Specific Binding of the Human T-Cell Leukemia Virus Type I Rex Protein to a Short RNA Sequence Located within the Rex-Response Element

HAL P. BOGERD,¹ LAURENCE S. TILEY,^{1†} AND BRYAN R. CULLEN^{1,2*}

*Howard Hughes Medical Institute¹ and Section of Genetics, Department of Microbiology,²
Duke University Medical Center, Durham, North Carolina 27710*

Received 24 July 1992/Accepted 22 September 1992

Expression of the structural proteins of human T-cell leukemia virus type I is dependent upon the interaction of the viral Rex *trans* activator with its highly structured *cis*-acting RNA target sequence, the 254-nucleotide Rex-response element. Nucleotides critical for Rex binding in vitro have been mapped by modification interference analysis to a discrete 12-nucleotide RNA sequence that is predicted to form a stem-bulge-stem structure. This minimal RNA binding site was sufficient to mediate specific Rex binding in vitro when analyzed in the context of a short RNA probe. The critical importance of this short RNA sequence in mediating Rex function in vivo is supported by its complete conservation among all primate T-cell leukemia virus isolates.

Human T-cell leukemia virus type I (HTLV-I) gene expression is regulated by two virally encoded nuclear proteins termed Tax and Rex (for a review, see reference 7). Tax is a DNA sequence-specific transcriptional *trans* activator that markedly enhances the expression of genes linked to the HTLV-I long terminal repeat promoter element (10, 13, 21, 23). In contrast, Rex acts posttranscriptionally to induce the cytoplasmic expression of incompletely spliced HTLV-I transcripts that encode the viral structural gene products (1, 9, 11, 13, 21, 25). The *cis*-acting RNA target for Rex coincides with a highly structured 254-nucleotide (254-nt) RNA sequence, termed the Rex-response element (RxRE), that has been mapped to the U3/R region of the HTLV-I long terminal repeat (Fig. 1) (1, 11, 25). Several groups have now demonstrated that Rex can bind the RxRE with high affinity and specificity in vitro (2, 4, 5, 26).

Rex is closely analogous in function to the Rev protein of human immunodeficiency virus type 1 (HIV-1) and can, in fact, partly rescue a Rev-defective HIV-1 provirus (7, 20). Like Rex, Rev must directly interact with a highly structured RNA target, the 234-nt Rev-response element (RRE), to exert its effect on viral gene expression (6, 8, 16, 18, 28). Recently, it has been demonstrated that the latter interaction is dependent on a structured ~13-nt RNA sequence that forms the critical high-affinity binding site for Rev within the larger RRE (3, 12, 15, 24). In this report, we describe the identification of a structured ~12-nt RNA sequence that appears both necessary and sufficient for specific Rex binding in vitro.

Previous attempts to define RxRE sequences required for Rex binding in vitro have relied on mutational analysis (1, 2, 5). These studies have demonstrated that RNA sequences located between RxRE positions 1 and 146, 181 and 199, and 230 and 255 are dispensable for Rex binding in vitro (Fig. 1) (1, 2, 5). In contrast, two ~10-nt substitution mutations introduced either between positions 170 and 179 or between

positions 201 and 210 were found to abrogate not only Rex binding in vitro but also RxRE function in vivo (1, 5).

While these earlier studies demonstrated that sequences located within the RNA stem-loop extending from positions 149 to 226 within the RxRE (Fig. 1) were necessary for Rex binding, it had not been demonstrated that they were also sufficient for binding. To examine this question, we isolated the RxRE segment between positions 149 and 226, using the polymerase chain reaction (17). The oligonucleotide primers used also introduced flanking unique *Kpn*I and *Bam*HI sites at the 5' and 3' ends, respectively, of this truncated RxRE sequence. These sites facilitated the cloning of this DNA fragment into the *Kpn*I and *Bam*HI sites present in the polylinker of the in vitro RNA expression vector pGEM-3ZF(+) (Promega). After linearization of the resultant plasmid with *Bam*HI, the RxRE(149-226) RNA probe was transcribed by using T7 RNA polymerase, 3' end labelled with [5'-³²P]pCp and T4 RNA ligase, and purified by electrophoresis through a denaturing polyacrylamide gel (5, 24). The ability to bind Rex was assessed as previously described by using a purified recombinant fusion protein consisting of full-length Rex attached to the carboxy terminus of the enzyme glutathione-S-transferase (GST-Rex) (5). It has been demonstrated that GST-Rex binds the RxRE RNA in vitro with an affinity and a specificity that are indistinguishable from those of the recombinant wild-type Rex protein (5).

As shown by the RNA gel retardation analysis presented in Fig. 2, the RxRE(149-226) probe bound the GST-Rex protein efficiently (lane 2). Specificity was demonstrated by the observation that binding could be effectively inhibited by an excess of the unlabelled full-length RxRE RNA or by the unlabelled RxRE(149-226) RNA itself but not by an equivalent amount of tRNA (Fig. 2, lanes 3 to 5). We therefore conclude that the RxRE region between positions 149 and 226 is not only necessary but also sufficient for specific binding to Rex in vitro.

Modification interference analysis has previously proven highly effective in defining the minimal RNA binding site for the HIV-1 Rev protein (15, 24). We therefore used limited chemical modification of bases within the RxRE(149-226) probe to identify specific nucleotides within the RxRE that

* Corresponding author.

† Present address: Department of Virology, Bristol Myers Squibb, Princeton, NJ 08543.

FIG. 1. Predicted secondary structure of the HTLV-I RxE. The validity of this 254-nt RNA secondary structure prediction is supported by analysis of *in vitro* RNase cleavage patterns (11, 25). The RxE is located within the HTLV-I long terminal repeat and is divided by the site of transcription initiation (cap site) into U3 (5') and R (3') sections. The major splice donor site is indicated. The border of RxE(149–226) probe is indicated by a dashed line.

contribute to Rex binding *in vitro*. Modification of RNA with diethyl pyrocarbonate (DEPC) leads to carboxyethylation of the N⁷ position of the adenosine residue and, to a lesser extent, the N⁷ position of the guanosine residue (19). In contrast, treatment of RNA with hydrazine in the presence of 375 mM sodium chloride modifies the N³ position of both cytosine and uridine (19). These modifications can interfere with protein-RNA interactions that depend upon the specific recognition of particular purine or pyrimidine bases and also render the modified bases sensitive to specific cleavage by aniline (15, 19, 24).

The purified 3'-end-labelled RxE(149–226) probe was modified *in vitro* by using either DEPC or hydrazine-NaCl under conditions that introduced an average of one modification per RNA molecule (19, 24). After gel purification, the modified probes were incubated with sufficient GST-Rex protein to bind ~30% of the input probe. The bound RxE(149–226) probe was then separated from the free probe by nondenaturing gel electrophoresis (5, 24). Free and bound probes were recovered from the gel and cleaved by treatment with aniline (19, 24) prior to electrophoresis through a 10% acrylamide-urea sequencing gel. The DEPC-treated (Fig. 3, lanes 1 and 2) and hydrazine-NaCl-treated (Fig. 3, lanes 3 and 4) probes were run in parallel, thus generating a chemical RNA sequencing ladder (19).

If the chemical modification of a particular base alters the affinity of binding of the probe to Rex *in vitro*, then that base will be identified by a decrease or, less commonly, an increase in the intensity of the appropriate specific band in the bound-probe lanes (Fig. 3, lanes 2 and 4) compared with the free-probe lanes (Fig. 3, lanes 1 and 3). In fact, this analysis clearly demonstrated a marked interference with GST-Rex binding upon DEPC modification of purine bases located between positions 174 and 180 in the RxE as well as upon hydrazine modification of pyrimidine bases located

FIG. 2. Analysis of *in vitro* interaction of Rex with subdomains of RxE. (Lanes 1 to 5) The specific interaction of ³²P-labelled RxE(149–226) RNA probe with GST-Rex was visualized by RNA gel retardation analysis as previously described (5). Briefly, ~500 ng of the partially purified GST-Rex fusion protein was suspended in 9 μ l of binding buffer together with 1 μ g of *Escherichia coli* tRNA and 4 μ g of *E. coli* tRNA. Additional specific or nonspecific unlabelled competitor RNA was added at this time. Lanes contain competitor RNAs as follows: lane 3, 1.25 μ g of full-length RxE RNA; lane 4, 1.25 μ g of RxE(149–226) RNA; lane 5, 1.25 μ g of *E. coli* tRNA. The reaction mixture was preincubated for 10 min before addition of labelled RxE(149–226) probe (~4 ng at ~10⁴ cpm/ng) and was then incubated for a further 10 min prior to analysis on a nondenaturing 5% polyacrylamide gel. The result with the probe incubated in the absence of GST-Rex is shown in lane 1. (Lanes 6 to 11) Analysis of the interaction of the predicted minimal RxE probe (Fig. 4B) with GST-Rex was performed exactly as described above except that two different levels of unlabelled RxE(149–226) RNA and *E. coli* tRNA were tested. Lanes 8 and 10, 1.25 μ g; lanes 9 and 11, 0.25 μ g. No Comp., no additional competitor RNA.

between positions 200 and 204 (Fig. 3). Some interference was also detected at the uridine at position 176. Interestingly, we also consistently observed a significant enhancement in binding due to hydrazine modification of cytosines at positions 172, 177, and 198. Although the two RxE RNA sequence elements identified by this analysis as important to Rex binding are significantly separated in terms of primary sequence, computer analysis (11) indicates that they are precisely aligned in the RNA secondary structure formed by the RxE (Fig. 1 and 4A). In fact, this analysis suggests that the core RxE binding site forms a 12-nt stem-bulge-stem structure that appears somewhat analogous in terms of structure, but not primary sequence, to the 13-nt stem-bulge-stem RNA structure recognized by the HIV-1 Rev protein (3, 12, 15, 24).

To test whether the Rex RNA core binding site is indeed both necessary and sufficient for Rex binding *in vitro*, we synthesized a short RNA probe (24) that retains a mere 24 nt of the RxE sequence flanked by heterologous RNA sequences that were designed to maintain a similar structural context (Fig. 4B). This predicted minimal Rex binding site probe indeed retained the ability to bind the GST-Rex protein effectively. Binding was specific, as demonstrated by the observation of competition upon challenge with unlabelled RxE(149–226) RNA but not with equivalent amounts of tRNA (Fig. 2, lanes 6 to 11).

We have used modification interference analysis (19) to

FIG. 3. Chemical interference analysis of Rex binding to an RxE RNA probe. A 3'-end-labelled RxE(149–226) RNA probe was subjected to limited chemical modification (19) by using either DEPC (lanes 1 and 2), which affects purine bases, or hydrazine-NaCl (lanes 3 and 4), which is specific for pyrimidines. After purification, the modified RNA probes were incubated with sufficient GST-Rex protein to bind ~30% of the input probe. The free probe (lanes 1 and 3) and bound probe (lanes 2 and 4) were then isolated by native gel electrophoresis (Fig. 2) and cleaved by treatment with aniline (19, 24). A total of 4×10^5 cpm derived from each RNA sample was subjected to electrophoresis through a denaturing 10% polyacrylamide sequencing gel, and cleavage products were visualized by autoradiography. DEPC treatment modifies adenine residues significantly more extensively than guanine residues, while hydrazine-NaCl treatment modifies pyrimidine bases equivalently (19). The pattern of aniline cleavage observed in the free-probe lanes therefore facilitated the unambiguous identification of RxE nucleotides that displayed chemical interference. These are indicated on both sides of the gel and are compiled in Fig. 4A.

define a structured 12-nt RNA sequence located within the HTLV-I RxE as critical to specific Rex binding in vitro (Fig. 3 and 4A). A short RNA probe containing this 12-nt Rex core binding site (Fig. 4B) was able to bind Rex specifically in vitro (Fig. 2), thus suggesting that this sequence is not only necessary but also sufficient. In fact, the Rex core binding site overlaps with a 20-nt sequence previously identified by mutational analysis as critical both for in vivo RxE function and for in vitro Rex binding (1, 5). This sequence is also fully contained within a 37-nt RxE segment shown by Toyoshima et al. (25) to be necessary for detectable in vivo RxE activity. It is therefore likely that

FIG. 4. Definition of the core RNA binding site for HTLV-I Rex. (A) Computer-predicted secondary structure of the high-affinity RNA binding site for Rex (11, 25). Nucleotides showing strong (>80%) or significant (50 to 80%) interference with Rex binding after chemical modification are indicated by continuous or dashed circles, respectively. Nucleotides that give rise to enhanced (twofold or greater) binding after chemical modification are indicated (+). Data were compiled from the experiment whose results are shown in Fig. 3 as well as three other independent experiments. (B) Primary sequence and predicted RNA secondary structure of the minimal Rex binding site validated by the results shown in Fig. 2 (lanes 6 to 11). The probe contains the 12-nt predicted core Rex binding site (boxed) determined by chemical interference (panel A) together with an additional 12 nt derived from the RxE (uppercase letters). Heterologous sequences added to maintain the predicted RNA secondary structure of this part of the RxE or required for efficient in vitro transcription are indicated by lowercase letters.

this 12-nt sequence forms the single high-affinity binding site for Rex within the 254-nt RxE just as HIV-1 Rev has been shown to have only a single 13-nt high-affinity binding site within the 234-nt RRE (3, 12, 15, 24).

The Rex proteins of HTLV-I (Rex-I) and the related HTLV-II (Rex-II) are able to fully cross-complement in vivo (14), thus suggesting that these proteins have retained similar or identical RNA sequence specificities. Mutational analysis of the RxE of HTLV-II has, in fact, suggested that the primary Rex-II RNA binding site is in an equivalent position within this structurally similar RNA sequence element (4, 14). However, the Rex-I binding site defined in this analysis excludes the conserved splice donor site that has been proposed by others to be critical for Rex-II binding in vitro (4). We note that this exclusion is consistent with data demonstrating that this splice site is fully dispensable for HTLV-I RxE function in vivo (25). As shown in Fig. 5, alignment of the relevant portions of the HTLV-I and HTLV-II RxE sequences demonstrates complete conservation of the proposed Rex-I core binding site. In contrast, other parts of the indicated RxE sequences (excluding the highly conserved major splice donor sequence) display less than 50% sequence identity. The fact that the in vitro defined Rex-I core binding site has been fully conserved during the evolutionary divergence of HTLV-I and HTLV-II is consistent with the hypothesis that this sequence plays a critical

FIG. 5. Sequence alignment of predicted Rex binding sites. The sequence around the Rex core binding site defined here for HTLV-I is shown aligned to equivalent sequences in the closely related simian T-cell leukemia virus (STLV) (27) and in the more distantly related HTLV-II (22). Identity in sequence to HTLV-I is indicated by a dash, while a deletion is indicated by a delta. The major splice site and the Rex core binding site (Fig. 4A) are highly conserved among these distinct viruses compared with flanking sequences. Although the *in vivo* target specificity of the Rex protein of the distantly related bovine leukemia virus is not identical to that of Rex-I (9), the bovine leukemia virus RxE does contain a structured sequence similar to the Rex-I core recognition element (i.e., 5'-GGGtGAG-3' and 5'-CTGtgaCC-3') (25). The Rex-I protein has also been shown to bind the HIV-1 RRE specifically *in vitro*, albeit with a significantly lower affinity than to the homologous RxE RNA (5, 26). However, inspection of the RRE has failed to reveal an RNA sequence element closely similar to the Rex-I core binding site defined here.

role in mediating the interaction of Rex with the RxE in both the HTLV-I and the HTLV-II systems.

We thank Sharon Goodwin for secretarial assistance.
Supported by the Howard Hughes Medical Institute.

REFERENCES

- Ahmed, Y. F., S. M. Hanley, M. H. Malim, B. R. Cullen, and W. C. Greene. 1990. Structure-function analyses of the HTLV-I Rex and HIV-1 Rev RNA response elements: insights into the mechanism of Rex and Rev action. *Genes Dev.* 4:1014-1022.
- Ballaun, C., G. K. Farrington, M. Dobrovnik, J. Rusche, J. Hauber, and E. Böhnlein. 1991. Functional analysis of human T-cell leukemia virus type I *rex*-response element: direct RNA binding of Rex protein correlates with *in vivo* activity. *J. Virol.* 65:4408-4413.
- Bartel, D. P., M. L. Zapp, M. R. Green, and J. W. Szostak. 1991. HIV-1 Rev regulation involves recognition of non-Watson-Crick base pairs in viral RNA. *Cell* 67:529-536.
- Black, A. C., C. T. Ruland, M. T. Yip, J. Luo, B. Tran, A. Kalsi, E. Quan, M. Aboud, I. S. Y. Chen, and J. D. Rosenblatt. 1991. Human T-cell leukemia virus type II Rex binding and activity require an intact splice donor site and a specific RNA secondary structure. *J. Virol.* 65:6645-6653.
- Bogerd, H. P., G. L. Huckaby, Y. F. Ahmed, S. M. Hanly, and W. C. Greene. 1991. The type I human T-cell leukemia virus (HTLV-I) Rex trans-activator binds directly to the HTLV-I Rex and the type 1 human immunodeficiency virus Rev RNA response elements. *Proc. Natl. Acad. Sci. USA* 88:5704-5708.
- Cochrane, A. W., C.-H. Chen, and C. A. Rosen. 1990. Specific interaction of the human immunodeficiency virus Rev protein with a structured region in the *env* mRNA. *Proc. Natl. Acad. Sci. USA* 87:1198-1202.
- Cullen, B. R. 1992. Mechanism of action of regulatory proteins encoded by complex retroviruses. *Microbiol. Rev.* 56:375-394.
- Daly, T. J., K. S. Cook, G. S. Gray, T. E. Maione, and J. R. Rusche. 1989. Specific binding of HIV-1 recombinant Rev protein to the Rev-responsive element *in vitro*. *Nature (London)* 342:816-819.
- Felber, B. K., D. Derse, A. Athanassopoulos, M. Campbell, and G. N. Pavlakis. 1989. Cross-activation of the Rex proteins of HTLV-I and BLV and of the Rev protein of HIV-1 and nonreciprocal interactions with their RNA responsive elements. *New Biol.* 1:318-330.
- Fujisawa, J.-I., M. Seiki, M. Sato, and M. Yoshida. 1986. A transcriptional enhancer sequence of HTLV-I is responsible for *trans*-activation mediated by p40^{rev} of HTLV-I. *EMBO J.* 5:713-718.
- Hanly, S. M., L. T. Rimsky, M. H. Malim, J. H. Kim, J. Hauber, M. Duc Dodon, S.-Y. Le, J. V. Maizel, B. R. Cullen, and W. C. Greene. 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., J. T. Finch, M. J. Gait, J. Karn, and M. Singh. 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.
- Inoue, J.-I., M. Yoshida, and M. Seiki. 1987. Transcriptional (p40^{rev}) and post-transcriptional (p27^{rev}) regulators are required for the expression and replication of human T-cell leukemia virus type I genes. *Proc. Natl. Acad. Sci. USA* 84:3653-3657.
- Kim, J. H., P. A. Kaufman, S. M. Hanly, L. T. Rimsky, and W. C. Greene. 1991. Rex transregulation of human T-cell leukemia virus type II gene expression. *J. Virol.* 65:405-414.
- Kjems, J., B. J. Calnan, A. D. Frankel, and P. A. Sharp. 1992. Specific binding of a basic peptide from HIV-1 Rev. *EMBO J.* 11:1119-1129.
- Malim, M. H., L. S. Tiley, D. F. McCarn, J. R. Rusche, J. Hauber, and B. R. Cullen. 1990. HIV-1 structural gene expression requires binding of the Rev *trans*-activator to its RNA target sequence. *Cell* 60:675-683.
- Mullis, K. B., and F. A. Faloona. 1987. Specific synthesis of DNA *in vitro* via a polymerase-catalyzed chain reaction. *Methods Enzymol.* 155:335-350.
- Olsen, H. S., P. Nelbock, A. W. Cochrane, and C. A. Rosen. 1990. Secondary structure is the major determinant for interaction of HIV *rev* protein with RNA. *Science* 247:845-848.
- Peattie, D. A. 1979. Direct chemical method for sequencing RNA. *Proc. Natl. Acad. Sci. USA* 76:1760-1764.
- Rimsky, L., J. Hauber, M. Dukovich, M. H. Malim, A. Langlois, B. R. Cullen, and W. C. Greene. 1988. Functional replacement of the HIV-1 *rev* protein by the HTLV-I *rex* protein. *Nature (London)* 335:738-740.
- Rosenblatt, J. D., A. J. Cann, D. J. Slamon, I. S. Smalberg, N. P. Shah, J. Fujii, W. Wachsmann, and I. S. Y. Chen. 1988. HTLV-II transactivation is regulated by the overlapping *tax/rex* nonstructural genes. *Science* 240:916-919.
- Shimotohno, K., W. Wachsmann, Y. Takahashi, D. W. Golde, M. Miwa, T. Sugimura, and I. S. Y. Chen. 1984. Nucleotide sequence of the 3' region of an infectious human T-cell leukemia virus type II genome. *Proc. Natl. Acad. Sci. USA* 81:6657-6661.
- Sodroski, J., C. Rosen, W. C. Goh, and W. Haseltine. 1985. A transcriptional activator protein encoded by the *x-lor* region of the human T cell leukemia virus. *Science* 228:1430-1434.
- Tiley, L. S., M. H. Malim, H. K. Tewary, P. G. Stockley, and B. R. Cullen. 1992. Identification of a high-affinity RNA-binding site for the human immunodeficiency virus type 1 Rev protein. *Proc. Natl. Acad. Sci. USA* 89:758-762.
- Toyoshima, H., M. Itoh, J.-I. Inoue, M. Seiki, F. Takaku, and M. Yoshida. 1990. Secondary structure of the human T-cell leukemia virus type 1 *rex*-responsive element is essential for *rex* regulation of RNA processing and transport of unspliced RNAs. *J. Virol.* 64:2825-2832.
- Unge, T., L. Solomin, M. Mellini, D. Derse, B. K. Felber, and G. N. Pavlakis. 1991. The Rex regulatory protein of human T-cell lymphotropic virus type I binds specifically to its target site within the viral RNA. *Proc. Natl. Acad. Sci. USA* 88:7145-7149.
- Watanabe, T., M. Seiki, Y. Hirayama, and M. Yoshida. 1986. Human T-cell leukemia virus type I is a member of the African subtype of simian viruses (STLV). *Virology* 148:385-388.
- Zapp, M. L., and M. R. Green. 1989. Sequence-specific RNA binding by the HIV-1 Rev protein. *Nature (London)* 342:714-716.