1-2002

Transcriptional Interference by Independently Regulated Genes Occurs in Any Relative Arrangement of the Genes and Is Influenced by Chromosomal Integration Position

Susan K. Eszterhas
Dartmouth College

Eric E. Bouhassira
Albert Einstein College of Medicine

David I. K. Martin
Victor Chang Cardiac Research Institute

Steven Fiering
Dartmouth College

MOLECULAR AND CELLULAR BIOLOGY, Jan. 2002, p. 469–479 Vol. 22, No. 2
0270-7306/02/$04.00+0 DOI: 10.1128/MCB.22.2.469–479.2002

Transcriptional Interference by Independently Regulated Genes Occurs in Any Relative Arrangement of the Genes and Is Influenced by Chromosomal Integration Position

Susan K. Eszterhas,[1] Eric E. Bouhassira,[2] David I. K. Martin,[3] and Steven Fiering[1]*

Department of Microbiology and Immunology, Dartmouth Medical School, Lebanon, New Hampshire 03756[1]; Division of Hematology, Department of Medicine, Albert Einstein College of Medicine, Bronx, New York 10461[2]; and Victor Chang Cardiac Research Institute, Sydney, New South Wales, Australia[3]

Received 11 June 2001/Returned for modification 20 July 2001/Accepted 15 October 2001

Transcriptional interference is the influence, generally suppressive, of one active transcriptional unit on another unit linked in *cis*. Its wide occurrence in experimental systems suggests that it may also influence transcription in many loci, but little is known about its precise nature or underlying mechanisms. Here we report a study of the interaction of two nearly identical transcription units juxtaposed in various arrangements. Each reporter gene in the constructs has its own promoter and enhancer and a strong polyadenylation signal. We used recombinase-mediated cassette exchange (RMCE) to insert the constructs into previously tagged genomic sites in cultured cells. This strategy also allows the constructs to be assessed in both orientations with respect to flanking chromatin. In each of the possible arrangements (tandem, divergent, and convergent), the presence of two genes strongly suppresses expression of both genes compared to that of an identical single gene at the same integration site. The suppression is most severe with the convergent arrangement and least severe in total with the divergent arrangement, while the tandem arrangement is most strongly influenced by the integration site and the genes' orientation within the site. These results suggest that transcriptional interference could underlie some position effects and contribute to the regulation of genes in complex loci.

Within higher eukaryotic genomes, independently regulated transcription units are often apposed and intermingled. The means by which transcriptional activity is partitioned to achieve precise control of an individual gene are poorly understood, but it is possible that genes are adapted to function while their expression is influenced by nearby genes. Clues to the existence of such mechanisms are provided by the numerous examples in which gene regulation is disrupted by changes in the arrangement of genomic elements; these examples suggest that an active promoter may influence other genes over large distances.

Naturally occurring chromosomal rearrangements and experimental manipulations have provided numerous examples in which the introduction of an active transcription unit into a locus changes the transcriptional activity of a gene or genes native to the locus. The earliest examples, described by McClintock and others, were numerous maize strains in which transposable elements altered normal patterns of gene expression, sometimes suppressing expression, sometimes altering its level or pattern (17, 25). These effects were frequently dependent not on the simple presence of the transposon, but rather on its transcriptional activity. Other examples of this general phenomenon, which is termed transcriptional interference (TI), have been described in a variety of experimental systems. Studies in viruses, bacteria, plants, flies, and mammals show that transcription of a gene can have a strong influence on the level of transcription of another gene linked in *cis* (1, 5–7, 9–11, 16, 18, 19, 28–31).

TI may be defined broadly as the perturbation of one transcription unit by another. Although it has often been encountered experimentally, the mechanisms have been difficult to characterize. Mechanisms likely include the following: (i) promoter occlusion, in which transcription from one gene through a regulatory region of another gene disrupts DNA binding factors; (ii) steric or topological changes induced by transcription, such as changes in supercoiling; (iii) competition for *cis*- or *trans*-acting factors, including enhancers or their binding proteins, or general transcription apparatus; (iv) production of antisense RNA associated with inhibition of translation (or RNAi); (iv) epigenetic mechanisms in which aspects of chromatin structure, like methylation or histone modifications, spread from one gene to another; and (vi) direct transcription of one gene by the promoter of another, changing the pattern of expression.

It is clear from a variety of studies that TI will occur when two transcription units overlap (i.e. when they share some portion of the mature mRNA in common, either sense or antisense). Most experimental demonstrations of TI in mammals have focused on situations of transcriptional overlap in which two genes or two promoters for one gene are in tandem and the upstream gene or promoter is not separated from the downstream gene or promoter by a polyadenylation site (7, 9, 10, 18, 28, 29, 31). This situation, however, rarely occurs naturally. There are situations in mammals in which TI has been demonstrated without overlapping transcription, such as the β-globin (15, 16, 20), α-globin (11, 22), and granzyme (26) loci. Each of these loci contains multiple nonoverlapping genes, the expression of which is influenced by shared regulatory ele-

* Corresponding author. Mailing address: Dartmouth Medical School, 6 West Borwell, Dartmouth Hitchcock Medical Center, Lebanon, NH 03756. Phone: (603) 650-6601. Fax: (603) 650-6223. E-mail: fiering@dartmouth.edu.

ments, and in this situation, expression of one gene influences the expression of other genes at the locus.

Missing from our understanding of TI is the influence of one independently regulated transcription unit on another, when there is no obvious transcription overlap or sharing of *cis* regulatory elements. The observation that the expression of a transgenic selectable marker used to make a gene knockout can influence nearby genes (reviewed in reference 14) is an example of TI occurring between independent transcription units. In another such example, in vivo gene targeting studies of N-Ras and its upstream tandem partner, UNR, show that abolition of expression of UNR increases expression of N-Ras (3) in its endogenous location. These studies suggest that each of two independent genes could influence expression of the other. We have developed an experimental system to explore this possibility further.

Perhaps because of the broad and varying nature of TI, it is poorly understood. A better understanding of TI might be gained from a system in which it could be studied in detail. We have constructed a simple system in which two transcription units can be juxtaposed in a variety of ways, and their expression can be assayed simultaneously. Position effects are controlled by use of recombinase-mediated cassette exchange (RMCE) to place different constructs into the same genomic site. This method also allows us to assess the effect of the construct's orientation with respect to the integration site. Since the genomic location and orientation at that location are controlled, the system permits comparison to a single gene in the same location, which provides novel insights. We find that regardless of arrangement, TI strongly suppresses expression of each gene in a linked pair and that the orientation of the units with respect to each other and to flanking chromatin significantly affects their transcription.

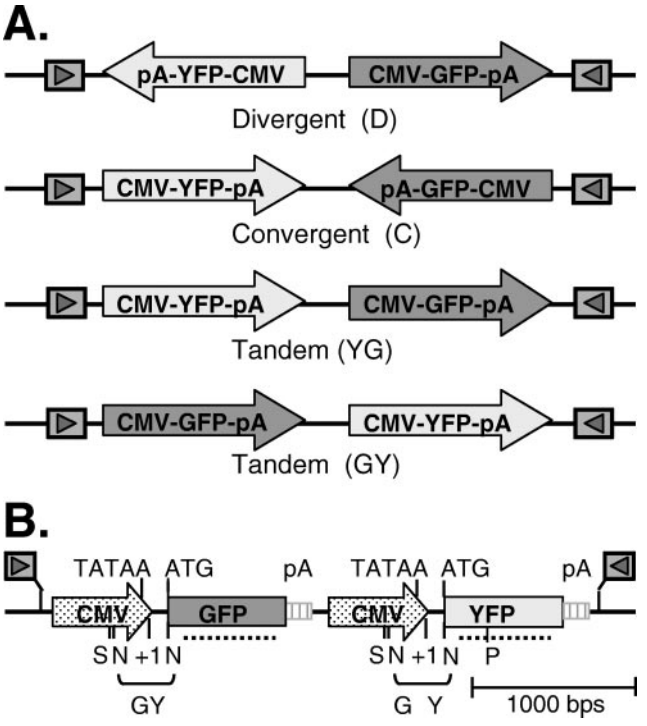

FIG. 1. Diagrammatic representations of the constructs developed for this study. (A) The arrangement of the two reporter transcription units is depicted by arrows, which indicate the direction of transcription. Each transcription unit contains the CMV promoter or enhancer driving GFP or YFP and the SV40 large T-antigen polyadenylation signal (pA). The dark gray arrow is the GFP-containing transcription unit, and the light arrow is the YFP-containing transcription unit; they are shown in the four possible orientations: divergent (D), convergent (C), and tandem (YG and GY). Thin lines represent vector sequences derived from pGEM. The two dark squares bearing triangles denote the two inverted 34-bp *loxP* sites, which are recognized by the CRE recombinase. (B) Tandem GY construct. The construct has been drawn to scale with the CMV promoters (stippled arrows), TATA boxes (TATAA), the starts of transcription (+1), translation initiation sites (ATG) of the fluorescent proteins (gene bodies are dark and light rectangles, GFP and YFP, respectively), and polyadenylation signals (pA; striped blocks) indicated. S, *Sna*B1; N, *Nco*I; and P, *Pst*I. The brackets delimit the basal promoter deletions for the two constructs, ΔGY and GΔY, and the dotted lines show the regions of hybridization to the radiolabeled probe used for Southern analyses.

MATERIALS AND METHODS

Plasmid constructs. The plasmid constructs were produced by standard techniques. The 740-bp *Bam*HI-*Not*I fragment, carrying the reporter gene, was isolated from pEGFP-N1 or pEYFP-N1 (Clontech, Palo Alto, Calif.). The *Bam*HI-*Not*I fragments were placed between the *Bgl*II sites of pL1EGFP1L (12). These constructs retained the cytomegalovirus (CMV) promoter, the SV40 large T antigen polyadenylation signal, as well as the two *loxP* sites. *Bam*HI was used to invert the 1,668-bp CMV-yellow fluorescent protein (YFP) fragment relative to the flanking restriction sites. For ease of cloning, fragments from these plasmids (a 1,735-bp *Cla*I-*Sph*I fragment or a 1,831-bp *Eco*RI-*Sph*I fragment) were moved into pGEM3Zf(+) (Promega); these plasmids were sequenced. The final two-gene constructs were assembled by isolating 1,737-bp *Eco*RI-*Nae*I fragments from the sequenced plasmids and inserting them into *Eco*RI-*Sma*I-digested complementary plasmids. This produced two-gene plasmid constructs in which the transcription units for green fluorescent protein (GFP) and YFP were divergent, convergent, or in tandem. The two-gene plasmids were checked by restriction analyses and limited sequencing. Deletion constructs were prepared with partial *Nco*I digests of the tandem GY construct. The deletion plasmids lack the basal promoter (including the TATAA box) of either the upstream promoter or the downstream promoter, but reconstitute the translational start site ATG (Fig. 1B). These plasmids were also checked by restriction analyses and sequencing.

Cell culture. Mouse erythroleukemia (MEL) cells and the derivatives RL5 and RL6 (12) were maintained under subconfluent conditions in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% calf serum, 50 U of penicillin per ml, 50 μg of streptomycin per ml, and 2 mM glutamine at 37°C in 5% CO_2.

Transient transfections. A total of 5×10^6 MEL cells were electroporated with 25 μg of plasmid DNA at 260 V, 1.0 mF. The cells were incubated for 48 h in 5 ml of tissue culture medium, centrifuged at $100 \times g$ for 3 min, and washed with phosphate-buffered saline, and 20,000 cells were analyzed by fluorescence-activated cell sorting (FACS). Transfection efficiency was controlled by cotransfecting a β-galactosidase expression construct and assaying the percentage of transfected cells by X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) staining.

RMCE. RL5 and RL6 (12) are MEL-derived cell lines carrying an exchange site consisting of a pair of inverted *loxP* sites flanking the fusion gene HyTK (23). Expression of HyTK can be selected for with hygromycin or selected against with ganciclovir. Prior to electroporation, these cell lines were cultured for at least 2 weeks with 1 mg of hygromycin per ml in the tissue culture medium. A total of 4×10^6 cells were electroporated with 25 μg of construct plasmid DNA, 15 μg of pCMV-CRE, and 150 μg of salmon sperm DNA at 260 V, 1.0 mF. Half of the cells were transferred to 20 ml of culture medium and grown for 2 to 3 days, diluted 1:10, and placed in selection medium (medium plus 10 μM ganciclovir). Cells were grown for approximately 1 week in selection medium and then dilution cloned into 96-well plates and grown for 3 days in selection medium. Surviving clones were expanded and analyzed by FACS and Southern blotting approximately 3 weeks posttransfection.

FACS. Cells were analyzed on a FACStar flow cytometer (Becton-Dickenson) as described previously (2). Forward and side scatter measurements were used to differentiate live and dead cells; dead cells were removed from the analyses. Compensation was empirically adjusted by using cells transfected with single-

gene constructs and nontransfected cells. Data on 10,000 cells were collected, and further analyses were performed with FloJo Software (TreeStar, Inc., San Carlos, Calif.). The mean fluorescence and standard deviation of green fluorescence and yellow fluorescence were calculated from a minimum of three clones derived from at least two separate transfections. The orientations of the clones within the site were determined by Southern analyses, and the mean fluorescences were determined by a single FACS analysis done on each clone.

Orientation analyses. Genomic DNA was isolated from 1.5 ml of rapidly growing cultured cells by standard techniques. Twenty micrograms of genomic DNA was digested with *Pst*I to give the orientation of the two-gene constructs relative to the flanking sequences in the RL5 and RL6 loci. Single-gene constructs were oriented with *Acc*I. The Southern blots were probed with a radiolabeled 595-bp PCR product made with oligo1 (24mer; AAACGGCCACAAG TTCAGCGTGTC) and oligo 2 (23mer; CAGGACCATGTGATCGCGCTTC T). The probe detects both GFP and YFP equivalently because of their nearly identical sequences. Southern analyses were done on all clonal lines, constructs, and locations: for RL5 ($n = 195$), 60% were assigned to orientation A and 40% were assigned to orientation B; and for RL6 ($n = 74$) the assignments were 56 and 44%, respectively. A radiolabeled 1-kb ladder (Life Sciences, Bethesda, Md.) was used as a size marker.

Transcriptional analyses. RNA was harvested from 10-cm-diameter plates of rapidly growing clonal cell lines and isolated by using RNAzol B (Tel-Test, Inc., Friendswood, Tex.) as directed by the manufacturer. Thirty micrograms of total RNA was used for Northern analysis; the blot was probed with a radiolabeled 595-bp probe that detects both GFP and YFP equivalently. The blot was stripped and reprobed with a radiolabeled *Pst*I fragment of murine glyceraldehyde-3-phosphate dehydrogenase (GAPDH). Autoradiography in both cases was performed overnight, and quantitation was performed with a Molecular Dynamics (Sunnyvale, Calif.) PhosphorImager 445 SI and ImageQuant version 1.2. The GFP or YFP mRNA signal was normalized to the signal from GAPDH (arbitrary units) to account for loading variation. RNA MW Markers I (Roche Molecular Biologicals, Mannheim, Germany) were used to confirm size (data not shown).

Poly(A) RNA was purified from approximately 100 μg of total RNA with an affinity matrix consisting of oligo$(dT)_{25}$ cellulose beads (New England BioLabs, Beverly, Mass.). Five hundred nanograms of poly(A) RNA was subject to One-Step reverse transcription-PCR (RT-PCR) as directed by manufacturer (Roche Molecular Biologicals, Mannheim, Germany) by utilizing 300 nM oligo1 and oligo2 (described above under "Orientation analyses"), 400 nM each deoxynucleoside triphosphate (dNTP), and 1.0 μCi of [^{32}P]dCTP, at 62°C for 30 min for RT (0°C for the no-RT step), and cycling at 94°C for 30 s, 62°C for 30 s, and 71°C for 1 min (25 repetitions). The RT-PCR product was separated from unincorporated nucleotides and oligonucleotides by spin column (Edge Biosystems, Gaithersburg, Md.), digested with *Pst*I, visualized on an ethidium bromide-stained 1.2% agarose gel, and quantitated on a 4.5% polyacrylamide gel with a Molecular Dynamics PhosphorImager 445 SI and ImageQuant version 1.2. The YFP fraction was determined as a percentage of cleaved RT-PCR product (459 bp) of the total RT-PCR products (uncleaved 595 bp plus cleaved 459 bp), adjusting for the reduction in radiolabel in the shorter, cleaved YFP product.

RESULTS

Constructs in which two nearly identical, but distinguishable reporters are juxtaposed. The aim of this study was to establish the basic rules governing TI in a simple system in which transcription units, and their arrangement with respect to each other, can be altered conveniently. This system can be used to generate models that can be tested in more complex systems. When designing constructs for these TI experiments, we used the following principles. (i) The two reporter gene expression cassettes should be as similar as possible, and this similarity should include the regulatory regions and the polyadenylation signals. Differences in their expression will thus reflect experimental variables rather than differences in the expression cassettes. (ii) The reporter genes should be distinguishable with a convenient single-cell assay, so that variant subpopulations can be recognized if they exist. (iii) Each transcription unit should contain all elements needed for independent regulation, without the need for any interaction with unknown elements at the integration sites. (iv) The cassettes should use control elements with well-characterized functions common to many mammalian genes.

To further these objectives, we used GFP and the spectral variant YFP, which differ by only 4 amino acids (Fig. 1). The emission spectra of GFP and YFP can readily be differentiated by flow cytometry (2). The coding sequences are optimized for mammalian codon usage, and the two reporters have identical consensus translational start signals. Transcription is regulated by the powerful promoter and upstream regulatory elements from the human CMV immediate-early gene. Polyadenylation is mediated by the SV40 T-antigen polyadenylation sequences.

As diagrammed in Fig. 1A, the GFP and YFP cassettes were arranged with their transcription oriented either convergently, divergently, or tandemly. In the divergent and convergent constructs, the two genes are transcribed from opposite strands of the DNA, whereas the tandem constructs are transcribed from the same strand of DNA. In the tandem YG construct, the YFP transcription unit is upstream of the GFP unit; this arrangement is reversed in the tandem GY construct. The cassettes are closely spaced, with 40 bp of pGEM multiple cloning site separating each cassette. We also made tandem constructs in which either the upstream or the downstream promoter was disabled (Fig. 1B).

Transient transfections. As a first step, we assessed expression of the two reporters in transient transfections in MEL cells (Fig. 2). In this system, because the constructs are not integrated, expression occurs without the influence of flanking chromatin structure due to a genomic integration site. Flow cytometry distinguishes individual cells and scores them for the intensity of each reporter. Yellow fluorescent intensity is shown on the *y* axis, and green is shown on the *x* axis. Thus, the fluorescence of transfected cells expressing both reporters at equal levels will fall on a line with a slope of 1, starting at the origin.

In the majority of cells, there is no expression of the reporters: either the cells have not taken up DNA, or the transfected plasmid is not transcriptionally active. Of the cells with fluorescence above that of the negative cells, most express the GFP and YFP reporters at comparable levels, as indicated by the diagonal distribution of the dots. In comparison with the other three constructs, the convergent construct produces lower levels of both green and yellow fluorescence. Significantly, the divergent and the two tandem constructs do not produce different patterns of expression: the distributions of the cells transfected with these constructs are very similar. When the mean levels of fluorescence of the cells transfected with the tandem or divergent constructs were compared to expression of transiently transfected single-gene constructs, there was no significant difference. These results show that in transient transfections, transcriptional interference does affect convergent constructs, but has little or no effect on the other arrangements.

Integration of the constructs into defined genomic sites by site-specific recombination. Integration into chromatin affects the transcriptional regulation of transgene constructs in ways that vary among integration sites. These position effects might be expected to confound an assessment of interactions between transcription units. To eliminate this problem, we used RMCE, a site-specific integration method that exchanges a preintegrated HyTK cassette with the transgene to be tested.

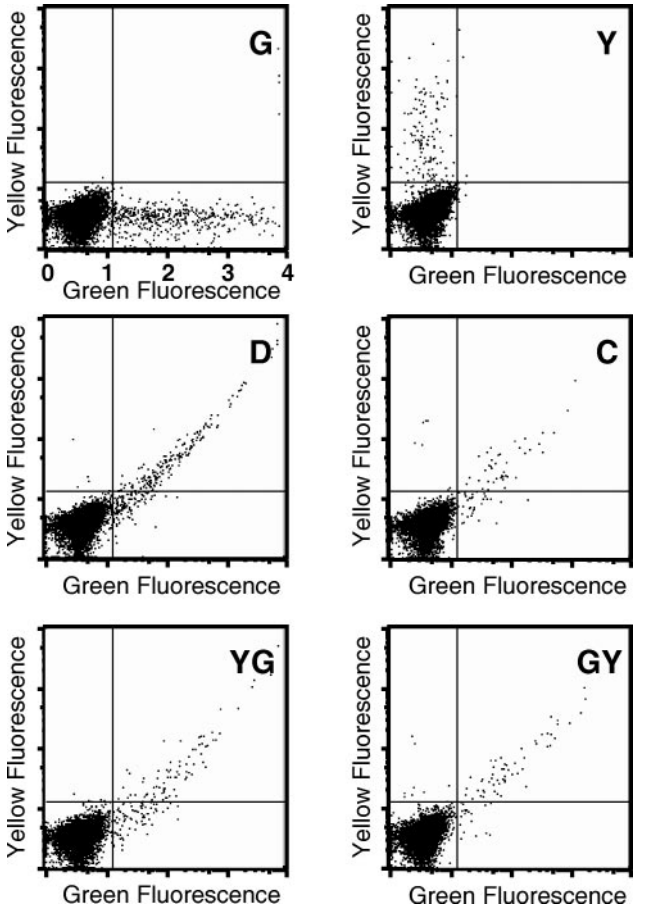

FIG. 2. Transiently transfected MEL cells do not reveal transcriptional interference. Shown are dot plots; green fluorescent intensity is on the *x* axis on a 4-decade log scale, and the yellow fluorescent intensity is on the *y* axis, also on a 4-decade log scale. Each dot represents one of 20,000 cells assayed; >95% remain untransfected, a percentage confirmed by cotransfection with pCMV-βgal (data not shown). Quadrants divide expressing and nonexpressing cells; quadrant lines are drawn to be above the level of 99% of cells with no construct. G, single-gene construct containing a single GFP-containing transcription unit under the same promoter and polyadenylation signal as diagrammed in Fig. 1; Y, construct with single YFP-containing transcription unit; D, two-gene divergent construct; C, two-gene convergent construct; YG, two-gene tandem construct; GY, two-gene tandem construct.

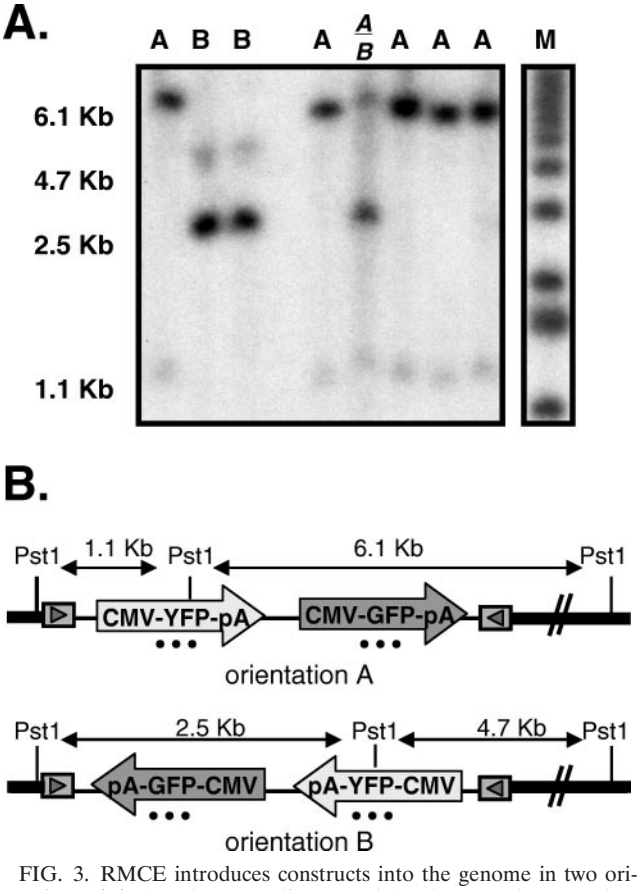

FIG. 3. RMCE introduces constructs into the genome in two orientations. (A) Sample autoradiogram of Southern analyses used to determine the orientation of two-gene constructs integrated into the RL5 genomic location. Shown are the *Pst*I-digested genomic DNAs of eight tandem YG clones that were probed with a radiolabeled PCR product recognizing both GFP and YFP equivalently. M denotes the lane containing a radiolabeled 1-kb ladder. Lanes A show clones assigned to orientation A based on the pattern of the fragments (1.1 and 6.1 kb, as indicated on the left). Lanes B show clones assigned to orientation B (with 2.5- and 4.7-kb fragments). A/B are mixed clones that appear to have cells with both orientations; only approximately 11% of the clones were found to be either aberrant or mixed, and these were excluded from further analyses. (B) Diagrammatic representation of the derivation of the *Pst*I restriction fragments that are diagnostic for either orientation A or orientation B with tandem YG integrated into the RL5 genome (thick line). Other diagram conventions are the same as in Fig. 1.

RMCE is based on double reciprocal recombination between *loxP* sites by CRE recombinase; the variant used here has two inverted but identical *loxP* sites and results in insertion of a single copy of the construct in either orientation with respect to flanking genomic sequence (13) (Fig. 3). Because the cassette exchange is based on selection against a preintegrated negative selectable marker, the test cassette does not have to be expressed. We inserted each of the constructs at two previously described integration sites in MEL cells, RL5 and RL6 (12). RMCE at these sites is very efficient: approximately 90% of the clones surviving negative selection have a correct exchange of the HyTK with the introduced construct as determined by Southern analysis, which also determines the orientation of the construct (Fig. 3A). Following RMCE, we derived and characterized several clones carrying each construct in each of the two orientations. By this method, position effects on the constructs are controlled; in addition, we can determine whether orientation with respect to flanking chromatin affects the interaction of the two transcription units.

Interference between transcription units integrated into chromatin. We analyzed clones by FACS. In the centers of Fig. 4 and 5, the results from typical clones are shown with yellow fluorescent intensity on the *y* axis and green fluorescent intensity on the *x* axis. The logarithmic FACS plots show that each construct produces a single population of cells with a broad expression profile. Within a clone, the expression of the two genes is correlated, so cells with the highest expression levels of one gene are most likely to have the highest expression levels of the other gene. When high- or low-expressing cells were

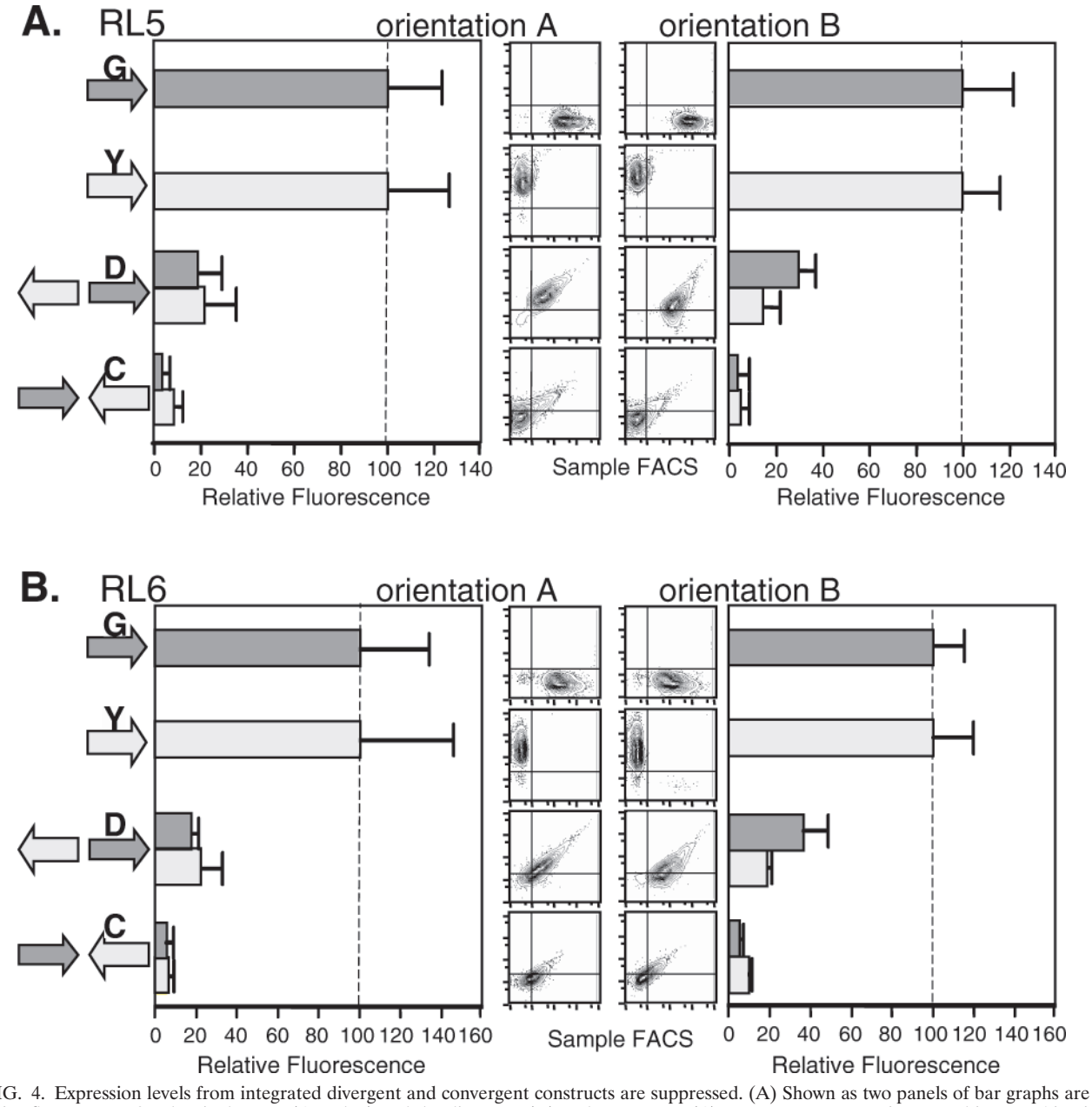

FIG. 4. Expression levels from integrated divergent and convergent constructs are suppressed. (A) Shown as two panels of bar graphs are the relative fluorescence for the single-gene (G and Y) and the divergent (D) and convergent (C) two-gene constructs integrated into RL5 in either orientation A or B. Dark gray bars are average green fluorescence, and light bars are average yellow fluorescence, expressed as a percentage of the mean fluorescence of single genes in the same orientation. Error bars represent the standard deviation of the mean fluorescence for three or more independent clones analyzed for each construct and orientation of that construct. All levels of expression are significantly different ($P < 0.001$) from the expression of the single gene. Shown between the two bar graph panels are contour FACS plots of representative clones for each construct in each orientation. The *x* axis is green fluorescent intensity on a 4-decade log scale, and the *y* axis is yellow fluorescent intensity also on the same log scale. Outliers are shown as small dots. Quadrants divide expressing and nonexpressing cells; quadrant lines are drawn to be above the level of 99% of cells with no construct. (B) Average of the fluorescence means for the same constructs integrated into RL6, again normalized to the mean fluorescence of the single genes. All levels of expression are significantly different ($P < 0.01$) from the expression of the single gene.

sorted from the populations and then expanded, the expanded population reproduced the pattern of the entire population from which the cloned cell was sorted (data not shown). Taken together, these observations suggest that each clonal population results from some variation of expression around a mean, but that expression of the two reporters is always correlated. The logarithmic scale of the FACS plots visually minimizes the differences between clones, but since the populations are continuous, expression information can be conveniently summarized by comparing the mean fluorescence levels.

In order to recognize TI on either gene, we compared the mean expression of each gene in two-gene constructs to expression of an identical single gene integrated at that site in the same orientation. The results are presented as bar graphs in

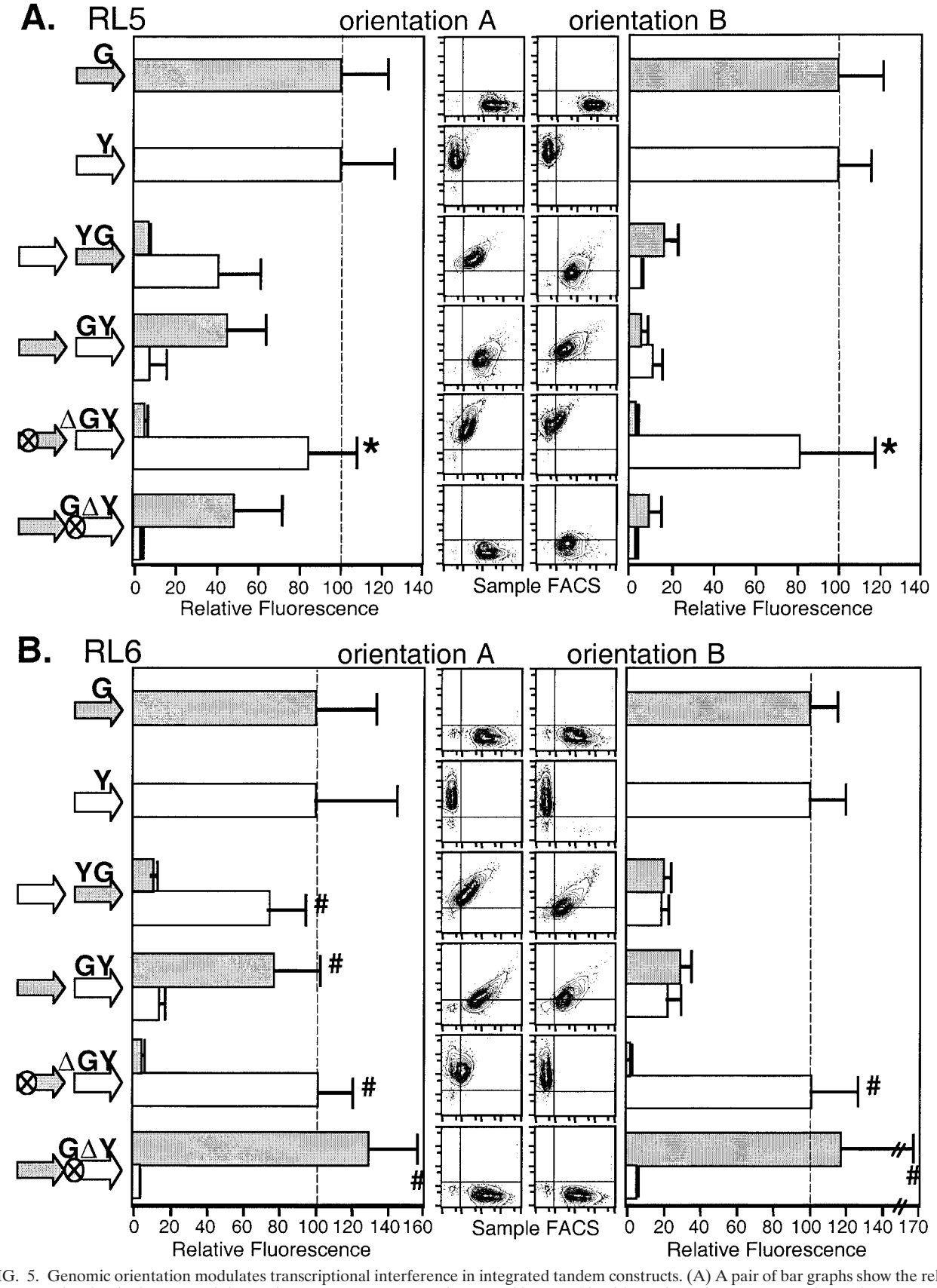

FIG. 5. Genomic orientation modulates transcriptional interference in integrated tandem constructs. (A) A pair of bar graphs show the relative fluorescence for the single gene and the tandem two-gene constructs integrated into RL5 in either orientation A or B. Dark gray bars are average green fluorescence, and light bars are average yellow fluorescence, expressed as a percentage of the mean fluorescence of single genes in the same

Fig. 4 and 5. It is obvious that, in all cases, expression from each of two active transcription units is less than expression from a single unit in the same integration site and orientation. However, when one of the two transcription units is disabled, expression often (but not always) approaches that of the single units (Fig. 5).

This consistent reduction in expression when two genes are closely linked in the genome could be due to an overall insufficiency of *trans*-acting factors or to a measurement problem in which measurement of one fluorophore reduces measurement of the other fluorophore. To address these questions, clonal cell lines carrying a single CMV-YFP at the RL5 locus were transiently transfected with a plasmid carrying a CMV-GFP transcription unit. After 48 h, cells were analyzed by FACS. Roughly 5% of the transfected cells express the transiently transfected CMV-GFP plasmid. Expression of YFP in these cells was not different from expression in cells in which the GFP plasmid was not transfected or was transfected and not expressed (data not shown). Cells that expressed GFP at any level did not show any reduction in YFP expression. The same experiment was done with the integrated GFP and the transiently transfected YFP, with the same result. Thus, the genes influence each other's expression only when they are integrated together. That this occurs at the transcriptional level and is not due to a fluorescence measurement artifact was confirmed independently by Northern analysis (see below).

When the reporter genes are arranged so that their transcription is convergent, expression of both genes is barely above the background level of fluorescence (Fig. 4A and B). Neither the integration site nor the orientation within that site has a significant influence.

The divergent constructs also showed a clear suppression of expression from each gene (Fig. 4). Both sites showed a three- to sixfold reduction in expression of each gene in comparison with that of a single gene in that site and orientation. Orientation within the site has a weak influence on the ratio of yellow and green fluorescence. The suppression is not as strong as that seen with the convergent constructs, but the sum of the fluorescence levels from both genes taken together is well below the level of a single gene. Thus, TI can occur under conditions in which readthrough transcription is not a possibility.

As with the convergent and divergent constructs, the expression from two genes in tandem is lower than the expression of a single gene (Fig. 5A and B). However, the integration sites and orientation strongly influence expression, and in three of the four configurations, the upstream unit is also significantly suppressed. The patterns of expression of YG and GY are mirror images of each other, supporting the view that the units are essentially identical.

At RL5 (Fig. 5A), the upstream gene almost completely suppresses the downstream gene in orientation A, but the upstream gene is itself also suppressed to <50% of the single gene. In orientation B, both genes are strongly suppressed, and the upstream gene is actually lower than the downstream unit. This suggests a strong suppressive effect by chromatin on one flank of the integration site (which, however, does not act on a single gene driven by the CMV promoter). At RL6 (Fig. 5B), suppression of the upstream gene is not as severe. In orientation A, expression of the upstream gene is not significantly different from that of a single gene, but the downstream gene is strongly suppressed. This was the only situation examined in which both genes were not strongly suppressed by their proximity. However, in orientation B, the expression of both genes is strongly suppressed, as at RL5, and both genes are suppressed equally.

Since we observed suppression of the upstream gene in a tandem pair at RL5, but not at RL6 (compare Fig. 5A and B, orientation A), we hypothesized that this effect is not inherent in the arrangement of the genes, but rather involves an interaction with the integration site. We therefore tested the importance of transcription of either gene in suppression of its partner. The CMV promoter was deleted from either the upstream (ΔGY) or downstream (GΔY) gene of the tandem GY pair (Fig. 1B). These constructs were assayed at RL5 and RL6 in both orientations (Fig. 5). As expected, deletion of the basal promoter reduced expression of a gene to background levels regardless of the integration site, orientation, or position relative to the other reporter. At both sites and in both orientations, deletion of the upstream promoter fully restored expression of the downstream gene to levels that are not significantly different from those with a single gene. Thus, suppression of the downstream gene by the upstream gene is dependent on transcription of the upstream gene. Deletion of the downstream promoter had no statistically significant effect on expression of the upstream gene at RL5, in either orientation (compare GY and GΔY for each orientation in Fig. 5A). This demonstrates that suppression of the upstream gene at this site is not mediated by transcription of the downstream gene and suggests other influences on gene expression.

However, the apparent independence of upstream suppression from downstream transcription was not reproduced at RL6. The strong suppression of the upstream gene in orientation B was fully relieved by deletion of the promoter of the downstream gene (Fig. 5B, compare GY and GΔY). In orientation A, suppression of upstream expression is slight when the downstream promoter is intact. Expression of the upstream gene is increased by deletion of the downstream promoter, but in neither case is expression significantly different from the expression of a single gene.

mRNA analysis confirms transcriptional basis. The fluorescence data show that total fluorescence is reduced in each case

orientation. G and Y are the single genes, YG and GY are the tandem two-gene constructs, ΔGY is the tandem GY construct with the deletion of the upstream CMV basal promoter, and GΔY is the tandem GY construct with the downstream CMV basal promoter deleted. Error bars represent the standard deviation of the mean fluorescence for six or more independent clones analyzed for each construct and orientation of that construct. All levels of expression are significantly different ($P < 0.001$) from the expression of the single gene, except the ones labeled with an asterisk ($P > 0.1$). Shown between the two bar graph panels are contour FACS plots of representative clones for each construct in each orientation, as in Fig. 4. (B) Average of the fluorescence means for the same constructs integrated into RL6, normalized to the mean fluorescence of the single gene. All levels of expression are significantly different ($P < 0.01$) from the expression of the single gene, except the ones labeled with # ($P > 0.2$).

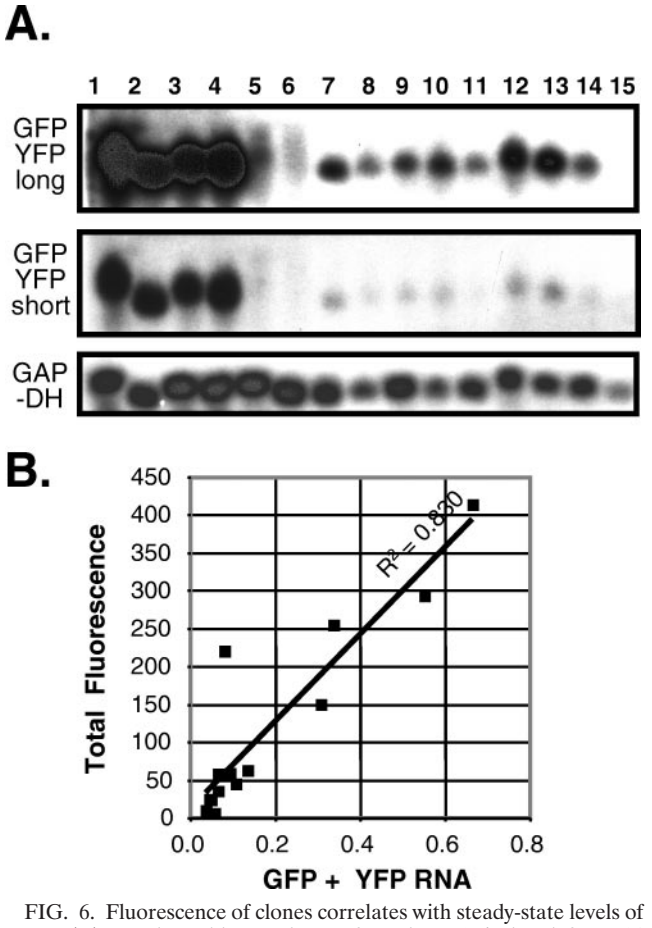

FIG. 6. Fluorescence of clones correlates with steady-state levels of RNA. (A) Northern blot analyses of total RNA isolated from 14 independent RL5 clones hybridized to a radiolabeled probe that recognizes both GFP and YFP (in the upper and middle panels, showing long and short exposures, respectively), and to a 1.2-kb fragment of GAPDH, as a loading control (in the lower panel). Lanes: 1, RNA isolated from a clone carrying a single GFP in orientation A; 2, GFP alone in orientation B; 3 and 4, YFP alone in orientations B and A, respectively; 5 and 6, convergent in orientations A and B, respectively; 7 and 8, divergent in orientations B and A, respectively; 9 and 10, tandem YG in orientations B and A, respectively; lanes 11, 12, 13, and 14, tandem GY in orientations B, A, B, and B, respectively; 15, RNA isolated from untransfected RL-5. (B) Total fluorescence (sum of mean GFP and YFP fluorescences, determined by FACS as in Fig. 4) of each of the clones shown above as a function of the normalized content of fluorescent protein transcripts detected in the Northern blot shown in panel A. The line is the trend line determined by regression analysis with a correlation coefficient (R^2) of 0.830.

in which there are two genes in a single locus. Since our assumption is that transcriptional interference is occurring, it is important to demonstrate that the fluorescence levels we measure are directly correlated with RNA expressed by the reporter genes. To address this question, the steady-state levels of total GFP plus YFP mRNA from a full series of clones at RL5 were assayed by Northern blotting. Using a probe that recognizes both GFP and YFP, the total amount of mRNA was compared to the total fluorescence from both proteins. As shown in Fig. 6, total GFP plus YFP RNA is well correlated with total GFP plus YFP fluorescence.

The fluorescence data also include striking differences between GFP and YFP fluorescence levels. To demonstrate that the ratio of GFP/YFP fluorescence correlates with the ratio of specific mRNA produced by each of these genes, we developed an RT-PCR-based system to measure the relative amounts of GFP and YFP. PCR primers that amplify the mRNA from either gene with equal efficiency were used to amplify the reverse-transcribed product from both genes from the series of RL5 clones examined in Fig. 6. After amplification, the relative amounts of each product were revealed by cutting the amplified DNA with a restriction enzyme that recognizes the YFP product, but not the GFP product. The ratio of the cut and uncut products was quantitated and compared to the ratio of fluorescence. As shown in Fig. 7, there is a strong correlation between YFP mRNA and yellow fluorescence.

DISCUSSION

We have devised a simple system in which the interaction of independently regulated transcription units can be analyzed; this system is amenable to expansion and will be useful in modeling the interaction of genes in more complex loci. We find that two transcription units interfere with each other in ways that cannot be explained fully by simple models of readthrough transcription and promoter occlusion. Furthermore, the genomic context exerts an influence that modulates the transcriptional interference, perhaps because other genes are located nearby. In eukaryotic genomes, the juxtaposition of active transcription units is common; thus, the mechanisms causing interference probably participate in regulation of many genes.

Although transcriptional interference has been quite widely described, there is no systematic understanding of the parameters or mechanisms involved. The system used for these studies is the first reported effort to evaluate TI by comparing expression of each gene in a linked pair of genes to expression of an otherwise identical single gene at a given genomic site. The results show that two closely linked RNA polymerase II-transcribed genes with robust polyadenylation sites suppress each other's expression. This was true for each case of convergent or divergent transcription, for the downstream gene in every tandem arrangement, and for the majority of the upstream genes in tandemly arranged constructs. The relative orientation of the genes influences the extent of suppression, but in every instance, the total expression from two genes was less than or roughly equal to the expression from a single gene in that site.

It is interesting that TI in the integration sites used here did not result in silencing of either gene, which would manifest as a distinct population of cells expressing one reporter and not the other. In some reported cases, TI does appear to result in transcriptional silencing (17). Our assay detects expression in single cells and would reveal subpopulations that differ in the G/Y ratio within a single clone. We did not see distinct subpopulations in FACS analyses: instead the FACS plots reveal continuous populations with a correlation of GFP and YFP expression. Although we cannot exclude the possibility that the long half-life of GFP and YFP (>24 h) masks short-term effects, we did not detect a series of metastable states in which a cell preferentially expresses one gene or the other. In every case, both genes appear to be active at all times. Silencing is an extremely site-dependent phenomenon (24), and the two sites

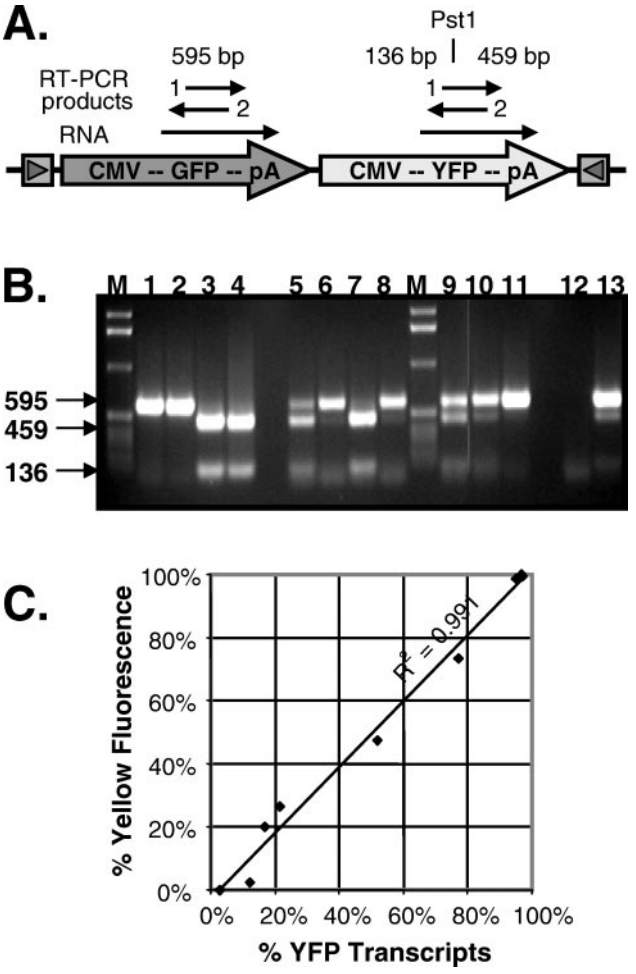

FIG. 7. Fraction of total fluorescence due to YFP correlates with the fraction of mRNA from YFP. (A) Diagrammatic representation of the RT-PCR and restriction fragment length polymorphism analysis used to distinguish GFP and YFP transcripts. The oligomers used amplify YFP and GFP with equal efficiency. The PCR products for both GFP and YFP are identical, with the exception of a *Pst*I site present in products derived from YFP transcripts that is lacking in GFP. (B) Sample ethidium bromide-stained agarose gel illustrating RT-PCR products digested with *Pst*I to differentiate GFP (uncut at 595 bp) and YFP (cleaved to give 459- and 136-bp products). Poly(A) RNAs from 11 independent clones were subjected to RT-PCR followed by restriction fragment length polymorphism analysis. Lanes 1 to 4 are derived from the single-gene clones G and Y as in Fig. 6. Lanes 5 to 8 are derived from two-gene convergent and divergent clones as in Fig. 6. Lane 9 is from a YG orientation B clone. Lanes 10 and 11 are GY orientations B and A, respectively. Lane 12 is the same as lane 1, but without the reverse transcriptase step, and lane 13 represents a mixture of poly(A) RNAs in a GFP/YFP transcript ratio of 4:1. Lane M contains the 1-kb ladder from Life Sciences. (C) The proportion of yellow fluorescence is plotted against the proportion of YFP transcripts. The proportion of yellow transcripts expressed as a percentage of total (GFP plus YFP) transcripts in that clone was determined radiometrically (in duplicate) on polyacrylamide gels by the RT-PCR and restriction fragment length polymorphism technique illustrated in panels A and B. The line is the trend line determined by regression analysis with a correlation coefficient (R^2) of 0.991.

used here are ones in which silencing of integrated reporters is not prominent (12).

The genomic site influences TI: origin of stable position effect? Our results demonstrate that the genomic site can strongly influence transcriptional interference. The genomic sites used here, RL5 and RL6, had previously been thoroughly examined for an influence of orientation on expression of a single CMV-GFP reporter gene (12). Orientation did not have an influence on the level of expression of CMV-GFP at either of these sites, although other promoters did show such an orientation effect at both RL5 and RL6 and one orientation at RL6 was prone to silencing over time with CMV-GFP (12). At both sites, however, there is a clear effect of orientation when two genes driven by CMV are transcribed from the same strand of the asymmetric tandem constructs. These studies show that transcriptional interference of independent transcription units can be influenced by genomic orientation, even when an identical single gene in the same site is not sensitive to genomic orientation.

Position effects on transgenes in mammals are ubiquitous (24) and imply that genes in their native contexts are strongly influenced by the characteristics of flanking chromatin. Variegating position effects, which cause stochastic silencing, have been well characterized and are often the result of proximity to heterochromatin. "Stable position effect" is a term coined by Lewis (21) to describe position-dependent changes in the level of gene expression without silencing. Based on the data presented here, we hypothesize that the stable position effect results in part from transcriptional interference by active genes around the integration site. As this study suggests, the strength of those effects will be strongly influenced by the precise orientation of the units with respect to each other.

Practical implications. These studies prompt some conclusions concerning the architecture of constructs designed for expression in mammalian cells, whether for experimental or gene therapeutic purposes. Our studies suggest that two closely linked transcription units will always interfere with each other. If two genes must be expressed, TI could be avoided by using a single promoter and an internal ribosome entry site, or by placing the gene on different plasmids and selecting clones in which the two are integrated in different sites. If the genes are placed on the same plasmid, a divergent arrangement appears to be the best option for expressing both.

Mechanisms of transcriptional interference. It has been found that suppression of downstream expression in tandem constructs is relieved when a polyadenylation and pause site separate the genes (18). This suggests that readthrough transcription is one mechanism of TI and that a poly(A) site will potentially limit that phenomena. However, our studies document interference with expression of the downstream gene even when the upstream gene has a robust polyadenylation site. Thus, either the poly(A) site used here [SV40 poly(A)] is not terminating transcription efficiently, or other mechanisms are involved, or both.

The nature of transcriptional termination in mammalian systems is poorly defined. Unlike prokaryotes or unicellular eukaryotes, higher eukaryotes do not have an identified consensus site for transcription termination. Current models postulate that the polymerase II elongation complex is modified at the polyadenylation site to a less processive form that eventually dissociates from the DNA (reviewed in reference 27). Suppression of the downstream gene in a tandem pair could therefore be attributed to promoter occlusion due to the disruption of the architecture of the downstream promoter region

by polymerase II complexes coming from the upstream gene and proceeding past the polyadenylation site. This readthrough transcription and promoter occlusion mechanism could also be involved in the convergent constructs, but could not mediate the suppression seen with divergent constructs or the suppression of the upstream gene by expression of the downstream gene in a tandem pair.

Convergent constructs showed the most significant interference. There is one mechanism that would be unique to the convergent arrangement, and that is antisense inhibition. In the convergent arrangement, readthrough transcription that extends a few hundred bases past the polyadenylation site of one gene will begin to transcribe the antisense strand of the other gene. Although this antisense RNA would only come from the nascent transcript before polyadenylation, it may still have an effect. It is also possible that the use of two nearly identical reporter genes exacerbates the antisense effect. In the convergent situation with two such similar genes, the readthrough from one convergent gene into the other would produce nascent antisense that could anneal to the mRNA of either gene.

Our finding that the upstream gene in a tandem pair is also subject to a suppressive influence from transcription of the downstream gene is novel. At one site (RL6), this effect occurs only in one orientation, implying that this suppression is not entirely inherent in the arrangement, but is also influenced by the genomic site. The finding of interference between divergent transcription units in mammals has also not been reported previously, although there is a report of this effect in a mammalian herpesvirus (5). Promoter occlusion or antisense from readthrough transcription cannot explain suppression of divergent genes or the upstream gene of a tandem pair. Suppression could be due to changes in supercoiling caused by transcription of the partner gene (8). Such torsional or topological effects could be transmitted in either direction from the site of transcription and therefore could be involved with any of the TI effects documented in this report.

Another less obvious mechanism that would affect every two-gene arrangement is suggested by recent insights into the importance of subnuclear structure in transcriptional regulation. Transcription may require formation of a large complex including the promoter and an immobile transcription factory (4). Formation of a promoter complex might inhibit formation of a second promoter complex within a given physical distance, and this would suppress expression of any two promoters in close proximity.

These studies show that linked genes with apparently independent regulatory elements generally suppress each other's expression. This view implies that transcriptional interference can be a significant factor in determining expression levels, both of transgenes and of genes in their normal positions in the genome. Proximity to other genes might therefore be one important factor influencing the level of endogenous gene transcription; presumably gene regulatory elements have adapted to the influence of flanking or intermingled elements. Further support for this hypothesis will come from studies that detail the extent and mechanisms of transcriptional interference in mammals.

ACKNOWLEDGMENTS

We acknowledge the support of the Burroughs-Wellcome Career Development Award (S.N.F.); National Institutes of Health grants DK54071 (S.N.F.), HL38655, HL55435, and DK56845 (E.B.); and the Dartmouth Flow Cytometry Facility and its director, Alice Longobardi Givan.

REFERENCES

1. **Bateman, E., and M. Paule.** 1988. Promoter occlusion during ribosomal RNA transcription. Cell **54:**985–992.
2. **Beavis, A., and R. Kalejta.** 1999. Simultaneous analysis of the cyan, yellow and green fluorescent proteins by flow cytometry using single-laser excitation at 458 nm. Cytometry **37:**68–73.
3. **Boussadia, O., F. Amiot, S. Cases, G. Triqueneaux, H. K. Jacquemin-Sablon, and F. Dautry.** 1997. Transcription of UNR (upstream of N-ras) downmodulates N-ras expression in vivo. FEBS Lett. **420:**20–24.
4. **Cook, P. R.** 1999. The organization of replication and transcription. Science **284:**1790–1795.
5. **Cook, W., K. Wobbe, J. Boni, and D. Coen.** 1996. Regulation of neighboring gene expression by the herpes simplex virus type 1 thymidine kinase gene. Virology **218:**193–203.
6. **Corbin, V., and T. Maniatis.** 1989. Role of transcriptional interference in the *Drosophila melanogaster Adh* promoter switch. Nature **337:**279–282.
7. **Cullen, B., P. Lomedico, and G. Ju.** 1984. Transcriptional interference in avian retroviruses—implications for the promoter insertion model of leukaemogenesis. Nature **307:**241–245.
8. **Dunaway, M., and E. Ostrander.** 1993. Local domains of supercoiling activate a eukaryotic promoter in vivo. Nature **361:**746–748.
9. **Eggermont, J., and N. Proudfoot.** 1993. Poly(A) signals and transcriptional pause sites combine to prevent interference between RNA polymerase II promoters. EMBO J. **12:**2539–2548.
10. **Emerman, M., and H. M. Temin.** 1984. Genes with promoters in retrovirus vectors can be independently suppressed by an epigenetic mechanism. Cell **39:**449–467.
11. **Esperet, C., S. Sabatier, M. Deville, R. Ouazana, E. Bouhassira, J. Godet, F. Morle, and A. Bernet.** 2000. Non-erythroid genes inserted on either side of the human HS-40 impair activation of its natural alpha-globin gene targets without being themselves preferentially activated. J. Biol. Chem. **275:**25831–25839.
12. **Feng, Y.-Q., M. C. Lorincz, S. Fiering, J. M. Greally, and E. E. Bouhassira.** 2001. Position effects are influenced by the orientation of a transgene with respect to flanking chromatin. Mol. Cell. Biol. **21:**298–309.
13. **Feng, Y.-Q., J. Seibler, R. Alami, A. Eisen, K. Westerman, P. Leboulch, S. Fiering, and E. Bouhassira.** 1999. Site-specific chromosomal integration in mammalian cells: highly efficient CRE recombinase-mediated cassette exchange. J. Mol. Biol. **292:**779–785.
14. **Fiering, S., M. Bender, and M. Groudine.** 1999. Analysis of mammalian cis-regulatory DNA elements by homologous recombination. Methods Enzymol. **306:**42–66.
15. **Fiering, S., E. Epner, K. Robinson, Y. Zhuang, A. Telling, M. Hu, D. I. K. Martin, T. Enver, T. Ley, and M. Groudine.** 1995. Targeted deletion of 5′HS2 of the murine β-globin locus reveals that it is not essential for proper regulation of the β-globin locus. Genes Dev. **9:**2203–2213.
16. **Fiering, S., C. G. Kim, E. E. Epner, and M. G. Groudine.** 1993. An "in-out" strategy using gene targeting and flp recombinase for the functional dissection of complex DNA regulatory elements: analysis of the β-globin locus control region. Proc. Natl. Acad. Sci. USA **90:**8469–8473.
17. **Fincham, J., and G. Sastry.** 1974. Controlling elements in maize. Annu. Rev. Genet. **8:**15–50.
18. **Greger, I., F. Demarchi, M. Giacca, and N. Proudfoot.** 1998. Transcriptional interference perturbs the binding of Sp1 to the HIV-1 promoter. Nucleic Acids Res. **26:**1294–1300.
19. **Greger, I., A. Aranda, and N. Proudfoot.** 2000. Balancing transcriptional interference and initiation on the GAL7 promoter of Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. USA **97:**8415–8420.
20. **Hug, B. A., R. L. Wesselschmidt, S. Fiering, M. A. Bender, E. Epner, M. Groudine, and T. J. Ley.** 1996. Analysis of mice containing a targeted deletion of the β-globin locus control region 5′ hyperactive site 3. Mol. Cell. Biol. **16:**2906–2912.
21. **Lewis, E. B.** 1950. The phenomenon of position effect. Adv. Genet. **3:**73–115.
22. **Liebhaber, S., F. Cash, and D. Main.** 1985. Compensatory increase in α-1 globin gene expression in individuals heterozygous for the α-thalassemia 2 deletion. J. Clin. Investig. **76:**1057–1064.
23. **Lupton, S. D., L. L. Brunton, V. A. Kalberg, and R. W. Overell.** 1991. Dominant positive and negative selection using a hygromycin phosphotransferase-thymidine kinase fusion gene. Mol. Cell. Biol. **11:**3374–3378.
24. **Martin, D. I. K., and E. Whitelaw.** 1996. The vagaries of variegating transgenes. Bioessays **18:**919–923.
25. **McClintock, B.** 1968. The states of a gene locus in maize. Carnegie Inst. Wash. Year Book **66:**20–28.
26. **Pham, C., D. MacIvor, B. Hug, J. Heusel, and T. Ley.** 1996. Long-range

disruption of gene expression by a selectable marker cassette. Proc. Natl. Acad. Sci. USA **93:**13090–13095.
27. **Proudfoot, N.** 2000. Connecting transcription to messenger RNA processing. Trends Biol. Sci. **25:**290–293.
28. **Proudfoot, N.** 1986. Transcriptional interference and termination between duplicated α-globin gene constructs suggests a novel mechanism for gene regulation. Nature **322:**562–565.
29. **Shaw-White, J., N. Denko, L. Albers, T. Doetschman, and J. Stringer.** 1993. Expression of the *lacZ* gene targeted to the HPRT locus in embryonic stem cells and their derivatives. Transgenic Res. **2:**1–13.
30. **Vales, L., and J. Darnell.** 1989. Promoter occlusion prevents transcription of adenovirus polypeptide IX mRNA until after DNA replication. Genes Dev. **3:**49–59.
31. **Wu, J., G. Grindlay, P. Bushel, L. Mendelsohn, and M. Allan.** 1990. Negative regulation of the human ε-globin gene by transcriptional interference: role of an Alu repetitive element. Mol. Cell. Biol. **10:**1209–1216.