The 5′ Boundary of the Human Apolipoprotein B Chromatin Domain in Intestinal Cells†

Travis J. Antes,‡,§ Stephanie J. Namciu,|| R. E. K. Fournier,|| and Beatriz Levy-Wilson*,‡,§

Research Institute, Palo Alto Medical Foundation, Palo Alto, California 94301, Department of Medicine, Stanford University, Palo Alto, California 94303, and Division of Basic Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington 98109-1024

Received January 11, 2001; Revised Manuscript Received April 5, 2001

ABSTRACT: The 5′ boundary of the chromosomal domain of the human apolipoprotein B (apoB) gene in intestinal cells has been localized and characterized. It is composed of two kinds of boundary elements; the first, functional boundary is an insulator activity exhibited by a 1.8 kb DNA fragment located between −58 and −56 kb upstream of the human apoB promoter. In this region, an enhancer-blocking activity has been mapped to a CTCF binding site that is located upstream of two apoB intestinal enhancers (IEs), the 315 IE and the 485 IE. The CTCF site represents a boundary between two types of chromatin structure: an open, DNaseI-sensitive region 3′ of the CTCF site containing the intestinal regulatory elements and a closed, DNaseI-resistant region 5′ of the CTCF site. The 1.8 kb fragment harboring the CTCF site also insulated *mini-white* transgenes against position effects in *Drosophila melanogaster*. The second, structural boundary is represented by a nuclear matrix attachment region (MAR), situated about 3 kb 5′ of the CTCF site. This MAR may represent the 5′ anchorage site for a chromosomal loop that functions to bring the intestinal regulatory elements closer to the apoB promoter.

The transcriptional regulation of tissue- and stage-specific eukaryotic genes is intimately tied to the nuclear and genomic organization of the expressing or nonexpressing tissue or cell type. Active and inactive genes exist in loci with independent programs of expression and are interspersed along the chromosome. Therefore, boundaries between active and inactive loci are a necessary feature of eukaryotic chromatin. The genome is arranged in chromosomal loops that vary in length from 20 to >100 kb. These loops are anchored at their bases by DNA sequences called matrix attachment regions (MARs)[1] (*1*) or scaffold association regions (SARs) (*2*), which interact with the nuclear matrix. The attachment points serve to delineate the boundary of the looped domain and may also help to insulate the genes contained within that loop from regulatory elements residing in neighboring loops. The looped domains by themselves might suffice to insulate genes in cases where there is only one gene within the looped domain, or when the loop contains a cluster of genes that are coordinately regulated, and therefore share regulatory elements. In these cases, adjacent loops would harbor inactive genes or spacer DNA. However, when several genes that are neither coordinately regulated nor expressed in the same tissues or cell types are located in proximity to each other, either sharing a looped domain or in adjacent domains, the need for functional barriers or boundaries with insulating capability becomes essential to guaranteeing correct patterns of gene expression.

This situation is well-exemplified by the chromatin domain encompassing the chicken β-globin gene, which is flanked at its 5′ end by a folate receptor gene and at its 3′ end by an olfactory receptor gene (*3*, *4*). The globin gene locus is separated from those two other gene loci by insulators, thus ensuring that the correct patterns of gene expression are preserved for each locus. The first examples of boundary insulators were the specialized chromatin structures (*scs* and *scs*′) from the *Drosophila* 87A7 heat shock locus. These elements conferred position-independent expression upon a reporter gene that determines eye color, thus protecting the transgenes introduced into flies against position effects at the sites of integration (*5*).

In recent years, vertebrate insulators have been described (*6*), and a second assay has been used to characterize them, namely, the enhancer-blocking assay. Insulators block the activity of an enhancer in a position-dependent manner, when placed between the enhancer and its target promoter. Understanding regarding the nuclear proteins with insulator activity is limited, although a protein named CCCTC-binding factor (CTCF) plays a key role as an insulator in the chicken β-globin locus as well as in the human T cell receptor α/δ locus (*7*), the insulin growth factor locus (*8*, *9*), the mycFV

† This work was supported by grants from the National Institutes of Health: HL-54775 to B.L.-W. and GM-26449 to R.E.K.F.

* To whom correspondence should be addressed: Research Institute, Palo Alto Medical Foundation, Ames Building, 795 El Camino Real, Palo Alto, CA 94301. Phone: (650) 853-4764. Fax: (858) 759-4641 or (650) 329-9114. E-mail: blwilson2000@yahoo.com.

‡ Palo Alto Medical Foundation.

§ Stanford University.

|| Fred Hutchinson Cancer Research Center.

[1] Abbreviations: MAR, matrix attachment region; SAR, scaffold association region; DH, DNaseI hypersensitive; LDL, low-density lipoprotein; apoB, apolipoprotein B; CAT, chloramphenicol acetyltransferase; IE, intestinal element; ICR, intestinal control region; *scs*, specialized chromatin structures; CTCF, CCCTC-binding factor; BEAD, blocking element alpha/delta-1; BEAF, boundary element-associated factor; F, forward; R, reverse.

10.1021/bi0100743 CCC: $20.00 © 2001 American Chemical Society
Published on Web 05/16/2001

gene (*10*), the amyloid precursor protein β locus (*11*), and the chicken lysozyme gene (*12*).

The human apolipoprotein B (apoB) gene encodes the protein of low-density lipoprotein (LDL) that plays a key role in cardiovascular disease and lipid metabolism (*13*). This gene is expressed mainly in liver and intestine (*14*). The liver-derived HepG2 cell line and the intestine-derived CaCo-2 cell line have provided good model systems for studying apoB synthesis and gene regulation. In hepatic cells, the gene resides in a ~50 kb DNaseI-sensitive, looped domain that is flanked by two nuclear matrix association regions (MARs). This looped domain contains all of the regulatory elements required for correct hepatic expression in vivo (*15*−*17*). In contrast, the intestinal elements capable of conferring expression upon transgenic mice are localized more than 53 kb 5′ from the structural gene (see ref *18*−*20*). In this paper, we present evidence that a DNA fragment encompassing a CTCF-binding site blocks the activity of the human apoB gene intestinal enhancer, when placed between this enhancer and the promoter. Additionally, we demonstrate that a 1.8 kb segment from the apoB intestinal control region harboring the CTCF site protects *mini-white* genes against position effects in transgenic *Drosophila.* Furthermore, the CTCF-binding site represents a chromatin boundary separating an open, DNaseI-sensitive chromatin region containing the intestinal control region (ICR) from a transcriptionally inactive, DNaseI-resistant chromatin region. Upstream of the insulator resides a MAR. Whereas the CTCF-binding site most likely represents the functional 5′ boundary of the apoB gene in intestinal cells, the MAR probably represents the physical or structural 5′ boundary of the apoB gene by providing the 5′ anchorage site for a chromosomal loop that would bring the intestinal control elements in close aposition to the apoB promoter, thus allowing transcriptional activation in the intestine.

EXPERIMENTAL PROCEDURES

Plasmid Construction. Construction of plasmids −85CAT, 2.1-F, 2.1-R, 315-F, and 315-R has been previously described (*20*). Plasmid 800-R was made by digesting plasmid 2.1-F with *Eco*RI, and the resulting 485 bp fragment was then treated with Klenow DNA polymerase and blunt-end ligated into 315R−85CAT. Plasmid 400-R was made by blunt-end ligation of a 400 bp *Sty*I−*Hin*dIII fragment derived from plasmid 2.1-F into −85CAT. Transformants harboring the 800 and 400 bp segments in the reverse orientation were selected by restriction enzyme analysis. To make 315/CTCF/−85CAT, a double-stranded CTCF oligonucleotide was blunt-end ligated into phosphatased 315R−85CAT that had been cut with *Xba*I. The number and orientation of the CTCF oligomers were determined by DNA sequencing. Plasmid 485RCATΔCTCF was made by PCR using oligonucleotides 485ΔCTCF and m485.2 as primers and the 1.8 IE as a template, followed by blunt-end ligation of the purified PCR product into −85CAT and selection for the reverse orientation of the 450 bp insert.

Cell Culture and Transfections. Intestinal CaCo-2 cell cultures were maintained as described previously (*20*). Transfections were performed using the calcium phosphate transfection kit from Eppendorf, Inc. Typically, 10 μg of the plasmid DNA along with 5 μg of an internal reference plasmid (pRSVβ-gal) were cotransfected into cells in duplicate plates. Cell lysates were prepared after 48 h with three freeze−thaw cycles and lysate clarification by centrifugation at 15000*g* for 5 min at 5 °C. The levels of β-galactosidase activities, as well as levels of chloramphenicol acetyltransferase (CAT) activity, were assessed as previously described (*20*). The levels were quantitated with PhosphorImage analysis and Image Quant software (Molecular Dynamics, Inc., Sunnyvale, CA). All CAT activity values represent averages of at least three independent transfection experiments and are corrected for transfection efficiencies between plates by dividing the CAT activity levels by the β-galactosidase levels.

Gel Retardation Assays. COS cells were transfected with 10 μg of an expression plasmid for human CTCF, and cellular lysates enriched in human CTCF were prepared as previously described (*20*). Binding reactions and electrophoretic analysis of DNA−protein complexes has been reported (*20*). A large excess of poly(dI/dC) was added to all reaction mixtures as a nonspecific competitor (*20*).

Oligonucleotides were as follows: chicken CTCF, TAATTACGTCCCTCCCCCGCTAGGGGGCAGCA; apoB CTCF, CAAATTATCCTGCCCCCTAGACATAACCTCCC; 485Δ-CTCF, GGCTGTTGAACATTCAAATTTATG; and m485.2, GCTCTAGAGCAGGAGAATTGGAATTCC.

Transgenic Fly Assay. Construction of the *P*-element vectors mwS′ and apoB3′MARmwS′ has been described (*21*). ApoB1.8IEmwS′ was made by ligating a 1.8 kb DNA fragment located −56 to −58 kb from the transcription start site of the human apoB gene into the *Kpn*I−*Not*I sites upstream of *mini-white* in the *P*-element vector mwS′. Transformation of *Drosophila* w^{-1118}, establishment of lines, and determination of transgene copy number were carried out as described by Namciu et al. (*21*).

Preparation of Nuclear Matrices. Nuclear matrices were prepared according to the procedure of Cockerill and Garrard (*1*). In brief, nuclei (~1 mg of nucleic acids/mL) in RSB [10 mM NaCl, 3 mM $MgCl_2$, and 10 mM Tris-HCl (pH 7.4)], 0.25 M sucrose, and 1 mM $CaCl_2$ were digested with 100 μg of DNaseI/mL (Worthington) for 1−2 h at 23 °C. After centrifugation at 750*g* for 10 min at 4 °C, the pellets were resuspended in an RSB/0.2 M sucrose mixture, and an equal volume of a cold solution containing 4 M NaCl, 20 mM EDTA, and 20 mM Tris-HCl (pH 7.4) was added. After 10 min at 4 °C, followed by centrifugation at 1500*g* for 15 min, the pellets were extracted twice by suspension in a cold solution of 2 M NaCl, 10 mM EDTA, 10 mM Tris-HCl (pH 7.4), 0.5 mM phenylmethanesulfonyl fluoride, and 0.25 mg of BSA/mL and then centrifuged for 15 min at 4500*g* at 4 °C. The nuclear matrices were then washed with an RSB/0.25 M sucrose mixture containing 0.25 mg of BSA/mL at 4 °C and resuspended in the same solution. Although fresh preparations were used in most experiments, occasionally matrices were stored for up to 3 months at −20 °C in the presence of 50% glycerol.

DNA Binding to Nuclear Matrices. DNA binding assays were performed by the protocol of Cockerill and Garrard (*1*). In brief, 60 μg of nuclear matrix preparations in 10 μL was added to 90 μL of an assay solution prepared to yield final concentrations of 50 mM NaCl, 10 mM Tris-HCl (pH 7.8), 2 mM EDTA, 0.25 M sucrose, 0.25 mg of BSA/mL, 20 ng of ^{32}P-end labeled DNA fragments/mL, and sheared,

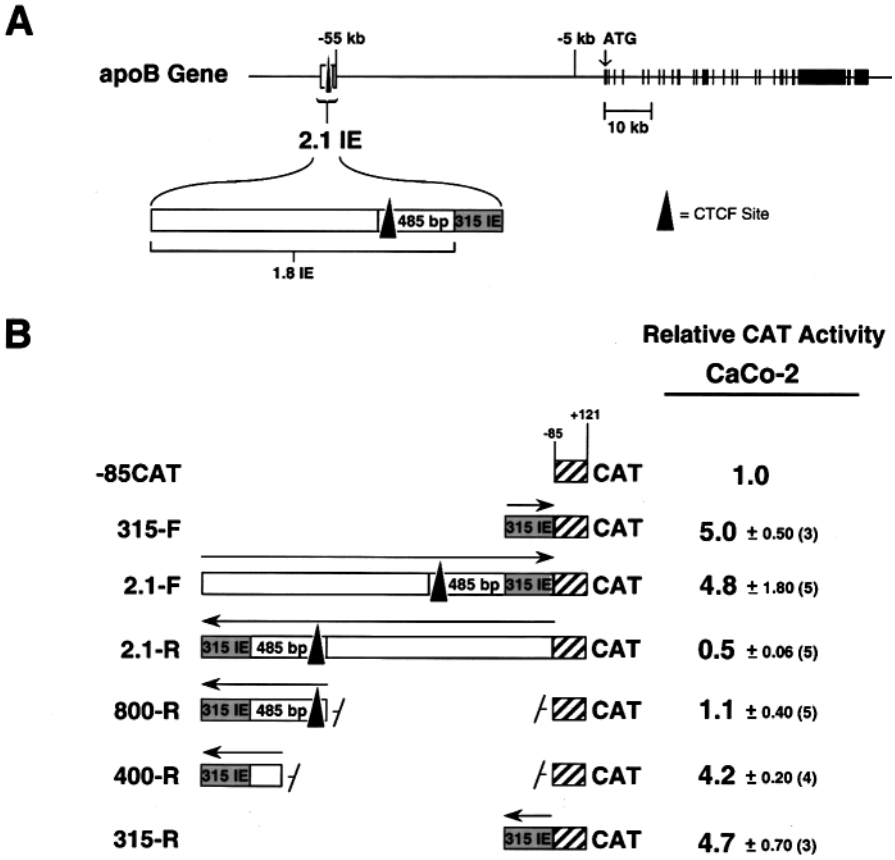

FIGURE 1: The 1.8 kb segment immediately upstream of the 315 bp apoB intestinal enhancer blocks enhancer activity when placed between the enhancer and the promoter. Panel A shows a diagram of the structure of the human apoB gene, with the exons shown as black boxes. The position of the ATG codon is also shown. A scale in kilobases is shown below the gene. At −55 kb is the intestinal control region. A 2.1 kb segment (2.1 IE) from the ICR is depicted, with its two components, namely, the 1.8 IE and the 315 IE clearly marked. A 485 bp segment within the 1.8 IE is also shown; the black triangle within it represents the CTCF binding site. The left side of panel B shows a scheme of the CAT reporter constructs used in the transfections, and the CAT activities are shown on the right ± the standard deviations. The number in parentheses represents the number of independent transfections performed. The locations of the apoB promoter and enhancer and the sizes of the fragments are indicated for each construct. The orientations of the segments of the ICR are indicated by arrows above each segment.

denatured *Escherichia coli* DNA (50−1000 μg/mL) as the unlabeled competitor. Under our experimental conditions, the ratio of labeled DNA fragments to matrices was such that 20−50% of the input DNA bound specifically to them. Incubation was for 1−2 h on a shaker at room temperature (23 °C) followed by the addition of 500 μL of assay buffer without DNA and centrifugation at 10000*g* for 1 min at 4 °C to recover the matrices. The pellets were washed in 1 mL of assay buffer without carrier DNA, followed by solubilization of matrix-bound DNA in 0.5% SDS, overnight treatment with 0.4 mg of proteinase K/mL, phenol extraction, and ethanol precipitation after addition of 10 μg of unlabeled carrier DNA. The purified matrix-bound DNA fragments were resolved on 1% agarose gels [40 mM Tris-acetate and 2 mM EDTA (pH 7.8)], and the dried gels were autoradiographed.

DNaseI Sensitivity Assays. DNaseI sensitivity studies with nuclei from CaCo-2 and HepG2 cells were performed as previously described (*15*). Following digestion with increasing amounts of DNaseI, DNA samples were purified by phenol extraction and ethanol precipitation. Five micrograms of DNA from each sample was vacuum-loaded onto a nylon membrane using the dot blot apparatus. The membrane was UV cross-linked for 1 min using the Stratalinker UV cross-linker (Stratagene, Inc.) and then prehybridized for 3 h at 42 °C in 5 mL of Ultrahybe buffer (Ambion, Inc.). Each of the five probe fragments was used as a template in a random-prime labeling reaction using [α-^{32}P]dCTP and Klenow DNA polymerase. Approximately 2×10^7 cpm of each probe was added separately to each membrane in Ultrahybe buffer and was allowed to hybridize overnight at 42 °C. The membranes were washed briefly in a solution of 2× SSC and 0.1% SDS at room temperature and then autoradiographed and imaged using a PhosphorImager. Hybridization signals were quantitated using ImageQuant software (Molecular Dynamics, Inc.).

RESULTS

Detection of an Insulator Activity within a 1.8 kb Segment Located Immediately Upstream of the Human ApoB 315 bp Intestinal Enhancer. A linear map representing the structure of the human apoB gene is shown in Figure 1A. The structural gene extends over 43 kb and contains 29 exons and 28 introns. The intestinal control region (ICR) is localized more than 54 kb upstream of the transcriptional start site and contains three separate elements capable of driving expression of apoB transgenes in the intestines of mice (see ref *18*).

A 2.1 kb DNA segment (2.1-F) extending from −58 to −56 kb of the human apoB gene, representing the 5′ portion

of the ICR, and containing the 485 IE and the 315 IE (see ref *18*) exhibits a 5-fold enhancer effect upon the apoB promoter in intestine-derived CaCo-2 cells (Figure 1B). Most of the enhancer activity of 2.1-F can be recovered in the 315 bp IE fragment, situated at the 3′ end of the 2.1 kb fragment. The fact that no cooperative or additive effects by the 485 IE and the 315 IE are observed when they are part of the larger 2.1-F segment suggests that they may be in competition for a limiting pool of coactivators. Enhancer activity of the 315 bp segment is virtually identical whether in the forward (F) or reverse (R) orientation (constructs 315 IE-F and -R). However, when the 2.1 kb fragment harboring the 315 bp enhancer is placed in the reverse orientation upstream of the promoter (2.1-R), a 10-fold decrease in transcriptional activity is observed (Figure 1B).

The position dependence of this large decrease in transcriptional activity observed with construct 2.1-R suggests that the activity of the 315 bp enhancer may be blocked by an insulator present within the 1.8 kb fragment. This explanation is consistent with the observation that the enhancer activity is only blocked when the 1.8 kb fragment is interposed between the promoter and the enhancer as is the case in construct 2.1-R. When the activity of the enhancer is completely blocked, the value of CAT activity should be 1.0 or the value obtained with the promoter alone. The fact that the CAT activity of the 2.1-R construct is 0.5 means that, in addition to the enhancer-blocking effect, we are detecting the effect of a negative regulatory element present within the 1.8 kb segment. To further localize the enhancer-blocking elements, construct 800-R was made. It contains the 315 IE in addition to the 3′-most 485 bp of the 1.8 kb fragment (485 IE) (Figure 1B). Transfection experiments revealed that the insulator activity was retained in construct 800-R. However, when 400 bp from the 5′ end of the 485 bp segment are removed, as in construct 400-R, the enhancer-blocking activity is obliterated. Therefore, we conclude that the enhancer-blocking activity of the insulator resides in the 5′-most 400 bp of the 485 IE segment, which is situated immediately upstream of the 315 bp IE.

CTCF Binds to the Segment of the Human ApoB Gene That Exhibits the Enhancer-Blocking Activity. Analysis of the DNA sequence of the 485 bp fragment revealed sequence similarities with the binding site for a transcription factor with known insulator function named CCCTC-binding factor (CTCF) (*22*). In Figure 2A, we compare the human apoB CTCF site with a chicken β-globin FII CTCF site (*23*), a mycFV site (*10*), an AP-β site (*11*), and a human Igf2-H19 site (*8, 9*). There is DNA sequence similarity among the various CTCF sites. To elucidate whether the apoB CTCF site would be bound by CTCF, gel shift experiments were performed. The left portion of Figure 2B shows a control experiment in which we used an oligonucleotide probe from the chicken β-globin gene that is known to bind CTCF. Two major complexes were obtained with the chicken CTCF oligonucleotide and human CTCF protein derived from an extract from COS cells that was transfected with an expression vector for human CTCF (lane 1). Specificity of binding was determined by competition with an excess of the unlabeled chicken CTCF oligonucleotide (lane 2). The apoB CTCF oligonucleotide effectively competed with the chicken CTCF probe for binding to CTCF (lane 3). Furthermore, the apoB CTCF probe formed the same two complexes with

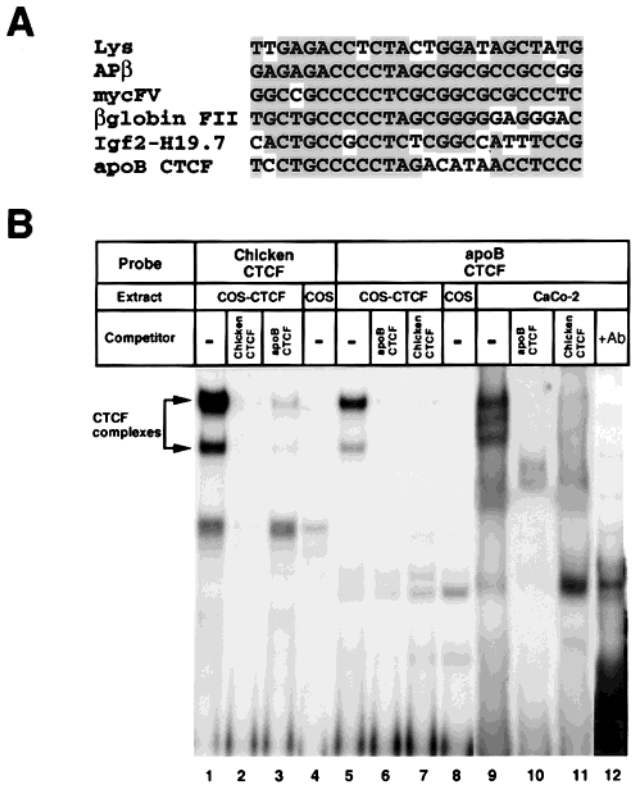

FIGURE 2: A binding site for CTCF 5′ of the human apoB gene intestinal control region. Panel A shows the sequence of CTCF sites from various genes, as well as that of the apoB CTCF site. Nucleotides in common among the various CTCF sites are shaded in gray. Panel B shows the results of gel shift analysis using as probes either the chicken CTCF or the apoB CTCF oligonucleotides.

CTCF that were observed with the chicken CTCF probe (lane 5), and these complexes were equally challenged by an excess of chicken CTCF (lane 7) or apoB CTCF (lane 6) oligonucleotides used as competitors, but not by an excess of PBR322 DNA used as a nonspecific competitor (data not shown). The chicken and human CTCF complexes were not detected with a COS cell extract that had been transfected with an empty plasmid (lanes 4 and 8). Additionally, to verify the presence of CTCF in intestinal cells, the apoB probe was challenged with a nuclear extract from CaCo-2 cells. Complexes similar to those observed with the COS CTCF extract were evident (lanes 9−11). When antibodies against CTCF were added to the binding reaction, they interfered with binding and abrogated the formation of the CTCF−DNA complex (lane 12). Therefore, the data in Figure 2 demonstrate that a CTCF binding site is present within a 485 bp *Eco*RI segment from the apoB gene intestinal control region that exhibits enhancer-blocking capabilities (see construct 800-R in Figure 1B).

The CTCF Binding Site Alone Is Responsible for the Enhancer-Blocking Activity of the 485 bp EcoRI Segment. To determine whether binding of CTCF to its apoB site leads to the blocking of the activity of the 315 IE, two additional constructs were made. In the first construct, 315F/CTCF1/CAT, one copy of the double-stranded CTCF oligonucleotide was inserted between the apoB promoter and the intestinal enhancer. The second construct, 315F/CTCF6/CAT, incorporated six copies of the CTCF site between the enhancer and the promoter. The results of transfection experiments with these constructs are illustrated in Figure 3A. One copy

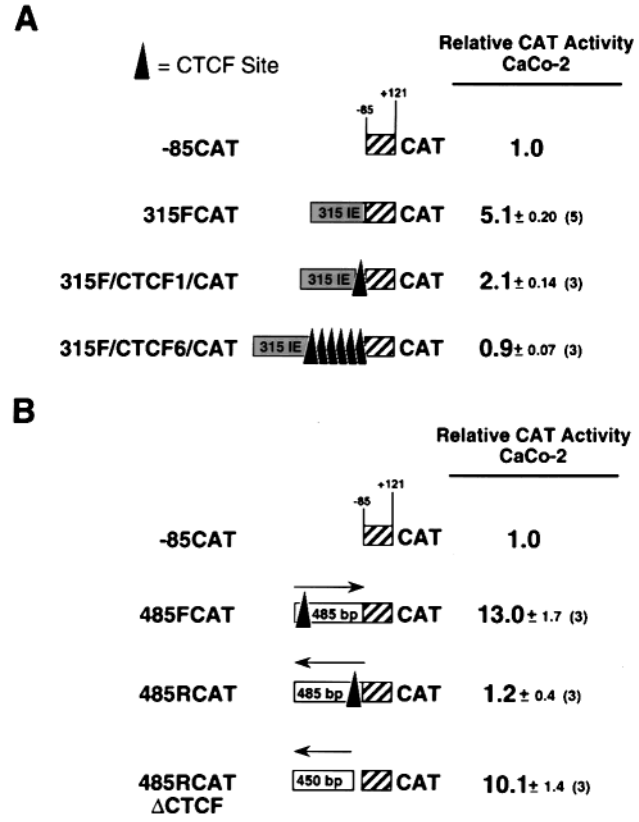

FIGURE 3: The apoB CTCF site blocks the activity of the 315 bp IE when placed between this enhancer and the promoter. In both panels, the CAT constructs, with their names, are shown on the left and the transfection results are shown on the right. The layout is similar to that of Figure 1. The CTCF site is shown as a triangle. The arrows in panel B reflect the orientation of those segments.

of the CTCF site was sufficient to block the activity of the enhancer by 70%, and six copies of the CTCF site completely obliterated enhancer activity. Therefore, we conclude that CTCF is responsible for the enhancer-blocking activity previously described within the larger fragment. Since this activity of CTCF has been classified as an insulator activity, we conclude that the CTCF site within the human apoB gene indeed reflects the presence of an insulator.

To further strengthen our argument that the apoB CTCF site plays an insulator role, additional experiments were performed. In Figure 3B, we demonstrate that the CTCF site, situated at the 5′ end of the 485 bp IE, also blocks the activity of a second intestinal enhancer present in the 3′ portion of the 485 bp IE. The 485 IE stimulates transcription from the apoB promoter in CaCo-2 cells by 13-fold, when placed in the forward orientation upstream of the promoter, as shown with constructs −85CAT and 485FCAT (Figure 3B). Notice that in 485FCAT, the CTCF site is located at the 5′ end of the 485 bp segment. However, when the orientation of the 485 IE is reversed, as in construct 485RCAT, the strong enhancer effect is blocked. In 485RCAT, the CTCF site is located between the enhancer and the promoter and thus obliterates enhancer activity.

Additional proof of the role of the CTCF site in enhancer blocking is demonstrated in experiments with construct 485RCATΔCTCF. In this construct, a 35 bp segment containing the CTCF site has been deleted. Transfections of this construct uncovered the strong enhancer, present within the 485 IE, thus confirming that the enhancer activity per se of the 485 IE is orientation-independent, as is the case for the enhancer within the 315 IE (Figure 1B) and many other classical enhancers. More importantly, deletion of the CTCF site within the 485 IE abolished the enhancer-blocking effect, again providing strong evidence that the CTCF site plays a key function, by blocking the activity of enhancers, thus insulating the promoter.

The 1.8 kb Fragment Flanking the Human ApoB Gene Intestinal Control Region Functions as an Insulator Element in Drosophila. The second assay used to test for insulator function of the DNA segment upstream of the 315 bp enhancer was the *mini-white* position effect assay in *Drosophila melanogaster* (*5*). In this assay, we tested the ability of the 1.8 kb IE fragment to shield *mini-white* reporters from chromosomal position effects in vivo. The *mini-white* gene contains a minimal promoter that results in low levels of *white* gene expression. However, expression of *mini-white* is sensitive to chromosomal position effects, which can either enhance or repress *white* gene expression. Thus, when integrated into *white*$^-$ flies, transgenic lines show a range of eye color phenotypes that range from pale yellow to red (*5*, *24*). Insulator elements protect *mini-white* from these effects, resulting in collections of transgenic flies whose eye color phenotypes are similar.

The *P*-element transformation vectors used in this study are shown in Figure 4. Three different constructs were used: a negative control with the *scs*′ downstream of *mini-white* but no insulator upstream (mwS′), a positive control for insulator function in which the 3′ apoB MAR was upstream of *mini-white* and *scs*′ was downstream (apoB3′MARmwS′), and the test construct, apoB1.8IEmwS′, with one copy of the 1.8 kb IE segment upstream and *scs*′ downstream. As shown previously (*21*), transgenic lines containing mwS′ displayed a range of eye color phenotypes (Figure 4A), whereas lines containing apoB3′MARmwS′ had predominantly yellow or light orange eyes, indicating that transgene expression was shielded from position effects in these lines (Figure 4B).

To determine whether the 1.8 kb apoB fragment harboring the CTCF site could insulate transgene expression from position effects in *Drosophila*, w^{-1118} flies were transformed with a *P*-element containing *mini-white* flanked by the CTCF-containing 1.8 kb *Eco*RI fragment and *scs*′ (apoB1.8IEmwS′) (Figure 4C). Twenty-four independent, single-copy transformants were obtained. Eighteen of these lines (75%) had yellow or light orange eyes (Figure 4C). Three lines had light yellow eyes, and three had orange eyes; however, no transgenic flies with red eyes were obtained. This distribution was similar to that of apoB3′MARmwS′, in which 20 of 24 lines (83%) had yellow or light orange eyes (Figure 4B), and it was markedly different from flies containing mwS′, whose phenotypes ranged from pale yellow to red (Figure 4A). These results suggest that the 1.8 kb element has insulator activity in *Drosophila*, but that the activity is lower than that exhibited by the apoB 3′ MAR.

A MAR Resides 5′ of the Intestinal Control Region of the Human ApoB Gene. The insulator described so far represents a functional boundary for the human apoB gene in intestinal cells, because it exhibits enhancer-blocking capabilities and is capable of protecting transgenes against position effects in flies. The question arises as to whether there is an additonal

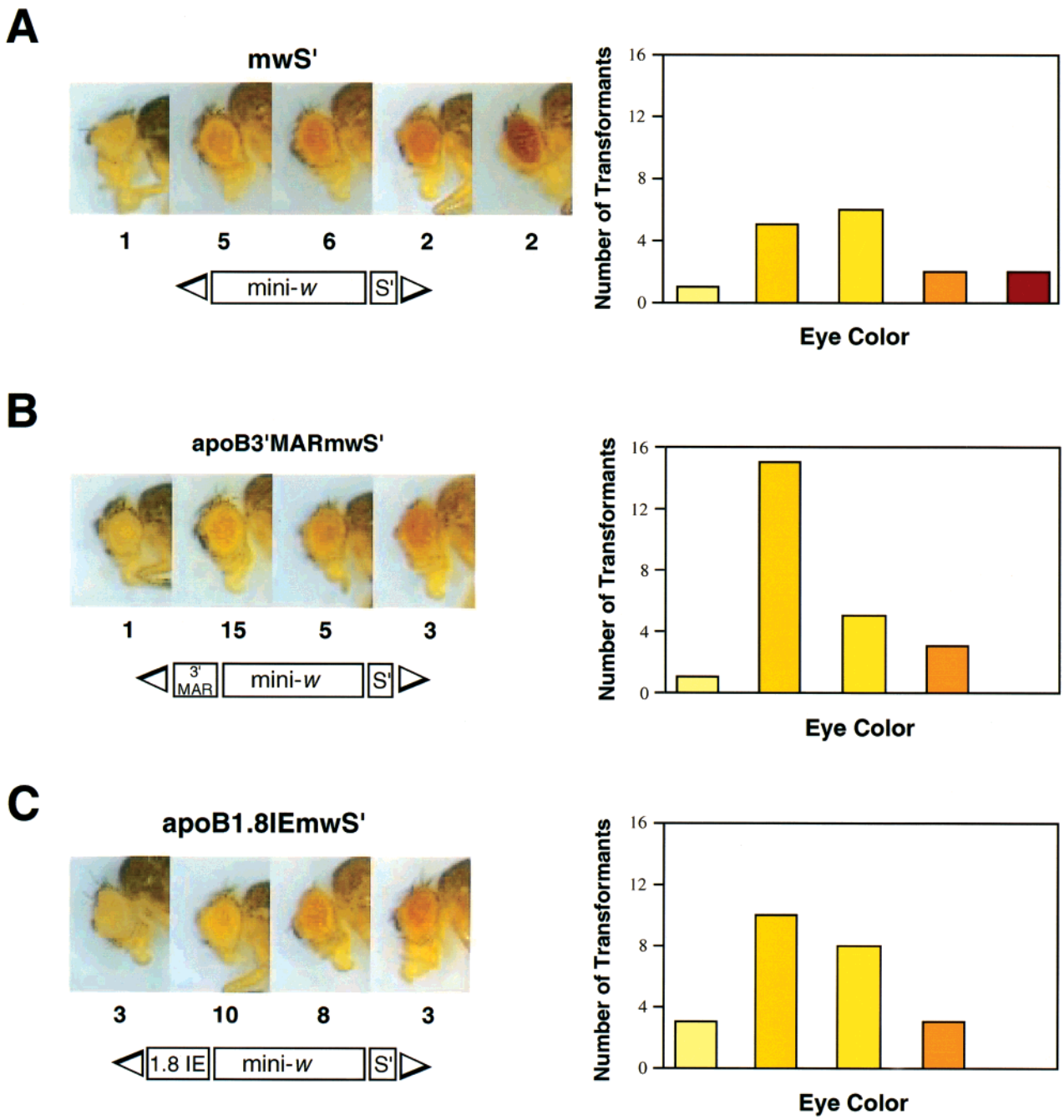

FIGURE 4: Distribution of eye color phenotypes of w^{-1118} *Drosophila* transformants expressing *mini-white*. Photographs of 4-day-old heterozygous females containing one copy of the *P*-element transformation vector. (A) *P*-element containing *mini-white* placed adjacent to the *Drosophila* specialized chromatin structure (*scs*′). (B) *P*-element containing *mini-white* flanked by the apo B 3′ MAR at the 5′ end and *scs*′ at the 3′ end. (C) *P*-element containing *mini-white* flanked by a 1.8 kb DNA fragment containing a CTCF binding site and *scs*′ at the 3′ end. The distribution of eye color phenotypes is presented in bar graph form.

structural boundary upstream of the intestinal control region, such as a MAR, as is the case at the 3′ boundary of the gene (*15*). Because a MAR can represent an anchorage site for a chromosomal loop (*15*) and, therefore, provide a physical or structural boundary for a gene locus, we asked whether there were any MARs upstream of the intestinal control region. To this end, we first determined the DNA sequence of a 6 kb region situated immediately upstream of the ICR, because in our model this would be the most probable location of the MAR. We then searched for potential MARs with the aid of a computer program called MAR finder, using default parameters. This algorithm focuses on AT-rich sequences that harbor topoisomerase II sites, and other classic MAR motifs (*25*). When a segment from the human apoB gene encompassing the 6 kb region 5′ of the 315 bp IE was analyzed in this manner, the MAR finder program predicted a MAR in a 1.3 kb *Hpa*I−*Nde*I region designated fragment B in Figure 5A.

To determine experimentally whether there was a MAR in fragment B or in its vicinity, the protocol of Cockerill and Garrard (*1*) was used. In this protocol, nuclear matrix preparations from CaCo-2 cells are incubated with one or more end-labeled DNA fragments and allowed to bind either in the presence or in the absence of a large excess of *E. coli*

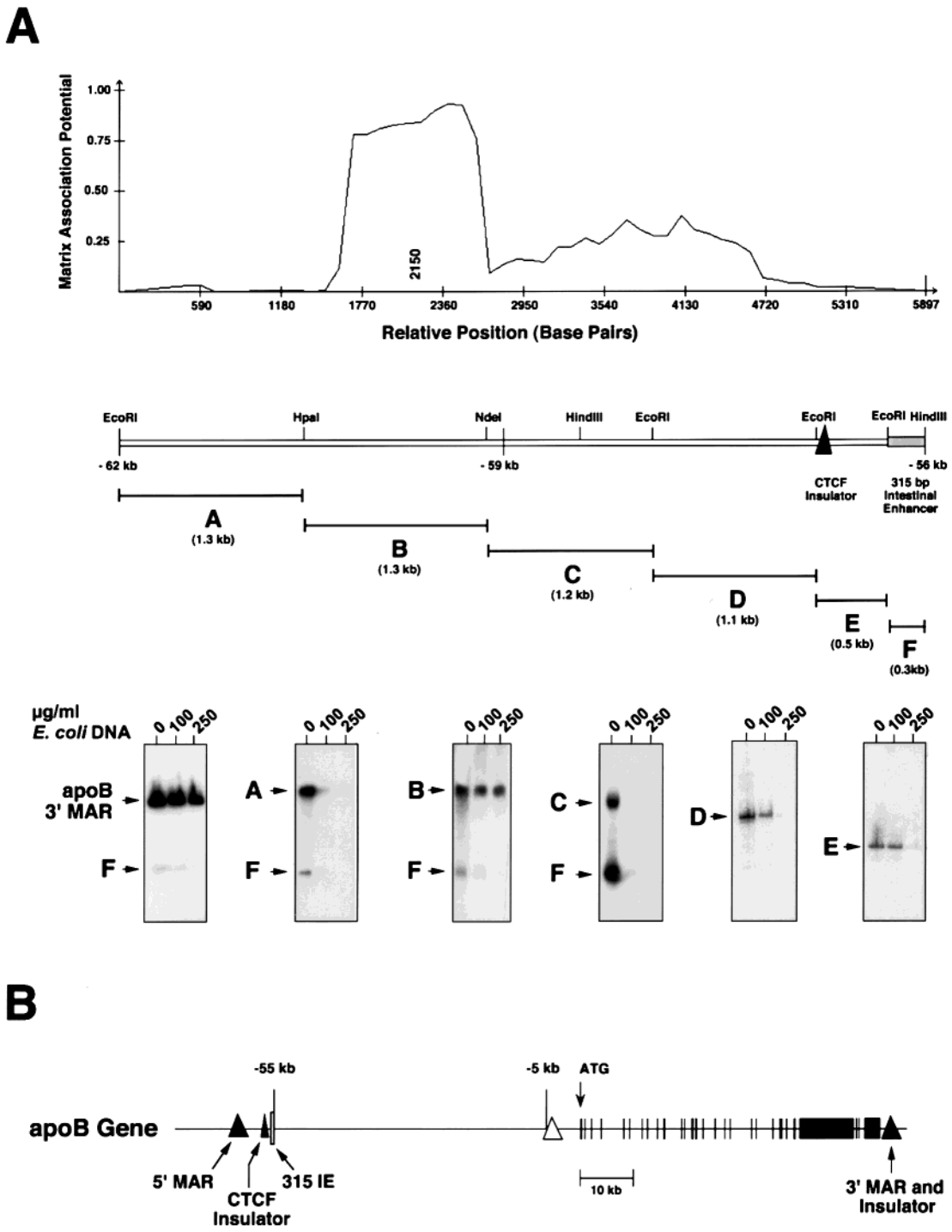

FIGURE 5: A MAR upstream of the human apoB 5′ insulator. Panel A shows a graph representing a profile of the potential MAR regions within the 6 kb segment extending from −56 to −62 kb of the human apoB gene whose map is shown below the graph. The locations of the CTCF binding site and of the 315 IE are clearly indicated in the map. Segments A−F below the map represent the fragments tested for MAR binding. Autoradiograms from the MAR binding studies with segments A−F and with the apoB 3′ MAR are shown. The quantities of *E. coli* DNA competitor are indicated at the top of the autoradiograms. Panel B shows a linear representation of the apoB gene locus in intestinal cells, indicating the positions of the 5′ and 3′ MARs and insulators in relation to the apoB structural gene. The white triangle at −5 kb reflects a 5′ distal MAR previously reported (*15*).

DNA included as a nonspecific competitor to eliminate nonspecific DNA−matrix interactions. After centrifugation to remove the unbound probe, the matrix-bound DNA fragments are separated on agarose gels and visualized by autoradiography.

Such an analysis was performed with fragments A−F, spanning the entire 6 kb region of interest (Figure 5A). As a positive control, we used the strong 3′ MAR from the human apoB gene, shown on the left side. It bound strongly to the matrix, as previously demonstrated (*15*), even in the presence of 250 µg/mL *E. coli* DNA competitor. In the same experiment, we included probe F, the 315 bp IE. In the absence of *E. coli* DNA competitor, probe F bound very weakly to the matrix, and this weak binding was eliminated with 100 µg of *E. coli* DNA. Thus, the 315 bp IE does not bind to CaCo-2 matrix proteins, and therefore, it provides

us with a negative control.

Next, we examined binding of fragment B to CaCo-2 matrix preparations. As shown in Figure 5, fragment B did bind to the matrix, and this binding was not affected by a large excess (250 μg/mL) of *E. coli* DNA. This result demonstrates that a MAR resides in fragment B, in good agreement with the prediction by MAR finder. Matrix binding activity was also examined in segments A and C−E, localized between the MAR and the 315 bp IE. As seen in Figure 5A, segment A did not bind specifically to the matrix, whereas segments D and E exhibited some weak binding. From our combined data in Figure 5, we conclude that a MAR is indeed present in the segment extending from −59 to −60.5 kb 5′ of the human apoB gene. Figure 5B illustrates the locations of the 5′ MAR and insulator described above, as well as the location of the 3′ MAR (*15*) and insulator (*21*) previously described. These elements represent the structural and functional boundaries of the human apoB gene in intestinal cells.

A Transition from a Closed to an Open Chromatin Structure in the Vicinity of the ApoB CTCF Site. Insulators also separate transcriptionally active and transcriptionally inactive chromatin regions (*6*). A transcriptionally active segment is characterized by an open chromatin region that exhibits an enhanced sensitivity to DNaseI. In contrast, a closed or transcriptionally inactive chromatin region has a reduced sensitivity to DNaseI. Previous work in our laboratory had established a broad DNaseI hypersensitive (DH) site within the 315 bp IE in CaCo-2 cells (*20*). More recent work in our laboratory has revealed a second DH site 3′ of the CTCF site within a 485 bp *Eco*RI segment, and there are no DH sites 5′ of the CTCF site in transcriptionally active CaCo-2 cells (see ref *18*). On the other hand, two additional DH sites are found about 1.5 kb 3′ of the 315 bp IE, indicating the presence of additional intestinal elements in that region (*18*).

To expand this work, the DNaseI sensitivity of the regions flanking the CTCF site were examined. This assay differs from the hypersensitivity assay in that a larger amount of digestion of the DNA is sought and the quantitation is based on the amount of DNaseI required to destroy 50% of the fragment of interest to fragments too small to be retained by the nylon membranes. The hypersensitive assay, on the other hand, focuses on the first cut made by the smallest amount of DNaseI, generally reflecting a transcription factor binding site, due to an opening within a regulatory region. DNaseI sensitive regions can include both regulatory and coding regions of the gene. Shown in Figure 6A is a map of the area under study, with the segments studied indicated below the map. Nuclei from intestine-derived CaCo-2 cells and liver-derived HepG2 cells were digested with various quantities of DNaseI, followed by purification of the resulting DNA fragments and analysis using the dot blot method. The results are illustrated in Figure 6B. Several comparisons were made. First, the DNaseI sensitivity among the various fragments in CaCo-2 cells was evaluated, and these data are depicted in Figure 6C. The 315 IE exhibited the highest sensitivity to DNaseI, followed by the 485 bp segment containing the CTCF site. The DNaseI sensitivity continued to decrease in the 1.1 and 1.2 kb *Eco*RI segments and was the lowest in the 5′-most 1.3 kb *Hpa*−*Nde* segment harboring the MAR.

The second comparison was of the DNaseI sensitivity of each of these fragments between CaCo-2 and HepG2 cells. The intestinal control region does not influence liver expression of the apoB gene and is located outside of the functional hepatic domain for this gene, whose 5′ boundary is at −5 kb (*16*). Therefore, one would predict that the region analyzed here would not display a high DNaseI sensitivity in HepG2 cells. That this is the case is shown in panels B and D of Figure 6. Each of the fragments that was studied exhibited a lower sensitivity to DNaseI in HepG2 than in CaCo-2 cells, and the differences were most evident with the 315 IE and the 485 IE, which were most sensitive in CaCo-2 cells. Most significant are the large differences seen with the 315 IE in CaCo-2 and HepG2 cells and the similarities observed with the 5′ MAR-containing fragment. The 5′ MAR-containing fragment exhibits a low DNaseI sensitivity in both CaCo-2 and HepG2 cells, consistent with its location in a transcriptionally inactive segment of the gene. In intestine-derived CaCo-2 cells, the boundary between the open, DNaseI-sensitive region and the closed, DNaseI-resistant region is located at or immediately 5′ of the CTCF site, supporting the view that the CTCF site represents a boundary insulator. Figure 6E summarizes the transition from DNaseI resistance to DNaseI sensitivity observed at the 5′ boundary of the apoB gene in CaCo-2 cells.

DISCUSSION

Insulators are characterized by two main functional properties. The first is their ability to block enhancer (or repressor) activity upon a promoter. The second characteristic of insulators is their ability to protect transgenes from position effects at their sites of integration. The *gypsy* insulator as well as the *scs* and *scs*′ elements flanking the hsp 70 genes at the 87A7 locus in *Drosophila* exhibit both of these properties (*26*−*28*).

In this study, we demonstrate that a 1.8 kb segment from the human apoB gene situated immediately upstream of the intestine-specific regulatory region of this gene exhibits the two main functional properties of insulators. This insulator blocks the activity of the 315 bp intestinal enhancer (315 IE) in transiently transfected CaCo-2 cells and protects *Drosophila mini-white* genes from position effects in transgenic flies. The enhancer-blocking activity was localized to an 85 bp segment within the 485 IE, situated upstream of the 315 IE (Figure 1B). CTCF bound to a site within the 85 bp segment (Figure 2). One copy of the CTCF site placed between the 315 IE and the promoter is sufficient to block the activity of the intestinal enhancer by 70%, and when six copies of the CTCF site are present, the enhancer is completely blocked (Figure 3A). Furthermore, deletion of the CTCF site (construct 485RCATΔCTCF) abolished the enhancer-blocking activity (Figure 3B). The 5′ apoB insulator is very similar to the chicken β-globin insulator in that it also functions via a CTCF site, in enhancer blocking assays in transient transfections with embryonic erythrocytes (*29*). Furthermore, like the chicken β-globin insulator (*30*), the apoB 5′ insulator also protects the *mini-white* gene against position effects in transgenic *Drosophila* (Figure 4). The differences in insulating strength observed between the apoB 3′ and 5′ boundary insulators may be due to the fact that at the 3′ end of the gene, the properties of the 3′ insulator and 3′ boundary of the looped domain in which the apoB gene

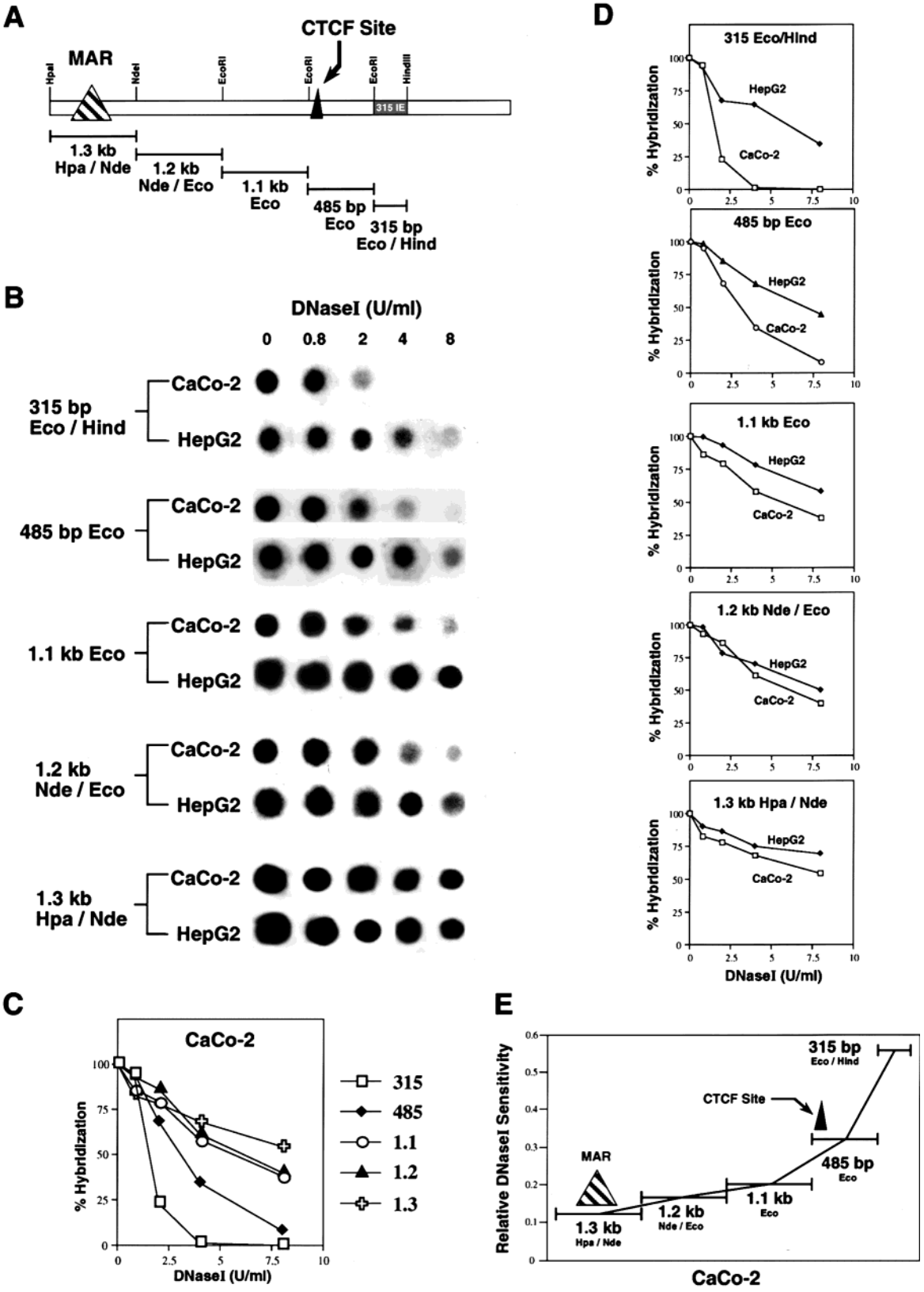

FIGURE 6: The apoB CTCF site marks a chromatin structure transition zone. Panel A shows a restriction map of the region of interest, with the location of the MAR shown as a hatched triangle, the CTCF site shown as a black triangle, and the 315 IE shown as a gray box. Relevant restriction sites are indicated which were used to generate the five fragments used as probes in the DNaseI sensitivity analysis. Autoradiographs of the dot blots are shown in panel B. The amounts of DNaseI used in the analyses are indicated above the autoradiographs. The probe used in the hybridization and the source of DNA, whether from CaCo-2 or HepG2 cells, are shown to the left of each autoradiograph. The amount of DNA detected by each probe was quantitated using PhosphorImage analysis and expressed as a percentage of the amount of DNA detected in the first position of each blot for each probe. These percentages as a function of DNaseI concentration for the CaCo-2 samples are graphically depicted in panel C for each of the five probes. The comparisons of the CaCo-2 and HepG2 DNaseI sensitivity profiles for each individual probe are shown in panel D. Panel E illustrates the transition from DNase resistance to DNase sensitivity in the 5′ boundary of the apoB gene in CaCo-2 cells.

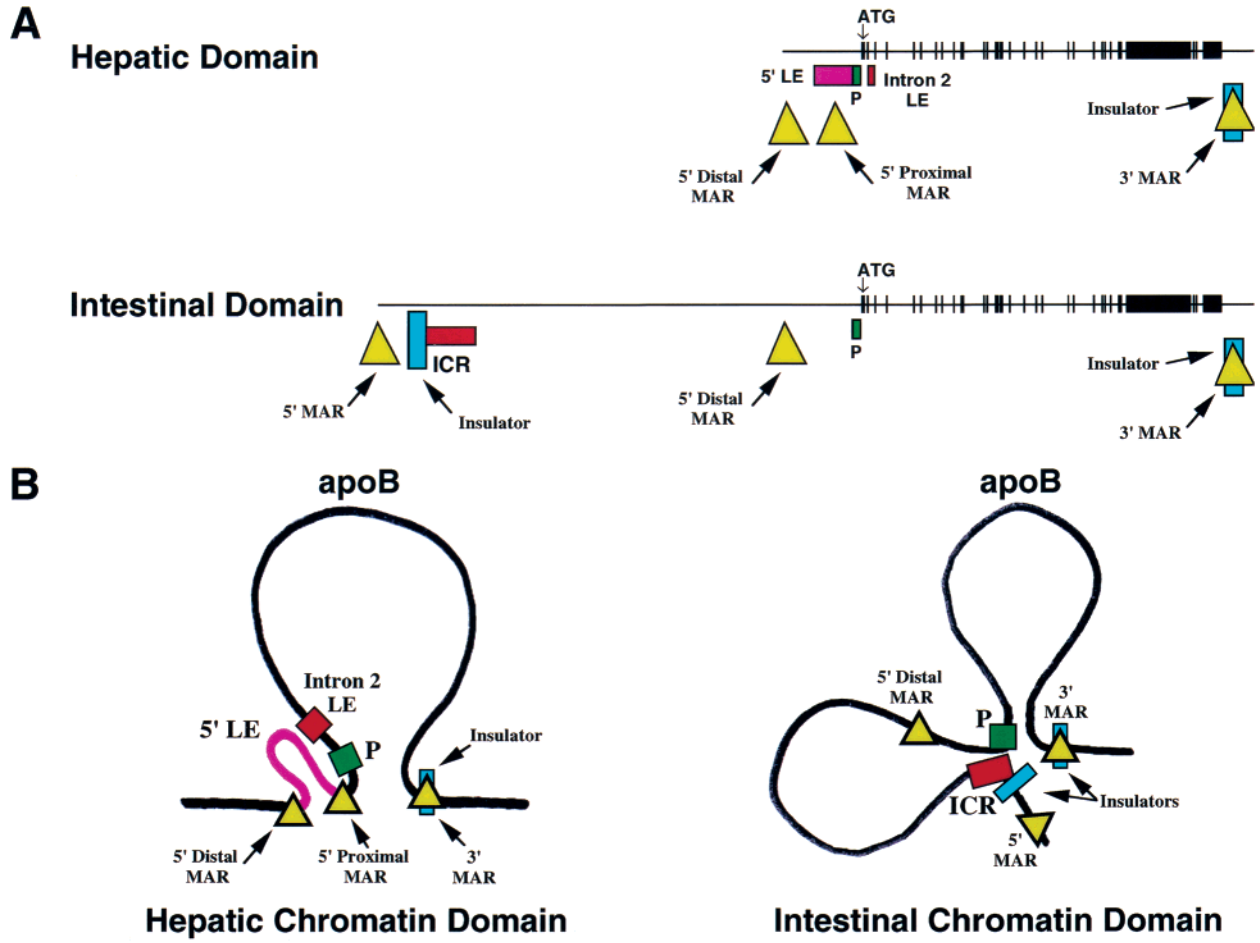

FIGURE 7: Chromatin loop model for the apoB gene in hepatic and intestinal cells. Panel A shows linear maps of the apoB gene locus with the locations of regulatory elements required for its expression in either the liver or intestine in transgenic mice (*16*−*18*, *20*). The black horizontal line represents the DNA with exons shown as black boxes. The location of the first ATG is also shown with an arrow. The positions of the promoter (P, green box), the second intron enhancer (Intron 2 LE) or the intestinal control region (ICR) (red boxes), the 5′ liver enhancer (5′ LE, pink box), and the insulators (blue rectangles) are indicated. The matrix association regions appear as gold triangles. Chromatin loop models for either the hepatic domain (left) or intestinal domain (right) are depicted in panel B. The apoB DNA is shown as a solid black line with the coding region positioned in the vertical loop and labeled *apoB* in both models. The shape and color designations for the various elements are identical to those in panel A.

resides (*15*) coexist within the 3′ MAR (*21*) (Figure 7A), whereas at the 5′ end of the apoB gene intestinal domain, the insulating activity provided by the 1.8 kb fragment is reinforced by the existence of a MAR ∼2 kb upstream of the insulator (Figure 7A). This MAR probably represents the 5′ boundary of the second chromosomal loop that brings the intestinal control region closer to the promoter; therefore, the looped domain itself may serve to insulate the gene contained within it (Figure 7B). The relative strength of an insulator may also vary depending on the relative proximity of regulatory elements from neighboring genes. Thus, one might expect that the further the distance to neighboring regulatory influences, the weaker the insulator that is required.

Recently, Dillon and Sabbattini (*31*) have discussed functional domains of eukaryotic gene expression in terms of two models: the strong and weak domain models. In strong domain models, a region of open or accessible chromatin containing the gene and associated regulatory elements is spatially isolated by insulators, which have evolved to act as functional boundaries. On the other hand, in weak domains, the structure of the domain is determined by the distribution of binding sites for transcriptional activators, without a requirement for functional boundaries.

On the basis of our current and previous studies with the apoB gene locus, we suggest that, in intestinal cells, the locus is found in a strong chromatin domain, flanked by insulators at both the 5′ and 3′ ends (Figure 7A). The MARs present at each end may represent the physical boundaries of the strong domain by providing anchorage points for chromosomal loops (Figure 7B). In hepatic cells, all the regulatory elements required for expression in transgenic mice are positioned relatively close to the structural gene and no insulator has been detected at the 5′ end of the apoB liver domain. Therefore, in hepatic cells, the apoB gene locus may reside in a weak chromatin domain, particularly since there appear to be no open reading frames as far as 80 kb 5′ of the apoB transcriptional start site. A physical boundary for the 5′ end of the domain may be provided by the 5′ distal MAR.

Other mammalian insulators with enhancer-blocking capabilities have been described. One of them, BEAD-1, is a 1.6- b element from the human T cell receptor α/δ locus (*7*) that also harbors a CTCF site. Its function is to ensure correct

developmental expression of the TCR α and δ loci by blocking E δ from opening the chromatin of the adjacent TCR α gene early in T cell development. The protein that binds to the *Drosophila scs*′ boundary element of the 87A7 hsp70 locus has been identified and named BEAF. It binds with high affinity to clustered, variably arranged CGATA motifs. These DNA sequence elements confer position-independent expression upon the *mini-white* reporter gene in transgenic flies (*24*). On the other hand, these motifs do not function in enhancer-blocking assays with cultured D1 cells, but they are effective in plasmid-based microinjection assays with *Xenopus laevis* oocytes (*32*).

Our DNaseI sensitivity studies suggest that the CTCF site may separate two structurally different chromatin regions; 5′ of the CTCF site, there is a region refractory to DNaseI digestion characteristic of nontranscribed or inactive chromatin (Figure 6), whereas 3′ of the CTCF site, a more open chromatin structure is observed, exhibiting DNaseI sensitivity throughout the 485 bp segment as well as the 315 bp IE and beyond. Several DH sites are also present in this region (*18*, *20*). This is a third characteristic shared by the chicken β-globin insulator and the apoB 5′ insulator. In the chicken β-globin locus, the boundaries are also identified by transitions from an open, DNaseI-sensitive chromatin structure to an inactive, DNaseI-resistant chromatin structure (*4*).

Insulator function has also been attributed to nuclear matrix attachment sites (MARs). For example, it has been demonstrated that the 3′ MAR of the human apoB gene as well as a MAR from the human α1-antitrypsin locus protect *mini-white* transgenes from position effects in *Drosophila* (*21*). Additionally, earlier work demonstrated that the apoB 5′ distal and 3′ MARs not only increase the expression level of a single-copy apoB-lacZ reporter gene in stable HepG2 cell transfectants by 200-fold but also shield it from position effects (*33*). Interestingly, the su(*Hw*)-binding portion of the *gypsy* insulator contains a MAR (*34*). Nevertheless, it is not clear whether MARs per se contain an insulator activity or whether they simply protect transgenes against position effects by providing an anchoring site for a loop, thereby isolating one or more genes within a separate looped domain. At least in the case of the apoB 3′ MAR, it does function like the 1.8 kb fragment in the enhancer blocking assay (data not shown), and it enhances the expression of human apoB constructs that contain this MAR in transgenic mice (*16*). The discovery of a MAR 5′ of the apoB insulator provides an anchorage site for a chromosomal loop whose likely function in vivo is to bring the intestinal control region, localized 53 kb upstream of the transcriptional start site, in close proximity to the promoter. Interestingly, there are no open reading frames between the ICR and the apoB promoter, indicating an absence of other genes in this region. However, some 36 kb 5′ of the MAR resides a *tudor*-like gene. *Tudor* is a *Drosophila* gene that is expressed only from the maternal chromosome, and it is required for normal abdominal segmentation (*35*). Mammalian *tudor*-like proteins have been identified as tumor antigens for colonic tumors (*36*). Therefore, the boundary insulator segments described above may play a physiological role by preventing activation of the expression of the colonic tumor antigen by the apoB intestinal enhancer.

An example of this class of regulation by boundary insulators has recently emerged in the case of the H19 and Igf2 genes, two neighbors that are expressed from either the maternal or the paternal chromosomes (*37*). H19 is maternally imprinted, while Igf2 is paternally imprinted. Both genes have an enhancer in common, situated downstream of H19. The imprinting control region of H19 is a CTCF boundary element whose activity is controlled by DNA methylation. The CTCF binds to the unmethylated, maternal boundary region, thereby preventing the promoter from the Igf2 gene from interacting with the enhancer, resulting in silencing of Igf2. The paternal boundary region is methylated, and CTCF cannot bind to it, thus permitting the enhancer to interact with the promoter of the paternal Igf2 and turn on transcription. On the other hand, the paternal H19 gene is silenced by methylation (*8*, *9*). By analogy to this example, it is possible that methylation of the apoB CTCF boundary element may play a role in modulating the expression of the colonic tumor antigen gene. Should this hypothesis be true, mutations in the CTCF site that promote methylation may increase the probability of developing colonic tumors. This can be experimentally tested by introducing human DNA containing the tumor antigen gene and the apoB IE region into transgenic mice and examining the susceptibility of the resulting mice to develop intestinal tumors.

ACKNOWLEDGMENT

We thank Stephen G. Young for providing BAC subclones containing the apoB intestinal control region, Victor Lobanenkov for the human CTCF expression vector and antibodies, and Allen D. Cooper, Brian J. McCarthy, and Kenneth C. Yu for their comments on the manuscript.

REFERENCES

1. Cockerill, P. N., and Garrard, W. T. (1986) *Cell 44*, 273−282.
2. Mirkovitch, J., Mirault, M.-E., and Laemmli, U. K. (1984) *Cell 39*, 223−232.
3. Prioleau, M. N., Mony, P., Simpson, M., and Felsenfeld, G. (1999) *EMBO J. 18*, 4035−4048.
4. Saitoh, N., Bell, A. C., Recillas-Targa, F., West, A. G., Simpson, M., Pikaart, M., and Felsenfeld, G. (2000) *EMBO J. 19*, 2315−2322.
5. Kellum, R., and Schedl, P. (1991) *Cell 46*, 941−950.
6. Bell, A. C., and Felsenfeld, G. (1999) *Curr. Opin. Genet. Dev. 9*, 191−198.
7. Zhong, X. P., and Krangel, M. S. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 5219−5224.
8. Bell, A. C., and Felsenfeld, G. (2000) *Nature 405*, 482−485.
9. Hark, A. T., Schoenherr, C. J., Katz, D. J., Ingram, R. S., Levorse, J. M., and Tilghman, S. M. (2000) *Nature 405*, 486−489.
10. Filippova, G. N., Fagerlie, S., Klenova, E. M., Myers, C., Dehner, Y., Goodwin, G., Neiman, P. E., Collins, S. J., and Lobanenkov, V. V. (1996) *Mol. Cell. Biol. 16*, 2802−2813.
11. Vostrov, A. A., and Quitschke, W. W. (1997) *J. Biol. Chem. 272*, 33353−33359.
12. Burcin, M., Arnold, R., Lutz, M., Kaiser, B., Runge, D., Lottspeich, F., Filippova, G. N., Lobanenkov, V. V., and Renkawitz, R. (1997) *Mol. Cell. Biol. 17*, 1281−1288.
13. Young, S. G. (1990) *Circulation 82*, 1574−1594.
14. Knott, T. J., Rall, S. C., Jr., Innerarity, T. L., Jacobson, S. F., Urdea, M. S., Levy-Wilson, B., Powell, L. M., Pease, R. J., Eddy, R., Nakai, H., Byers, M., Priestley, L. M., Robertson, E., Rall, L. B., Betsholtz, C., Shows, T. B., Mahley, R. W., and Scott, J. (1985) *Science 230*, 37−43.
15. Levy-Wilson, B., and Fortier, C. (1989) *J. Biol. Chem. 264*, 21196−21204.

16. Brooks, A. R., Nagy, B. P., Taylor, S., Simonet, W. S., Taylor, J. M., and Levy-Wilson, B. (1994) *Mol. Cell. Biol. 14*, 2243−2256.
17. Levy-Wilson, B. (1995) in *Progress in Nucleic Acids Research* (Cohn, W., and Moldave, K., Eds.) pp 161−190, Academic Press, San Diego.
18. Antes, T. J., Goodart, S. A., Chen, W., and Levy-Wilson, B. (2001) *Biochemistry 40*, 6720−6730.
19. Nielsen, L. B., Kahn, D., Duell, T., Weier, H. U. G., Taylor, S., and Young, S. G. (1998) *J. Biol. Chem. 273*, 21800−21807.
20. Antes, T. J., Goodart, S. A., Huynh, C., Sullivan, M., Young, S. G., and Levy-Wilson, B. (2000) *J. Biol. Chem. 275*, 26637−26648.
21. Namciu, S. J., Blochlinger, K. B., and Fournier, R. E. (1998) *Mol. Cell. Biol. 18*, 2382−2391.
22. Lobanenkov, V., Nicolas, R., Adler, V., Paterson, H., Klenova, E., Polotskaja, A., and Goodwin, G. (1990) *Oncogene 5*, 1743−1753.
23. Chung, J. H., Bell, A. C., and Felsenfeld, G. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 575−580.
24. Cuvier, O., Hart, C. M., and Laemmli, U. K. (1998) *Mol. Cell. Biol. 18*, 7478−7486.
25. Singh, G. B., Kramer, J. A., and Krawetz, S. A. (1997) *Nucleic Acids Res. 25*, 1419−1425.
26. Gerasimova, T. I., and Corces, V. G. (1998) *Cell 82*, 511−521.
27. Roseman, R. R., Pirrotta, V., and Geyer, P. K. (1993) *EMBO J. 12*, 435−442.
28. Udvardy, A. (1999) *EMBO J. 18*, 1−8.
29. Recillas-Targa, F., Bell, A. C., and Felsenfeld, G. (1999) *Proc. Natl. Acad. Sci. U.S.A. 96*, 14354−14359.
30. Chung, J. H., Whiteley, M., and Felsenfeld, G. (1993) *Cell 74*, 505−514.
31. Dillon, N., and Sabbattini, P. (2000) *BioEssays 22*, 657−665.
32. Dunaway, M., Hwang, J. Y., Xiong, M., and Yuen, H. L. (1997) *Mol. Cell. Biol. 17*, 182−189.
33. Kalos, M., and Fournier, R. E. K. (1995) *Mol. Cell. Biol. 15*, 198−207.
34. Nabirochkin, S., Ossokina, M., and Heidmann, T. (1998) *J. Biol. Chem. 273*, 2473−2479.
35. Golumbeski, G. S., Bardsley, A., Tax, F., and Boswell, R. E. (1991) *Genes Dev. 5*, 2060−2070.
36. Scanlan, M. J., Chen, Y. T., Williamson, B., Gure, A. O., Stockert, E., Gordan, J. D., Tureci, O., Sahin, U., Pfreundschuh, M., and Old, L. J. (1998) *Int. J. Cancer 76*, 652−658.
37. Reik, W., and Murrell, A. (2000) *Nature 405*, 408−409.

BI0100743