Distinct *cis*-acting regions in U3 regulate *trans*-activation of the human spumaretrovirus long terminal repeat by the viral *bel*1 gene product

L.K.Venkatesh, P.A.Theodorakis and G.Chinnadurai*

Institute for Molecular Virology, Saint Louis University Medical Center, 3681 Park Avenue, Saint Louis, MO 63110, USA

Received March 8, 1991; Revised and Accepted June 3, 1991

ABSTRACT

The human spumaretrovirus (HSRV) genome contains, in addition to coding information for the structural proteins, open reading frames (ORFs) for at least three additional genes termed *bel*1, *bel*2 and *bel*3. We report here the localization of the transcriptional activator of HSRV to the *bel*1 ORF. In reporter-based transient expression assays in COS cells utilizing the bacterial CAT gene linked to HSRV LTR sequences between –710 and +309 with respect to the transcriptional initiation site, co-expression of the *bel*1 gene product alone caused an over 100 to 300-fold increase in the level of LTR activity. High-level *trans*-activation by *bel*1 was specific for the HSRV LTR, as relatively minor positive and negative regulatory effects were observed on HIV-1 LTR and RSV LTR expression, respectively, whereas HTLV-I LTR activity remained unaffected. Distinct regions of the HSRV LTR were found to be involved in *bel*1-induced *trans*-activation. Specifically, deletions between –500 and –389 and between –136 and –62 in the U3 region resulted in a 4- and 30 to 35-fold decline, respectively, in the response to *bel*1. Limited mutagenesis of the *bel*1 ORF indicated that most of the *bel*1 coding region, except for the carboxy-terminal 27 residues, is essential for the activation function.

INTRODUCTION

Human spumaretroviruses (HSRVs) belong to a distinct group of retroviruses, the spumaretroviruses or foamy viruses, and have been isolated from humans with a variety of diseases although proven pathogenicity for any isolate has yet to be established (1,2). The genome of an HSRV isolate from a nasopharyngeal carcinoma (3) has been molecularly cloned and characterized and these studies have revealed that HSRV has a genomic complexity similar to other human retroviruses of the lentivirus and oncovirus families (4,5,6). Genome sequence analysis and structural mapping of HSRV transcripts in infected human cells have indicated that, in addition to coding for the virion structural

proteins Gag, Pol and Env, HSRV has the potential to code for several novel genes, collectively termed *bel*, that are located between the env gene and the 3' long terminal repeat (LTR). It has been suggested that the *bel* open reading frames (ORFs), by virtue of their genomic location and limited homology to certain regulatory genes of human lentiviruses, may serve regulatory functions that qualitatively and quantitatively affect HSRV gene expression and the life cycle (4,7,8). In agreement with this prediction, it has been reported that the level of HSRV LTR-directed gene expression markedly increases in HSRV-infected cells, suggesting that HSRV encodes a transcriptional activator (9).

In the present study, using a replication-competent plasmid vector to express various segments of the HSRV genome under the control of the powerful cytomegalovirus immediate early (CMV-IE) promoter in monkey COS cells, we have localized the transcriptional activator function of HSRV to the *bel*1 ORF. Additionally, we have identified *cis*-acting regions within the HSRV LTR that regulate *bel*1-mediated *trans*-activation. Finally, limited mutagenesis of the *bel*1 ORF has identified the putative translation initiator codon as well as the minimal region required for *trans*-activation.

MATERIALS AND METHODS

HSRV expression plasmids

The plasmids utilized for the isolation and expression of various segments spanning the HSRV genome were pHSRV-B52 which carries HSRV sequences from nucleotide position 9883 in the *bel*1 ORF (BamHI site) to the 3' end of the viral genome and pHSRV-H-C55 which harbors HSRV sequences between the HindIII sites at nt 5335 and 10699 (4). Progressively larger segments originating from the right end of the viral genome (EcoRI site in pHSRV-B52) were cloned into the expression vector pgTAT (10) between the SalI (blunt-ended) and the EcoRI sites thereby placing them under the control of the CMV-IE promoter and the HSRV poly A addition site in the 3' LTR. The various HSRV genomic segments expressed in this manner are illustrated in the lower-half of Fig. 1A. Plasmids HSRV-C1, -C2,

* To whom correspondence should be addressed

Figure 2. Specificity of *bell* trans-activation. COS cells were transfected with 2 μ g of the various retroviral LTR-linked CAT reporter plasmids, as indicated, and 1 μ g of either pCMV (-*bell*) or pCMV-*bell* (+*bell*). CAT activity for a given LTR in the presence of *bell*, as an average of three independent experiments, is expressed relative to its activity in the absence of *bell* assigned a value of 1.0.

human immunodeficiency virus type 1 long terminal repeats, respectively, have been described (12,13,14).

Cells, DNA transfection and CAT assay

The monkey kidney cell line, COS-7, was maintained in Dulbecco's modified Eagle medium containing 10% fetal bovine serum. For DNA transfections, 2×10^5 COS cells per 60mm dish were transfected with the indicated plasmid DNAs as a calcium phosphate coprecipitate. The total amount of plasmid DNA per transfection was maintained at $12 \mu\text{g}$ with carrier pUC13 DNA. A plasmid expressing *E. coli* β -galactosidase was invariably cotransfected as a control for normalizing transfection efficiencies. Cell extracts were prepared 48 hours after transfection and CAT activity was assayed as described (15) utilizing extract amounts normalized for equal β -galactosidase activity. CAT activity was quantitated by scintillation counting of excised acetylated chloramphenicol spots after thin-layer chromatography. Reproducibility in the patterns of CAT expression was confirmed by repeating each experiment at least three times. Levels of CAT activity in different experiments ranged between 5–15% of the average value indicated for each transfection.

RESULTS

Identification of the HSRV *trans*-activator

To identify the genomic region that encodes the *trans*-activator of the HSRV LTR, an overlapping set of genomic fragments, C1–C4 (Fig. 1A), originating at the 3' LTR and thus retaining the viral 3' poly A addition signal, was cloned downstream of

the powerful CMV-IE promoter in a plasmid vector that is replication-competent in the monkey kidney cell line COS (see Materials and Methods), thereby permitting high-level expression of the viral genes contained in these HSRV regions. These plasmids were examined for *trans*-activation potential in transient expression assays of CAT activity by transfection into COS cells along with the reporter plasmid –710/+309 HSRV LTR CAT (see Materials and Methods for description; also Fig. 3A). Expression of the C1 region (*i.e.* pHSRV-C1) did not induce significant activation of HSRV LTR CAT expression above that in cells transfected with the negative control vector pCMV. In contrast, expression of the C2 region, which retains coding potential for all *bel* genes, resulted in an approximately 60-fold increase in CAT expression (Fig. 1B). Inclusion of additional 5' sequences from the *pol-env* region in C3 and C4 did not cause any further increase in CAT expression; rather, a 2–3 fold decline was observed in the ability to activate the HSRV LTR upon expression of these additional regions in C3 and C4, compared to C2. When an XbaI fragment containing the *bel* region was deleted from C4, the activation capability of this clone was lost (lane C4 Δ B), thereby suggesting that the region expressing the *bel* ORFs encodes the HSRV *trans*-activator. Since the HSRV sequences in C2, but not C1 alone, were found to be sufficient for *trans*-activation, it appeared that the *bel1* ORF that comprises sequences from both the C1 and C2 regions may encode the *trans*-activator. We therefore examined the effect of a BamHI deletion, between nt 9675 and 9883, exclusively in the *bel1* ORF of C2 on HSRV LTR CAT expression. The resulting clone, C2 Δ B, was unable to activate the HSRV LTR, thereby predicting the involvement of the *bel1* ORF in HSRV *trans*-

Figure 3. Effect of 5' deletions on HSRV LTR activity. **A**, Organization of the HSRV LTR in –710/+309 LTR CAT. The U3, R and U5 regions are shown. Restriction sites abolished during cloning, including a BglII site (not shown) coincident with the MstII site, are indicated in parentheses. **B**, CAT expression in COS cells transfected with 4 μg of the 5' LTR deletion mutants alone (**a**, –*bel1*) or with 1 μg of the indicated LTR mutant and 1 μg of pCMV-*bel1* (**b**, +*bel1*). CAT reaction in panel **a** was performed for a period 6 times longer than in panel **b** to permit visualization of the very low levels of uninduced LTR expression. CAT activity in each transfection (average of four experiments) is expressed relative to the –710/+4 LTR-directed level assigned a value of 1.0.

activation. In agreement with this prediction, the expression of a 1048 bp SspI fragment which contains the *bel1* ORF alone was found to *trans*-activate the HSRV LTR $\approx$ 150-fold. Again, an internal BamHI deletion (identical to that in C2ΔB) of the *bel1* gene (in pCMV *bel1*) abolished its capacity to *trans*-activate (lane *bel1*ΔB). Thus, these results suggest that the *bel1* ORF codes for a powerful *trans*-activator of the HSRV LTR and that the *bel2* and *bel3* ORFs are dispensable for *trans*-activation.

Specificity of *bel1* *trans*-activation

To determine whether *trans*-activation by *bel1* was specific for the HSRV LTR, we examined the effects of *bel1* on the expression of CAT genes linked to LTRs of various other retroviruses such as the Rous sarcoma virus (RSV), human T cell leukemia virus type I (HTLV-I) and human immunodeficiency virus type 1 (HIV-1). As shown in Fig. 2, *bel1* consistently induced a 2 to 3-fold reduction in RSV LTR expression. Whereas HTLV-I LTR expression remained essentially unchanged, the HIV-1 LTR was found to be activated about 4-fold, in the presence of *bel1*. Thus, the effects of *bel1* expression on the activity of heterologous retroviral LTRs appear to be relatively marginal, thereby suggesting that *bel1*-mediated *trans*-activation is specific for the HSRV LTR.

Effect of 5' deletions in HSRV LTR

The parental plasmid for studying the effects of sequences upstream from the transcriptional initiation site on HSRV LTR expression was -710/+309 HSRV LTR CAT that contains 710 bp of the U3 region upstream of the cap site (+1), in addition to downstream sequences comprising 189 bp of the R and 120 bp of U5 regions of the LTR (Fig. 3A). During initial experiments (results not shown), we found that the -710/+309 LTR was *trans*-activated to the same extent as the -915/+309 LTR in pHSRV-CAT (9) and that the deletion of R-U5 sequences between +4 and +309 resulted in an $\approx$ 10-fold increase in induced levels of HSRV LTR activity. Therefore, all deletions introduced upstream of the cap site (in the U3 region) were examined in the +4 background to facilitate resolution of the effects of various 5' deletion mutants. As shown in Fig. 3B (panel a), deletion of U3 sequences between -710 and -15 did not have any significant effect on the uninduced level of LTR expression. In the presence of *bel1*, deletion between -710 and -500, which resulted in the removal of two sites for the transcription factor AP-1, caused only a marginal drop in LTR activity (Fig. 3B, panel b). A significant reduction, about 4 to 5-fold, in induced LTR expression occurred upon further deletion to -389. Thereafter, CAT expression remained relatively constant upon removal of U3 sequences between -389 and -136. Further deletion of HSRV sequences located between -136 and -62 resulted in a 30 to 35-fold reduction in *bel1*-induced CAT activity, thereby suggesting that the -136/-62 region constitutes a major target site for *bel1* *trans*-activation. Further deletion of sequences up to -16, inclusive of the TATA box sequence at -31, resulted in a very marginal decline in induced LTR expression with the corresponding level of CAT activity approaching that observed with the uninduced -710/+4 HSRV-LTR construct (data not shown). A similar pattern of CAT expression was observed when the 5' LTR deletions were examined for activity in the +309 background. The data thus indicates two distinct target sites for *bel1* interaction that reside in the U3 region of the HSRV LTR; a minor site between -500 and -389 and a major site between -136 and -62.

Figure 4. Effect of chain-termination mutants of *bel1* on HSRV LTR activity. **A**, the complete amino acid sequence of the *bel1* ORF is shown. For each truncation mutant, an arrow indicating the terminal *bel1* residue coded for, the missense amino acids (in lower case) introduced by linker insertion, and the designation of the corresponding mutant are indicated. Highly basic regions of potential nuclear localization/nucleic acid binding function are shown in italics. Met codons are underlined. **B**, CAT activity in COS cells cotransfected with 1 μ g of -710/+4 HSRV LTR CAT and 1 μ g of either pCMV (negative control) or the various *bel1*-expressing plasmids as indicated. CAT activities (average of three experiments) are expressed relative to the uninduced (lane pCMV) -710/+4 LTR-directed level taken as 1.0.

Chain termination mutants of *bel1* ORF

The *bel1* ORF contains a number of Met residues (Fig. 4A) and it is not known, at present, as to which of these Met residues is utilized as the initiator ATG codon. A series of nonsense mutations was generated at several sites within the *bel1* ORF in the plasmid pCMV-*bel1* (as described under Materials and Methods) to identify the putative translation initiator codon and to delineate the minimal protein region required for *trans*-activation (Fig. 4A). Plasmids expressing wt or truncated forms of *bel1* were examined for their ability to *trans*-activate -710/+4 HSRV LTR CAT. The wt *bel1* gene was found to enhance CAT expression about 300-fold (lane *bel1*) above the uninduced level (lane pCMV). *bel1*-16 which retains coding potential for the amino-terminal 16 residues was found to be defective for *trans*-activation as were mutants that could code for the amino-terminal 135, 173, 208, 225 or 238 residues. Among the mutants examined, only *bel1*-273, which lacks the carboxy-terminal 27 amino acids, was found to *trans*-activate the HSRV LTR efficiently. The fact that the *bel1*-16 mutation, which terminates translation immediately downstream of the first Met codon in the *bel1* ORF, was functionally defective suggests that this ATG codon functions as the *bel1* translation initiator and that initiation from downstream Met codons, if it occurs, is unlikely to yield

functionally active *bel1* proteins. Further, it appears that the amino-terminal 273 residues of *bel1* are sufficient for *trans*-activation of the HSRV LTR.

DISCUSSION

We have demonstrated conclusively that the *bel1* ORF of human spumaretrovirus codes for a potent *trans*-activator of the viral LTR. A 1048 bp region of the HSRV genome between nucleotides 9335 and 10383 that contains complete coding information for the *bel1* gene alone, when expressed under the control of the powerful CMV-IE promoter, was found to activate HSRV LTR expression at least 100 to 300-fold in transient expression assays. Furthermore, a deletion within the *bel1* ORF that leaves only the amino-terminal 82 residues intact was found to abolish its activator function. The consistently lower levels of *trans*-activation induced by the C2, C3 and C4 regions relative to the 1048 bp *bel1* region observed in our experiments may possibly be due to lower levels of *bel1* expression from plasmids containing the C2, C3 and C4 regions; alternatively, it might reflect the expression of additional virus-encoded modulators of LTR expression. It has been recently demonstrated that the closely related simian foamy virus (SFV) also encodes a transcriptional activator of its LTR (16). *bel1* *trans*-activation was specific for the HSRV LTR as HTLV-I LTR expression was not significantly affected. Similarly, the SFV *trans*-activator has been reported to have no effect on HTLV-I LTR activity (16). Marginal suppression and activation effects were observed on RSV LTR and HIV-1 LTR expression, respectively, and it remains to be determined if *bel1* is pleiotropic in function like the adenovirus E1A *trans*-activator (17).

Analysis of progressive 5' deletions within the U3 region on HSRV LTR expression confirmed that this region is primarily responsible for the activated response to *bel1*. The major site for *bel1* *trans*-activation appears to be located within the -136/-62 region. Uninduced levels of LTR expression remain unaltered between the -136 and -62 deletions whereas a 30 to 40-fold decline occurred in the *trans*-activation response. This strongly suggests that *bel1* may function primarily by increasing transcription initiation. A version of the CCAAT consensus binding site for the transcription factor C/EBP, CCACT, that has been found to be functionally active in the herpes virus *tk* and the Moloney murine sarcoma virus promoters (18) is located at -99 and its role in HSRV LTR expression remains to be ascertained. An additional, but relatively less prominent target, for *bel1* interaction was the -500/-389 region. No significant difference in the activated response was observed upon deletion of the -389/-201 region. This region of the HSRV 3' LTR contains 139 nucleotides, postulated to be of cellular origin, that is absent from the 5' LTR. The -710/-500 region contains two completely homologous binding sites (TGACTCA) for the transcription factor AP-1 that mediates transcription induction by phorbol ester tumor promoters such as TPA (reviewed in ref. 19). Only a marginal reduction in activated LTR expression occurred upon the deletion of this region and the role, if any, of these AP-1 sites in HSRV LTR expression remains to be investigated. Interestingly, AP-1 binding sites in the U3 region of the visna virus LTR are critical for mediating the response to the viral *trans*-activator (20). The HTLV-I Tax *trans*-activator acts through a 21 bp repeat in the U3 region of the HTLV-I LTR (21) that also appears to contain an AP-1-like motif (22). Similar patterns of CAT expression were also observed when the above-

described HSRV LTR 5' deletions were examined for *bel1* responsiveness in human HeLa cells, thereby ruling out any effects of the SV40 T-antigen that is constitutively expressed in COS cells on HSRV LTR expression. After the submission of this manuscript, Rethwilm *et al.* (23) and Keller *et al.* (24) also reported the involvement of the U3 region of the HSRV LTR in *trans*-activation by *bel1*, in agreement with our observations.

Since HSRV LTR sequences between -136 and +4 are sufficient for *trans*-activation by *bel1*, it would appear from our studies that the R-U5 region of the HSRV LTR is dispensable for *trans*-activation by *bel1*. Preliminary studies with progressive deletions in the R-U5 region have indicated an ≈ 10 -fold increase in the level of *bel1*-induced HSRV LTR activity upon deletion of the +4/+309 region in -710/+309 HSRV LTR CAT (L.V. and G.C., unpublished observations) and the basis for this negative regulation is currently being examined. About 85% homology has been reported between the R and U5 regions of SFV and HSRV (16) and it remains to be seen if a common negative regulatory element of spumavirus LTRs exists in this region. These results are in contrast to those observed for the HIV-1 (reviewed in ref. 25) and infectious anemia virus (26,27) promoters, where sequences present in the R region (in the case of HIV, a nascent RNA stem-loop structure (28,29) termed TAR) constitute the primary target for *trans*-activation and whose removal abolishes the response to the cognate activator.

The amino terminal 273 residues of the *bel1* ORF were found to be sufficient for the *trans*-activation function. Present in this region are two highly basic stretches with the potential for functioning as nuclear targeting and/or nucleic acid binding motifs. Our analysis of nonsense mutants suggests that translation initiates at the first Met codon in the *bel1* ORF. Detailed mutagenesis to identify the functional domains of the *bel1* gene product should clarify if the organization of this activator conforms to that of the typical eukaryotic transcriptional activator with distinct nucleic-acid binding and activating regions (30).

ACKNOWLEDGEMENTS

We thank R.Flugel, A.Rethwilm, P.Luciw and I.S.Y.Chen for various plasmids. We also thank R.Flugel and P.Luciw for helpful discussions. This work was supported by research grants AI-29541 and AI-29000.

REFERENCES

- Teich, N. (1984) in Weiss, R.A., Teich, N., Varmus, H., and Coffin, J. (eds.)—RNA Tumor Viruses, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, Vol. 1, pp. 157–163.
- Maurer, B. and Flugel, R.M. (1988) AIDS Res. Human Retrovir., 4, 467–473.
- Achong, G., Mansell, P.W.A., Epstein, M.A. and Clifford, P. (1971) J. Natl. Cancer Inst., 42, 299–307.
- Flugel, R.M., Rethwilm, A., Maurer, B. and Darai, G. (1987) EMBO J., 6, 2077–2084.
- Maurer, B., Bannert, H., Darai, G. and Flugel, R.M. (1988) J. Virol., 62, 1590–1597.
- Muranyi, M., and Flugel, R.M. (1990) J. Virol., 65, 727–735.
- Maurer, B. and Flugel, R.M. (1987) FEBS Lett., 222, 286–288.
- Rethwilm, A., Baunach, G., Netzer, K., Maurer, B., Borisch, B. and ter Meulen, V. (1990) Nucleic Acids Res., 18, 733–738.
- Rethwilm, A., Mori, A., Maurer, B. and ter Meulen, V. (1990) Virology, 175, 568–571.
- Malim, M.H., Hauber, J., Le, S., Maizel, J.V. and Cullen, B.R. (1989) Nature, 338, 254–257.
- Lillie, J.W. and Green, M.R. (1989) Nature, 338, 39–44.

12. Gorman, C.M., Merlino, G.T., Willingham, M.C., Pastan, I. and Howard, B.H. (1982) *Proc. Natl. Acad. Sci. USA*, 79, 6777–6781.
13. Chen, I.S.Y., Cann., A.J., Shah, N.P. and Gaynor, R.B. (1985) *Science*, 230, 570–573.
14. Peterlin, B.M., Luciw, P.A., Barr, P.J. and Walker, M.D. (1986) *Proc. Natl. Acad. Sci. USA*, 83, 9734–9738.
15. Gorman, C.M., Moffat, L.F. and Howard, B.H. (1982) *Mol. Cell. Biol.*, 2, 1044–1051.
16. Mergia, A., Shaw, K., Pratt-Lowe, E., Barry, P.A. and Luciw, P.A. (1990) *J. Virol.*, 64, 3598–3604.
17. Moran, E. and Mathews, M.B. (1987) *Cell*, 48, 177–178.
18. Graves, B.J., Johnson, P.F. and McKnight, S.L. (1986) *Cell*, 44, 565–571.
19. Lewin, B. (1991) *Cell*, 64, 303–312.
20. Hess, J.L., Small, J.A. and Clements, J.E. (1989) *J. Virol.*, 63, 3001–3015.
21. Jeang, K.-T., Boros, I., Brady, J., Radonovich, M. and Khoury, G. (1988) *J. Virol.*, 62, 4499–4509.
22. Jeang, K.-T., Chiu, R., Santos, E. and Kim, S.-J. (1991) *Virology*, 181, 218–227.
23. Rethwilm, A., Erlwein, O., Baunach, G., Maurer, B. and ter Meulen, V. (1991) *Proc. Natl. Acad. Sci. USA*, 88, 941–945.
24. Keller, A., Partin, K.M., Lochelt, M., Bannert, H., Flugel, R.M. and Cullen, B.R. (1991) *J. Virol.*, 65, 2589–2594.
25. Cullen, B.R. (1989) *Cell*, 63, 655–657.
26. Dorn, P.L. and Derse, D. (1988) *J. Virol.*, 62, 3522–3526.
27. Sherman, L., Gazit, A., Yaniv, A., Kawakami, T., Dahlberg, J.E. and Tronick, S.R. (1988) *J. Virol.*, 62, 120–126.
28. Rosen, C.A., Sodroski, J.G. and Haseltine, W.A. (1985) *Cell*, 41, 813–823.
29. Berkhout, B., Silverman, R.H. and Jeang, K.-T. (1989) *Cell*, 59, 273–282.
30. Ptashne, M. (1988) *Nature*, 335, 683–689.