The long terminal repeats of the HERV-H human endogenous retrovirus contain binding sites for transcriptional regulation by the Myb protein

Nathalie de Parseval, Hanan Alkabbani and Thierry Heidmann

Unité des Rétrovirus Endogènes et Éléments Rétroïdes des Eucaryotes Supérieurs, CNRS UMR 1573, Institut Gustave Roussy, 39 rue Camille Desmoulins, 94805 Villejuif Cedex, France

Binding sites for transcriptional regulation by the Myb protein have been identified in the long terminal repeats (LTRs) of the H-type human endogenous retrovirus-like elements (HERV-H). Transfection assays using reporter plasmids containing the luciferase gene under the control of a HERV-H LTR disclosed a sevenfold increase in promoter activity in human teratocarcinoma cells when cotransfected with an expression vector for the Myb protein. Binding sites for Myb were unambiguously identified within the LTR by both DNase I footprinting experiments and mobility shift assays using a bacterially expressed purified Myb recombinant protein. Possible roles of these Myb-responsive elements are discussed.

The human genome contains distinct families of endogenous retrovirus-like sequences (HERV, for human endogenous retrovirus), which constitute up to 1% of the genome and are distributed over all chromosomes. Like retroelements in other species, HERVs can transpose and must therefore be considered as potential mutagens (reviewed in Wilkinson et al., 1994; Löwer et al., 1996; Urnovitz & Murphy, 1996). The HERV-H elements are one of the most represented human endogenous retroviruses, with approximately 1000 elements disclosing high sequence identity and 1000 additional solitary long terminal repeats (LTRs) per haploid genome (Mager & Henthorn, 1984). Many of the HERV-H elements are partially deleted, but a small subset (about 10%) are full-length (Hirose et al., 1993; Wilkinson et al., 1993). Expression of HERV-H elements has been described in several tissues and tumour cell lines (Hirose et al., 1993; Wilkinson et al., 1990; Medstrand et al., 1992; Lindeskg et al., 1993; Kelleher et al., 1996). In vivo, HERV-H LTRs can activate the transcription of adjacent cellular genes, since chimeric cDNAs initiating in HERV-H LTRs have been reported (Liu et al., 1991; Feuchter et al., 1992; Feuchter-Murthy et al., 1993; Kowalski et al., 1997; Kowalski & Mager, 1998). Cloned LTRs have been assayed ex vivo in transient transfection experiments and shown to promote transcription of reporter genes in several cell lines (Feuchter & Mager, 1990). Their sequence discloses a consensus TATA box and binding sites for factors involved in transcription initiation. The Sp1 transcription factor was shown to stimulate transcription from a cloned HERV-H LTR (Nelson et al., 1996; Sjottem et al., 1996). Differences have been reported in the organization of HERV-H LTRs. They have been classified into three subtypes, based on variations in repeated patterns found within the U3 region, which contains the sequences necessary for transcription initiation (Mager, 1989; Goodchild et al., 1993).

In a search for sequence elements within the HERV-H LTRs involved in transcriptional regulation, we performed a computer analysis on the sequences of 18 previously characterized HERV-H LTRs (Feuchter & Mager, 1990, 1992; Anderssen et al., 1997). It revealed four DNA sequences clustered in the U3 region 5’ to the TATA box that conform to the general consensus sequence [AAC(T/G)G] for Myb-induced transcriptional activation of a variety of promoters. These sequences are referred to as myb1, myb2, myb3 and myb4 in this report. Sequences myb1 and myb3 are present in the strict consensus form in the majority of the LTRs studied, myb4 in one third of the LTRs, and myb2 only in the H6 LTR. In the H6 LTR, which was previously shown to have the strongest promoter activity in transient assays (Feuchter & Mager, 1990), the Myb consensus binding sequences are located at positions –252, –233, –208 and –164 relative to the proviral RNA start site (Fig. 1). An additional Myb binding site sequence was identified 140 nt downstream of myb4, but was not included in our study as it overlaps with both the TATA box at its 5’ end, and the GC/ GT box (which has been shown to be necessary for promoter activity; Sjottem et al., 1996) at its 3’ end.

To test for the relevance of these consensus sequences to Myb regulatory sites, we used a transient transfection assay with constructions containing the H6 LTR (Mager, 1989) – or part of it – linked to the luciferase reporter gene [(HERV-
Fig. 1. Structure and organization of the HERV-H LTR used in this study and position of the Myb consensus binding sites. The positions of the U3, R and U5 domains of the LTR are indicated above the sequence. The U3 and R domains are from the H6 cDNA LTR and the U5 domain is from the RTVL-H2 genomic clone (see text). The TATA box and the transcription start site are indicated in bold. Consensus Myb binding sites are in tinted boxes and the previously identified Sp1 binding sites are in untinted boxes. Restriction sites used for the construction of the deleted LTRs and for the fragments used in the gel shift experiments are indicated.

Fig. 2. Transactivation by a Myb expression vector of HERV-H LTR reporter genes and control in human NTera2D1 cells. Transactivation by Myb was determined by measuring luciferase activity of the indicated reporter plasmids upon cotransfection of NTera2D1 cells with expression vectors for Myb (tinted bars) or no protein (untinted bars). The reporters used comprise the HERV-H LTR either complete [(HERV-H)LTR–luc] or internally deleted (ΔAccI–NcoI and ΔNcoI–DraIII). The Myb binding site containing HIV-1 LTR [(HIV)LTR–luc] is used as control. Transfections were performed in 30 mm diameter dishes, with 10 ng of reporter plasmid, 100 ng of expression vector and 25 ng of cytomegalovirus promoter–lacZ indicator plasmid, the total amount of transfected DNA being maintained constant (2–5 µg) with PGEM-T (Promega). Luciferase activity was measured using a luminometer and was normalized for transfection efficiency by measuring β-galactosidase activity using a spectrophotometric assay with CPRG (Boehringer) as a β-galactosidase substrate. Data from four independent experiments were pooled. The error bars indicate standard errors of the mean.
expression vector resulted in a sevenfold stimulation of luciferase activity, compared to cotransfections using (HERV-H)LTR–luc together with the same expression vector without the c-myb cDNA insert. To assess the role of the LTR domain containing the consensus Myb sites, deletion mutants of the (HERV-H)LTR–luc plasmid were similarly assayed. Internally deleted HERV-H LTR plasmids were obtained by Acc–NcoI digestion or Nco–DraIII digestion of pH6-LTR, followed by Klenow treatment and direct ligation. Deletion of the 90 nt region encompassing the first three putative Myb sites (ΔAcl–NcoI) almost completely abolished the enhancing effect of Myb. In contrast, deletion of the 90 nt region encompassing the fourth site alone (ΔNcoI–DraIII) had a limited effect, only reducing the stimulation of luciferase activity from seven- to fivefold (Fig. 2). Myb was previously demonstrated to be a transcriptional activator of the human immunodeficiency virus type 1 (HIV-1) LTR promoter in Hela cells (10- to 20-fold transactivation level; Dasgupta et al., 1990). In order to compare the transactivation capacity of Myb on the two LTRs (HERV-H and HIV), we assayed a construct containing the HIV-1 LTR linked to the luciferase reporter gene [(HIV)L-LTR–luc, a gift from H. Jacquemin-Sablon, UPR 9044, Villejuif, France]. As shown in Fig. 2, the stimulation of the HERV-H LTR by Myb was found to be about twice that of the HIV-1 LTR.

The cotransfection experiments above do not indicate whether Myb acts directly or indirectly (i.e. via other intermediates) on the responsive elements identified within the HERV-H LTR. Therefore, we assayed for the presence of Myb binding sites specifically located in the functionally defined regulatory region. In vitro assays (DNase I protection and mobility shift) were carried out using a purified, bacterially produced Myb protein as a specific ligand (a gift from J. Ghysdael). DNase I protection experiments were performed with a 270 bp NolI (in the polylinker cloning site) to DraIII fragment from the HERV-H LTR, encompassing the four previously identified Myb consensus sites. About 3 μg of end-labelled DNA was incubated with increasing concentrations of purified Myb protein, then submitted to DNase I digestion and run on a sequencing gel as described by Cavarec et al. (1997). As illustrated in Fig. 3(a), four regions of the HERV-H LTR are protected from DNase I digestion in the presence of the recombinant Myb protein. Each of these four protected regions contains the expected Myb binding motif. The first three sites should correspond to high affinity sites, since protection takes place at high Myb dilution. The fourth identified Myb consensus site is protected to a lesser extent, even at high Myb concentration.

To further assess that the protected regions are bona fide Myb binding sites and to evaluate their relative affinity for Myb, we performed mobility shift assays with the bacterially produced Myb recombinant protein and sequences containing the Myb responsive elements. Two fragments of the HERV-H LTR were generated. Fragment 1 is a BglII–NcoI fragment spanning the first three consensus sites, and fragment 2 is a Nol–DraIII fragment containing the fourth site (Fig. 1). The agarose gel-purified fragments were tested by mobility shift assays as described by Cavarec et al. (1997). As shown in Fig. 3(b, upper part), fragment 1 exhibits a strong Myb binding activity with a three band pattern, the first retarded band (band ‘a’) appearing in the 40–400 nM range of Myb concentrations. This affinity is close to that observed with the MRE-A oligonucleotide (Fig. 3b, lower part), which corresponds to the high affinity binding site identified in the promoter region of mim-1, a well-characterized v-myb regulated gene (Ness et al., 1989). The appearance and the intensity of the two other bands obtained with fragment 1 (bands ‘b’ and ‘c’) correlate with the increase in Myb concentration. These bands most probably correspond to Myb/ADN complexes formed upon Myb binding to two and then three Myb sites. This interpretation is supported by the fact that the intensity of the first appearing band (band ‘a’) decreases when Myb concentration increases. The pattern observed with fragment 2 is simpler, with a single band appearing in the 40–400 nM range of Myb concentrations, the intensity of which increases in a dose-dependent manner with Myb (Fig. 3b, upper part). Myb binding to fragments 1 and 2 is specific, since binding sites are efficiently competed for by the addition of a 100-fold molar excess of the MRE-A oligonucleotide (Fig. 3b, upper part, last lane). In contrast, a 500-fold molar excess of a mutated, non-binding form of MRE-A (MRE-M) failed to compete, as expected (data not shown).

We have therefore demonstrated that the HERV-H LTRs possess binding sites for the Myb protein. We further show that Myb can activate transcription of at least one LTR of the HERV-H family (the H6 LTR). Interestingly, functional Myb binding sites have also been identified in the U3 region of the HIV-1 and human T-lymphotropic virus 1 (HTLV-1) LTRs (Dasgupta et al., 1990, 1992; Bosselut et al., 1992). The similarities for the transcriptional regulations of an endogenous retrovirus (HERV-H) and exogenous retroviruses (HIV-1 and HTLV-1) most probably reflect a close phylegetic relationship between these two classes of genetic elements. It also suggests that the strategy used by exogenous retroviruses to replicate at a high rate in their host might be used in a similar way by endogenous retroviruses to actively transpone, and in so doing disperse regulatory sequences throughout the genome.

Myb expression is preferentially observed in the haematopoietic system, where it plays a key role in cell proliferation and differentiation. Expression of c-myb is also found in non-haematopoietic tissues as well as in a variety of tumours and cell lines (reviewed in Thompson & Ramsay, 1994). Yet, there is no evidence of a clear-cut correlation between levels of expression of HERV-H and c-myb, except in NTera2D1 cells, which actually disclose high levels of both HERV-H elements (Wilkinson et al., 1990) and c-myb (Miller et al., 1990). This lack of correlation could be accounted for by the fact that activation
of these endogenous elements, which have to be severely repressed in normal tissues to prevent their transposition-mediated mutagenic effects, most likely requires the coordinate action of several transcription factors. Such factors might bind to sites on the HERV-H LTR and/or to each other, and only then act as cooperative transactivators, as commonly observed for genes with complex patterns of expression. The myb3 site and the first Sp1 site, which are close to each other in the HERV-H LTR, could be likely candidates for such cooperative interactions.
We are very grateful to J. Ghysdael for the gift of the pΔEB-c-myb expression vector and the bacterially expressed Myb protein and for suggestions, to D. Mager for the gift of the HERV-H-containing plasmids, and to M. P. Loireau and O. Heidmann for assistance and helpful discussions. This work was supported by the CNRS and by grants from the ARC (contract 6552 to T.H.) and from the Ligue Nationale contre le Cancer (fellowship to N.P.).

References


Received 12 November 1998; Accepted 23 December 1998