The human endogenous retrovirus K Rev response element coincides with a predicted RNA folding region

J. Yang, H. Bogerd, S. Y. Le and B. R. Cullen

RNA 2000 6: 1551-1564

RNA (2000), *6*:1551–1564. Cambridge University Press. Printed in the USA.

The human endogenous retrovirus K Rev response element coincides with a predicted RNA folding region

JIN YANG,[1] HAL BOGERD,[1] SHU-YUN LE,[2] and BRYAN R. CULLEN[1]

[1]Department of Genetics and Howard Hughes Medical Institute, Duke University Medical Center, Durham, North Carolina 27710, USA
[2]Laboratory of Experimental and Computational Biology, Frederick Cancer Research Center, National Cancer Institute, National Institutes of Health, Frederick, Maryland 21702, USA

ABSTRACT

Human endogenous retrovirus K (HERV-K) is the name given to an ~30-million-year-old family of endogenous retroviruses present at >50 copies per haploid human genome. Previously, the HERV-K were shown to encode a nuclear RNA export factor, termed K-Rev, that is the functional equivalent of the H-Rev protein encoded by human immunodeficiency virus type 1. HERV-K was also shown to contain a *cis*-acting target element, the HERV-K Rev response element (K-RRE), that allowed the nuclear export of linked RNA transcripts in the presence of either K-Rev or H-Rev. Here, we demonstrate that the functionally defined K-RRE coincides with a statistically highly significant unusual RNA folding region and present a potential RNA secondary structure for the ~416-nt K-RRE. Both in vitro and in vivo assays of sequence specific RNA binding were used to map two primary binding sites for K-Rev, and one primary binding site for H-Rev, within the K-RRE. Of note, all three binding sites map to discrete predicted RNA stem-loop subdomains within the larger K-RRE structure. Although almost the entire 416-nt K-RRE was required for the activation of nuclear RNA export in cells expressing K-Rev, mutational inactivation of the binding sites for K-Rev resulted in the selective loss of the K-RRE response to K-Rev but not to H-Rev. Together, these data strongly suggest that the K-RRE, like the H-RRE, coincides with an extensive RNA secondary structure and identify specific sites within the K-RRE that can recruit either K-Rev or H-Rev to HERV-K RNA transcripts.

Keywords: nuclear export; retrovirus; Rev; RNA binding; RNA structure

INTRODUCTION

All retroviruses must express both unspliced and spliced forms of their initial, genome-length RNA transcript during the viral replication cycle (reviewed by Pollard & Malim, 1998; Cullen, 2000). Therefore, retroviral replication is dependent on the nuclear export of not only fully spliced viral mRNAs that are comparable to cellular mRNAs, but also of incompletely spliced viral mRNAs that are essentially analogous to cellular pre-mRNAs. However, eukaryotic cells have evolved mechanisms to block the access of such pre-mRNAs to cellular mRNA export factors until processing is complete. Retroviruses have therefore had to evolve mechanisms for the nuclear export of their incompletely spliced transcripts that are at least in part distinct from the canonical cellular mRNA export pathway.

Reprint requests to: Dr. Bryan R. Cullen, Howard Hughes Medical Institute, Duke University Medical Center, Room 426 CARL Building, Research Drive, Durham, North Carolina 27710, USA; e-mail: culle002@mc.duke.edu.

At least two different mechanisms utilized by retroviruses for the nuclear export of their unspliced transcripts have been identified. Several simple retroviruses, including simian type D viruses and avian leukemia viruses, contain a structured RNA element, termed a constitutive transport element (CTE), that directly recruits a cellular RNA export factor (Bray et al., 1994). In the case of the simian type D viruses, this cellular cofactor has been identified as Tap, a protein that is also believed to play a critical role in the final stages of cellular mRNA export (Grüter et al., 1998; Kang & Cullen, 1999; Katahira et al., 1999). Unlike simple retroviruses, several more complex retroviruses utilize a virally encoded adaptor protein, termed Rev in human immunodeficiency virus type 1 (HIV-1) and Rex in human T-cell leukemia virus type I (HTLV-I), to recruit a distinct cellular nuclear export factor, termed Crm1, to incompletely spliced viral transcripts (Fornerod et al., 1997; Neville et al., 1997; Stade et al., 1997). Thus, Rev and Rex bind not only to a highly structured viral RNA element termed the Rev response element (RRE) in HIV-1

and the Rex response element (RxRE) in HTLV-I (Hanly et al., 1989; Malim et al., 1989), but also, via a short leucine-rich sequence, to the Crm1 nuclear export factor. Although Crm1 is unrelated to Tap and does not play a direct role in cellular mRNA export, Crm1 and Tap share the ability to target viral RNAs to the cytoplasm by directly binding to components of the nuclear pore complex (Neville et al., 1997; Katahira et al., 1999).

Recently, a third protein functionally analogous to HIV-1 Rev (H-Rev) and HTLV-I Rex has been identified in human endogenous retrovirus K (HERV-K) (Magin et al., 1999; Yang et al., 1999). The HERV-K are a family of endogenous viruses, estimated to be present at between 50 and 170 copies per haploid human genome, that first entered the human germ line ~30 million years ago (Medstrand & Mager, 1998; Tristem, 2000). The HERV-K are not closely related to any currently known exogenous retroviruses, and bear little similarity to either HIV-1 or HTLV-I (Griffiths et al., 1997). Nevertheless, the original exogenous form of HERV-K faced the same problem of how to get genome-length, unspliced viral transcripts out of the nucleus in the face of nuclear retention induced by unused splice sites. The solution achieved by the HERV-K proved to be very similar to that seen in HIV-1 and HTLV-I, that is, the HERV-K encode a small adaptor protein, here termed K-Rev, that binds to both the Crm1 nuclear export factor and to a *cis*-acting RNA element, termed the K-RRE (Magin et al., 1999; Yang et al., 1999). However, little is known about the K-RRE except that it is located within the U3 region of the HERV-K long terminal repeat (LTR) and is ≤434 nt in length.

In this article, we describe a computational and mutational analysis of the K-RRE RNA element. We have found that the K-RRE coincides with a highly significant unusual RNA folding region that is conserved across known HERV-K genomes, and propose an RNA stem-loop structure for the K-RRE that is 416 nt in length. Mutational analysis, combined with in vitro and in vivo assays for K-Rev binding as well as functional assays, was used to support the existence of this predicted RNA structure and to identify two primary binding sites for K-Rev and, interestingly, one primary binding site for H-Rev. Of note, each of these binding sites coincides with a discrete predicted stem-loop subdomain of the K-RRE. These data reveal that K-Rev function, like H-Rev and Rex function, is dependent on the recruitment of several K-Rev molecules to discrete RNA-binding sites that form part of a viral RNA sequence that is predicted to fold into a complex RNA secondary structure.

RESULTS

The biological activity of the orthologs of K-Rev observed in exogenous retroviruses, such as H-Rev and HTLV-I Rex, requires the recruitment of these viral regulatory proteins to a discrete primary RNA-binding site that is presented in the context of a larger region of tightly folded RNA termed a "response element" (Hanly et al., 1989; Malim et al., 1989, 1990; Zapp & Green, 1989; Bartel et al., 1991; Heaphy et al., 1991; Bogerd et al., 1992). Although the primary binding site is clearly essential for biological activity, it is not sufficient and the H-Rev- or Rex-induced export of a target viral RNA therefore requires an almost full-length RNA response element located in *cis* (Malim et al., 1989; Huang et al., 1991). Because the biological activity of Rev or Rex response elements is dependent on their folding, these RNA elements can be effectively identified, and their structure predicted, using computer analysis. Indeed, structures for several retroviral response elements, including both the ~234-nt HIV-1 RRE and the ~254-nt HTLV-I, were first predicted by computer analysis and only subsequently validated by mutational analysis (Hanly et al., 1989; Malim et al., 1989; Le et al., 1990).

The K-RRE coincides with a predicted unusual RNA folding region

Previously, we and others have used functional assays to roughly map the extent of the HERV-K Rev response element (K-RRE) to between residues 8719 and 9152 according to the prototypic HERV-K10 proviral sequence reported by Ono et al. (1986) (Magin et al., 1999; Yang et al., 1999). This 434-nt RNA sequence, which is entirely located in the U3 region of the HERV-K LTR, is considerably larger than either the H-RRE or the HTLV-I RxRE, and it therefore appeared likely that the actual extent of the K-RRE would be smaller than this initial estimate.

To gain insight into the true size of the K-RRE, and to determine whether the K-RRE indeed coincides with an unusual folding region (UFR), we used the same computational approach that previously allowed the definition of not only the H-RRE and the HTLV-I RxRE but also the HIV-2 RRE and the Visna-Maedi Virus RRE (V-RRE) (Hanly et al., 1989; Malim et al., 1989; Le et al., 1990; Tiley & Cullen, 1992). Initially, we focused our attention on sequences extending from 8504 to 9250 within the HERV-K108 proviral clone (Barbulescu et al., 1999), a sequence that fully includes, but extends significantly beyond, the functionally defined 8719–9152 K-RRE sequence. Monte Carlo simulations were used to search for any UFR within this HERV-K sequence by calculation of the statistical significance and thermodynamic stability of local secondary structures in the HERV-K108 sequence (Le et al., 1990; 1993). UFR were searched for in the 8504–9250 HERV-K sequence by calculating the significance score (Sigscr) and the stability score (Stbscr) for window sizes varying from 100 nt to 450 nt with a window step size of 2 nt. For each window, the calculation was carried out by sliding 1 nt along the search sequence. The

optimal window size obtained in this simulation was 416 nt and this gave rise to the most stable and highly significant UFR detected for all window sizes tested (Table 1). Most importantly, this 416-nt segment extends from 8721 to 9136 within the HERV-K 108 proviral genome, a sequence that is equivalent to residues 8720–9135 in the prototypic HERV-K 10 sequence described by Ono et al. (1986). Therefore, this predicted RNA folding region is fully contained within the 8719–9152 sequence previously shown to display full K-RRE function (Magin et al., 1999; Yang et al., 1999).

To further examine the significance of the detected UFR, we next slid the 416-nt window along the entire HERV-K108 sequence and computed the relevant Sigscr and Stbscr. As may be seen in Figure 1, the unusual RNA folding region that coincides with the K-RRE is highly significant both in absolute terms and by comparison to the remainder of the HERV-K genome. Of note, the fact that the K-RRE is located within the U3 region of the viral LTR means that there are two predicted unusual folding regions in the integrated HERV-K provirus. However, only the more 3′ K-RRE would be predicted to be transcribed.

The HERV-K108 provirus was chosen for this computational analysis because it is full-length, because it contains U3 sequences that are very similar to the biologically active K-RRE used in our previous work (98.3% identical), and because the 5′ and 3′ LTR in the K108 endogenous provirus differ at only six positions, thus implying minimal genetic drift in the HERV-K108 K-RRE sequence over time (Barbulescu et al., 1999). In general, the U3 region of retroviral LTR is however the least highly conserved part of the retroviral genome, and we therefore wished to ask if the unusual folding region present in HERV-K108 is also present in other HERV-K proviruses that have entered the human genome over several million years of primate evolution. In Table 1, we present the calculated significance and stability scores of unusual RNA folding regions identified by Monte Carlo simulations in a series of distinct HERV-K proviruses initially characterized and sequenced by the Lenz laboratory (Barbulescu et al., 1999). As may be seen, all of these HERV-K proviruses also contain predicted unusual RNA folding regions, in more or less the same location in both LTR U3 regions. In comparison, the 416-nt K-RRE sequence used in our previous biological analysis of K-RRE function (Yang et al., 1999) is predicted to have a significance score of −3.59 standard deviation units and a free energy of folding of −170.2 kcal/mol (Table 1).

TABLE 1. A conserved unusual RNA folding region in HERV-K is both stable and highly significant.

HERV-K provirus (accession number)	Location in sequence	Significance score	Stability score	Free energy
HERV-K108	217–632	−3.05	−3.32	−163.8
(AF164614)	8721–9136	−3.10	−3.35	−162.8
HERV-K101	217–632	−2.44	−3.18	−159.0
(AF164609)	8428–8843	−2.44	−3.18	−159.0
HERV-K102	217–632	−3.01	−3.60	−162.7
(AF164610)	8428–8843	−2.22	−3.07	−155.6
HERV-K103	217–632	−3.46	−3.20	−164.1
(AF164611)	8430–8845	−2.95	−3.14	−161.3
KERV-K104	209–614	−2.90	−2.95	−153.0
(AF164612)	8694–9109	−2.70	−3.00	−159.6
HERV-K109	209–624	−3.03	−3.42	−162.0
(AF164615)	8671–9086	−2.94	−3.36	−160.7
K-RRE	2–417	−3.59	ND[a]	−170.2

Computed significance and stability scores are given in standard deviation units and the free energy is given in kcal/mol and computed by MFOLD (version 3.1). The K-RRE listed here is the biologically active K-RRE sequence shown in Figure 2.

[a]ND: not done.

Given that the K-RRE, like other RREs and RxREs, indeed coincides with a highly significant and conserved RNA folding region, we next wished to predict the likely folding of the biologically active K-RRE reported previously (Yang et al., 1999). The candidate K-RRE structure shown in Figure 2 was computed by EFFOLD (Le et al., 1993) and checked by Zucker's MFOLD (Mathews et al., 1999). In this calculation, EFFOLD generated 50 simulated energy rules in which the uncertainties of the free-energy parameters in the Turner energy rules (Freier et al., 1986; Jaeger et al., 1989) were fluctuated by a normal approximation within the range of experimental errors for these parameters. For each simulated energy rule, the lowest free-energy structure was searched and the optimal prediction assessed based on the computed 50 structures. The RNA structural model of the K-RRE that was obtained (Fig. 2) is highly stable and only a small number of reasonable local folding possibilities were identified. In fact, 46 out of the 50 lowest free-energy structures computed by EFFOLD are essentially identical with the structural model shown in Figure 2. In this K-RRE structure, and in all subsequent experiments, the indicated 5′ nucleotide, which coincides with residue 8719 in the HERV-K10 sequence reported by Ono et al. (1986), is referred to as position 1.

Deletion analysis of the K-RRE

Although the candidate K-RRE structure given in Figure 2 provides a good template for the design of mutants of the K-RRE, we initially chose to construct nested random 5′ and 3′ deletion mutants of the K-RRE and thereby avoid any bias in mutant design. A set of 13 deletion mutants of the K-RRE was derived and assayed for their ability to mediate RNA export (Fig. 3) or to bind to K-Rev in vitro (Fig. 4A) or in vivo (Fig. 4B). These data are summarized in Table 2.

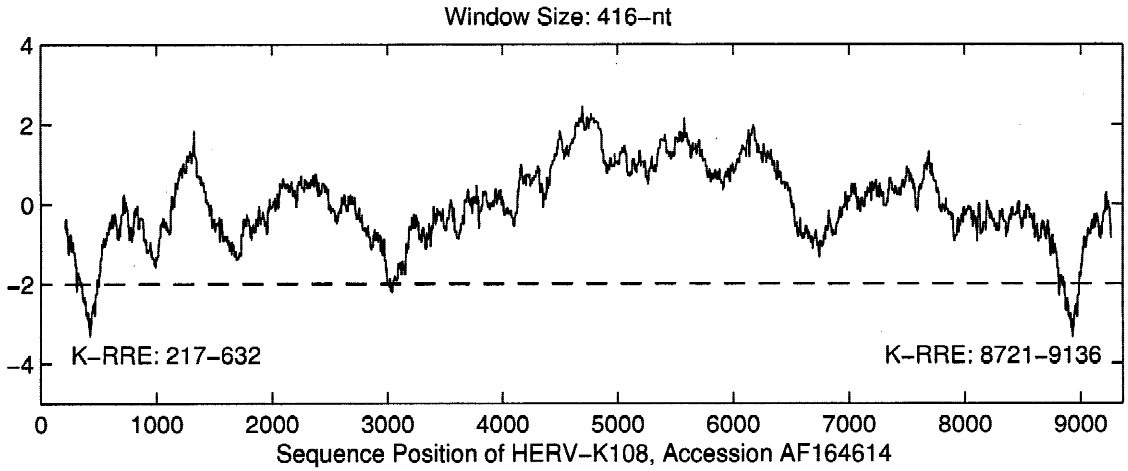

FIGURE 1. Computational detection of unusual RNA folding regions in the HERV-K108 provirus. The ability of transcribed segments of the HERV-K 108 genome to fold was calculated using a 416-nt window that was moved along the entire proviral genome in 1-nt increments. The statistical likelihood of RNA folding of each segment is calculated as a stability score (Stbscr) and is given in standard deviation units. A folding potential that exceeds −2.33 is considered significant.

The experiment shown in Figure 3 utilizes indicator constructs based on the previously described pDM128/PL plasmid, which contains a *cat* gene sequestered in an intron flanked by a 3′ polylinker that can be used to insert candidate response element sequences (Hope et al., 1990; Fridell et al., 1993). Only if the inserted sequence is able to effectively recruit an RNA export factor can the unspliced *cat* mRNA be exported to the cytoplasm and CAT protein expression induced. Using this assay, we have previously reported that the full-length K-RRE is responsive not only to K-Rev but also to H-Rev, and this result is reproduced in Figure 3 (Yang et al., 1999). In this experiment, we also demonstrate that the K-RRE can effectively mediate target mRNA export induced by the HTLV-I Rex protein. This induction is specific, as K-Rev, H-Rev, and Rex all fail to induce expression of CAT from the parental pDM128/PL plasmid or from a derivative containing the V-RRE. Whereas pDM128/PL derivatives bearing residues 1–383 or 51–433 of the K-RRE remained partially responsive to the K-Rev protein, and fully responsive to both H-Rev and Rex, all 11 of the more extensive deletions of the K-RRE were largely nonresponsive to K-Rev, although many retained at least some responsiveness to H-Rev and/or Rex (Fig. 3). Therefore, we conclude that almost the entire K-RRE sequence given in Figure 2 is required for K-RRE function in response to K-Rev expression.

The functional data presented in Figure 3 suggest that the 5′ border for response of the K-RRE to HTLV-I Rex lies at around position 251, whereas the 3′ border appeared to map to around 310. This suggests that K-RRE sequences in stem-loop IIf (SLIIf) and/or SLIII might be particularly important for Rex function. More importantly, the continued ability of most of these K-RRE deletion mutants to respond to H-Rev and/or Rex (Fig. 3) not only demonstrates that RNAs containing these K-RRE sequences were expressed in the nucleus of transfected cells but also implies that many of the K-RRE deletions retain the ability to fold, at least in part, into a conformation comparable to that adopted by the full-length K-RRE. Therefore, the inability of the more extensive K-RRE deletion mutants to respond to K-Rev appears not to be due to a complete misfolding of the K-RRE.

We next wished to ask whether these same deletion mutants of the K-RRE retained the ability to bind to K-Rev. We have previously reported that a recombinant glutathione-S-transferase (GST)/K-Rev fusion protein can bind to a radiolabeled K-RRE probe in vitro and that this interaction can be specifically competed by addition of excess unlabeled K-RRE, but not by addition of an irrelevant RNA competitor, such as the H-RRE (Yang et al., 1999). This result is reproduced in Figure 4A, lanes 1–4. Analysis of the ability of the K-RRE deletion mutants to compete for binding showed that residues 51–360, as well as all larger K-RRE fragments, competed effectively for K-Rev binding even though this same 51–360 segment of the K-RRE displayed little or no biological activity in vivo (Fig. 3). Nested deletions of the 51–360 fragment from the 5′ end (101–360, 151–360, 201–360, 251–360, and 301–360) showed little or no competition by the 101–360 and 151–360 fragments, partial competition by the 201–360 fragment, efficient competition by the 251–360 fragment, and no competition for binding to K-Rev, once again, by the 301–360 fragment. Thus, although the 251–360 fragment can bind to K-Rev effectively, larger fragments containing this sequence, that is, 151–360 and 201–360, fail to bind. We surmise that these larger fragments must either fold in such a way as to block binding by K-Rev to a site present in the 251–360 se-

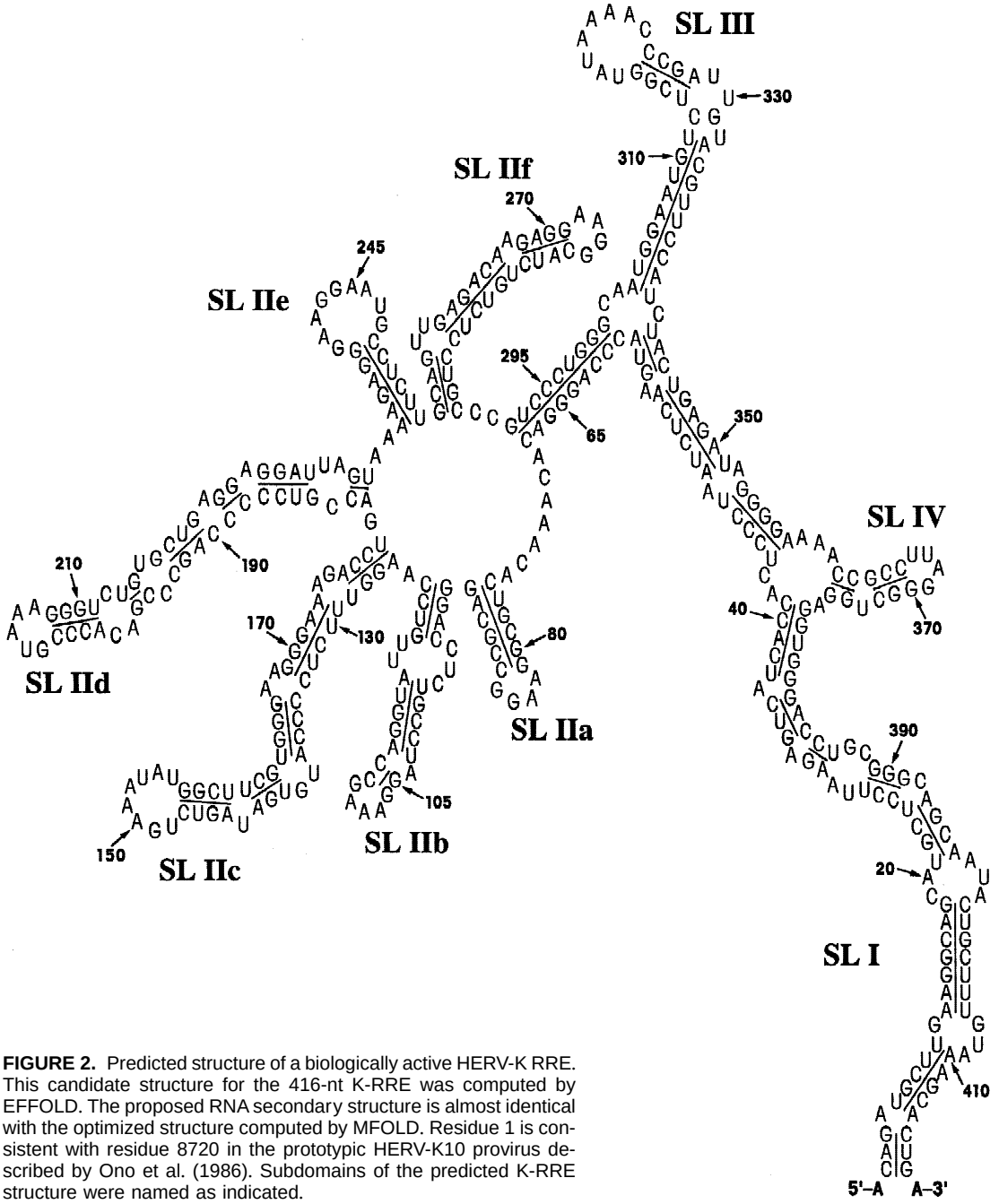

FIGURE 2. Predicted structure of a biologically active HERV-K RRE. This candidate structure for the 416-nt K-RRE was computed by EFFOLD. The proposed RNA secondary structure is almost identical with the optimized structure computed by MFOLD. Residue 1 is consistent with residue 8720 in the prototypic HERV-K10 provirus described by Ono et al. (1986). Subdomains of the predicted K-RRE structure were named as indicated.

quence or, less probably, that the 251–360 sequence folds to give an artifactual K-Rev-binding site when flanking 5′ residues have been deleted. A more clear-cut picture emerged when 3′ nested deletions of the 51–360 fragment were analyzed (Fig. 4A). Specifically, fragments 51–310, 51–260, and 51–210 all bound K-Rev effectively and 51–160 gave partial binding. In summary, these in vitro studies imply the existence of two discrete K-Rev-binding sites in the K-RRE, one located between 251 and 360 that is coincident with predicted RNA stem-loops SLIIf and/or SLIII, and a second located between 51 and 160 that is coincident with the predicted SLIIa and/or SLIIb RNA stem-loops.

We next wished to confirm the in vitro RNA binding data presented in Figure 4A using an in vivo assay, and therefore asked whether K-Rev would bind the K-RRE specifically in the yeast three-hybrid assay (SenGupta et al., 1996). As shown in Figure 4B, K-Rev indeed proved able to bind to the sense K-RRE but not to the antisense K-RRE or to the equine infectious anemia virus (EIAV) TAR (eTAR) RNA element, an RNA that we have previously shown can effectively bind to the EIAV Tat protein in the three-hybrid assay (Bieniasz et al., 1999).

Analysis of the K-RRE deletion mutants using this in vivo assay (Fig. 4B) largely confirmed the in vitro data

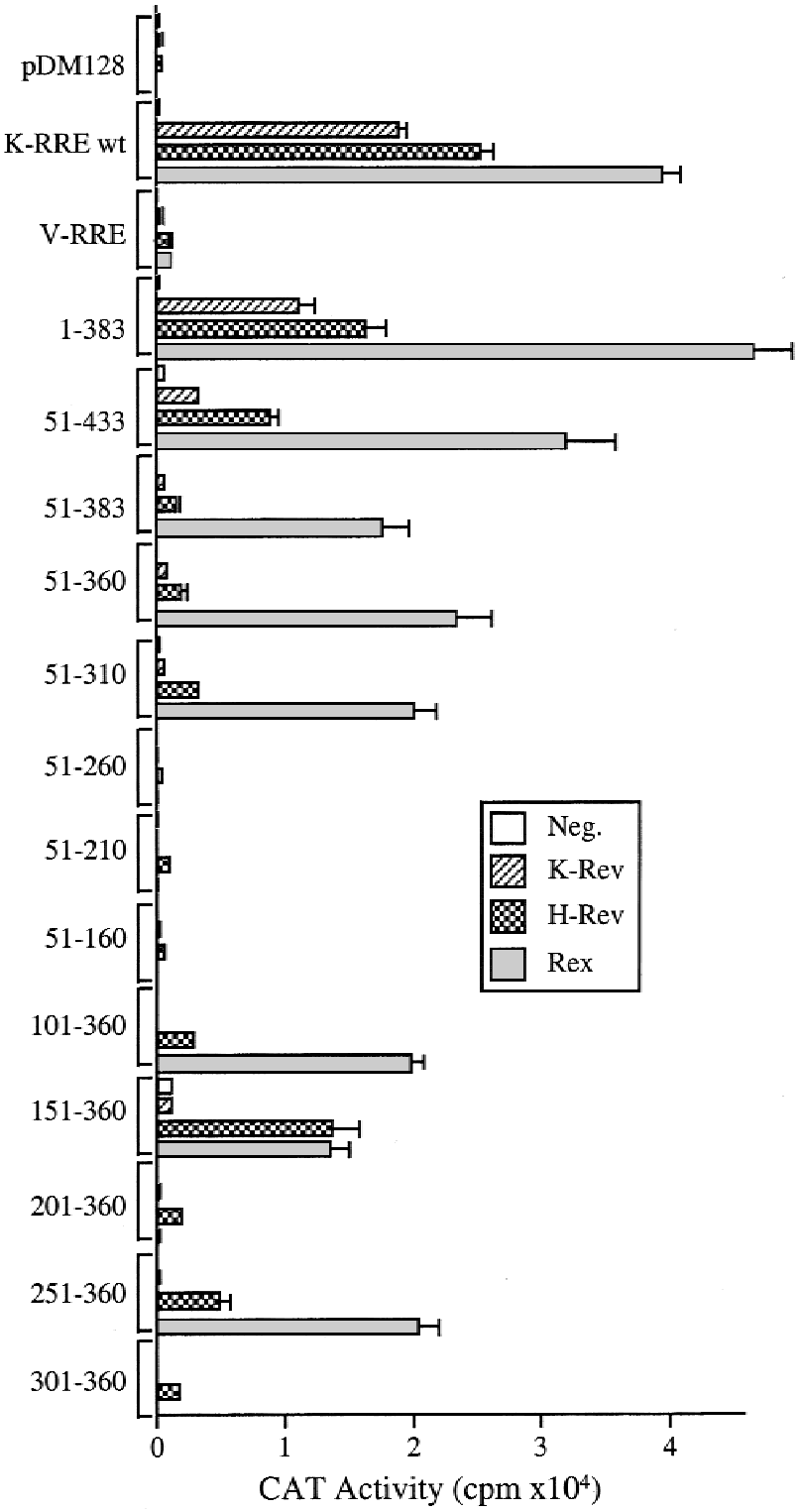

FIGURE 3. Biological activity of 5′ or 3′ deletions of the K-RRE. The K-RRE segments indicated by the given coordinates (based on the K-RRE structure shown in Fig. 2) were cloned into the polylinker present in the pDM128/PL indicator plasmid. These derivatives, as well as the parental pDM128/PL plasmid and a pDM128 derivative containing the wild-type 1–433 K-RRE sequence, were transfected into 293T cells together with expression plasmids encoding K-Rev, H-Rev or HTLV-I Rex, or the parental pBC12/CMV plasmid as a negative control (Neg.). After ~48 h, cultures were harvested and induced CAT activities determined. CAT activity is given in counts per minute. These data represent the average of three independent transfection experiments, with standard deviation indicated.

presented in Figure 4A. Specifically, both the 51–160 and the 251–360 K-RRE sequences, which do not overlap, showed strong binding to K-Rev. Again, the larger 151–360 and 100–360 sequences, which include the 251–360 binding site, failed to bind to K-Rev. Therefore, both in vitro and in vivo, the folding undergone by the 251–360 K-RRE sequence when present alone must differ from the folding seen when it is part of these larger transcripts. There are only two minor discrepancies between the in vitro and the in vivo RNA binding data (Fig. 4; Table 2). One relates to the 201–360 sequence, which failed to bind in vivo even though it dis-

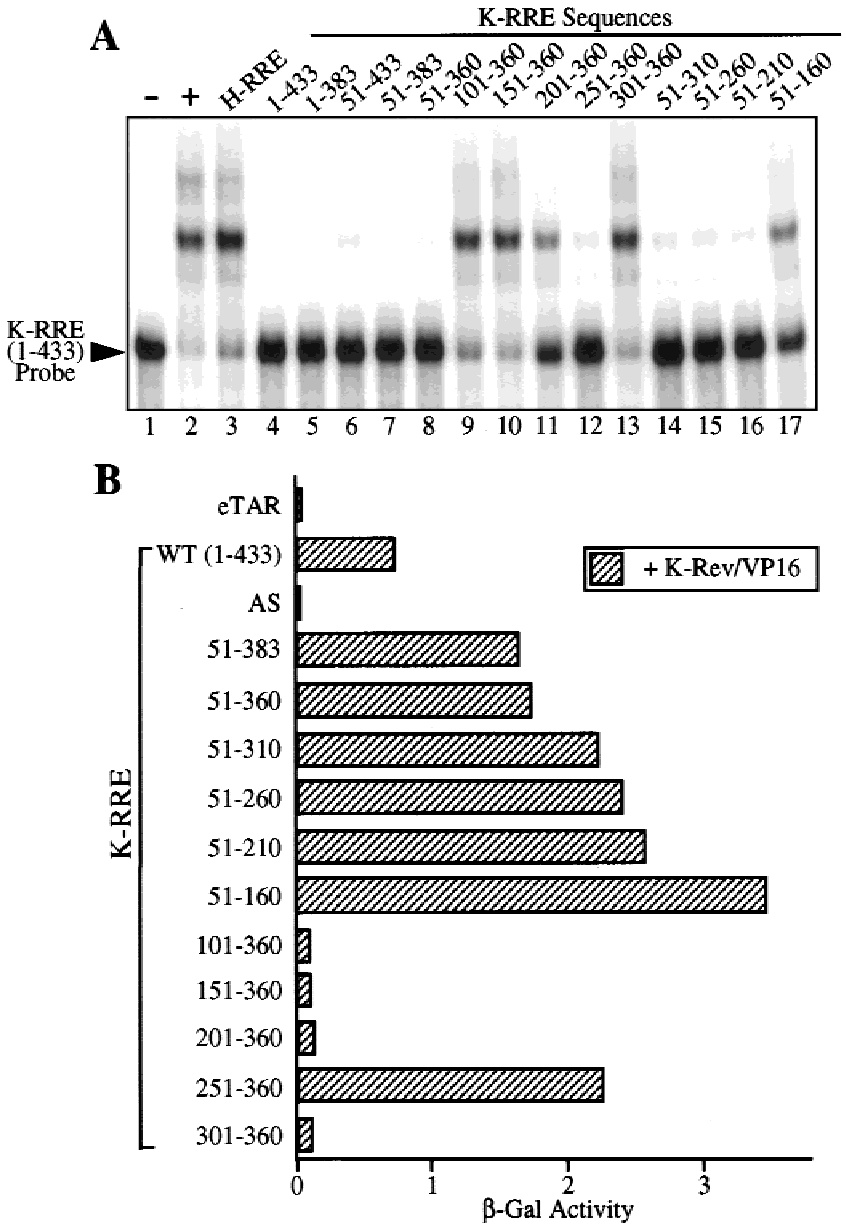

FIGURE 4. K-Rev binding by 5′ or 3′ deletion mutants of the K-RRE. **A**: An RNA gel-shift assay was used to compare the ability of different K-RRE deletion mutants to bind to K-Rev. In vitro binding to K-Rev was assessed by measuring the ability of a 200-fold molar excess of an unlabeled RNA transcript of the indicated K-RRE segment to compete binding of recombinant GST-K-Rev protein to a ^{32}P-labeled full-length K-RRE RNA probe. **B**: The yeast three-hybrid assay was used to measure the binding of a wild-type K-Rev/VP16 fusion protein to the indicated K-RRE segments. Binding results in induction of the β-gal indicator gene, whose activity was measured in optical density units per milliliter of culture. These data are representative of several independent experiments.

played partial binding in vitro, and the second involves the 51–160 sequence, which again was partial in vitro but was fully active for binding in vivo. Presumably, these discrepancies reflect differences in the folding of these two RNA fragments in these very different experimental contexts. In any event, these small differences do not change the major conclusion of both Figures 4A and 4B, which is that there are two discrete binding sites for K-Rev on the K-RRE that map between 51 and 160 and 251 and 360.

Individual predicted stem-loop subdomains within the K-RRE bind K-Rev specifically

As noted above, the K-Rev binding site present between residues 51 and 160 in the K-RRE overlaps with predicted stem-loops SLIIa, SLIIb, and, in part, SLIIc, whereas the 251–360 sequence overlaps with SLIIf and SLIII (Fig. 2). We therefore next asked if any of these predicted stem-loops, if present on their own, would function as specific K-Rev-binding sites. For this purpose, we initially precisely cloned each of the above single predicted stem-loops, as well as both SLIIa and SLIIb in tandem, into an in vitro transcription plasmid used to generate RNAs that could be tested as specific competitors of K-Rev binding to the full-length K-RRE. Of note, during the cloning process, each stem-loop was flanked both 5′ and 3′ by a *Not*I restriction enzyme site (5′-GCGGCCGC-3′). The stem at the base of each predicted stem-loop is therefore expected to be extended by 8 G:C base pairs beyond what is predicted in Figure 2, and should therefore be quite stable.

TABLE 2. Biological activity and K-Rev binding potential of nested deletion mutants of the K-RRE.

K-RRE sequence	In vitro binding	In vivo binding	Biological activity
1–433 (full length)	+	+	+
1–383	+	ND	+
51–433	+	ND	±
51–383	+	+	−
51–360	+	+	−
101–360	−	−	−
151–360	−	−	−
201–360	±	−	−
251–360	+	+	−
301–360	−	−	−
51–310	+	+	−
51–260	+	+	−
51–210	+	+	−
51–160	±	+	−

This table summarizes the data presented in Figures 3 and 4. +: fully active; ±: partially active; −: inactive; ND: not done.

As indicated in Figure 5A, we observed effective competition for binding to the full-length K-RRE by not only the wild-type K-RRE but also by the isolated SLIIb and SLIIf RNA stem-loops, as well as by a combination of SLIIb and SLIIa. In contrast, SLIIa alone, SLIIc, and SLIII all proved unable to compete for binding by K-Rev. To confirm this result in vivo, we next asked whether individual predicted stem-loops derived from the K-RRE would also bind to K-Rev in the yeast three-hybrid assay (Fig. 5B). In fact, we detected efficient binding of K-Rev to both SLIIb and SLIIf, whereas SLIIc proved unable to support K-Rev binding. We therefore conclude that the two discrete K-Rev-binding sites mapped by deletion analysis of the K-RRE to between residues 51 and 160 and 251 and 360 actually coincide with the predicted stem-loops SLIIb (residues 92–123) and SLIIf (residues 255–289).

The binding data presented thus far clearly demonstrate the existence of two discrete binding sites for K-Rev on the K-RRE when assayed in vitro or in the

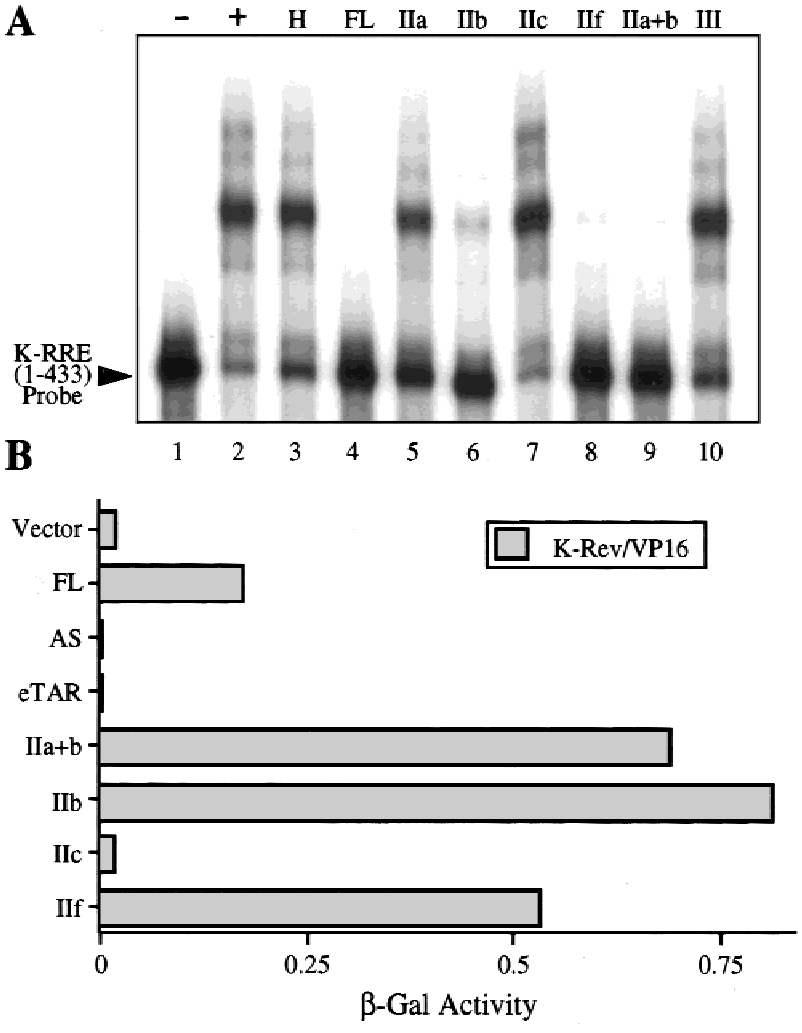

FIGURE 5. K-Rev binding to predicted K-RRE stem-loop subdomains. **A**: Binding to K-Rev in vitro was measured as described in Figure 4A. The indicated individual K-RRE subdomains, as defined in Figure 2, were used as unlabeled competitors in the illustrated RNA gel-shift assay. **B**: The yeast three-hybrid assay was used to measure binding of K-Rev to the indicated K-RRE stem-loop subdomains, as described in Figure 4B.

yeast three-hybrid assay. However, these data do not reveal whether these same sites can function as binding sites for K-Rev in the mammalian cell nucleus. To address this question, we used a previously described (Tiley et al., 1992; Kang & Cullen, 1999) mammalian assay for specific protein:RNA interactions that utilizes the HIV-1 Tat (H-Tat) protein. H-Tat is an unusual transcription factor in that it activates gene expression from the HIV-1 LTR after recruitment to an RNA stem-loop structure, the H-TAR element, encoded within the viral LTR (reviewed by Taube et al., 1999). Importantly, H-Tat can also effectively activate HIV-1 LTR-dependent gene expression if recruited to a heterologous RNA target sequence, substituted in place of H-TAR, upon fusion to the appropriate RNA-binding domain. Thus, an HIV-1 LTR bearing the H-RRE-derived SLIIB primary binding site for H-Rev in place of H-TAR is effectively activated by a coexpressed H-Tat/H-Rev fusion protein but not by either wild-type H-Tat or wild-type H-Rev (Tiley et al., 1992).

To test whether K-Rev could interact with specific stem-loop subdomains of the K-RRE in human nuclei, we therefore constructed indicator plasmids in which sequences encoding the SLIIb, SLIIf, or SLIII subdomains of the K-RRE were substituted for the apical region of the H-TAR RNA stem-loop in the *cat*-based indicator plasmid pHIV/H-TAR/CAT. Controls included the parental pHIV/H-TAR/CAT, which should respond to all Tat fusion proteins, and a derivative containing the heterologous bovine immunodeficiency virus TAR (B-TAR) element in place of H-TAR, which should be nonresponsive (Bogerd et al., 2000). A construct containing the SLIIB subdomain of the H-RRE was predicted to be responsive to the H-Tat/H-Rev fusion protein but not to the H-Tat/K-Rev fusion.

The results obtained (Fig. 6) reveal strong activation of the pHIV/H-TAR/CAT protein by both tested fusion proteins, thus showing that these are functionally expressed. The H-Tat/H-Rev fusion did not activate the pHIV/B-TAR/CAT indicator plasmid but strongly activated pHIV/SLIIB/CAT, as predicted. Although the H-Tat/H-Rev fusion failed to activate the indicator construct bearing K-RRE SLIIb, and was at most slightly active with K-RRE SLIIf, a strong activation of the indicator plasmid bearing K-RRE SLIII was observed (Fig. 6). Therefore, these data unexpectedly demonstrated that the predicted SLIII stem-loop structure can serve as a specific binding site for the H-Rev protein in vivo.

In contrast to the H-Tat/H-Rev fusion, the H-Tat/K-Rev fusion failed to significantly activate expression of a *cat* gene linked to either the H-RRE SLIIB sequence or the K-RRE SLIII sequence. However, significant (~17-fold) activation of indicator plasmids containing either subdomain SLIIb or subdomain SLIIf of the K-RRE was consistently detected (Fig. 6). Therefore, these data confirm that both SLIIb and SLIIf can indeed serve as specific binding sites for K-Rev in the human cell nucleus.

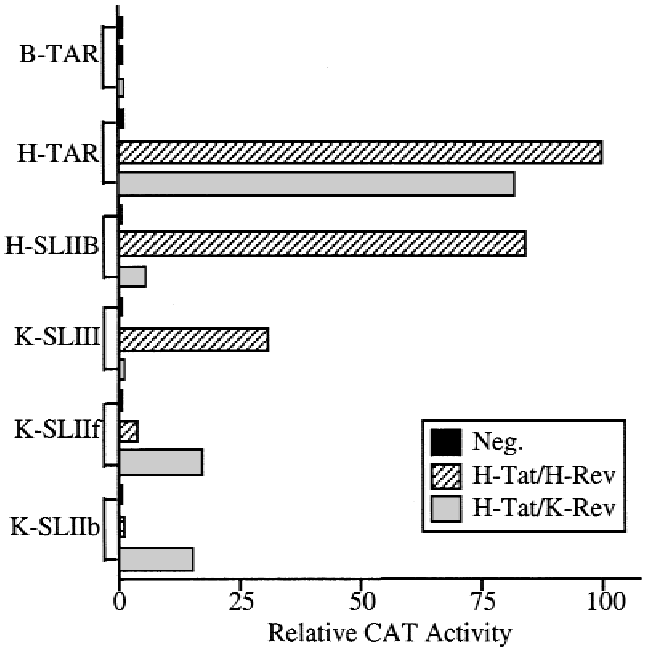

FIGURE 6. Both H-Rev and K-Rev bind to specific subdomains of the K-RRE in the human cell nucleus. This experiment measures the ability of fusion proteins consisting of H-Tat fused to either H-Rev or K-Rev to activate HIV-1 LTR-based indicator constructs bearing precise substitutions of the indicated RNA element for the H-TAR element. The parental pBC12/CMV plasmid served as a Negative control (Neg). 293T cells were transfected as previously described (Blair et al., 1998) and induced CAT enzyme levels measured ~48 h after transfection. The indicated data represent the average of four independent experiments and were normalized to the CAT activity observed in cells cotransfected with pBC12/H-TAR/CAT and pH-Tat/H-Rev, which was arbitrarily set at 100.

Deletion of specific predicted stem-loop subdomains of the K-RRE

Like the K-RRE (Fig. 2), the H-RRE is predicted to consist of a long basal RNA stem topped by an RNA bulge from which extends a series of smaller RNA stem-loop structures termed SLII, III, IV, and V (Malim et al., 1990). In the case of the H-RRE, there is a single primary binding site for H-Rev that maps to SLII. The precise, individual deletion of stem-loop III, IV, or V does not markedly inhibit H-RRE function, although larger deletions, which in principle also leave SLII intact, do block H-RRE activity (Malim et al., 1989, 1990).

By analogy, we predicted that precise deletion of SLIIb and/or SLIIf in the K-RRE might block the ability of the K-RRE to respond to K-Rev, whereas the exact deletion of other predicted stem-loop subdomains, such as SLIId, might leave K-RRE function unimpaired. We therefore precisely and individually deleted each of the predicted SLII stem-loop subdomains indicated in Figure 2, as well as SLIII, and asked whether the resultant K-RRE mutants would respond to K-Rev expression in vivo using the pDM128 assay (Fig. 7). Again, H-Rev and HTLV-I Rex, which presumably bind to a different target than K-Rev in the K-RRE, were used as controls

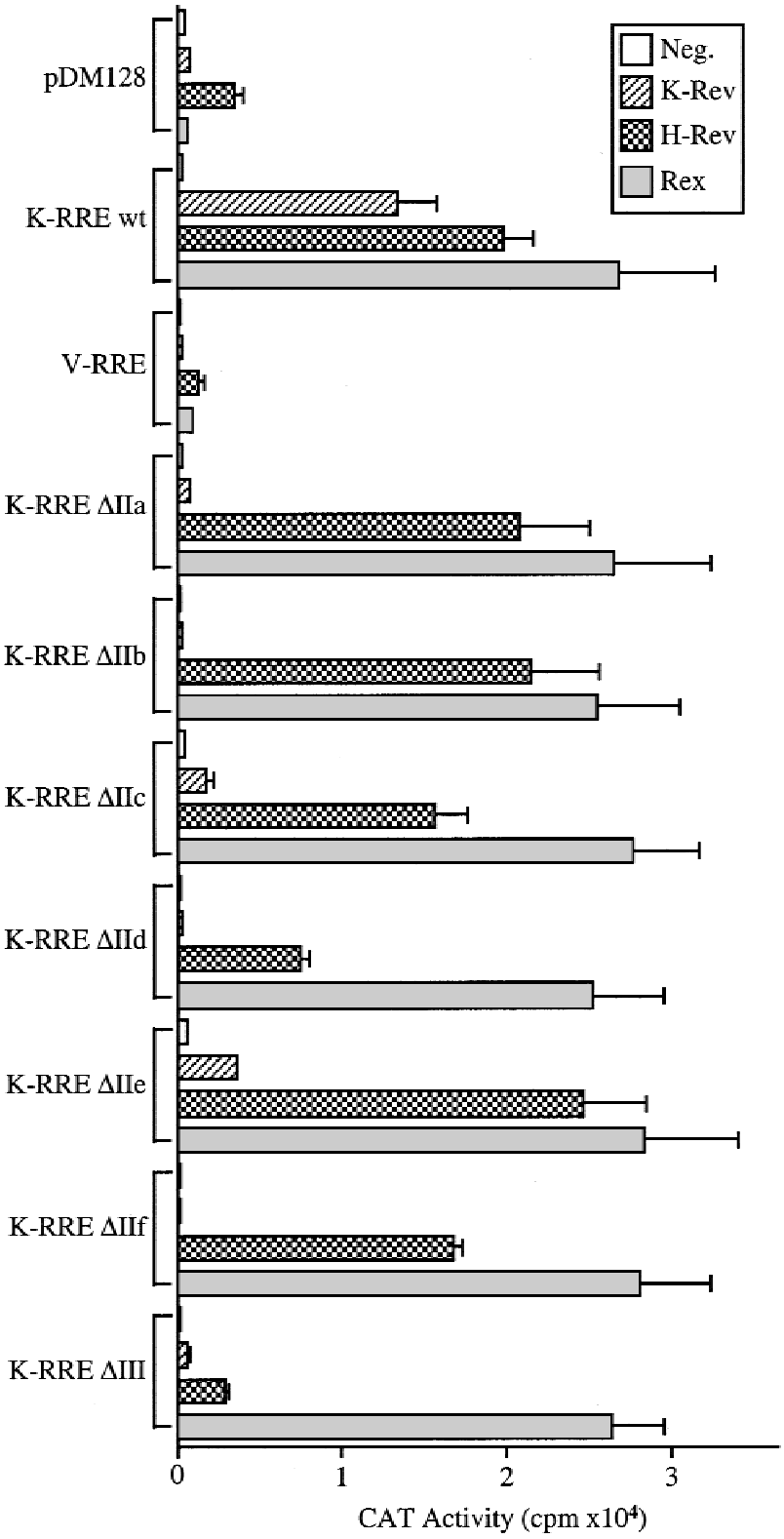

FIGURE 7. Biological activity of targeted internal deletion mutants of the K-RRE. The biological activity of deletion mutants of the K-RRE, precisely lacking the indicated stem-loop subdomains as defined in Figure 2, was measured as described in Figure 3. The parental pDM128/PL plasmid and a derivative containing the visna virus RRE (V-RRE) served as controls. These data represent the average of three independent transfection experiments, with the standard deviation indicated.

to test whether individual deletions would result in destabilization and/or complete misfolding of RNAs bearing the K-RRE, which should abrogate the response to all three retroviral export factors. We also hoped that this analysis might help to further define the binding site(s) for either H-Rev or Rex on the K-RRE, if this indeed coincided with a single predicted stem-loop subdomain.

Unfortunately, the results obtained from this quite extensive analysis were only moderately informative, as

all the K-RRE mutants obtained were either entirely or largely nonresponsive to K-Rev (Fig. 7). Only in the case of the ΔIIc and ΔIIe mutants was any biological activity above background observed in the presence of the K-Rev protein. In contrast, none of these deletion mutations had any significant effect on the response to HTLV-I Rex, thus implying the existence of two functional target sites or of a single Rex-binding site that is not affected by any of these K-RRE deletions (Fig. 7). Finally, although all the tested SLII subdomain deletion mutants of the K-RRE remained responsive to the H-Rev protein, the SLIII deletion was, like the pDM128/PL negative control, essentially refractory to activation by H-Rev. In combination with the previous observation that K-RRE SLIII can serve as an effective binding site for H-Rev in the human cell nucleus (Fig. 6), these data strongly support the hypothesis that the H-Rev-induced nuclear export of K-RRE-containing transcripts is largely due to recruitment of H-Rev to a primary binding site located in the predicted SLIII domain of the K-RRE.

DISCUSSION

Although retroviruses normally infect somatic cells, infection of germ line cells can also occur. If the infected germ cell then participates in the formation of progeny, the resultant integrated provirus will be vertically inherited and can then become a fixed part of the host, or even the species, genome. Although comparatively rare, the difficulty of reversing such proviral integrations has led to their accumulation over time and endogenous proviruses now constitute ~0.1% of the entire human genome (reviewed by Patience et al., 1997).

Vertical, as opposed to horizontal, transmission of retroviruses does not require virally encoded gene products and the selective pressure to maintain these in an intact form is therefore lost. Indeed, retroviral proteins can be deleterious to the cell and their functional expression may therefore be actively selected against over time. In fact, the large majority of endogenous retroviruses bear obvious frame-shift or deletion mutations and are clearly defective. Although this applies also to the HERV-K family of endogenous retroviruses, several individual HERV-K proviruses appear to be only minimally defective (Barbulescu et al., 1999) and at least some proviruses encode functional forms of the enzymes critical for the retroviral life cycle, such as reverse transcriptase (Berkhout et al., 1999). Nevertheless, no HERV-K endogenous provirus has been shown to be replication competent and it is probable that HERV-K proviruses that are not obviously defective have nevertheless accumulated deleterious point mutations over time. The lack of a replication-competent HERV-K means that it is not possible to demonstrate that the K-Rev protein or K-RRE target sequence used in the present analysis are fully biologically active by, for example, demonstrating their ability to support viral replication in culture, as would be the case if an exogenous retrovirus was being analyzed.

Although this is a significant concern, it is nevertheless clear that K-Rev does interact with the K-RRE in a manner that is similar, albeit not identical, to what is seen for the equivalent nuclear RNA export factors encoded by exogenous retroviruses such as HIV-1 and HTLV-I. Specifically, the K-RRE, like the H-RRE and the RxRE (Hanly et al., 1989; Malim et al., 1989), clearly coincides with a region of the HERV-K genome that is very likely to fold into an extensive RNA stem-loop structure (Fig. 1; Table 1). Computational analysis of the secondary structure likely to be adopted by the ~416-nt K-RRE predicts a highly stable RNA structure that, like the H-RRE (Malim et al., 1989, 1990), can be subdivided into a set of stem-loop subdomains (Fig. 2). Efforts to identify the specific binding site for K-Rev on the K-RRE, using a series of in vitro and in vivo assays (Figs. 4–6), revealed the existence of two specific binding sites for K-Rev in the K-RRE, coincident with predicted stem-loop subdomains SLIIb and SLIIf. Interestingly, a single specific binding site for H-Rev coincident with the predicted SLIII subdomain of the K-RRE (Fig. 2) was also observed. Importantly, each of these RNA sequences proved able to specifically recruit either the K-Rev or the H-Rev protein when assayed in the human nucleus, the milieu where this interaction would normally be predicted to occur, and specific binding of K-Rev to the SLIIb and SLIIf subdomains of the K-RRE was also readily detectable in vitro and in the yeast three-hybrid assay. Together, these data convincingly identify sequences within the K-RRE that have the potential to recruit either the K-Rev or the H-Rev nuclear export factor to the HERV-K viral genome and hence induce its export from the human cell nucleus.

Inspection of predicted K-RRE stem-loops IIb and IIf reveals a number of interesting structural similarities, including the presence in both structures of a 4-nt purine-rich terminal loop flanked 5′ by a bulged A residue (Fig. 2). Whether these shared features are indeed critical for K-Rev binding remains to be determined. Also of interest is the absence in SLIII of any predicted RNA bulge similar to the H-Rev primary binding site previously defined in the H-RRE (Bartel et al., 1991; Heaphy et al., 1991). As the H-Rev binding site in K-RRE SLIII is able to specifically bind H-Rev in the human nucleus (Fig. 6) and is also critical for H-Rev dependent nuclear export of a K-RRE containing mRNA (Fig. 7), this may imply that H-Rev can functionally interact with more than one type of RNA target site.

Unfortunately, it has not proven possible to fully correlate the RNA-binding data reported in this manuscript with the in vivo biological activity of K-RRE mutants. Specifically, it appears that essentially the entire 416-nt K-RRE is required for biological function in the presence of K-Rev as both 5′ and 3′ deletions (Fig. 3) and targeted internal deletions (Fig. 7) almost

always result in a complete loss of activity. Although this could result from the complete misfolding of these mutant K-RRE sequences, we believe this is improbable, as most K-RRE mutants do retain the ability to respond to HTLV-I Rex or to H-Rev (Figs. 3 and 7), both of which are known to bind to specific RNA sequences presented in the context of RNA secondary structures (Zapp & Green, 1989; Malim et al., 1990; Bartel et al., 1991; Heaphy et al., 1991; Bogerd et al., 1992). It therefore seems possible that the functional interaction of the K-RRE with K-Rev may require binding of multiple K-Rev molecules to not only the two primary RNA-binding sites but also, perhaps, to discrete secondary binding sites in the K-RRE that were not detected in this analysis. In fact, K-Rev is known to multimerize (Yang et al., 1999) and multimerization is known to be critical for both H-Rev and HTLV-I Rex function (Malim & Cullen, 1991; Zapp et al., 1991; Hakata et al., 1998). Further, for both H-Rev and HTLV-I Rex, the primary binding site in the RRE or RxRE is essential but not sufficient for biological activity, which instead requires the recruitment of multiple Rev or Rex molecules to a larger structured RNA target (Malim et al., 1989; Malim & Cullen, 1991; Huang et al., 1991; Iwai et al., 1992; Gröne et al., 1994). Thus, the inability of K-Rev to induce the nuclear export of RNAs containing deletion mutants of the K-RRE that retain intact K-Rev-binding sites has also been reported for equivalent H-RRE deletion mutants in the presence of H-Rev, although the H-RRE is clearly not inactivated as readily as is the K-RRE (Malim et al., 1990).

In conclusion, we have demonstrated that the previously functionally defined HERV-K RRE (Magin et al., 1999; Yang et al., 1999) coincides with a sequence that is highly likely to fold into an RNA stem-loop structure that is similar to, but larger than, equivalent H-RRE and RxRE sequences (Figs. 1 and 2; Table 1). Within this K-RRE sequence, we have defined candidate RNA stem-loop subdomains that specifically bind to K-Rev both in vitro and in vivo (Figs. 4–6). In addition, we have identified a binding site for H-Rev on the K-RRE (Fig. 6) that is critical for the nuclear export of K-RRE-containing transcripts induced by H-Rev (Fig. 7). Together, these data reinforce the remarkable functional similarity between K-Rev, a nuclear RNA export factor encoded by a 30-million-year-old human endogenous retrovirus and H-Rev, an essential regulatory protein encoded by the pathogenic exogenous retrovirus HIV-1 that remains very much alive today.

MATERIALS AND METHODS

Computational analysis of unusual RNA folding regions

UFR in the complete proviral sequence of the human endogenous retrovirus HERV-K were detected using the program SEGFOLD (Le & Maizel, 1989; Le et al., 1990). The program SEGFOLD was used to perform statistical simulations by computing significance and stability scores for successive segments by sliding a fixed window along the RNA sequence from the 5′ to the 3′ end. The Sigscr was evaluated by comparison of the computed lowest free energy of the segment in the window with the mean of the lowest free energies from a large number of randomly shuffled, corresponding segments. The Stbscr was similarly assessed by comparing the window segment at a given place with the average of all others having the same size resulting from sliding the fixed window along the tested sequence. Both Sigscr and Stbscr were standard z-scores and their units were measured by the standard deviation from the corresponding samples.

The potential size of the UFR in the tested sequence was assessed by exhaustive statistical simulations in which the size of the fixed window was changed over a wide range. In this study, the window size was continuously varied from 100 to 450 nt with a window step size of 2 nt. For each fixed window, Sigscr and Stbscr were computed by sliding 1 nt along the sequence. The global minima of Sigscr and Stbscr, as well as the location of the corresponding segment in the sequence, were recorded. Using the distribution plot of Sigscr and Stbscr, we can detect UFR that are highly significant and/or very stable in the test sequence. The RNA secondary structure of the HERV-K UFR was computed by the programs EFFOLD (Le et al., 1993) and MFOLD version 3.1 (Mathews et al., 1999).

To speed up the computation of Sigscr, the mean and standard deviation of the lowest free energies from 300 randomly shuffled sequences were computed from the tabulated coefficients based on the window size and the base composition in the window sequence, if the percent content of base G+C in the window was less than 75% and each base percentage was larger than 3%. Otherwise, the random mean and standard deviation of the lowest free energies were computed from 100 randomly shuffled sequences. These tabulated coefficients were derived from the Monte Carlo simulation and least-squares fit (Chen et al., 1990) by means of Turner energy rules (Freier et al., 1986; Jaeger et al., 1989).

Plasmid construction

The following plasmids have been previously described: the mammalian expression plasmids pBC12/CMV, pBC12/CMV/β-Gal, pcRev, pcK-Rev, and pcRex (Malim et al., 1989; Bogerd et al., 1998; Yang et al., 1999); a derivative of the pDM128/CMV indicator plasmid containing an intronic polylinker (pDM128/PL) (Hope et al., 1990; Fridell et al., 1993), as well as pDM128/PL derivatives containing the V-RRE or the full-length K-RRE (Yang et al., 1999); the HIV-1 LTR-based indicator plasmid pHIV/H-TAR/CAT and derivatives in which the H-TAR element has been substituted by the B-TAR sequence (pHIV/B-TAR/CAT) or the SLIIB minimal RNA-binding site for H-Rev (pHIV/SLIIB/CAT) (Tiley et al., 1992; Bogerd et al., 2000); the bacterial expression plasmid pGST/K-Rev (Yang et al., 1999) and the yeast expression plasmids pVP16, pK-Rev/VP16, and pIII/MS2/eTAR (Bieniasz et al., 1999; Yang et al., 1999).

In vitro transcription plasmids, encoding K-RRE deletion mutants, were derived by PCR amplification of the appropriate regions from the full-length K-RRE, with DNA primers

containing *Not*I sites, and subsequent insertion into pBluescript(KS) cleaved with *Not*I. Individual K-RRE stem-loop expression plasmids were made by annealing two DNA oligonucleotides precisely corresponding to each predicted stem-loop structure, with *Not*I overhangs, and also inserting into pBluescript(KS). The previously published (Yang et al., 1999) full-length K-RRE in vitro transcription plasmid contained the K-RRE cloned into pGEM-3ZF(+) as a *Bam*HI-*Xba*I fragment. For consistency with the K-RRE truncation plasmids, the full-length K-RRE was also cloned into pBluescript(KS) as a *Not*I fragment.

The same *Not*I-digested PCR fragments or annealed K-RRE oligonucleotides were also inserted into pIII/MS2 to generate equivalent K-RRE-containing RNA expression plasmids for yeast three-hybrid assays. To generate K-RRE 5′ or 3′ deletions in the context of the pDM128/PL indicator plasmid, the corresponding fragments from the pBluescript(KS)/K-RRE plasmid series were excised by cleavage with *Sac*II and *Xba*I and inserted into the polylinker present in pDM128/PL. The QuickChange (Stratagene) method was applied to delete individual K-RRE stem-loops using wild-type pDM128/K-RRE as the template. Derivatives of pHIV/H-TAR/CAT, in which the H-TAR element was replaced by K-RRE stem-loop subdomains SLIIb, SLIIf, or SLIII, were generated by deletion of the apical region of H-TAR by cleavage with *Bgl*II and *Sac*I followed by insertion of a double-stranded synthetic oligonucleotide encoding the relevant K-RRE sequence, with appropriate *Bgl*II and *Sac*I overhangs. All plasmids used in this study were sequenced to verify their orientation and integrity.

Cell culture and transfection

Human 293T cells were maintained as previously described (Bogerd et al., 1998). For each transfection, 25 ng of pBC12/CMV/β-Gal was included as internal control. Reporter plasmids that measure RNA export (i.e., pDMV128/PL, pDMV128/V-RRE, or one of the pDMV128/K-RRE series) were included at 50 ng per transfection, whereas effector plasmids (i.e., the pcRev, pcV-Rev, or pcRex) were present at 500 ng per transfection. The parental plasmid pBC12/CMV served as the negative control and was also added to the DNA mixture to maintain the same amount of total DNA (2 μg) per transfection. All transfections were performed using FuGene 6 (Roche), according to the manufacturer's protocol, in 35-mm 6-well plates. Induced CAT activities were determined at ~48 h after transfection and were normalized using the β-Gal internal control (Bogerd et al., 1998). Mammalian RNA binding assays were also performed in 293T cells, as previously described (Blair et al., 1998).

In vitro transcription and gel shift assays

Except for pBluescript/IIa, which was linearized with *Sac*II and transcribed using T3 RNA polymerase, all other pBluescript/K-RRE derivatives were linearized by *Bam*HI digestion and transcribed by T7 RNA polymerase (Promega). The full-length K-RRE probe was labeled with 50 μCi of α-^{32}P-CTP, using the Promega Riboprobe protocol, and purified from an 8% denaturing polyacrylamide gel. The unlabeled K-RRE RNA competitors were made using the RiboMax Large Scale RNA system (Promega) according to the manufacturer's instructions. The GST-K-Rev fusion protein was expressed in bacteria and purified on glutathione affinity resin as described previously (Bieniasz et al., 1999). Gel-shift reactions were carried out as described (Yang et al., 1999) with 4×10^4 cpm (~1 ng) of probe and 10 ng of GST-K-Rev in 20 μL buffer containing 150 mM KCl, 10 mM HEPES (pH 7.6), 0.5 mM EGTA, 2 mM $MgCl_2$, 1 mM DTT, 10% glycerol, 1 μg of yeast tRNA, 4 μg of *Escherichia coli* rRNA and appropriate competitor RNAs at an ~200-fold molar excess. The GST-K-Rev protein was first mixed with both the nonspecific RNA competitor and the cold competitor RNA in buffer and the mixture incubated on ice for 10 min. The labeled probe was then added to the mixture and the reaction incubated for a further 10 min on ice. Reaction products were then resolved on a 5% native polyacrylamide gel and visualized by autoradiography.

Yeast three-hybrid assays

The *Saccharomyces cerevisiae* indicator strain L40-coat (SenGupta et al., 1996) was transformed with the pVP16 or pVP16/K-Rev expression plasmid together with a pIII/MS2/K-RRE derivative or pIII/MS2 or pIII/MS2/eTAR as a negative control. After selection for transformants on Ura$^-$ Leu$^-$ plates, pooled transformants were suspended in β-Gal assay buffer, normalized to optical density, and β-Gal activities assayed as previously described (Bieniasz et al., 1999).

Received May 18, 2000; returned for revision
July 11, 2000; revised manuscript received July 24, 2000

REFERENCES

Barbulescu M, Turner G, Seaman MI, Deinard AS, Kidd KK, Lenz J. 1999. Many human endogenous retrovirus K (HERV-K) proviruses are unique to humans. *Curr Biol 9*:861–868.

Bartel DP, Zapp ML, Green MR, Szostak JW. 1991. HIV-1 Rev regulation involves recognition of non-Watson–Crick base pairs in viral RNA. *Cell 67*:529–536.

Berkhout B, Jebbink M, Zsíros J. 1999. Identification of an active reverse transcriptase enzyme encoded by a human endogenous HERV-K retrovirus. *J Virol 73*:2365–2375.

Bieniasz, PD, Grdina TA, Bogerd HP, Cullen BR. 1999. Highly divergent lentiviral Tat proteins activate viral gene expression by a common mechanism. *Mol Cell Biol 19*:4592–4599.

Blair WS, Parsley TB, Bogerd HP, Towner JS, Semler BL, Cullen BR. 1998. Utilization of a mammalian cell-based RNA binding assay to characterize the RNA binding properties of picornavirus 3C proteinases. *RNA 4*:215–225.

Bogerd HP, Echarri A, Ross TM, Cullen BR. 1998. Inhibition of human immunodeficiency virus Rev and human T-cell leukemia virus Rex function, but not Mason–Pfizer monkey virus constitutive transport element activity, by a mutant human nucleoporin targeted to Crm1. *J Virol 72*:8627–8635.

Bogerd HP, Tiley LS, Cullen BR. 1992. Specific binding of the human T-cell leukemia virus type I Rex protein to a short RNA sequence located within the Rex-response element. *J Virol 66*:7572–7575.

Bogerd HP, Wiegand HL, Bieniasz PD, Cullen BR. 2000. Functional differences between human and bovine immunodeficiency virus Tat transcription factors. *J Virol 74*:4666–4671.

Bray M, Prasad S, Dubay JW, Hunter E, Jeang K-T, Rekosh D, Hammarskjöld M-L. 1994. A small element from the Mason–Pfizer monkey virus genome makes human immunodeficiency virus type 1 expression and replication Rev-independent. *Proc Natl Acad Sci USA 91*:1256–1260.

Chen J-H, Le S-Y, Shapiro B, Currey KM, Maizel JV Jr. 1990. A computational procedure for assessing the significance of RNA secondary structure. *Comput Appl Biosci 6*:7–18.

Cullen BR. 2000. Nuclear RNA export pathways. *Mol Cell Biol 20*:4181–4187.

Fornerod M, Ohno M, Yoshida M, Mattaj IW. 1997. CRM1 is an export receptor for leucine-rich nuclear export signals. *Cell 90*: 1051–1060.

Freier SM, Kierzek R, Jaeger JA, Sugimoto N, Caruthers MH, Neilson T, Turner DH. 1986. Improved free-energy parameters for prediction of RNA duplex stability. *Proc Natl Acad Sci USA 83*: 9373–9377.

Fridell RA, Partin KM, Carpenter S, Cullen BR. 1993. Identification of the activation domain of equine infectious anemia virus Rev. *J Virol 67*:7317–7323.

Griffiths DJ, Venables PJW, Weiss RA, Boyd MT. 1997. A novel exogenous retrovirus sequence identified in humans. *J Virol 71*: 2866–2872.

Gröne M, Hoffmann E, Berchtold S, Cullen BR, Grassmann R. 1994. A single stem-loop structure within the HTLV-I Rex response element is sufficient to mediate Rex activity in vivo. *Virology 204*: 144–152.

Grüter P, Tabernero C, von Kobbe C, Schmitt C, Saavedra C, Bachi A, Wilm M, Felber BK, Izaurralde E. 1998. TAP, the human homolog of Mex67p, mediates CTE-dependent RNA export from the nucleus. *Mol Cell 1*:649–659.

Hakata Y, Umemoto T, Matsushita S, Shida H. 1998. Involvement of human CRM1 (Exportin 1) in the export and multimerization of the Rex protein of human T-cell leukemia virus type I. *J Virol 72*:6602–6607.

Hanly SM, Rimsky LT, Malim MH, Kim JH, Hauber J, Duc Dudon M, Le S-Y, Maizel JV, Cullen BR, Greene WC. 1989. Comparative analysis of the HTLV-I rex and HIV-1 rev trans-regulatory proteins and their RNA response elements. *Genes & Dev 3*:1534–1544.

Heaphy S, Finch JT, Gait MJ, Karn J, Singh M. 1991. Human immunodeficiency virus type 1 regulator of virion expression, rev, forms nucleoprotein filaments after binding to a purine-rich "bubble" located within the rev-responsive region of viral mRNAs. *Proc Natl Acad Sci USA 88*:7366–7370.

Hope TJ, Huang X, McDonald D, Parslow TG. 1990. Steroid-receptor fusion of the human immunodeficiency virus type 1 Rev transactivator: Mapping cryptic functions of the arginine-rich motif. *Proc Natl Acad Sci USA 87*:7787–7791.

Huang X, Hope TJ, Bond BL, McDonald D, Grahl K, Parslow TG. 1991. Minimal Rev-response element for type 1 human immunodeficiency virus. *J Virol 65*:2131–2134.

Iwai S, Pritchard C, Mann DA, Karn J, Gait MJ. 1992. Recognition of the high affinity binding site in Rev-response element RNA by the human immunodeficiency virus type-1 Rev protein. *Nucleic Acids Res 20*:6465–6472.

Jaeger JA, Turner DH, Zuker M. 1989. Improved predictions of secondary structures for RNA. *Proc Natl Acad Sci USA 86*:7706–7710.

Kang Y, Cullen BR. 1999. The human Tap protein is a nuclear mRNA export factor that contains novel RNA-binding and nucleocytoplasmic transport sequences. *Genes & Dev 13*:1126–1139.

Katahira J, Straßer K, Podtelejnikov A, Mann M, Jung JU, Hurt E. 1999. The Mex67p-mediated nuclear mRNA export pathway is conserved from yeast to human. *EMBO J 18*:2593–2609.

Le S-Y, Chen JH, Maizel JV Jr. 1993. Prediction of alternative RNA secondary structures based on fluctuating thermodynamic parameters. *Nucleic Acids Res 21*:2173–2178.

Le S-Y, Maizel JV. 1989. A method for assessing the statistical significance of RNA folding. *J Theor Biol 138*:495–510.

Le S-Y, Malim MH, Cullen BR, Maizel JV. 1990. A highly conserved RNA folding region coincident with the Rev response element of primate immunodeficiency viruses. *Nucleic Acids Res 18*:1613–1623.

Magin C, Löwer R, Löwer J. 1999. cORF and RcRE, the Rev/Rex and RRE/RxRE homologues of the human endogenous retrovirus family HTDV/HERV-K. *J Virol 73*:9496–9507.

Malim MH, Cullen BR. 1991. HIV-1 structural gene expression requires the binding of multiple Rev monomers to the viral RRE: Implications for HIV-1 latency. *Cell 65*:241–248.

Malim MH, Hauber J, Le S-Y, 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.

Malim MH, Tiley LS, McCarn DF, Rusche JR, Hauber J, Cullen BR. 1990. HIV-1 structural gene expression requires binding of the Rev trans-activator to its RNA target sequence. *Cell 60*:675–683.

Mathews DH, Sabina J, Zuker M, Turner DH. 1999. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J Mol Biol 288*:911–940.

Medstrand P, Mager DL. 1998. Human-specific integrations of the HERV-K endogenous retrovirus family. *J Virol 72*:9782–9787.

Neville M, Stutz F, Lee L, Davis LI, Rosbash M. 1997. The importin-beta family member Crm1p bridges the interaction between Rev and the nuclear pore complex during nuclear export. *Curr Biol 7*:767–775.

Ono M, Yasunaga T, Miyata T, Ushikubo H. 1986. Nucleotide sequence of human endogenous retrovirus genome related to the mouse mammary tumor virus genome. *J Virol 60*:589–598.

Patience C, Wilkinson DA, Weiss RA. 1997. Our retroviral heritage. *Trends Genet 13*:116–120.

Pollard VW, Malim MH. 1998. The HIV-1 Rev protein. *Annu Rev Microbiol 52*:491–532.

SenGupta DJ, Zhang B, Kraemer B, Pochart P, Fields S, Wickens M. 1996. A three-hybrid system to detect RNA-protein interactions in vivo. *Proc Natl Acad Sci USA 93*:8496–8501.

Stade K, Ford CS, Guthrie C, Weis K. 1997. Exportin 1 (Crm1p) is an essential nuclear export factor. *Cell 90*:1041–1050.

Taube R, Fujinaga K, Wimmer J, Barboric M, Peterlin BM. 1999. Tat transactivation: A model for the regulation of eukaryotic transcriptional elongation. *Virology 264*:245–253.

Tiley LS, Cullen BR. 1992. Structural and functional analysis of the visna virus Rev-response element. *J Virol 66*:3609–3615.

Tiley LS, Madore SJ, Malim MH, Cullen BR. 1992. The VP16 transcription activation domain is functional when targeted to a promoter-proximal RNA sequence. *Genes & Dev 6*:2077–2087.

Tristem M. 2000. Identification and characterization of novel human endogenous retrovirus families by phylogenetic screening of the human genome mapping project database. *J Virol 74*:3715–3730.

Yang J, Bogerd HP, Peng S, Wiegand H, Truant R, Cullen BR. 1999. An ancient family of human endogenous retroviruses encodes a functional homolog of the HIV-1 Rev protein. *Proc Natl Acad Sci USA 96*:13404–13408.

Zapp ML, Green MR. 1989. Sequence-specific RNA binding by the HIV-1 Rev protein. *Nature 342*:714–716.

Zapp ML, Hope TJ, Parslow TG, Green MR. 1991. Oligomerization and RNA binding domains of the type 1 human immunodeficiency virus Rev protein: A dual function for an arginine-rich binding motif. *Proc Natl Acad Sci USA 88*:7734–7738.