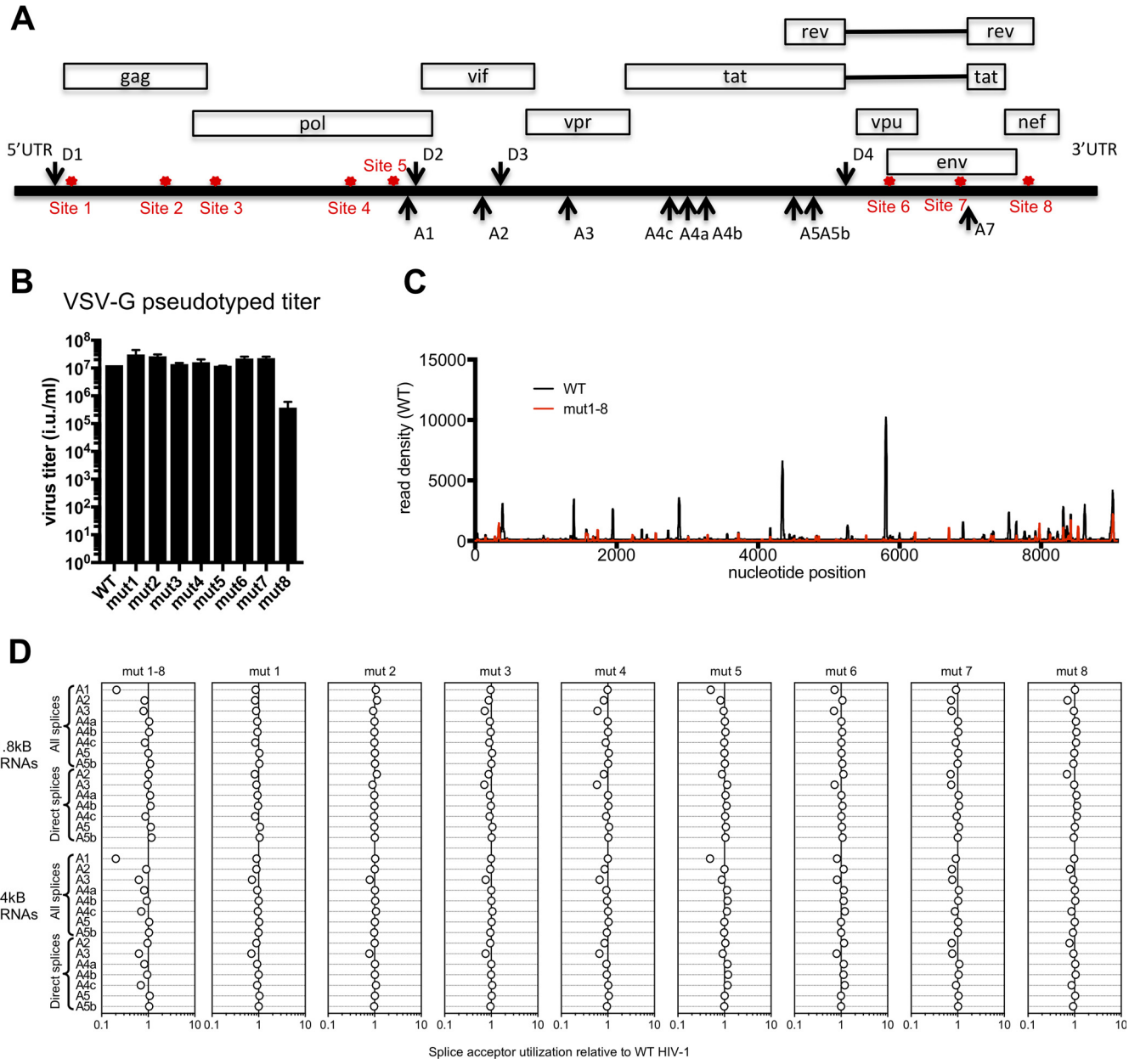

FIG 11 Mutation of hnRNP H1 binding sites alters HIV-1 alternative splicing. (A) Schematic diagram of the splice acceptors, donors, and hnRNP H1 binding sites (red asterisks) (not to scale). (B) Infectious virion yield for VSV-G-pseudotyped hnRNP H1 binding site mutants using TZM-bl target cells. (C) Frequency distribution of nucleotide occurrence (read density) of reads mapping to the HIV-1_{NL4-3} genome derived from CLIP-seq experiments conducted in hnRNP H1-expressing cells transfected with WT versus hnRNP H1 binding site mutant proviral clones. (D) Splice acceptor use for the 1.8-kb and 4-kb classes of RNAs is shown for viruses in which all eight hnRNP H1 binding sites were mutated (mut1-8) or viruses in which individual hnRNP H1 binding sites were mutated (mut1 to mut8).

levels were determined 2 days after infection. Mutations within the majority of sites did not have statistically significant effects on protein expression, but a mutation within site 5 (mut5) caused a specific decrease in Vif expression (Fig. 12A and B). This finding is concordant with the observed reduction in the use of A1, which controls Vif expression (Fig. 11D). Curiously, the mutation within site 8, within the 3′ polypurine tract, led to a significant decrease in Env and Nef expression (Fig. 12A and B). It is possible that reductions in Nef protein levels are a consequence of nonsynonymous substitutions in the underlying coding sequence (two Lys-Arg substitutions). Moreover, the mutations at site 8 resulted in the inadvertent introduction of a cluster of 5 CG dinucleotides within a 13-nucleotide stretch. This is

TABLE 1 Mutations of the hnRNP H1 binding sites assigned based on CLIP-seq experiments[a]

Site 1	WT- GGG GGA GAA TTA GAT AAA TGG GAA AAA ATT CGG TTA AGG CCA GGG GGA Mut-GGC GGA GAA TTA GAT AAG TTC GAC AAG ATT CGG TTA AGG CCA GGC GGC W→F E→D Gag
Site 2	WT- CAG GGA GTG GGG GGA CCC GGC CAT AAA Mut-CAA GGC GTC GGC GGC CCC GGC CAT AAG
Site 3	WT- AAA CCA AAA ATG ATA GGG GGA Mut-AAG CCC AAG ATG ATA GGC GGC
Site 4	WT- AAA TTA GTG GGA AAA TTG AAT TGG Mut-AAG TTA GTC GGC AAG CTC AAT TTC W→F Pol
Site 5	WT- AAA AGA AAA GGG GGG ATT GGG GGG TAC AGT GCA GGG GAA Mut-AAG AGG AAG GGC GGC ATT GGC GGC TAC AGT GCA GGC GAG
Site 6	WT- ATG GGG GTG GAA ATG GGG Mut-ATT CGG CTG GCG CTT CGG M→I G→R V→L E→A M→L G→R Vpu W→F K→R Env
Site 7	WT- CTG GGG ATT TGG GGT TGC TCT GGA AAA Mut-CTC GGC ATT TTC GGT TGC TCT GGC AAG W→F Env
Site 8	WT- TTT TTA AAA GAA AAG GGG GGA CTG GAA GGG Mut-TTT CTC CGC GAG CGC GGC GGA CTG GAA GGC K→R K→R Nef

[a]The mutations (in red) are derived from Fig. 7.

highly unusual in mammalian or viral mRNAs and could potentially result in their depletion (53). Note that control experiments conducted in the presence of the reverse transcriptase inhibitor nevirapine indicated that the viral proteins were synthesized *de novo* and do not represent input virus proteins (Fig. 12A).

Because the mut5 virus contained only synonymous mutations (Table 1) and the mutations in the mut5 virus induced a specific splicing perturbation at a proximal splice site (A1) and a lower level of Vif expression that could be ascribed to the loss of hnRNP binding, we tested whether the mut5 virus had reduced fitness in a competitive spreading replication assay. The mut5 virus replicated indistinguishably from WT HIV-1_{NL4-3} in the absence of competition, indicating that any fitness impact resulting from mutations at this site were modest (Fig. 12C). However, independent competition experiments in which H9 and CEMx174 cells, which express APOBEC3G (55), were infected with WT and mut5 virus revealed that the percentage of WT virus increased (Fig. 12D and E) over time. Overall, these experiments reveal that mutation of a novel *cis*-acting element that regulates A1 use and Vif expression imposes a small fitness deficit.

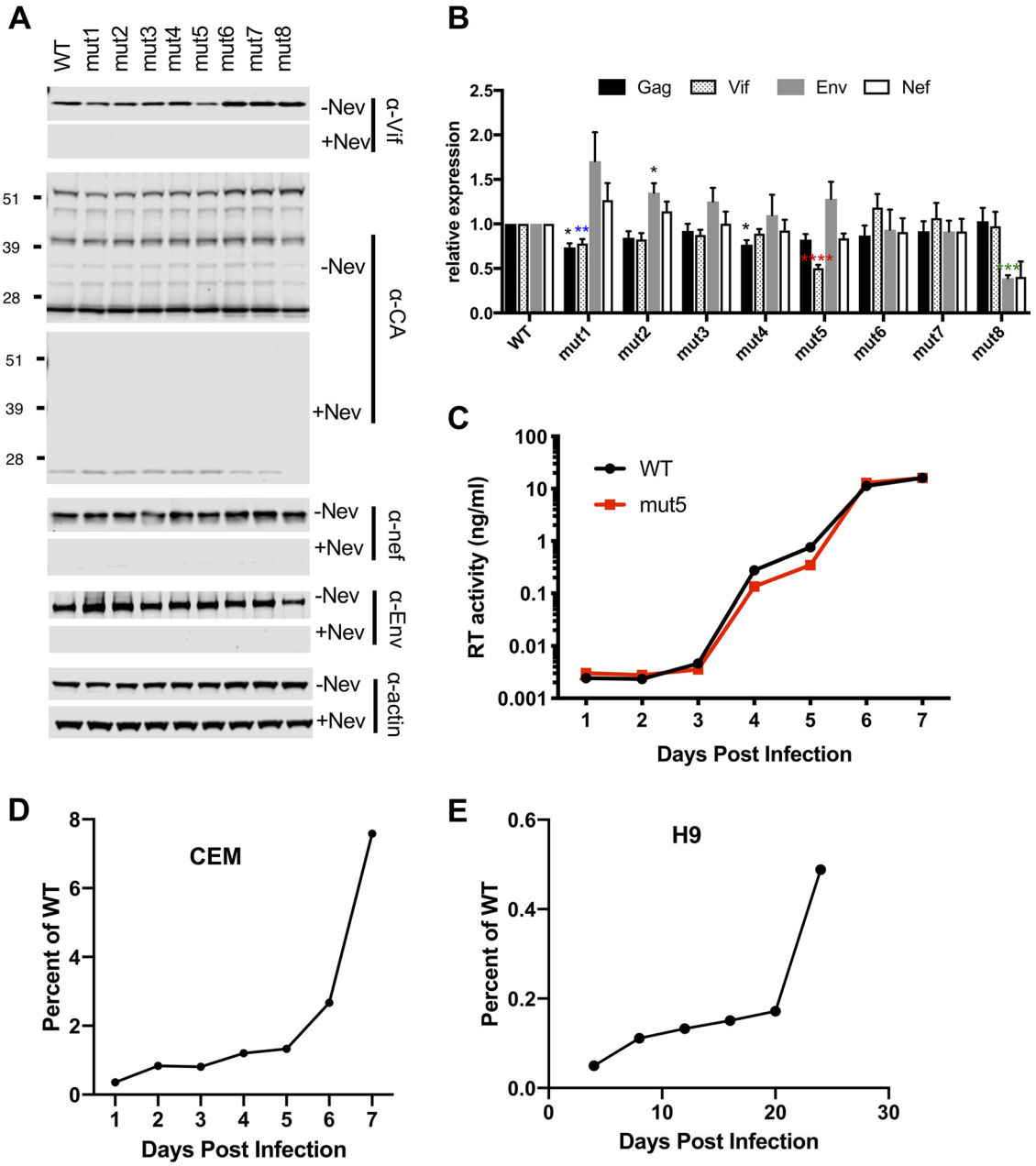

FIG 12 Mutation of hnRNP H1 binding sites alters HIV-1 gene expression and replication. (A, B) MT4 T cells were infected with the indicated VSV-G-pseudotyped hnRNP H1 mutant viruses at an MOI of 0.5 IU/cell in the absence and presence of 5 μM nevirapine (Nev). (A) Gag, Vif, Nef, Env, and actin protein levels in cell lysates were analyzed by immunoblotting 2 days later. (B) Mean values ± SEM from quantitation of Western blots from 4 independent experiments are shown. A paired *t* test was performed to assign statistical significance. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; ****, $P < 0.0001$. (C) MT4 cells were infected with WT and mut5 viruses at an MOI of 0.001 IU/cell, and virus spread was monitored using a quantitative PCR-based reverse transcriptase (RT) activity assay over 7 days. Data are representative of those from 4 independent replication experiments. (D) CEMx174 cells were infected with hnRNP H1 site 5 mutant virus. WT HIV-1_{NL4-3} was added to the established infection. Culture supernatant was harvested daily to sequence viral RNA and infect fresh cells. (E) H9 cells were infected as described in the legend to panel D. Supernatant was harvested every 4 days, and viral RNA was sequenced.

DISCUSSION

Alternative splicing of HIV-1 mRNAs has traditionally been studied using individual *cis*-acting sites and *trans*-acting proteins, sometimes in a nonnative context. In this study, we employed next-generation sequencing-based approaches to identify the binding targets of several hnRNP proteins on HIV-1 RNAs in cells and how these binding events directly affect alternative splicing of viral mRNAs in a native context. In contrast

to previous studies, CLIP-seq experiments revealed that members of the hnRNP A/B family bind broadly throughout viral RNA, in part by targeting frequently occurring AGG motifs. In contrast, hnRNP H1 bound to discrete G-rich sequence elements, often consisting of two or more copies of a GGGG sequence. As such, it would be expected that the viral genome contains far more numerous hnRNP A/B binding sites than hnRNP H1 binding sites, which is indeed what we observed.

Binding of hnRNP H1 at the most frequently bound site (site 5), near A1, enhanced splicing at this site. Indeed, mutations at this site led to a significant decrease in the use of A1 and Vif expression. While hnRNPs are generally thought to repress splicing at nearby sites, many act in a context-dependent manner and can exhibit splice enhancer or silencer activities (56). While we cannot exclude the possibility that the mutations employed herein inadvertently introduced an inhibitor element at site 5, the fact that the mutations phenocopied the effect of hnRNP H1 depletion, at least in terms of A1 usage, argues against this possibility.

Our finding that the hnRNP A1 and A2/B1 proteins bound to many sites on HIV-1 seems at odds with the notion that specific binding of these proteins to ESS and ISS elements controls splicing. However, hnRNPs can repress the assembly of spliceosomes through multimerization along exons (57). Therefore, the broad binding pattern of hnRNP A1 and A2/B1 proteins on viral RNA may, in part, be the result of an initial interaction with high-affinity binding sites (e.g., ESS or ISS elements), followed by oligomerization at lower-affinity binding sites. It is also possible that not all binding events identified by CLIP-seq are relevant to splicing regulation but reflect other roles for hnRNPs. Indeed, hnRNP proteins have known regulatory roles in the retention of unspliced viral genomic RNA in the nucleus (12, 13); in the nuclear export of viral/cellular RNAs (14); in mRNA translation, storage, degradation; and in stress responses (14, 15, 17, 18, 58). While at steady state the majority of hnRNP/SR proteins are localized in the nucleus, many hnRNP/SR proteins shuttle between the nucleus and the cytosol (16, 59). Because the CLIP-seq experiments were done using whole-cell lysates, it is possible that some of the binding events documented herein may serve roles other than splicing regulation.

Several previous studies have identified hnRNP H1 binding sites on HIV-1 RNA. Notably, however, none of these previously reported sites was bound at high levels by hnRNP H1 in our CLIP-seq experiments. For example, the proposed hnRNP H1 binding sites within *tat* exon 2 (32) and *tev*-specific exon 6D in certain HIV-1 isolates (33) consists of only one GGG element and no surrounding A-rich sequences. Notably, *in vitro* binding experiments previously revealed a preference for the presence of multiple copies of three or more G-rich sequences interspersed with other nucleotides for high-affinity binding for all members of the hnRNP H family (35). We found that the affinity of HqRRM1,2 toward the CLIP-seq-derived binding site 5, which contains three high-affinity binding sites for hnRNP H1, was similar to that toward the isolated G-rich sequence elements studied *in vitro*. This perhaps suggests that the ability of hnRNP H1 to bind to target sequences is driven not only by the affinity of individual RRMs but also by the avidity gained by multivalent interactions, which might explain why site 5 within the viral genome constitutes a primary binding site for hnRNP H1. Although hnRNP H1 did not bind A-rich sequences *in vitro*, the surrounding A-rich sequences may serve to maintain the hnRNP H1 binding G nucleotides in a single-stranded conformation, allowing easier access to hnRNP H1. Future structural studies will be crucial in determining how G-rich repeats affect RNA structure and how hnRNP H1 gains specificity for target sequences, given the apparent modest affinity of individual RRM binding *in vitro*.

Mutations introduced within the majority of individual hnRNP H1 binding sites modestly affected splicing. However, we noted that simultaneous mutation of all eight prominent hnRNP H1 binding sites led to a greater decrease in splice acceptor use. It is thus possible that a single hnRNP H1 molecule might bind at multiple positions on a single mRNA. This notion is consistent with the model that hnRNPs can loop out entire exons (60), which would require a minimum of two binding sites surrounding the exon. Previously, we showed that the N-terminal HqRRM1,2 didomain of hnRNP H1

exists in a conformational equilibrium between compact and extended structures (54). The compact structure has both qRRMs juxtaposed, generating a continuous RNA-binding surface that would accommodate only a single isolated G-tract element. Conversely the extended conformer may bind transcripts with multiple G tracts separated by short linker sequences. In particular, site 5, which shows a robust contribution to splicing, contains tandem 6-mer G tracts connected by a 5′-ATT-3′ linker. Similarly, others have shown that hnRNP H1 binds transcripts containing bipartite G-tract cores separated by a spacer of 0 to 4 nucleotides (61).

Together, our studies reveal the precise locations of elements on HIV-1 RNAs that are bound by hnRNPs and how these binding events regulate HIV-1 splicing. In doing so, we have revealed an unexpected splicing enhancer activity for hnRNP H1 and its *cis*-acting regulatory target element that modulates Vif expression. Finally, these studies provide a framework to further investigate how numerous *trans*-acting proteins and *cis*-acting RNA elements regulate HIV-1 alternative mRNA splicing both in a comprehensive manner and in greater detail.

MATERIALS AND METHODS

Cell culture, transfection, and infections. HeLa cell-derived TZM-bl MT4, H9, and CEMx174 cells (NIH AIDS Reagent Program) and HEK293T cells (ATCC CRL-11268) cells were maintained under standard conditions. Full-length HIV-1 infectious DNA plasmid clones and derivatives thereof were transfected into HEK293T cells either alone or together with a VSV-G-expression vector at a ratio of 4:1 using polyethylenimine (PolySciences). To generate HEK293T cells stably expressing 3×HA-tagged hnRNPs, cells were transduced with VSV-G-pseudotyped retroviral vectors. Following selection in medium supplemented with 50 μg/ml hygromycin, single-cell clones were subsequently isolated. For CLIP experiments, HEK293T cells stably expressing 3×HA-tagged hnRNPs were grown in 15-cm dishes and transfected with 30 μg of proviral plasmids using polyethylenimine (PolySciences). In other experiments, HEK293T cells grown in 24-well cell culture dishes were transfected with siRNA duplexes against hnRNPs using the Lipofectamine RNAimax reagent (Life Technologies). At 1 day post-siRNA transfection, cells were infected with VSV-G-pseudotyped HIV_{NL4-3} or $HIV\text{-}1_{NDK}$, and viral gene expression was analyzed 2 days later.

Antibodies. Antibodies were as follows: mouse monoclonal anti-HA.11 (catalog number MMS-101R; Covance), mouse monoclonal anti-hnRNP A1 (catalog number SC-32301; Santa Cruz), mouse monoclonal anti-hnRNP A2/B1 (catalog number SC-374053; Santa Cruz), goat polyclonal anti-hnRNP H1 (catalog number SC-10042; Santa Cruz), mouse monoclonal hnRNP K (catalog number SC-28380; Santa Cruz), and mouse monoclonal anti-HIV-1 p24CA (catalog number 183-H12-5C; NIH AIDS Reagent Program).

Plasmids. Full-length viral DNA plasmids $pHIV\text{-}1_{NL4-3}$ (NIH AIDS Reagents Program) and $pHIV\text{-}1_{NDK}$ were described previously (62, 63). N- and C-terminally 3×HA-tagged hnRNPs were cloned into the pCR3.1 and pLHCX (Clontech) vector backbones. hnRNP H1 binding site mutations were generated on the $pHIV\text{-}1_{NL4-3}$ viral DNA backbone by overlap-extension PCR (Table 1).

CLIP-seq experiments. CLIP-seq experiments were conducted as described previously (44). Briefly, HEK293T cells expressing 3×HA-tagged hnRNPs were transfected with viral DNA plasmids and grown in the presence of 4-thiouridine. Cells were washed and UV cross-linked at 365 nm. Following lysis, hnRNP-RNA complexes were immunoprecipitated using antihemagglutinin (anti-HA) monoclonal antibodies. Cross-linked RNA was end labeled with [γ-^{32}P]ATP and T4 polynucleotide kinase. Then, protein-RNA adducts were separated by SDS-PAGE, transferred to nitrocellulose membranes, exposed to film, and excised from the membranes, and the RNA was purified following proteinase K digestion. The RNA was sequentially ligated to 3′ and 5′ adapters, reverse transcribed, and PCR amplified, and the resulting DNA library was subjected to sequencing using Illumina platforms. The remainder of the data analysis was followed as detailed before (44), using the FASTX toolkit (http://hannonlab.cshl.edu/fastx_toolkit), the Bowtie algorithm (64), and the PARalyzer tool (65). The resulting reads were mapped to the human genome (hg19) or viral genome ($pHIV\text{-}1_{NL4-3}$ or $pHIV\text{-}1_{NDK}$) containing a single repeat (R) region using the Bowtie algorithm by allowing two mismatches (-v 2 -m 10 –best strata). Reads that mapped to the human genome were further clustered by PARalyzer using default settings and annotated using an in-house script. Motif analysis within binding sites was assessed using the cERMIT algorithm (46).

Competition assays. One million CEMx174 cells were infected with the mut5 virus, and after infection was established, WT $HIV\text{-}1_{NL4-3}$ was added to constitute <1% of the total viruses. Every day the cells were washed and resuspended in 10 ml medium for an additional 2 h to harvest newly released virus. Five milliliters of this supernatant was frozen, while the remaining 5 ml was used to infect 1 million fresh CEMx174 cells. The cells were incubated for 4 h, washed, resuspended in 10 ml of medium, and incubated for 24 h. This procedure was repeated daily. At the conclusion of the experiment, collected samples were concentrated using a Lenti-X concentrator (TaKaRa), resuspended in 220 μl AVL buffer from the QIAamp viral RNA minikit, and RNA extracted. Illumina sequencing library preparation was as described previously (44), except that the reverse primer GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC TNNNNNNNNNNNNNNATTACTACTGCCCCTTCACC was used for the cDNA step and the primer GCCTCC CTCGCGCCATCAGAGATGTGTATAAGAGACAGNNNNATTTCACCAGTACTACAGTTAAGGC was used as the forward primer in the first PCR amplification step. Ten microliters of sample from amplification PCR 1 was used as input for amplification PCR 2. The final amplicons were size selected using a BluePipppin system

(Sage Science) in a range of from 300 to 500 bp. Eluted amplicons were concentrated using AMPure XP beads (Beckman Coulter), sequenced using 15% phiX, and overloaded at a cluster density of 1,475/mm.

Alternatively, 1 million H9 cells were infected first with the mut5 virus followed by WT HIV-1_{NL4-3}, as described above. At 4 days after the HIV-1_{NL4-3} inoculation and then every 4 days, cells were spun down, the supernatant was harvested, and infected cells were resuspended in 10 ml medium. At the conclusion of the competition, viral RNA was isolated as described above and prepared for sequencing as in the second competition, with the BluePippin system size selection range being 400 to 700 bp.

Protein expression and purification. His_6-tagged qRRM1, qRRM2, qRRM3, and qRRM12 domains were overexpressed in *Escherichia coli* strain BL21(DE3). Cells were harvested in lysis buffer containing 20 mM Na_2HPO_4, 20 mM imidazole, 500 mM NaCl, 4 mM Tris(2-carboxyethyl)phosphine hydrochloride (TCEP) at pH 8.0 and lysed by sonication at 4°C. Recombinant proteins in the cleared lysate were purified using Ni-nitrilotriacetic acid agarose, the His_6 tag was removed, and the proteins were further purified by size exclusion chromatography with NMR buffer containing 20 mM sodium phosphate, 20 mM NaCl, 4 mM TCEP at pH 6.2.

Isothermal titration calorimetry. The binding affinities of qRRM1, qRRM2, qRRM3, and qRRM12 with the G-tract RNA oligonucleotides were measured using a Microcal VP-ITC calorimeter. The protein and RNA buffer solutions were exchanged to 20 mM sodium phosphate, 20 mM NaCl, 4 mM TCEP at pH 6.2 before use. All titrations were performed at 25°C.

NMR experiments. The NMR experiments were performed on a Bruker 800-MHz spectrometer at 32°C equipped with a cryoprobe. The ^{15}N-labeled and $^{15}N/^{13}C$-labeled samples for NMR experiments were prepared in a buffer containing 20 mM sodium phosphate, 20 mM NaCl, 4 mM TCEP, and 10% D_2O at pH 6.2. Backbone assignments of qRRM12 were obtained from HNCA, HNCACB, HNCO, and HN(CA)CO. For NMR titrations, the uniformed ^{15}N-labeled HqRRM1,2 samples were prepared at a concentration of 90 μM in 20 mM sodium phosphate, 20 mM NaCl, 4 mM TCEP, and 10% D_2O at pH 6.2. The unlabeled 5′-AGGGC-3′ oligonucleotide was added to uniformly ^{15}N-labeled HqRRM1,2 at molar ratios of 1:0:33, 1:0:66, 1:1, and 1:2.

ACKNOWLEDGMENTS

This work was supported by NIH grants U54GM103297 (the Center for HIV RNA Studies, to P.D.B., S.B.K., and R.I.S.), K22 AI116258 (to S.B.K.), R01 AI50111 (to P.D.B.), GM101979 (to B.S.T.), and F31 AI116406 (to A.E.). The work was also supported by the UNC Center for AIDS Research (NIH award P30 AI50410) and the UNC Lineberger Comprehensive Cancer Center (NIH award P30 CA16068).

REFERENCES

1. Karn J, Stoltzfus CM. 2012. Transcriptional and posttranscriptional regulation of HIV-1 gene expression. Cold Spring Harb Perspect Med 2:a006916. https://doi.org/10.1101/cshperspect.a006916.
2. Stoltzfus CM. 2009. Chapter 1. Regulation of HIV-1 alternative RNA splicing and its role in virus replication. Adv Virus Res 74:1–40. https://doi.org/10.1016/S0065-3527(09)74001-1.
3. Purcell DF, Martin MA. 1993. Alternative splicing of human immunodeficiency virus type 1 mRNA modulates viral protein expression, replication, and infectivity. J Virol 67:6365–6378.
4. Ocwieja KE, Sherrill-Mix S, Mukherjee R, Custers-Allen R, David P, Brown M, Wang S, Link DR, Olson J, Travers K, Schadt E, Bushman FD. 2012. Dynamic regulation of HIV-1 mRNA populations analyzed by single-molecule enrichment and long-read sequencing. Nucleic Acids Res 40:10345–10355. https://doi.org/10.1093/nar/gks753.
5. Cullen BR. 2003. Nuclear mRNA export: insights from virology. Trends Biochem Sci 28:419–424. https://doi.org/10.1016/S0968-0004(03)00142-7.
6. Malim MH, Hauber J, Le SY, Maizel JV, Cullen BR. 1989. The HIV-1 rev trans-activator acts through a structured target sequence to activate nuclear export of unspliced viral mRNA. Nature 338:254–257. https://doi.org/10.1038/338254a0.
7. Holmes M, Zhang F, Bieniasz PD. 2015. Single-cell and single-cycle analysis of HIV-1 replication. PLoS Pathog 11:e1004961. https://doi.org/10.1371/journal.ppat.1004961.
8. Krummheuer J, Lenz C, Kammler S, Scheid A, Schaal H. 2001. Influence of the small leader exons 2 and 3 on human immunodeficiency virus type 1 gene expression. Virology 286:276–289. https://doi.org/10.1006/viro.2001.0974.
9. Schwartz S, Felber BK, Benko DM, Fenyo EM, Pavlakis GN. 1990. Cloning and functional analysis of multiply spliced mRNA species of human immunodeficiency virus type 1. J Virol 64:2519–2529.
10. Madsen JM, Stoltzfus CM. 2005. An exonic splicing silencer downstream of the 3′ splice site A2 is required for efficient human immunodeficiency virus type 1 replication. J Virol 79:10478–10486. https://doi.org/10.1128/JVI.79.16.10478-10486.2005.
11. Fu XD, Ares M, Jr. 2014. Context-dependent control of alternative splicing by RNA-binding proteins. Nat Rev Genet 15:689–701. https://doi.org/10.1038/nrg3778.
12. Gordon H, Ajamian L, Valiente-Echeverria F, Levesque K, Rigby WF, Mouland AJ. 2013. Depletion of hnRNP A2/B1 overrides the nuclear retention of the HIV-1 genomic RNA. RNA Biol 10:1714–1725. https://doi.org/10.4161/rna.26542.
13. Olsen HS, Cochrane AW, Rosen C. 1992. Interaction of cellular factors with intragenic cis-acting repressive sequences within the HIV genome. Virology 191:709–715. https://doi.org/10.1016/0042-6822(92)90246-I.
14. Roy R, Durie D, Li H, Liu BQ, Skehel JM, Mauri F, Cuorvo LV, Barbareschi M, Guo L, Holcik M, Seckl MJ, Pardo OE. 2014. hnRNPA1 couples nuclear export and translation of specific mRNAs downstream of FGF-2/S6K2 signalling. Nucleic Acids Res 42:12483–12497. https://doi.org/10.1093/nar/gku953.
15. Vincendeau M, Nagel D, Brenke JK, Brack-Werner R, Hadian K. 2013. Heterogenous nuclear ribonucleoprotein Q increases protein expression from HIV-1 Rev-dependent transcripts. Virol J 10:151. https://doi.org/10.1186/1743-422X-10-151.
16. Pinol-Roma S, Dreyfuss G. 1993. hnRNP proteins: localization and transport between the nucleus and the cytoplasm. Trends Cell Biol 3:151–155. https://doi.org/10.1016/0962-8924(93)90135-N.
17. Han TW, Kato M, Xie S, Wu LC, Mirzaei H, Pei J, Chen M, Xie Y, Allen J, Xiao G, McKnight SL. 2012. Cell-free formation of RNA granules: bound RNAs identify features and components of cellular assemblies. Cell 149:768–779. https://doi.org/10.1016/j.cell.2012.04.016.
18. Kato M, Han TW, Xie S, Shi K, Du X, Wu LC, Mirzaei H, Goldsmith EJ, Longgood J, Pei J, Grishin NV, Frantz DE, Schneider JW, Chen S, Li L, Sawaya MR, Eisenberg D, Tycko R, McKnight SL. 2012. Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers

within hydrogels. Cell 149:753–767. https://doi.org/10.1016/j.cell.2012.04.017.
19. Abrahamyan LG, Chatel-Chaix L, Ajamian L, Milev MP, Monette A, Clement JF, Song R, Lehmann M, DesGroseillers L, Laughrea M, Boccaccio G, Mouland AJ. 2010. Novel Staufen1 ribonucleoproteins prevent formation of stress granules but favour encapsidation of HIV-1 genomic RNA. J Cell Sci 123:369–383. https://doi.org/10.1242/jcs.055897.
20. Valiente-Echeverria F, Melnychuk L, Vyboh K, Ajamian L, Gallouzi IE, Bernard N, Mouland AJ. 2014. eEF2 and Ras-GAP SH3 domain-binding protein (G3BP1) modulate stress granule assembly during HIV-1 infection. Nat Commun 5:4819. https://doi.org/10.1038/ncomms5819.
21. Han SP, Tang YH, Smith R. 2010. Functional diversity of the hnRNPs: past, present and perspectives. Biochem J 430:379–392. https://doi.org/10.1042/BJ20100396.
22. Smith SA, Ray D, Cook KB, Mallory MJ, Hughes TR, Lynch KW. 2013. Paralogs hnRNP L and hnRNP LL exhibit overlapping but distinct RNA binding constraints. PLoS One 8:e80701. https://doi.org/10.1371/journal.pone.0080701.
23. Ray D, Kazan H, Cook KB, Weirauch MT, Najafabadi HS, Li X, Gueroussov S, Albu M, Zheng H, Yang A, Na H, Irimia M, Matzat LH, Dale RK, Smith SA, Yarosh CA, Kelly SM, Nabet B, Mecenas D, Li W, Laishram RS, Qiao M, Lipshitz HD, Piano F, Corbett AH, Carstens RP, Frey BJ, Anderson RA, Lynch KW, Penalva LO, Lei EP, Fraser AG, Blencowe BJ, Morris QD, Hughes TR. 2013. A compendium of RNA-binding motifs for decoding gene regulation. Nature 499:172–177. https://doi.org/10.1038/nature12311.
24. Jain N, Lin HC, Morgan CE, Harris ME, Tolbert BS. 2017. Rules of RNA specificity of hnRNP A1 revealed by global and quantitative analysis of its affinity distribution. Proc Natl Acad Sci U S A 114:2206–2211. https://doi.org/10.1073/pnas.1616371114.
25. Caputi M, Mayeda A, Krainer AR, Zahler AM. 1999. hnRNP A/B proteins are required for inhibition of HIV-1 pre-mRNA splicing. EMBO J 18:4060–4067. https://doi.org/10.1093/emboj/18.14.4060.
26. Del Gatto-Konczak F, Olive M, Gesnel MC, Breathnach R. 1999. hnRNP A1 recruited to an exon in vivo can function as an exon splicing silencer. Mol Cell Biol 19:251–260. https://doi.org/10.1128/mcb.19.1.251.
27. Si ZH, Rauch D, Stoltzfus CM. 1998. The exon splicing silencer in human immunodeficiency virus type 1 Tat exon 3 is bipartite and acts early in spliceosome assembly. Mol Cell Biol 18:5404–5413. https://doi.org/10.1128/mcb.18.9.5404.
28. Damgaard CK, Tange TO, Kjems J. 2002. hnRNP A1 controls HIV-1 mRNA splicing through cooperative binding to intron and exon splicing silencers in the context of a conserved secondary structure. RNA 8:1401–1415. https://doi.org/10.1017/s1355838202023075.
29. Bilodeau PS, Domsic JK, Mayeda A, Krainer AR, Stoltzfus CM. 2001. RNA splicing at human immunodeficiency virus type 1 3′ splice site A2 is regulated by binding of hnRNP A/B proteins to an exonic splicing silencer element. J Virol 75:8487–8497. https://doi.org/10.1128/JVI.75.18.8487-8497.2001.
30. Tange TO, Damgaard CK, Guth S, Valcarcel J, Kjems J. 2001. The hnRNP A1 protein regulates HIV-1 tat splicing via a novel intron silencer element. EMBO J 20:5748–5758. https://doi.org/10.1093/emboj/20.20.5748.
31. Asai K, Platt C, Cochrane A. 2003. Control of HIV-1 env RNA splicing and transport: investigating the role of hnRNP A1 in exon splicing silencer (ESS3a) function. Virology 314:229–242. https://doi.org/10.1016/s0042-6822(03)00400-8.
32. Jacquenet S, Mereau A, Bilodeau PS, Damier L, Stoltzfus CM, Branlant C. 2001. A second exon splicing silencer within human immunodeficiency virus type 1 tat exon 2 represses splicing of Tat mRNA and binds protein hnRNP H. J Biol Chem 276:40464–40475. https://doi.org/10.1074/jbc.M104070200.
33. Caputi M, Zahler AM. 2002. SR proteins and hnRNP H regulate the splicing of the HIV-1 tev-specific exon 6D. EMBO J 21:845–855. https://doi.org/10.1093/emboj/21.4.845.
34. Jablonski JA, Buratti E, Stuani C, Caputi M. 2008. The secondary structure of the human immunodeficiency virus type 1 transcript modulates viral splicing and infectivity. J Virol 82:8038–8050. https://doi.org/10.1128/JVI.00721-08.
35. Schaub MC, Lopez SR, Caputi M. 2007. Members of the heterogeneous nuclear ribonucleoprotein H family activate splicing of an HIV-1 splicing substrate by promoting formation of ATP-dependent spliceosomal complexes. J Biol Chem 282:13617–13626. https://doi.org/10.1074/jbc.M700774200.
36. Widera M, Hillebrand F, Erkelenz S, Vasudevan AA, Munk C, Schaal H. 2014. A functional conserved intronic G run in HIV-1 intron 3 is critical to counteract APOBEC3G-mediated host restriction. Retrovirology 11:72. https://doi.org/10.1186/PREACCEPT-1252941606128392.
37. Lund N, Milev MP, Wong R, Sanmuganantham T, Woolaway K, Chabot B, Abou Elela S, Mouland AJ, Cochrane A. 2012. Differential effects of hnRNP D/AUF1 isoforms on HIV-1 gene expression. Nucleic Acids Res 40:3663–3675. https://doi.org/10.1093/nar/gkr1238.
38. Woolaway K, Asai K, Emili A, Cochrane A. 2007. hnRNP E1 and E2 have distinct roles in modulating HIV-1 gene expression. Retrovirology 4:28. https://doi.org/10.1186/1742-4690-4-28.
39. Marchand V, Santerre M, Aigueperse C, Fouillen L, Saliou J-M, Van Dorsselaer A, Sanglier-Cianférani S, Branlant C, Motorin Y. 2011. Identification of protein partners of the human immunodeficiency virus 1 tat/rev exon 3 leads to the discovery of a new HIV-1 splicing regulator, protein hnRNP K. RNA Biol 8:325–342. https://doi.org/10.4161/rna.8.2.13984.
40. Bieniasz PD, Kutluay SB. 2018. CLIP-related methodologies and their application to retrovirology. Retrovirology 15:35. https://doi.org/10.1186/s12977-018-0417-2.
41. Lee FCY, Ule J. 2018. Advances in CLIP technologies for studies of protein-RNA interactions. Mol Cell 69:354–369. https://doi.org/10.1016/j.molcel.2018.01.005.
42. Hafner M, Landthaler M, Burger L, Khorshid M, Hausser J, Berninger P, Rothballer A, Ascano M, Jr, Jungkamp AC, Munschauer M, Ulrich A, Wardle GS, Dewell S, Zavolan M, Tuschl T. 2010. Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. Cell 141:129–141. https://doi.org/10.1016/j.cell.2010.03.009.
43. Emery A, Zhou S, Pollom E, Swanstrom R. 2017. Characterizing HIV-1 splicing by using next-generation sequencing. J Virol 91:e02515-16. https://doi.org/10.1128/JVI.02515-16.
44. Kutluay SB, Zang T, Blanco-Melo D, Powell C, Jannain D, Errando M, Bieniasz PD. 2014. Global changes in the RNA binding specificity of HIV-1 gag regulate virion genesis. Cell 159:1096–1109. https://doi.org/10.1016/j.cell.2014.09.057.
45. Huelga SC, Vu AQ, Arnold JD, Liang TY, Liu PP, Yan BY, Donohue JP, Shiue L, Hoon S, Brenner S, Ares M, Jr, Yeo GW. 2012. Integrative genome-wide analysis reveals cooperative regulation of alternative splicing by hnRNP proteins. Cell Rep 1:167–178. https://doi.org/10.1016/j.celrep.2012.02.001.
46. Georgiev S, Boyle AP, Jayasurya K, Ding X, Mukherjee S, Ohler U. 2010. Evidence-ranked motif identification. Genome Biol 11:R19. https://doi.org/10.1186/gb-2010-11-2-r19.
47. Hutchison S, LeBel C, Blanchette M, Chabot B. 2002. Distinct sets of adjacent heterogeneous nuclear ribonucleoprotein (hnRNP) A1/A2 binding sites control 5′ splice site selection in the hnRNP A1 mRNA precursor. J Biol Chem 277:29745–29752. https://doi.org/10.1074/jbc.M203633200.
48. Bruun GH, Doktor TK, Borch-Jensen J, Masuda A, Krainer AR, Ohno K, Andresen BS. 2016. Global identification of hnRNP A1 binding sites for SSO-based splicing modulation. BMC Biol 14:54. https://doi.org/10.1186/s12915-016-0279-9.
49. Caputi M, Zahler AM. 2001. Determination of the RNA binding specificity of the heterogeneous nuclear ribonucleoprotein (hnRNP) H/H′/F/2H9 family. J Biol Chem 276:43850–43859. https://doi.org/10.1074/jbc.M102861200.
50. Thisted T, Lyakhov DL, Liebhaber SA. 2001. Optimized RNA targets of two closely related triple KH domain proteins, heterogeneous nuclear ribonucleoprotein K and alphaCP-2KL, suggest distinct modes of RNA recognition. J Biol Chem 276:17484–17496. https://doi.org/10.1074/jbc.M010594200.
51. Okunola HL, Krainer AR. 2009. Cooperative-binding and splicing-repressive properties of hnRNP A1. Mol Cell Biol 29:5620–5631. https://doi.org/10.1128/MCB.01678-08.
52. van der Kuyl AC, Berkhout B. 2012. The biased nucleotide composition of the HIV genome: a constant factor in a highly variable virus. Retrovirology 9:92. https://doi.org/10.1186/1742-4690-9-92.
53. Takata MA, Goncalves-Carneiro D, Zang TM, Soll SJ, York A, Blanco-Melo D, Bieniasz PD. 2017. CG dinucleotide suppression enables antiviral defence targeting non-self RNA. Nature 550:124–127. https://doi.org/10.1038/nature24039.
54. Penumutchu SR, Chiu LY, Meagher JL, Hansen AL, Stuckey JA, Tolbert BS. 2018. Differential conformational dynamics encoded by the linker between quasi RNA recognition motifs of heterogeneous nuclear ribonucleoprotein H. J Am Chem Soc 140:11661–11673. https://doi.org/10.1021/jacs.8b05366.
55. Sheehy AM, Gaddis NC, Choi JD, Malim MH. 2002. Isolation of a human

gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. Nature 418:646–650. https://doi.org/10.1038/nature00939.
56. Busch A, Hertel KJ. 2012. Evolution of SR protein and hnRNP splicing regulatory factors. Wiley Interdiscip Rev RNA 3:1–12. https://doi.org/10.1002/wrna.100.
57. Zhu J, Mayeda A, Krainer AR. 2001. Exon identity established through differential antagonism between exonic splicing silencer-bound hnRNP A1 and enhancer-bound SR proteins. Mol Cell 8:1351–1361. https://doi.org/10.1016/S1097-2765(01)00409-9.
58. Swanson CM, Sherer NM, Malim MH. 2010. SRp40 and SRp55 promote the translation of unspliced human immunodeficiency virus type 1 RNA. J Virol 84:6748–6759. https://doi.org/10.1128/JVI.02526-09.
59. Caceres JF, Screaton GR, Krainer AR. 1998. A specific subset of SR proteins shuttles continuously between the nucleus and the cytoplasm. Genes Dev 12:55–66. https://doi.org/10.1101/gad.12.1.55.
60. Martinez-Contreras R, Fisette JF, Nasim FU, Madden R, Cordeau M, Chabot B. 2006. Intronic binding sites for hnRNP A/B and hnRNP F/H proteins stimulate pre-mRNA splicing. PLoS Biol 4:e21. https://doi.org/10.1371/journal.pbio.0040021.
61. Dominguez D, Freese P, Alexis MS, Su A, Hochman M, Palden T, Bazile C, Lambert NJ, Van Nostrand EL, Pratt GA, Yeo GW, Graveley BR, Burge CB. 2018. Sequence, structure, and context preferences of human RNA binding proteins. Mol Cell 70:854–867.e9. https://doi.org/10.1016/j.molcel.2018.05.001.
62. Spire B, Sire J, Zachar V, Rey F, Barre-Sinoussi F, Galibert F, Hampe A, Chermann JC. 1989. Nucleotide sequence of HIV1-NDK: a highly cytopathic strain of the human immunodeficiency virus. Gene 81:275–284. https://doi.org/10.1016/0378-1119(89)90188-1.
63. Adachi A, Gendelman HE, Koenig S, Folks T, Willey R, Rabson A, Martin MA. 1986. Production of acquired immunodeficiency syndrome-associated retrovirus in human and nonhuman cells transfected with an infectious molecular clone. J Virol 59:284–291.
64. Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10:R25. https://doi.org/10.1186/gb-2009-10-3-r25.
65. Corcoran DL, Georgiev S, Mukherjee N, Gottwein E, Skalsky RL, Keene JD, Ohler U. 2011. PARalyzer: definition of RNA binding sites from PAR-CLIP short-read sequence data. Genome Biol 12:R79. https://doi.org/10.1186/gb-2011-12-8-r79.