Deletion of a Negatively Acting Sequence in a Chimeric GATA-1 Enhancer-Long Terminal Repeat Greatly Increases Retrovirally Mediated Erythroid Expression*

Anna Testa‡§, Francesco Lotti‡, Linda Cairns‡, Alexis Grande, Sergio Ottolenghi‡, Giuliana Ferrari‡, and Antonella Ronchi‡**  

From the ‡Dipartimento di Biotecnologie e Bioscienze, Università di Milano-Bicocca, 20126 Milano, ¶H.S.Raffaele-Telethon Institute for Gene Therapy (HSR-TIGET), 20132 Milano, and the ‡Dipartimento di Scienze Biomediche, Università di Modena e Reggio Emilia, 41100 Modena, Italy  

The locus control region of the β-globin gene cluster has been previously directly to express erythroid expression of globin genes from retroviral vectors for the purpose of gene therapy. Short erythroid regulatory elements represent a potentially valuable alternative to the locus control region. Among them, the GATA-1 enhancer HS2 was used to replace the retroviral enhancer within the 3’-long terminal repeat (LTR) of the retroviral vector SFCM, converting it into an erythroid-specific regulatory element. In this work, we have functionally studied an additional GATA-1 enhancer, HS1. HS1 participates in the transcriptional autoregulation of GATA-1 through an essential GATA-binding site that is footprinted in vitro. In this work we identified within HS1 a new in vivo footprinted region, and we showed that this sequence indeed binds a nuclear protein in vitro. Addition of HS1 to HS2 within the LTR of SFCM significantly improves the expression of a reporter gene. The deletion of the newly identified footprinted sequence in the retroviral construct further increases expression up to a level almost equal to that of the wild type retroviral LTR, without loss of erythroid specificity, suggesting that this sequence may act in a negative regulatory element. An improved vector backbone, MΔN, allows even better expression from the new GATA cassette. These results suggest that substantial improvement of overall expression can be achieved by the combination of multiple changes in both regulatory elements and vectors.

GATA-1 is a zinc finger transcription factor that is expressed in a subset of early multipotent and lineage-committed hematopoietic progenitors, including erythroblasts, megakaryocytes, basophils, etc. (1, 2). GATA-1 regulates many transcription factors and lineage-specific genes, and its normal expression is essential for the correct development of several hematopoietic lineages, in particular the erythroid, megakaryocytic, and eosinophilic lineages (3–7).

The regulation of the expression of GATA-1 itself has been the subject of many investigations (7–16). Constructs including the mouse GATA-1 promoter up to a DNase I-hypersensitive site lying at about −700 nts (HS2) are expressed, at low efficiency, in adult hematopoietic cells in transgenic mice but not in yolk sac cells; however, constructs including an additional more upstream site (HS1) are much more efficient and are also active in primitive hematopoietic cells (12–16). GATA-binding sequences in HS1 and HS2 are essential for activity in a variety of constructs, suggesting GATA-1 auto-regulation (8–10, 14–16); however, a GATA-1 transgenic construct is active in mice lacking the endogenous GATA-1 gene, suggesting that other members of the GATA family of transcription factors (possibly GATA-2) may control GATA-1 transcription (13, 16). As GATA-2 is expressed in a wider range of cell types than GATA-1, this implies that other regulatory elements are necessary for appropriate GATA-1 regulation. Sequences relevant for this additional level of regulation have not yet been fully characterized.

Regulatory elements of erythroid genes have been employed in retroviral vectors to express, in model systems, globin genes for therapeutic purposes. In particular, the locus control region (LCR) of the human β-globin cluster, which confers high level, position-independent, erythroid expression of globin genes, gave encouraging initial results (17, 18).

More recently, short functional erythroid enhancers, such as the α-globin HS40 enhancer, the 5-aminolevulinate synthase intron 8, and the GATA-1 HS2 enhancer have been used instead of the β-globin LCR, in various combinations with different erythroid promoters (β and γ globin, spectrin and ankyrin) to drive gene expression in erythroid cells (19–22). These short elements might thus provide a valuable alternative to LCR for the development of new vectors (23). In this context, the characterization of additional control elements conferring high levels of specific expression is essential for vector optimization. Among the enhancers studied, the GATA-1 HS2 element showed significant activity both in retroviral and lentiviral vectors (19–21).

In this work we have further characterized functional sequences of the GATA-1 HS1 enhancer. By adding HS1 to HS2

---

*This work was supported by a core grant from the Italian Telethon Foundation (to G. F.) and cofin 2002 (to S. O. and F. Mavilio, head of the group at the Universita de Modena e Reggio Emilia). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.  
†Present address: Dipartimento di Scienze Biomediche, Università di Modena e Reggio Emilia, 41100 Modena, Italy.  
**To whom correspondence should be addressed: Dipartimento di Biotecnologie e Bioscienze, Università di Milano-Bicocca, Piazza della Scienza 2, 20126 Milano, Italy. Tel.: 3902-6448-3337; Fax. 3902-6448-3314; E-mail: antonella.ronchi@unimib.it.

1§ The abbreviations used are: nts, nucleotides; LTR, long terminal repeat; LCR, locus control region; DMS, dimethyl sulfate; EMSA, electrophoretic mobility shift assays; BM, bone marrow; Epo, erythropoietin; FCS, fetal calf serum; FACs, fluorescence-activated cell sorting; MEL, mouse erythroblasts; wt, wild type; IVF, in vitro fertilization; GFP, green fluorescent protein; EGFP, enhanced GFP, NGF, nerve growth factor receptor; ts, thermosensitive.
within the retroviralLTR and by deleting an inhibitory sequence within HS1, we obtained retroviruses that express downstream reporter genes at similar efficiencies as those retaining the wild-type LTR, without loss of erythroid specificity.

**EXPERIMENTAL PROCEDURES**

**In Vivo Footprinting—**In vivo DMS treatment of cells, DNA extraction, and piperidine treatment were according to Ref. 24; ligation-mediated PCR was according to Ref. 25. Primers used are as follows: P1 from nt 2 to 25 (GATCCAGAAGAGAAGCAGAGGAGAGCCGGACAGCACCACACCCCCGCCGTCGTAAGG3`; core mut, 5`-GTCAGACGAGGACATTAG); P2 from nt 2 to 25 (GATCCAGAAGAGAAGCAGAGGAGAGCCGGACAGCACCACACCCCCGCCGTCGTAAGG3`; core mut, 5`-GTCAGACGAGGACATTAG). Electrophoretic Mobility Shift Assay (EMSAP)—Nuclear extracts were prepared as in Refs. 26 and 27; in vitro binding and gel electrophoresis were as in Refs. 27 and 28.

The sequences of the oligonucleotides used are as follows: +46/+96 wt, 5`-CTGACAGGAGAAGGATTTAATTTGCTTCAACAAGCTCCGGACACCCCACCCCCGCCTGCAATGGG-3`; 5`-CTGACAGGAGAAGGATTTAATTTGCTTCAACAAGCTCCGGACACCCCACCCCCGCCTGCAATGGG-3`; core mut, 5`-GTCAGACGAGGACATTAG). Electrophoretic Mobility Shift Assay (EMSA)—Nuclear extracts were prepared as in Refs. 26 and 27; in vitro binding and gel electrophoresis were as in Refs. 27 and 28.

The sequences of the oligonucleotides used are as follows: +46/+96 wt, 5`-CTGACAGGAGAAGGATTTAATTTGCTTCAACAAGCTCCGGACACCCCACCCCCGCCTGCAATGGG-3`; core mut, 5`-GTCAGACGAGGACATTAG). Electrophoretic Mobility Shift Assay (EMSA)—Nuclear extracts were prepared as in Refs. 26 and 27; in vitro binding and gel electrophoresis were as in Refs. 27 and 28.

**RESULTS**

An *In Vivo* Footprint in GATA-1 Target DNA Fragments from the GATA-1 region of the mouse GATA-1 gene have been shown previously to have enhancer activity, when linked to the GATA-1 promoter alone or in combination with HS2 and other GATA-1 regulatory elements, in transient transfections, and in transgenic assays (13–16). The activity of HS1 is totally dependent on a strong GATA-1-binding site (14–16) that appears to be occupied in *in vivo*, as indicated by in *in vivo* footprinting (14).

To identify additional functional motifs within HS1, we further extended the dimethyl sulfate (DMS) *in vivo* footprinting analysis of HS1 (Fig. 1A).

**Transcription factors bound to specific DNA sequences protect guanines from in vivo DMS-induced methylation and thus from subsequent *in vitro* piperidine cleavage of the methylated sites. The following ligation-mediated PCR results in the absence of bands at the positions corresponding to the guanines protected in *in vivo* (24, 25).

### Experimental Procedures

Spinoculation was used as above. Human erythroblastic K562 and HEL cell lines (ATCC) were grown in RPMI 1640 medium supplemented with 10% FCS (Hyclone) and infected for 16 h with undiluted viral supernatants containing 8 μg/ml Polybrene (Sigma). Further analyses were carried out on transduced and control bulk cultures. For transduction of human cells viral stocks were produced by transient transfection of Phoenix-Ampho cells as described previously (32).

Human erythroblastic K562 and HEL cell lines (ATCC) were grown in RPMI 1640 (Invitrogen) supplemented with 10% FCS and transduced by spinoculation (33) in the presence of Polybrene (8 μg/ml). For transduction of mouse hematopoietic cells with LGSΔN and MγN viral stocks, spinoculation was used as above.

**Pharyngeal—**Expression of the reporter gene (ΔLNGFr) was monitored by flow cytomtery (FACSscan, BD Biosciences) using the murine anti-human p75-NGFr monoclonal antibody 20-4 (ATCC). A goat monoclonal anti-mouse IgG (Fab-specific) conjugated to fluorescein isothiocyanate or phycoerythrin was used as secondary antibody (Pharmingen).

### RESULTS

An *In Vivo* Footprint in GATA-1 Target DNA Fragments from the GATA-1 region of the mouse GATA-1 gene have been shown previously to have enhancer activity, when linked to the GATA-1 promoter alone or in combination with HS2 and other GATA-1 regulatory elements, in transient transfections, and in transgenic assays (13–16). The activity of HS1 is totally dependent on a strong GATA-1-binding site (14–16) that appears to be occupied in *in vivo*, as indicated by in *in vivo* footprinting (14).

To identify additional functional motifs within HS1, we further extended the dimethyl sulfate (DMS) *in vivo* footprinting analysis of HS1 (Fig. 1A).

**Transcription factors bound to specific DNA sequences protect guanines from in vivo DMS-induced methylation and thus from subsequent *in vitro* piperidine cleavage of the methylated sites. The following ligation-mediated PCR results in the absence of bands at the positions corresponding to the guanines protected in *in vivo* (24, 25).

Fig. 1B shows six guanines (lower strand) within an 11-nt sequence between positions 59 and 69 of HS1 are footprinted *in vivo* in hematopoietic mouse erythroleukemia cells (HMEC) grown in culture; a weaker footprint is also detected on the CACC motif between positions 78 and 82 (not shown).

**In Vitro Binding Studies of the Footprint Region**—To better characterize proteins responsible for the in *vivo* footprint, we performed EMSA. A 5`-P-labeled oligonucleotide (wt, see “Experimental Procedures” spanning from nt 46 to 96 generates a complex pattern consisting of an intense slow band and additional faster bands in the presence of nuclear extracts from the hematopoietic cell lines: human K562 erythroleukemia cells, murine GATA-1 to Epo BM, and GATA-1 to Epo YS cells (30, Fig. 2); however, when an oligonucleotide deleted between nucleotides 59 and 69 (Δ59–69) is used, one of these bands is clearly missing (Fig. 2B, arrow, compare lanes 1 and 2). The protein responsible for this band will be called from now on the GATA-1 HS1 in *in vivo* footprinting protein (IVFP). As expected, unlabeled oligonucleotides including a CACC box sequence from the human B-globin gene and the GC-rich region (positions 81–111) from the HS1 itself compete all bands, with the exception of the IVFP band (Fig. 2B, lanes 3 and 4 and lanes 7 and 8); an unrelated control oligonucleotide has no effect (lanes 5 and 6).

We further tested two additional oligonucleotides (Fig. 2): “core mutant” carries mutations centered on nucleotides involved in the *in vivo* footprint, whereas the “5’ mutant” is
mutated at the upstream border of the footprinted region (see "Experimental Procedures" for sequences). The mutation in oligonucleotide core mutant (Fig. 2B, lane 4) has the same effect as the deletion of nucleotides 59–69, i.e., loss of IVFP binding, whereas the upstream mutation in oligonucleotide 5’ mutant does not significantly affect IVFP binding (Fig. 2B, lane 3).

To better resolve the IVFP band, we used in EMSA experiments poly(dC-dG) as a competitor, instead of poly(dI-dC) (Fig. 2B, lanes 5–18). Under these conditions, only the IVFP band is visible (lane 5) with the wt oligonucleotide and is efficiently competed by the unlabeled wt oligonucleotide itself (Fig. 2B, lanes 6–8) and by the 5’ mutant oligonucleotide (lanes 12–14).
but not by the deleted (lanes 9–11) or core mutant (lanes 15–17) oligonucleotides.

The IVFP protein appears to be relatively widespread, although its amount varies between different cell types. In particular, it is expressed in both mouse and human cells, particularly in human K562 and mouse erythroleukemia MEL cells, in SV40-T immortalized bone marrow (GATA-1 ts Epo, BM) and yolk sac (GATA-1 ts Epo, YS) erythropoietin-dependent cells (11, 30),2 and in lymphoid cells (CH27, A20, and BAF3) (Fig. 2 C and data not shown). By using mutant oligonucleotides with these extracts, we obtained results similar to those shown for K562 (data not shown).

HS1 DNA Sequences Co-operate with HS2 in Retrovirally Mediated Gene Expression in Hematopoietic Cells—To analyze functional effects of the HS1 DNA sequences, we used a modified SFCM retroviral vector, as described previously (19). In this vector, the Moloney-leukemia virus enhancer in the 3′-LTR is deleted (ΔSFCM) and can be replaced by an exogenous enhancer; following retroviral infection of the target cells, the new enhancer is moved to the 5′-LTR from which it drives retroviral transcription and expression of a truncated form of nerve growth factor receptor (ΔLNGFr), a reporter gene. In addition, an internal SV40 promoter drives the expression of a neomycin resistance gene (Fig. 3A). Our strategy was therefore to infect in vitro hematopoietic cell lines with appropriate constructs and to select cells expressing the neo-resistance gene by G418 treatment. Following complete selection, the proportion of expressing cells and levels of expression were analyzed by fluorescence-activated cell sorting (FACS).

The copy number of integrated constructs was examined by Southern blotting using a neo-resistance probe (for the detection of the retroviral DNA) and a GATA-1 3′ genomic probe for detection of an endogenous gene and normalization. The copy numbers were very similar between the various constructs and close to 1 copy per cell, as expected for the relatively low probability of infection in these experiments (between 25 and 50% of cells transduced).

For these experiments, we primarily used two cell lines, GATA-1 ts Epo (BM) and GATA-1 ts Epo (YS) (11, 30),2 which represent erythropoietin-dependent multipotent cells derived...
SV40 T-antigen (driven by a transgenic GATA-1 promoter linked to HS2 or HS1 respectively). These cells are 95% positive for endogenous GATA-1 (based on nuclear immunofluorescence tests) (30) and are thus expected to express foreign constructs depending on GATA regulatory elements.

In a first series of experiments (Fig. 3B), we compared the activities of retroviral constructs (HS1-SFCM and HS2-SFCM) containing HS1 or HS2 as the foreign enhancer replacing the original Moloney enhancer. The mean fluorescence intensity (MFI) of cells infected with the wild type SFCM is given as 100% activity. The MFI of each construct is given below as percentage of activity relative to that of the wild type SFCM. The activity of HS1-SFCM (2 ± 1% in both BM and YS cells) and mutated derivatives of it (data not shown) did not differ from that of the inactive control HS2 construct (2 ± 1% of SFCM). On the other hand, HS2-SFCM was clearly active in both BM and YS cell lines (as expected, see Ref. 19) at moderate levels (10 ± 2 and 11 ± 2% of SFCM, respectively).

We then linked HS1 upstream to HS2 (HS1-HS2-SFCM); this resulted in a 2-fold increase (p < 0.05 by Student’s t test) of the activity as compared with HS2-SFCM (19 ± 3% of SFCM in BM and 21 ± 2% in YS), indicating that HS1 sequences synergize with HS2. Interestingly the same result was obtained when HS1 carrying a deletion of its 3’ sequences (nt 1–171, data not shown) or a point mutation in the GATA-1 motif (HS1GATA -HS2-SFCM, Fig. 3B) was used instead of the intact HS1. Similar results were obtained with GATA-1 ts Epo (BM) and YS) cells (Fig. 3B).

Deletion of the IVFP-binding Site Increases the Activity of the HS1-HS2-SFCM Construct—We tested the effects of the deletion of the in vivo footprinted elements (Fig. 3C). Although a vector carrying HS1 deleted at nucleotides 59–69 (∆HS1-SFCM) is completely inactive on its own (data not shown), the addition of the HS1Δ59–69 element to HS2 (ΔHS1-HS2-SFCM vector) caused a strong increase of reporter gene expression in both bone marrow and yolk sac cell lines. Most interesting, in the latter, the effect on activity of the addition of HS1 was more pronounced (56 ± 12% versus 100 ± 20%, respectively).

To ascertain whether the loss of the IVFP-binding site could contribute to this increase, we inserted into HS1 the same mutations as those tested in EMSA experiments. Introducing the core mutation in HS1-HS2-SFCM, increases the average level of activity of the construct (HS1core mutant-HS2-SFCM, 37 ± 14% of SFCM in BM and 40 ± 21% in YS), although not as much as the deletion. Note that in the core mutant-infected cells, many (80%) of the cells express ΔLNGFr at high levels, as in ΔHS1-HS2-SFCM, but a significant shoulder of non-expressing cells is present. In contrast, the vector (HS1 5’ mutant-HS2-SFCM) carrying the 5’ mutation that slightly affects IVFP binding (Fig. 2B) is active in a much smaller proportion of cells, although some still express at high levels.

Exogenously inserted genes are often silenced during long term cell propagation in culture, as a repressive chromatin structure adjacent to the integration site might extend into the foreign gene (34). When this event occurs, both the reporter gene and the neomycin-resistance gene are likely repressed (34). Under these conditions, the presence of G418 in the medium will cause loss of these cells, leading to an overestimate of the proportion of reporter-expressing cells. We thus tested whether releasing the cells from G418 selection affects reporter gene expression (Fig. 4). Cells grown in G418 for at least 3 weeks to kill all non-virus-infected cells were further propagated in the absence of G418 and tested for reporter gene expression. With all tested constructs, including the wild type Moloney enhancer-dependent SFCM construct, there was some degree of progressive inactivation, as expected (34–36); how-

![Figure 4](http://www.jbc.org/)
ever, after 25 days in culture without G418 (~30 cell cycles), the reporter gene was still active at high levels in 75–90% of cells with all constructs, except for HS15\(^{\prime}\) mutant-HS2-SFCM which remains active in only 25% of the cells (Fig. 4). This demonstrates that SFCM or HS1-HS2-SFCM, \(\Delta\)HS1-HS2-SFCM, and (to a lesser extent) HS1core mutant-HS2-SFCM-modified LTR are relatively well protected from position of integration effects. The mutation of the 5\(^{\prime}\) sequences within the HS1-HS2 context, however, clearly interferes with the maintenance of the activity.

**HS1 Increases Gene Expression in Erythroid Cells but Not in Fibroblasts**—HS2-SFCM has been shown previously to drive reporter gene expression in hematopoietic but not non-hematopoietic cells (19). To ascertain if the \(\Delta\)HS1:HS2 cassette retains hematopoietic specificity, we constructed an additional vector (\(\Delta\)HS1:HS2-LGS\(\Delta\)N) in which the modified LTR drives the activity of GFP, whereas a constitutive internal promoter expresses the \(\Delta\)LNGFr gene, acting as an internal reference for the level of ubiquitous expression in uninduced cells (Fig. 5A). This experimental arrangement allows us to measure LTR-dependent gene expression in the absence of any previous selection. We thus compared the transgene expression level in cells transduced with the HS2-LGS\(\Delta\)N vector and the \(\Delta\)HS1:HS2-LGS\(\Delta\)N vector (Fig. 5B). Mouse NIH3T3 fibroblasts, mouse hematopoietic GATA-1 ts Epo BM and GATA-1 ts Epo YS cells, and human

---

**Fig. 5. Erythroid restriction of GATA-1-modified LTR-dependent reporter EGFP gene expression.** A, the GATA-1-modified LGS\(\Delta\)N vector. SV\(\mu\) indicates SV40 promoter. B, FACS analysis of transduced hematopoietic mouse GATA-1 ts Epo BM, GATA-1 ts Epo YS cells, and mouse fibroblast 3T3 cells.
erythroid K562 cells were transduced with HS2-LGS\textsuperscript{N} and \(\Delta\text{HS1-HS2-LGS\textsuperscript{N}}\) viral supernatants, maintained in culture as bulk population, and assayed for \(\Delta\text{NGFr}\) and GFP expression by FACS analysis (Fig. 5B). Most (over 80\%) of the NIH3T3 cells transduced with either HS2-LGS\textsuperscript{N} or \(\Delta\text{HS1-HS2-LGS\textsuperscript{N}}\) expressed the \(\Delta\text{NGFr}\) gene, at similar levels (Fig. 5B). In contrast, very few cells also expressed GFP, at low levels, indicating that the GATA-1 cassettes are essentially inactive in NIH3T3 cells. Hematopoietic GATA-1-ts Epo BM and GATA-1-ts Epo YS cells were transduced with the same viral supernatants as above, to allow comparison. In contrast to NIH3T3 cells, most of the hematopoietic NGFr-positive cells were also strongly GFP-positive; the fact that some cells with very low levels of NGFr expression do express some GFP suggests that the erythroid cassette may be more easily activated within hematopoietic cells than the SV40 promoter driving NGFr. As expected, the proportion of GFP-expressing cells was much higher in \(\Delta\text{HS1-HS2-LGS\textsuperscript{N}}\) than in HS2-LGS\textsuperscript{N} transduced cells.

In similar experiments, in human erythroleukemic K562 cells transduced with \(\Delta\text{HS1-HS2-LGS\textsuperscript{N}}\), the GFP was expressed at MFI (2.5-fold higher than that in cells transduced with HS2-LGS\textsuperscript{N} (not shown)). These results confirm that erythroid restriction of transgene expression by the HS2 element of the GATA-1 gene is maintained also when this element is used in combination with the \(\Delta\text{HS1}\) element of the GATA-1 gene.

An Improved Vector Backbone Further Increases Gene Expression Activity—Another approach to increase the overall efficiency of transgene expression was to improve the vector characteristics at the post-transcriptional level. We compared the \(\Delta\text{SFCM}\) backbone (originally derived from the LXS\textsuperscript{N} vector) with that of the \(\text{M}\text{AN}\) vector (originally derived from MFG (37)), in which the primary genomic transcript is spliced from a donor and \(\text{gag}\) cassette is translated under the control of the highly efficient \(\text{env}\) promoter driving NGFr. As expected, the proportion of GFP-expressing cells was much higher in \(\Delta\text{HS1-HS2-LGS\textsuperscript{N}}\) than in HS2-LGS\textsuperscript{N} transduced cells.

In experiments on human erythroleukemic K562 cells transduced with \(\Delta\text{HS1-HS2-LGS\textsuperscript{N}}\), the GFP was expressed at MFI (2.5-fold higher than that in cells transduced with HS2-LGS\textsuperscript{N} (not shown)). These results confirm that erythroid restriction of transgene expression by the HS2 element of the GATA-1 gene is maintained also when this element is used in combination with the \(\Delta\text{HS1}\) element of the GATA-1 gene.

An Improved Vector Backbone Further Increases Gene Expression Activity—Another approach to increase the overall efficiency of transgene expression was to improve the vector characteristics at the post-transcriptional level. We compared the \(\Delta\text{SFCM}\) backbone (originally derived from the LXS\textsuperscript{N} vector) with that of the \(\text{M}\text{AN}\) vector (originally derived from MFG (37)), in which the primary genomic transcript is spliced from a donor and \(\text{gag}\) cassette is translated under the control of the highly efficient \(\text{env}\) promoter driving NGFr. As expected, the proportion of GFP-expressing cells was much higher in \(\Delta\text{HS1-HS2-LGS\textsuperscript{N}}\) than in HS2-LGS\textsuperscript{N} transduced cells.

In similar experiments, in human erythroleukemic K562 cells transduced with \(\Delta\text{HS1-HS2-LGS\textsuperscript{N}}\), the GFP was expressed at MFI (2.5-fold higher than that in cells transduced with HS2-LGS\textsuperscript{N} (not shown)). These results confirm that erythroid restriction of transgene expression by the HS2 element of the GATA-1 gene is maintained also when this element is used in combination with the \(\Delta\text{HS1}\) element of the GATA-1 gene. Negatively Acting Elements—In vivo footprinting studies have so far identified within HS1 only two sites that are footprinted in vivo, thus directly implying a significant functional role: the GATA-binding site (14) and the upstream 5' donor promoter region. Mouse GATA-1-ts Epo BM cells and human erythroid HEL cells were transduced, using the same conditions of infection, with \(\Delta\text{HS1-HS2-SFCM}\) and with \(\Delta\text{HS1-HS2-MAN}\) viral supernatants and grown in culture as bulk population. Transduction efficiency with the \(\Delta\text{HS1-HS2-SFCM}\) and \(\Delta\text{HS1-HS2-MAN}\) vectors was above 80\% in all cell lines. \(\Delta\text{NGFr}\) expression, in terms of MFI, was 2–5-fold higher in cells transduced with \(\Delta\text{HS1-HS2-MAN}\) vector than in those transduced with the HS2-SFCM vector and approached that of the wild type \(\text{MAN}\) vector. The increased expression observed with \(\text{MAN}\) vector was at the level of RNA accumulation (data not shown).

**DISCUSSION**

**DNA Sequences Regulating GATA-1 Expression**—The DNA region surrounding the HS1 site of the mouse GATA-1 gene appears to play a central role in GATA-1 expression in transgenic experiments, as it is essential for high level activity in hematopoietic cells in general, and in particular for expression in embryonic erythroblasts (12–16). It is likely that a certain degree of functional redundancy exists in the GATA-1 gene region; in fact, the deletion from the endogenous GATA-1 gene of a region comprising HS1, by homologous recombination, has little effect on GATA-1 expression in erythroid cells (7). Nevertheless, characterizing functional elements of the GATA-1 HS1 region may lead to the discovery of important regulatory factors for GATA-1 expression. In this regard, it is relevant that the single HS2 site, in the absence of any additional GATA-1 regulatory elements, was able to confer proper expression of a reporter gene within the SFCM retrovirus in stem cell-derived early hematopoietic progenitors and erythroblasts (19).

Cooperation between HS1 and HS2—Our results show that HS1, unexpectedly, is unable to direct hematopoietic cell expression of a reporter gene when inserted alone into the SFCM LTR. This is in contrast with evidence obtained in transient and stable transfection experiments by ourselves and others (14–16). However, the same element significantly stimulates the activity of a linked HS2 sequence. The HS1 element contains a GATA-binding motif that might bind either GATA-1 or GATA-2 and that might play a role in an autoregulatory loop and/or in the initial activation of the GATA-1 gene itself (14–16). In fact, the HS1 GATA-binding sequence was found to be essential in transient transfection assays of HS1 alone (14) as well as in transgenic constructs carrying HS2 and other regulatory elements (13, 15, 16) within wild type or even GATA-1 mutant mice. In our experiments, HS1 synergizes with HS2 (Fig. 5) even when the HS1 GATA-binding site is mutated; this suggests that sequences other than the sole GATA-binding site are important within HS1 for its functional cooperation with HS2. The discrepancy between the previously described requirement for the GATA-binding site and our present data might be explained by the different arrangement of HS1 and HS2 in the transgenic versus the retroviral constructs; in the latter one the HS1 and HS2 sites are immediately adjacent, and the strong GATA-1-binding element of HS2 might be sufficient for activity, whereas in the transgenic constructs HS1 and HS2 are separated by almost 3 kb and may thus act independently.

Negatively Acting Elements—In vivo footprinting studies have so far identified within HS1 only two sites that are footprinted in vivo, thus directly implying a significant functional role: the GATA-binding site (14) and the upstream 5'–69 footprint (see above). The latter footprint is immediately adjacent to a GC-rich region that binds SP1 in vitro and the erythroid EKLF factor weakly (data not shown) and yields a weaker footprint. Binding sites for other (uncharacterized) factors also lie upstream to the footprint. When HS1 deleted in the footprinted region is added to HS2, the resulting construct \(\Delta\text{HS1-HS2-SFCM}\) is more active than the undeleted HS1-HS2-SFCM.

It is interesting that \(\Delta\text{HS1-HS2}\) expression is stronger in the yolk sac-derived cell line than in the bone marrow line immortalized by a T gene dependent on HS2 activity. This is in keeping with the notion that HS1 is strictly required for GATA-1 expression in yolk sac cells (12, 13, 15, 16).

The functional effect of the deletion of the 59–69 region (Fig. 3) can be reproduced, at least in part, by a mutation (core mutant) that destroys the binding of IVFP (Fig. 2B). Thus a specific ubiquitous protein, IVPP, might negatively regulate HS1 activity (or its ability to cooperate with HS2) in our experimental system. However, it is possible that the stronger effect of the deletion (versus the core mutation) might also be due to a perturbation of the architecture of the region introduced by the deletion itself. The deletion might favor the functional cooperation of factors binding to the sequences flanking the footprinted region by joining them into a single DNA stretch (Fig. 1A). It is of further interest that the HS1 core mutation, while improving reporter expression in a large subset of cells, also significantly increases the number of cells that fail to express (Fig. 3B).
Similarly, the HS1 5’ mutation results in a bimodal distribution of reporter activities, with a large proportion of non-expressing cells (Fig. 3B). Intriguingly, HS2-SFCM, although active at relatively low levels, is expressed in the majority of the cells (Fig. 3A). These results indicate that sequences within HS1, when mutated, might interfere with the activity of the adjacent HS2 element. It is possible that the IVFP binding region and flanking sequences include elements affecting “chromatin opening” or boundaries shielding HS1 from the effects of neighboring sequences. Alternatively, the mutations introduced (but not the deletion) might have inadvertently generated negatively acting elements. Future methylation studies of integrated sequences carrying normal and mutated IVFP binding region might help to solve this problem.

We wish to stress that because of the differences between the relative arrangement of HS1 and HS2 sequences within the retroviral constructs and their in vivo location, we cannot infer from these data the in vivo role of the IVFP-binding sequence in GATA-1 regulation; however, the in vivo footprint points to a significant functional role within hematoepoietic cells. In future work we will try to identify additional binding sites for IVFP within the mouse GATA-1 gene, and possibly other erythroid genes, and to assess the functional role of the footprinted region by transgenic experiments. From the evolutionary point of view, the IVFP-binding site is very weak or lost in the human globin gene LCR-derived cassettes, recently achieving promising results (17, 18). The large size of the fragments necessary for optimal LCR activity (i.e. expression level and full insulation from position effects) has been an obstacle to the formation of high titer, high efficiency vectors which do not undergo structural rearrangements. As an alternative approach, other erythroid regulatory elements have been used instead of the LCR (19–23), including the GATA-1 HS2 enhancer. Its relatively small size allowed us to insert it directly into the LTR, replacing the viral enhancer. This element converts the ubiquitously active LTR into an erythroid enhancer-promoter, but the activity obtained with this vector is significantly lower than that of the wild type LTR (19, 21).

The present results show that simply adding a short sequence (HS1) to the GATA-1 HS2 and appropriately mutating a negatively acting element (Fig. 5) creates a much better cassette for hematopoietic expression that retains the hematopoietic specificity already shown for the HS2 cassette (19). The average activity conferred by the ΔHS1-HS2 element approaches that of the wild type LTR in both the SFCM (Fig. 3, B and C) and ΔMNa vectors (Fig. 6). The peak obtained with the ΔHS1-HS2 element is broader than that of SFCM, with a significant proportion of cells even exceeding the maximum expression level of the Moloney enhancer-based vector SFCM. The greater variance of the expression of the ΔHS1-HS2-SFCM vector versus the wild type SFCM vector is evidence that the former construct may not be fully insulated from position of integration effects. Nevertheless, the large majority of G418-resistant cells express the reporter gene, and the proportion of expressing cells remains fairly constant for tens of division cycles even in the absence of the selective agent. It will be of interest to test if these vectors can equally and efficiently drive globin gene expression in erythroid cells.

**Acknowledgment**—We thank Dr. G. Lamorte for help with FACS data analysis.

**REFERENCES**


---

**Fig. 6. Improved ΔLNGFr expression in the ΔMNa-based vector.** A, comparison of the SFCM and ΔMNa backbones. B, FACS analysis of transduced mouse GATA-1 to Epo BM and human erythroid HEL cells with different vectors. Untransduced BM and HEL cells are indicated by “c” in B and C, respectively.
Erythroid Expression from a GATA-1 Enhancer-dependent LTR

Deletion of a Negatively Acting Sequence in a Chimeric GATA-1 Enhancer-Long Terminal Repeat Greatly Increases Retrovirally Mediated Erythroid Expression
Anna Testa, Francesco Lotti, Linda Cairns, Alexis Grande, Sergio Ottolenghi, Giuliana Ferrari and Antonella Ronchi

doi: 10.1074/jbc.M313638200 originally published online December 29, 2003

Access the most updated version of this article at doi: 10.1074/jbc.M313638200

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 38 references, 20 of which can be accessed free at http://www.jbc.org/content/279/11/10523.full.html#ref-list-1