MOLECULAR AND CELLULAR BIOLOGY, Sept. 1992, p. 4170–4185
0270-7306/92/094170-16$02.00/0

Vol. 12, No. 9

Functional Analysis of the Human Adenosine Deaminase Gene Thymic Regulatory Region and Its Ability To Generate Position-Independent Transgene Expression

BRUCE J. ARONOW,* RICHARD N. SILBIGER, MARY R. DUSING, JEFFREY L. STOCK, KAREN L. YAGER, S. STEVEN POTTER, JOHN J. HUTTON, AND DAN A. WIGINTON

Department of Pediatrics, Division of Basic Science Research NRB 2047, University of Cincinnati College of Medicine, Cincinnati Children's Hospital, Cincinnati, Ohio 45229

Received 24 April 1992/Returned for modification 22 June 1992/Accepted 29 June 1992

We previously observed that human ADA gene expression, required for the intrathymic maturation of T cells, is controlled by first-intron sequences. Used as a *cis* activator, the intron generates copy-dependent reporter expression in transgenic thymocytes, and we here dissect its critical determinants. Of six DNase I-hypersensitive sites (HS sites) in the intron, only HS III was a transfection-active classic enhancer in T cells. The enhancer contains a critical core region, ACATGGCAGTTGGTGGTGGAGGGGAACA, that interacts with at least two factors, ADA-NF1 and ADA-NF2. Activity of the core is strongly augmented by adjacent elements contained within a 200-bp domain corresponding to the limits of HS III hypersensitivity. These core-adjacent sequences include consensus matches for recognition by the AP-1, TCF-1α, μE, and Ets transcription factor families. In contrast, considerably more extensive sequences flanking the enhancer domain were required for position-independent and copy-proportional expression in transgenic mouse thymocytes. The additionally required upstream segment encompassed the nonenhancer HS II site. The required downstream segment, composed largely of *Alu*-repetitive DNA, was non-DNase I hypersensitive. Transgenes that lacked either segment were subject to strong positional effects. Among these variably expressing lines, the expression level correlated with the degree of hypersensitivity at HS III. This finding suggests that formation of hypersensitivity is normally facilitated by the flanking segments. These results delineate a complex thymic regulatory region within the intron and indicate that a series of interactions is necessary for the enhancer domain to function consistently within chromatin.

Genetic deficiency of adenosine deaminase (ADA) in humans causes intrathymic arrest of T-cell development and severe combined immunodeficiency disease. ADA is expressed in essentially all tissues, but its specific activity in most is 10- to 1,000-fold less than in cortical thymocytes (3, 15, 51). Several extrathymic sites of high-level expression have also been observed (16, 74), but none correlates with known pathophysiological effects of the disease. Rather, absence of ADA causes a selective failure to develop cortical thymocytes (10, 35), with affected children exhibiting otherwise normal development and physiologic function. ADA may be essential at a specific stage of thymic T-cell development because mature T cells exhibit only 1 to 5% of the ADA activity found in cortical thymocytes (15). Furthermore, strong regulatory elements for cortical thymic expression are contained in the first intron of the human ADA gene (3). Other evidence that ADA deficiency is a disease of cell type specificity is that in several patients, engraftment of lymphoid lineages with histocompatible bone marrow has been curative (40). Taken together, these results suggest that the introduction of a normally regulated ADA gene into the genome of a hematopoietic stem cell will likely cure the disease. To test this hypothesis, we are first attempting to delineate human ADA gene regulatory elements capable of ensuring high-level thymic gene expression. Doing this offers the opportunity to better understand the molecular basis for its in vivo expression pattern, the cell type specificity of the disease, and the identification of segments of the ADA gene that could confer appropriate regulation to a replacement gene.

In contrast to the accumulating body of information demonstrating how individual *cis* elements and *trans* factors interact locally to alter gene activity, there is much less information available to explain how endogenous genes are able to attain high levels of expression. With limited exceptions, most transgenes are not expressed at as high a level as are their endogenous counterparts. Notably, the β-globin gene locus control region (LCR) is able to generate copy-proportional and position-independent transgene expression at the level of the endogenous gene (38, 59), doing so by utilization of *cis*-activating sequences located within a series of DNase I-hypersensitive sites (HS sites) that flank the β-globin-like gene cluster (64). While delimited hypersensitive sites from the globin LCRs appear capable of generating their own active domain (28, 50, 56, 58, 63), this does not appear to be a general case.

Two recently characterized classes of locus boundary elements have been shown to be critical for the attainment of position-independent transgene expression in *Drosophila* cells and differentiated cell lines (22, 46, 61). To function, the elements need to bilaterally flank minigenes that contain enhancers. In the absence of the flanking regions, minigene constructs exhibit position effects. Using a chloramphenicol acetyltransferase (CAT) reporter gene flanked with large segments of the human ADA gene, we have previously established the following lines of evidence suggesting that ADA expression in cortical thymocytes is controlled by the first intron of the gene. First, an array of HS sites is present in the first intron of the complete ADA gene in human

* Corresponding author.

thymus and thymic-derived high-expressing MOLT 4 and CCRF-CEM (CEM) human leukemic cell lines. Second, inclusion of a 12.8-kb fragment of the region in transgene constructions caused high-level reporter expression in mouse cortical thymocytes. Analysis for DNase I hypersensitivity detected HS sites in transgenic thymus identical to those detected in human thymus. Each copy of the transgene in lines with up to 100 copies of the transgene was hypersensitive to DNase I at HS III site. Finally, a 1.3-kb fragment encompassing HS II and III exhibited enhancer properties when combined in a reporter gene construction with the human ADA gene promoter and transfected into MOLT 4 cells.

In this study, in addition to dissecting enhancer elements of the 1.3-kb segment, we have also analyzed the ability of the intronic region to function in transgenic mice. Our results distinguish regulatory elements that are sufficient for enhancer function in transfection from those necessary in transgenic experiments. Combining these results has permitted us to define a thymic regulatory region in an extended segment of the ADA intron. Within the region, *cis*-regulatory information is hierarchically organized. *cis* elements in a localized domain are sufficient to form a cell-type-specific enhancer in transfection experiments, but additional *cis* information is essential to ensure that the region can become functional in transgenic mice. In the absence of the additional segments, the function of the enhancer is subject to strong position effects. An intriguing possibility is that some of the accessory information that flanks the T-cell enhancer is similar to that which allows introns to rather generally increase transgene transcriptional capability (13, 18, 19).

MATERIALS AND METHODS

Materials for transfection analyses. Plasmids were grown in *Escherichia coli* DH5α (Bethesda Research Laboratories) and purified by banding in cesium chloride gradients. Transfection transient expression was achieved by using a modified DEAE-dextran protocol (3, 36) with plasmid preparations less than 4 months old. The previously determined levels of endogenous ADA activity of the human lymphoid cell lines used were 1,650 (MOLT 4 [T cell]), 270 (CCRF-CEM [T cell]), and 14 (Raji [B cell]) nmol/min/mg of protein (3). Cell culture and CAT assays were performed as previously described (3). CAT activities are expressed as the average of two independent transfections performed on the same day, the values of which were typically within 10 to 20% of each other. When the same DNA samples are assayed on different days, absolute values of CAT activity vary up to twofold, but the relative expression of the constructs within a cell line is maintained.

Materials for transgenic analyses. Transgenic mice were made with the fertilized eggs of (C3H × C57BL/6)F_1 hybrid parents as described previously (3), using procedures from Hogan et al. (43). F_0 transgenic pups were identified and bred to other C3H × C57BL/6 mice to obtain nonmosaic pups, which were then analyzed between 4 and 9 weeks of age. Transgene copy numbers were determined by comparison of blot hybridizations of restriction endonuclease-digested DNA isolated from tails with known standards of human DNA and cloned plasmid DNA. For quantitative Southern analyses, autoradiograms were densitometrically scanned, using various exposures to optimize for differing signals. Some blots were also analyzed by surface emission scanning, using the Betascope 603 blot analyzer (Betagen Corp., Waltham, Mass.) and a phosphorimager device (Molecular Dynamics Inc., Mountain View, Calif.). Transgenic tissue CAT assays, protein concentration determinations, mouse dissections, and chromatin DNase I hypersensitivity assays were performed as previously described (3).

Footprinting. DNase I footprinting was based on methods described by Emerson et al. (24). Nuclear extracts utilized for footprinting were prepared by the method of Dignam et al. (21) except that nuclei were first isolated as previously described, using sucrose cushions and polyamine lysis buffers (3). Footprinting of the two strands was performed on DNA fragments isolated from BAL-31 deletion plasmid 3, derived from pADA CAT 4/I1.3 (see Fig. 2). For the upper strand, a 374-bp *Bam*HI-to-*Sac*I fragment was labeled at the *Bam*HI site by using Klenow polymerase; for the lower strand, the 511-bp *Hin*cII-to-*Eco*RI fragment was labeled at the *Eco*RI site, also by using Klenow polymerase.

Gel shifts and methylation interference assays were performed by using whole-tissue thymic extract prepared with thymic nuclear isolation buffer (3) that contained 0.4 M NaCl. Following 30 min of extraction at 4°C, the suspension was centrifuged at 50,000 × *g* for 20 min, dialyzed versus buffer D (21), aliquoted, and stored at −80°C. Radiolabeled fragments were partially methylated and analyzed for factor interactions essentially as described by Chodosh (17) and Baldwin (4), respectively. Protein-bound DNA fragments were eluted from the preparative acrylamide gels, using overnight crush elution into a solution that contained 0.5 M ammonium acetate, 10 mM magnesium acetate, 1 mM EDTA, and 0.1% sodium dodecyl sulfate. The eluate was extracted with phenol-chloroform, and the fragments were precipitated with ethanol and further purified with Elutip-D cartridges (Schleicher & Schuell) prior to piperidine cleavage. The fragment that was used was a 126-bp *Bam*HI-to-*Taq*I fragment isolated from BAL-31 deletion plasmid 3 (see Fig. 2) labeled at the top and bottom strands of the *Bam*HI site (residue 9229 plus the *Bam*HI half-site), using T4 polynucleotide kinase and Klenow DNA polymerase, respectively.

Plasmids. The following plasmids are diagrammed in Fig. 1: pADA CAT 4/12, pADA CAT 4/I1.3, and pADA CAT 4/5 (3); pADA CAT 4/8d1 [contains a 6.536-kb *Bam*HI-to-*Mlu*I fragment placed 3′ to the poly(A) site of pADA CAT 4.0 (3)]; pADA CAT 4/8d2 [contains a 4.2-kb *Sph*I (partial)-to-*Mlu*I fragment placed 3′ to the poly(A) site of pADA CAT 4.0]; pADA CAT 4/8d3 [contains a 2.292-kb *Kpn*I-to-*Mlu*I fragment placed 3′ to the poly(A) site of pADA CAT 4.0]; pADA CAT 4/8d4 [contains a 1.948-kb *Sph*I (complete)-to-*Mlu*I fragment placed 3′ to the poly(A) site of pADA CAT 4.0]; and pADA CAT 4/8d5 (from an 8.5-kb *Bss*HII-to-*Mlu*I fragment with an internally deleted *Kpn*I fragment that removed HS II and HS III [nucleotides 7560 to 10235]).

The following sets of plasmids are diagrammed in Fig. 2.

For 5′-deleted I1.3 plasmids (Fig. 2 constructions 1 to 7), pADA CAT 4/i3.5 (3) was cut at the unique *Sph*I site representing the 5′ end of the I1.3 fragment and digested for various lengths of time with BAL-31. The plasmid was then cut with *Bam*HI, the ends were repaired with T4 polymerase, and the large fragment self-ligated, producing the resulting unidirectional deletions of the I1.3 fragment.

For 3′-deleted I1.3 plasmids (Fig. 2, constructions 8 to 19), pADA CAT 4/i1.3 (3) was partially cut with *Sac*I, and the ends were repaired with T4 polymerase and ligated to *Sal*I linkers to place a unique *Sal*I site at the 3′ terminus of the I1.3 fragment. The resulting plasmid, pADA CAT 4/i1.3Sal, was cut with *Sal*I, purified, and digested for various lengths

of time with BAL-31. The total DNA was cut with *Aat*II, the ends were repaired with T4 polymerase, and the large fragment was self-ligated.

Secondary BAL-31 deletion constructions (Fig. 2, constructions 20 to 37) were made from selected 5′-deleted I1.3 constructs by digesting with *Kpn*I, and then with BAL-31 for various lengths of time, followed by cleavage with *Aat*II, repair, and self-ligation as described above.

The following items are diagrammed in Fig. 4.

Enhancer fragments. Fragment 1.3 is a 1.3-kb HS-II-plus-III fragment (GenBank ADA numbers 8332 to 9584), fragment A is a 355-bp HS III fragment (GenBank ADA numbers 9229 to 9584), fragment B is a 295-bp HS III fragment (GenBank ADA numbers 9289 to 9584), and fragment C is a 98-bp HS III fragment (GenBank ADA numbers 9289 to 9387) obtained from pADA CAT 4/I1.3 and BAL-31 deletants 3, 5, and 32, respectively.

Short promoter/long promoter constructs. Constructions designated pADA CAT 0.2 (X) refer to the placement of various intronic enhancer fragments into the *Bam*HI site 3′ to the poly(A) site of pADA CAT 0.2 (3). Constructs designated pADA CAT 4.0 (X) refer to placements into plasmid pADA CAT 4.0. The 0.2 promoter is a 232-bp ADA promoter fragment, and 4.0 is the 3,980-bp ADA promoter fragment identical at its 3′ end to the short promoter. X refers to the various combinations of the enhancer fragments. To demonstrate orientation and position independence, enhancer fragments were cloned into *Bgl*II-*Sac*I-cut pADA CAT 4.0 and *Sac*I-*Bss*HII-cut pADA CAT 1.0. These plasmids without inserts are d1 and d2, respectively.

Heterologous promoters. A series of heterologous promoter CAT constructs was made or obtained that contained the following promoters: simian virus 40 (SV40) early promoter, a 323-bp fragment in pSV2-CAT made by Gorman (32); the Rous sarcoma virus (RSV) long terminal repeat (LTR)-type promoter as a 524-bp fragment (including the enhancer) in pRSV-CAT, constructed by Gorman et al. (34); the thymidine kinase (TK) promoter as a 253-bp fragment that contains a TATA box, a CAAT box, and two Sp1 sites (53); and the mouse metallothionein I (MMT) promoter as a 336-bp fragment that contains a TATA box and a number of metal regulatory elements. The latter two promoters were placed into a backbone that was essentially identical to the pSV0-CAT vector (33) except that the *Hin*dIII-to-*Bam*HI fragment from pSV0-CAT was placed into similarly cut pUC18 so as to generate a polylinker after the 3′ *Bam*HI site. This allowed the putative enhancer fragment to be placed 3′ to the CAT coding/splice/polyadenylation signal of each construct.

The following constructions are used in Fig. 6 and 7.

Oligonucleotide enhancers. Complementary oligomers were synthesized with ligation-compatible *Xho*I and *Sal*I sticky ends. Oligonucleotides to be multimerized were ligated into a *Sal*I site that was previously placed into the *Bam*HI site of pADA CAT 4.0 (3). One to four copies (in tandem) were inserted into the *Sal*I site. Since a 32-bp oligomer (RC) that did not contain adjacent AP-1 or TCF-1α recognition motifs was as effective an enhancer in MOLT 4 cells as oligomers that did, site-directed mutagenesis was carried out on the 32-bp oligomer only. To do this, adjacent 2-base motifs were mutated C to G, A to T, and so on (Fig. 6, oligonucleotides M1 to M7).

All BAL-31 deleted or oligonucleotide-derivatized plasmids were subjected to sequence analysis of their modified regions. Plasmids were sequenced by using Sequenase (U.S. Biochemical) according to the protocols recommended by the manufacturer for double-stranded plasmids, using primers that flanked the cloning sites. Concatameric enhancer element plasmids usually required sequencing in both directions when inverted repeats of the element caused severe compression artifacts in the sequencing gels. A significant number of individual clones had random and (based on transfection) adversely acting mutations in one or more copies of their polymerized segments. These were not included in the data presented.

Fragments for transgenes. The fragments for transgenes are diagrammed in Fig. 8. Restriction fragments from various plasmids were isolated from agarose gels and prepared for microinjection, using CsCl essentially as described by Hogan et al. (43). The transgenes described below are designated by (5′-flanking DNA)/(hypersensitive sites or intron segment placed downstream of CAT).

0.2/III was obtained by *Bam*HI digestion (linearization) of pADA CAT 0.2/BAL-31 plasmid 3 that contains a small segment of the HS III site (GenBank numbers 9229 to 9584). Therefore, this construction contains the 0.2-kb promoter as does transgene 0.2/0-VI, but it also contains vector sequence from the cloning plasmid. It should be noted that transgene 4/0-VI also contains vector sequences as previously described (3), and this had no effect on expression. We have, however, avoided making any conclusions about the activity of the minimal enhancer because of these considerations. None of the other transgenes contain vector sequences.

3.7/II-III was from pADA CAT 4.0/i1.3 digested with *Nde*I. 3.7/I-III was from pADA CAT 4.0/i3.5 digested with *Nde*I. 0.3/IIIab was from pADA CAT 4.0/i1.6 digested with *Bss*HII and *Sph*I. The i1.6 fragment is an *Eco*RI-to-*Sph*I fragment (GenBank numbers 8919 to 10579). 0.3/II-IIIa was from pADA CAT 4.0/8d2 digested with *Bss*HII and *Dra*I. 0.3/II-IIIab was from pADA CAT 4.0/8d2 digested with *Bss*HII and *Sph*I. 0.3/II-V was from pADA CAT 4.0/8d2 digested with *Bss*HII and *Aat*II. 3.7/0-V was from pADA CAT 4/12 digested with *Nde*I and *Mlu*I. 0.2/0-VI was from pADA CAT 0.2/12 digested with *Nde*I. 3.7/0-I and 3.7/0-VI are listed for comparison and were previously designated transgenes ADA CAT 3.7/4.3 and ADA CAT 4/12, respectively (3).

RESULTS

Deletion analysis for enhancer activity in the human ADA gene first intron. We previously identified six HS in a 12.8-kb segment of the ADA gene first intron. HS I to V were sufficient to activate the expression of a CAT reporter gene in transgenic mouse thymus. In transfected T lymphoblasts, a 1.3-kb fragment that included only HS II and III exhibited strong enhancer activity. To localize the determinants for T-cell enhancer activity, a series of ADA CAT plasmid DNAs (4/8d series) was constructed. The plasmids were analyzed for enhancer function by the transient expression of CAT following DEAE-mediated transfection into human lymphoid cell lines MOLT 4, CEM, and Raji (Fig. 1). High-level CAT activity was found only when the plasmids were transfected into MOLT 4 and only from those plasmids which had intronic fragments containing HS III. Consistent with their reduced transcription of the ADA gene compared with MOLT 4 (49), CEM cells exhibited much less CAT expression from plasmids that contained HS III alone or in combination with the other HS sites of the first intron. In the Raji cell line, a B-lymphoid cell line that transcribes the ADA gene at rates even lower than does MOLT 4 or CEM (49), inclusion of the intron fragments led to little or no

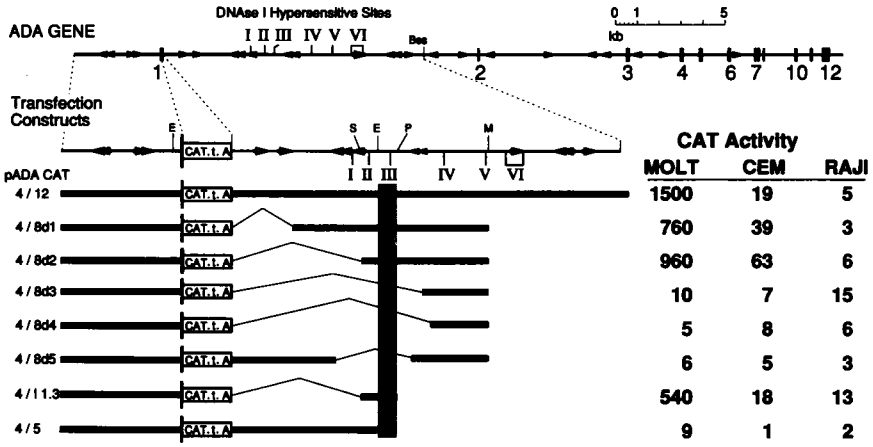


FIG. 1. Deletion analysis for enhancers in the first intron of the ADA gene. HS III is necessary for T-cell-specific enhancer activity. Supercoiled CAT reporter plasmids containing the indicated segments of the human ADA gene were transfected into lymphoid cell lines MOLT 4, CEM, and Raji. CAT activity was measured after 48 h. Results are expressed as the average of two independent transfections performed on the same day, the values of which were typically within 10 to 20% of each other. CAT activity values are picomoles of chloramphenicol acetylated per hour per picomole of plasmid DNA transfected into 3.3 × 10E6 cells. The entire human ADA gene is depicted at the top, with the incorporated segments shown below; vertical bars represent exons, arrowheads represent *Alu*-type repetitive sequence. HS sites are as determined in reference 3. The region necessary for enhancer activity is depicted as a gray bar surrounding HS III.

stimulation of CAT expression. In the absence of HS III, none of the other five HS sites of the first intron had detectable enhancer activity in any of the cell lines. Thus, HS III appears necessary for T-lymphoid cell-specific enhancer activity.

BAL-31 deletion analysis of the HS III enhancer domain. To locate and dissect the element(s) associated with the HS III enhancer region, a series of BAL-31 deletions was made into each end of the 1.3-kb HS II-plus-III segment of pADA CAT 4/i1.3 as indicated in Fig. 2. Transfection studies indicated that enhancer activity was essentially unchanged in MOLT 4 cells until deletion approached the center of the HS III region (Fig. 2) from either side. Enhancement was almost entirely abolished once the middle of the HS III domain was encroached. These constructs located a necessary core of the enhancer in the virtual center of HS III.

On the basis of a two- to threefold drop in CAT activity upon 5′ deletion from positions 9229 to 9262 (constructs 1 to 3 versus constructs 4 and 5), two other groups of 3′-deleted constructions that were truncated to 9229 and 9289 on the 5′ end were made. The 3′ deletions of the third and fourth groups of constructions indicated that there is clearly a fine structure to the sequences that flank the core. In the third set of BAL-31 deletions that contained 78 bp upstream of the core, progressive shortening of the 3′ end led to a significant loss of enhancer activity 25 to 46 bp downstream of the core (constructs 26 and 27 versus construct 28). Note that this effect differs from that seen when the additional 5′ sequence contained in constructs 13 and 14 is included. In the fourth series of BAL-31 deletions, the 5′ limit of the enhancer region was placed within 18 bp of the core. In this group, enhancement occurred with 93 bp of downstream sequence but not 53 bp (construct 31 versus construct 32). These results suggest that the enhancer domain associated with the HS III region is complex and at minimum includes about 150 bases of DNA sequence centered on the essential core region. At least one of the flanking segments must be completely intact for the core region to have high-level enhancer function. Careful inspection of the deletion analyses suggests that each flanking region must be composed of at least two separate segments. For example, comparison of constructs 13 and 14 versus construct 28 suggests that sequences 5′ of 9229 are also important. Comparison of constructs 24 to 28 with constructs 31 and 32 suggests that at least two complementing elements are present within 9359 to 9433. The motifs present in these active segments will be addressed below.

Footprint analysis of factors binding to the HS III region. BAL-31 deletion analysis suggested that a number of factors may interact with sequences within the HS III region. To evaluate this possibility and to approach identification of these factors, DNase I footprint analysis of the region was performed. To do this, total nuclear extract from MOLT 4 cells was incubated with DNA fragments that included the enhancer region identified by deletion analysis. As shown in Fig. 3, three distinct regions of the fragments were protected from DNase I-mediated attack on each strand. When the protected sequences are compared with the BAL-31 deletions (protected regions are indicated as boxes over the BAL-31 deletion data shown in Fig. 2), there is a good correlation with the functional activity determined by deletion of flanking complementary and core enhancer regions. At the central core region of HS III, where BAL-31 deletion analysis indicated the presence of critical motifs for enhancer activity, two adjacent motifs within the region (labeled ADA-NF1 and TCF-1α in Fig. 3A and B) were protected from DNase I. In the downstream flanking region, another area of significant DNase I protection (PU region in Fig. 3) was observed about 40 to 50 bp from the core region. Separate experiments designed to examine the upstream flanking region indicated sporadic protection of a small segment on one strand (data not shown; hatched bar in Fig. 3C). Both this upstream region and the downstream PU region correspond to segments where significant decreases in enhancer function were observed with BAL-31 deletions (Fig. 2). To examine the specificity of the DNase I protection observed, competition experiments were performed. Complementary 32-bp oligonucleotides that encompassed the entire ADA-NF1 site (RC oligonucleotide; see Fig. 6 and below) were synthesized. Preincubation of the MOLT 4

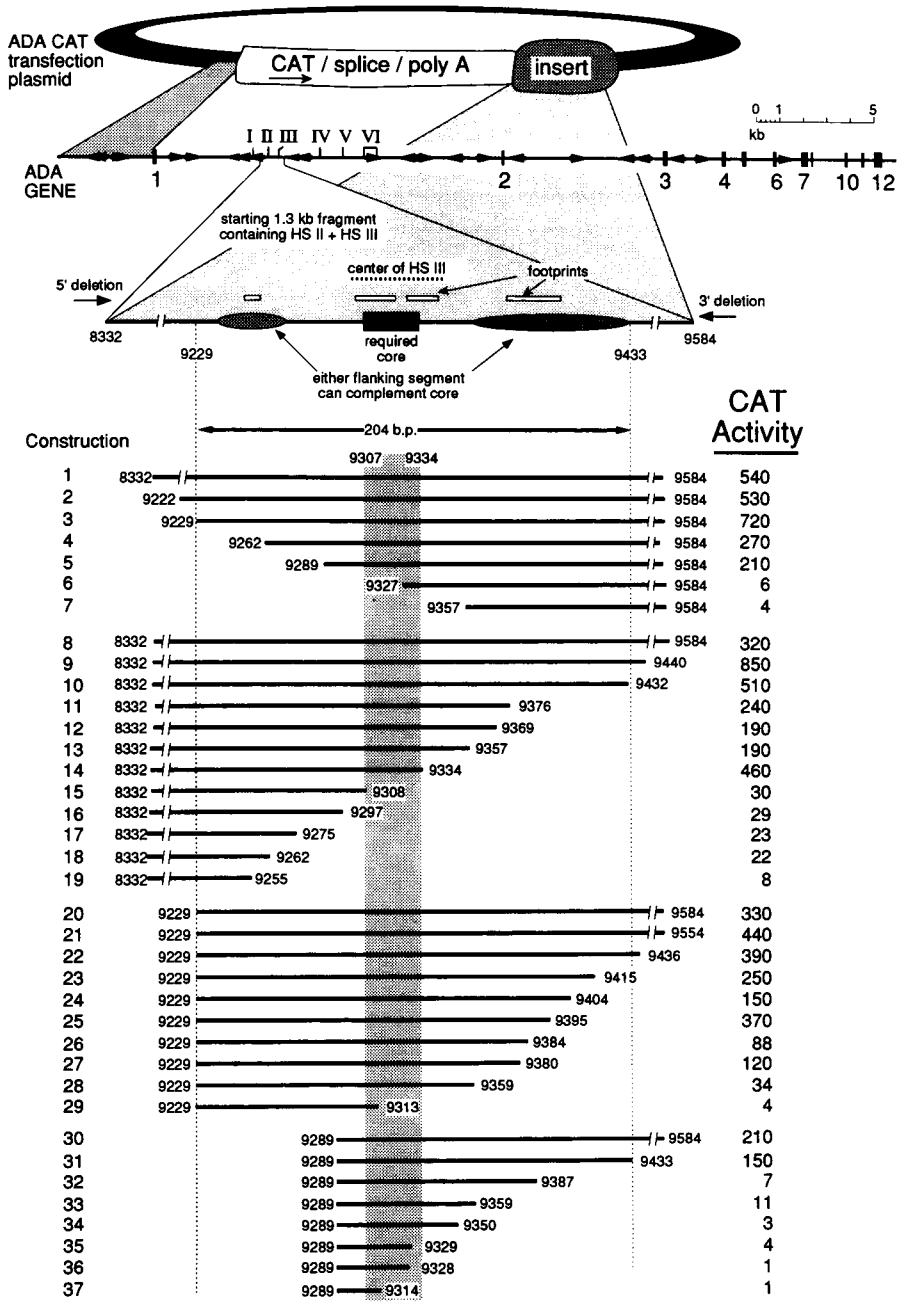


FIG. 2. Evidence that the region within HS III is sufficient for enhancer activity and has a fine structure of interactive *cis* elements that includes a required core. BAL-31 exonuclease deletions (5′, 3′, or both) were made of the 1.3-kb intronic fragment that includes HS II and HS III. The parental plasmid is pADA CAT 4/11.3 (Fig. 1). The resulting plasmids were sequenced and transfected in duplicate into MOLT 4 cells and assayed for CAT activity as described in the legend to Fig. 1. Numbering is according to the GenBank file of the human ADA gene. Stippled segments within the HS III region indicate the implicated domains of the region that have functional activity. Footprints are the locations of MOLT 4 nuclear extract protected DNA sequences (Fig. 3).

nuclear extract with an excess of RC oligonucleotide specifically abolished the ADA-NF1 footprint (Fig. 3B, lane 5). Preincubation with an excess of a fragment encompassing the whole HS III enhancer region partially abolished all three of the DNase I footprints (Fig. 3B, lane 6).

Conventional enhancer at HS III. (i) Orientation and position independence. Several plasmid constructions were prepared and tested by a transient transfection assay to evaluate the ability of the enhancer region to function in a location-, distance-, and orientation-independent manner (Fig. 4). We previously showed that a 12.8-kb intronic fragment can enhance the promoter of the ADA gene from either orientation when placed downstream (3). This fragment enhanced expression from either the short core promoter, composed of the first 132 bp 5′ of the major transcriptional start site, or a larger 4-kb promoter. Similar enhancer activity was ob-

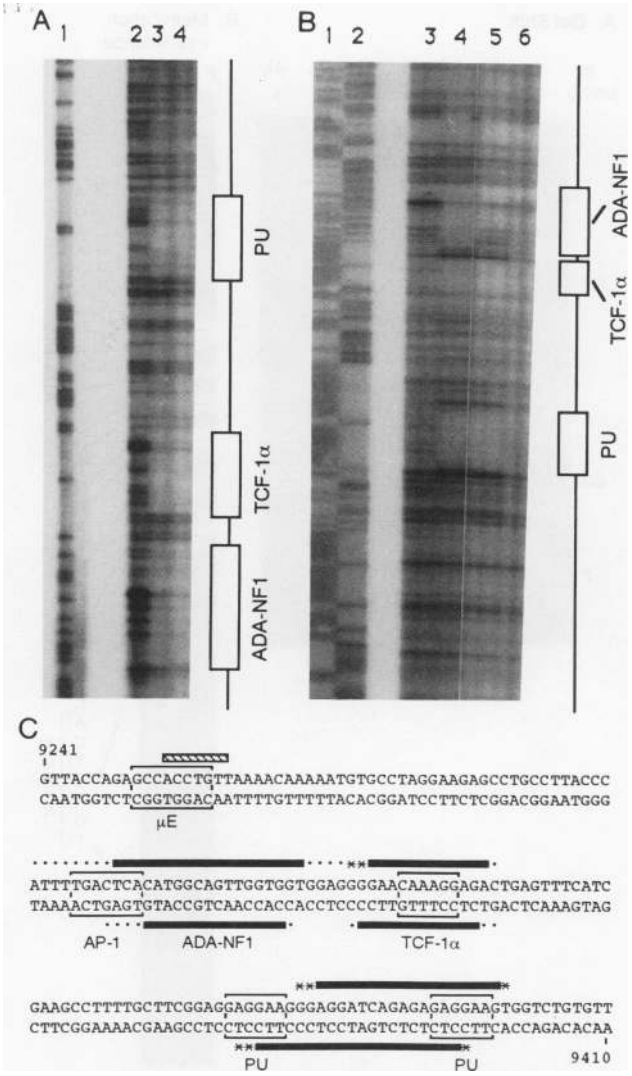


FIG. 3. (A and B) DNase I footprint analysis of the ADA HS III enhancer. Analyses were performed with MOLT 4 nuclear extract. Footprints on each strand were evaluated. Protected areas are indicated by boxes. (A) Lower-strand footprints with 0 (lane 2), 5 (lane 3), or 10 (lane 4) μg of nuclear extract. Lane 1 is a Maxam-Gilbert G+A purine sequencing reaction of the same fragment. (B) Upper-strand footprints with 0 (lane 3) or 5 (lanes 4 to 6) μg of nuclear extract. Lanes 5 and 6 also had 20 ng of RC oligomer and 30 ng of the 374-bp *Bam*HI-*Sac*I fragment, respectively, present in the incubation mixture as a competitor. Lanes 1 and 2 are, respectively, Maxam-Gilbert G+A and C+T sequencing reactions of the same fragment. (C) Sequence of the enhancer region, with numbering according to GenBank for the human ADA gene. Bars above and below the sequence indicate protected areas of the upper and lower strands, respectively. Matches to known recognition sequences are bracketed and labeled. Dots indicate residues where protection is undetermined because no banding is observed in the control; asterisks indicate sites of increased sensitivity to DNase I associated with a bound factor(s).

served with a smaller 1.3-kb fragment encompassing HS II and HS III. What is now shown is that this fragment also demonstrates high-level enhancer activity when placed upstream of the promoter in either orientation (Fig. 4). Small segments of the HS III enhancer, as defined by the BAL-31 deletion analysis, were tested downstream of either the small

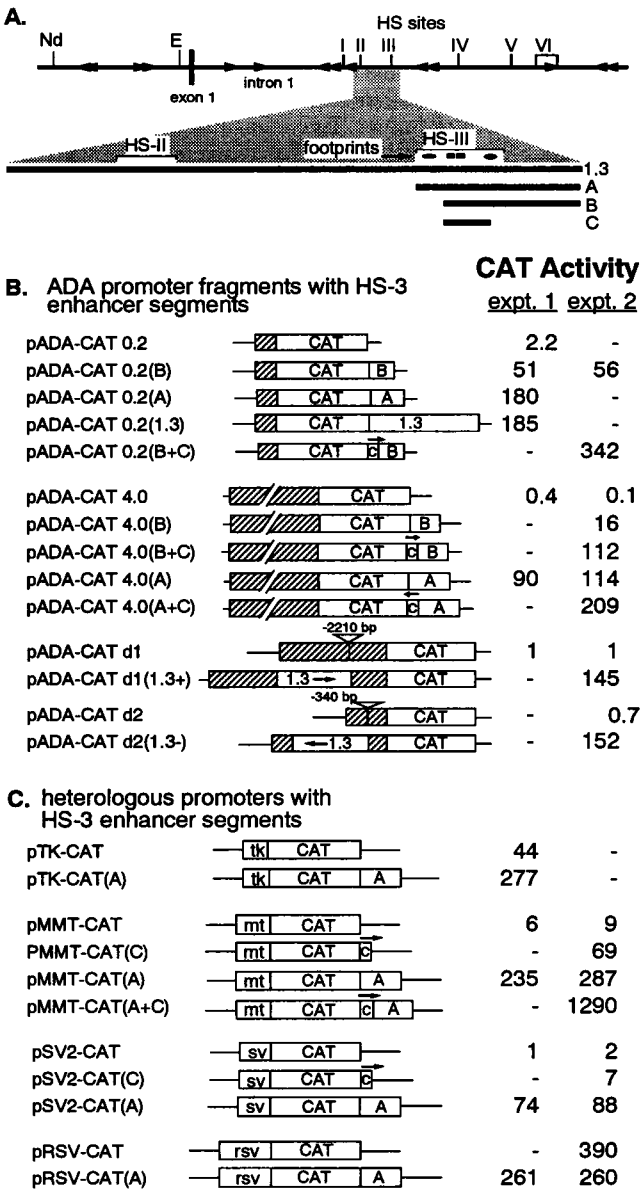


FIG. 4. Conventional enhancer characteristics of position, orientation, and promoter independence of segments encompassing the HS III region. (A) Intronic fragments used in the analysis. 1.3 is the 1.3-kb fragment that contains HS II and HS III; A, B, and C are subfragments described in Materials and Methods. CAT activity produced (average of duplicate, independent transfections in MOLT 4 cells) was normalized to that of pADA-CAT d1. (B) Relative effects of various HS III enhancer segments on the ADA promoter. Enhancer segments are used alone, in combination, or in altered positions and orientations. The cross-hatched segments are contiguous 5′-flanking/promoter/5′ nontranslated fragments from the human ADA gene. The enhancer effects on minimal and extended ADA promoters were evaluated by using the 0.2 and 4.0 promoters. The ability of the enhancer segments to function when positioned within 5′-flanking DNA was evaluated by using the d1 and d2 promoter fragments (see Materials and Methods). (C) Effects of enhancer fragments on heterologous promoters. The TK promoter is of the herpes simplex virus TK gene (contains a TATA box, a CAAT box, and two Sp1 sites). The MMT promoter contains a TATA box and a number of metal regulatory elements. The SV2 promoter is a 323-bp fragment from the SV40 early-region promoter. The RSV promoter is a 524-bp fragment of the RSV LTR (including the enhancer).

or large ADA promoter and were capable of significant stimulation of CAT expression in either case. Thus, the enhancer sequences were capable of exerting their effects independent of orientation at distances from 300 bp up to 9 kb from the core promoter. Even small segments of the HS III enhancer region were capable of acting on small segments of the ADA promoter. The long promoter was considerably less active than the short promoter, consistent with the presence of negatively acting sequences in the 5′-flanking DNA, but upon addition of enhancer segments, the level of expression was relatively independent of the presence or absence of the additional 5′-flanking DNA. The relative strengths of several enhancer segments on the core promoter were also similar: 1.3 = A > B. In some of the constructs, a copy of the C fragment was added (e.g., B plus C), and this led to a further stimulation of the apparent enhancer activity.

(ii) Activation of heterologous promoters. To examine the ability of the enhancer fragment found at HS III to *cis* activate heterologous promoters, A and/or C enhancer fragments were placed downstream of heterologous promoter-CAT constructs (Fig. 4) and tested by transfection into MOLT 4 cells. The enhancer (A fragment) stimulated CAT expression from the TK, MMT, and SV2 promoters by factors of 6-, 35-, and 55-fold, respectively. Expression from the RSV LTR was not increased by the enhancer and instead exhibited moderate inhibition (33%). This finding may relate to the impressive ability of the RSV promoter/enhancer sequences to be recognized by factors contained in MOLT 4 cells, as its CAT expression was very high. The C subfragment of the HS III enhancer also showed effects of 5- to 10-fold stimulation of the MMT and SV40 promoters by itself and caused a similar rise in the activity of the A-enhanced MMT construct. In results not shown, these constructs were also transfected into CEM and Raji cell lines. As could be expected from the results shown in Fig. 1, there were only small effects of the enhancer segments on reporter expression, regardless of the heterologous promoter or how well the heterologous promoter was itself expressed by CEM or Raji cell lines. This finding indicates that cell type specificity is determined by the intronic enhancer elements.

Methylation interference. To identify factors that interact with elements within the HS III domain, gel shift experiments were performed by using a 126-bp fragment that spanned the core enhancer region. Two major shifted complexes were observed (Fig. 5A). Both of the shifted complexes could be blocked by unlabeled DNA containing the same sequences, but heterologous DNA fragments also caused some diminution of each band (not shown). Methylation interference analysis was performed to determine which sequences of the fragment were necessary to form the complexes. The fragment was end labeled on either strand, partially methylated, incubated with extract, and applied to the gel, and the DNAs were isolated from complex 1 (C1, slower mobility), complex 2 (C2, faster mobility), and the freely migrating bands. As shown in Fig. 5, DNA purified from C2 exhibited a lack of methylation corresponding to 10 bp within the RC oligonucleotide. The factor responsible is thereby defined as ADA-NF2. Complex 1 showed only partly diminished intensities of G residues in extended sequence that included those implicated for complex 2 and extended into the AP-1 site (not shown). However, the pattern was too weak for a clear-cut interpretation. Complex 2 showed methylation interference only on the upper strand, with no potentially informative bases present on the lower strand in the ADA-NF2 region. Neither strand exhibited diminished G-residue intensities in any other region. Thus,

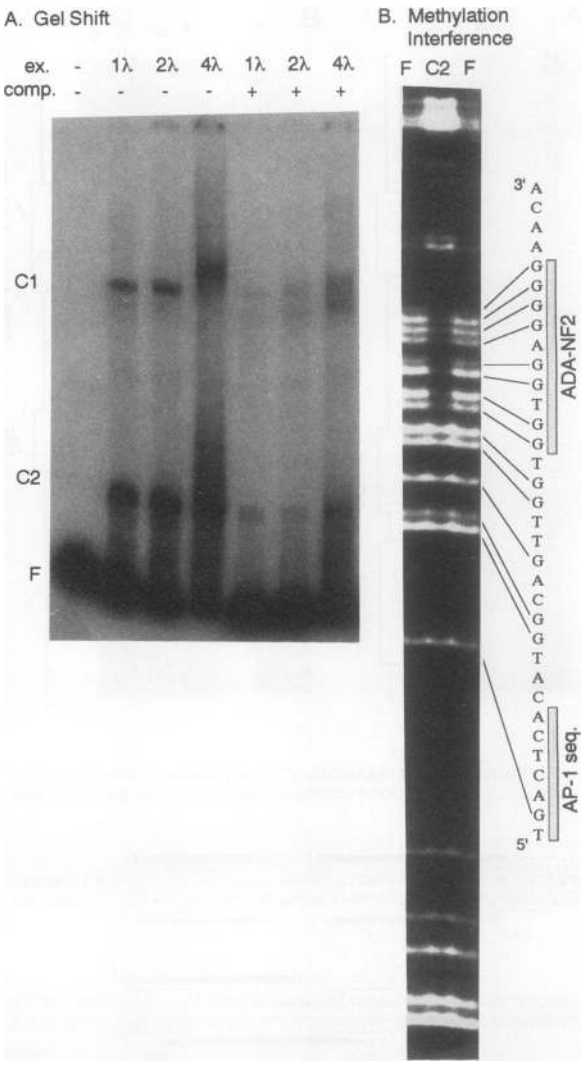
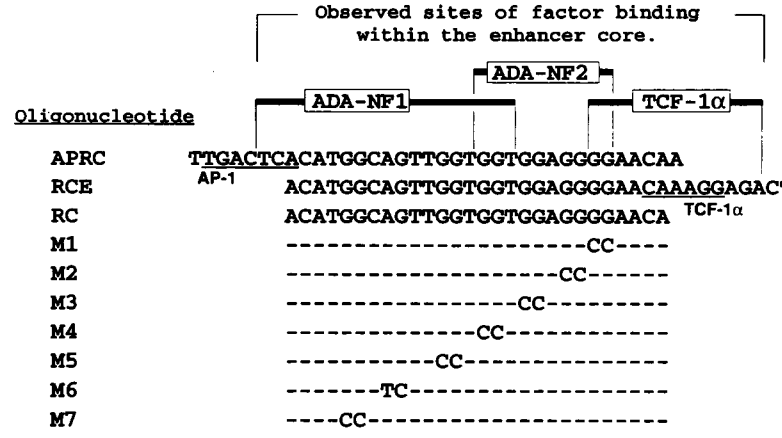


FIG. 5. Detection by gel shift methylation interference analysis of a factor binding to bases within the enhancer core that are distinct from those implicated by DNase I footprint analysis. (A) Electrophoretic mobility shift experiments were performed by using a partially methylated, end-labeled 126-bp *Bam*HI-to-*Taq*I fragment from BAL-31 deletion construct 3 (Fig. 2). The fragment was combined with the indicated volume (in microliters [λ]) of mouse thymic whole-cell extract (10 μg/μl of protein) in a final volume of 25 μl and subjected to electrophoresis through a 4% polyacrylamide gel that contained 1/2× Tris-borate-EDTA. Binding was challenged with unlabeled DNA (±40 ng of competitor, ~0.5 ng of labeled fragment). Labeled fragments that migrated at C1, C2, and F (free) were eluted from a separate preparative gel by using the equivalent of 2 μl of extract, subjected to piperidine cleavage, and applied to a sequencing gel (B). Only complex C2 with its upper strand labeled is shown.

methylation interference analysis suggests that a high-mobility complex did make eight G-base contacts along a 10-bp segment at the 3′ portion of the core enhancer region. As is inherent in the methylation interference technique, the specificity of the interaction is proven by the ability of any of the single G methylations to strongly inhibit the binding reaction. This 10-bp stretch occurs directly between the ADA-NF1 site and the TCF-1α sites, with a considerable overlap of the DNase footprint of each. To assess the relative role of



Oligonucleotide		CAT ACTIVITY 1-copy	2-copies	3-copies	4-copies
APRC	TTGACTCACATGGCAGTTGGTGGTGGAGGGGAACAA	5	53±15	250	-
RCE	ACATGGCAGTTGGTGGTGGAGGGGAACAAAGGAGACT	12±7	87	570	-
RC	ACATGGCAGTTGGTGGTGGAGGGGAACA	7	51±17	340	850±330
M1	----------CC----	-	-	11	-
M2	----------CC------	-	-	12	-
M3	----------CC---------	-	-	5	-
M4	----------CC----------	-	-	11	-
M5	----------CC----------	-	-	-	6
M6	-------TC----------	-	-	11	-
M7	----CC----------	-	2	1	-

FIG. 6. Polymeric oligonucleotide analysis of the essential core region of the HS III enhancer. The indicated oligonucleotides and their complementary strands were synthesized with four additional bases in order to ligate into a *Sal*I site placed 3′ to the poly(A) signal of the enhancerless plasmid pADA CAT 4.0. Plasmids were restriction and/or sequence analyzed and transfected in duplicate into MOLT 4 cells. Standard deviations are given when three or more independent clones of the indicated concatemerization were obtained. CAT activity is the fold increase over the activity of the parent vector, which was 0.1 nmol acetylated per h per 10E7 cell per pmol of DNA transfected. The relative orientation of the individual monomers had no effect, and the resulting constructions are listed together as dimer or trimer, etc. ADA-NF1 is defined from Fig. 3 as DNase I-protected bases within, and competed for by, the RC oligonucleotide. ADA-NF2 is defined from methylation interference analysis (Fig. 5) as the bases that interfered with the formation of complex 2. Underlined bases indicate sequence consensus match to sites for known *trans* factors. For oligonucleotides M1 to M7, only the bases that are altered compared with the wild-type sequence are indicated. -, no clones obtained.

the AP-1, ADA-NF1, ADA-NF2, and TCF-1α motifs within the HS III enhancer core, reporters were constructed by using synthetic oligonucleotides.

Oligonucleotide analysis of the enhancer core sequence. Since the BAL-31 data indicated that the core had to be complemented by flanking sequences in order for strong enhancer activity to be observed, we sought to overcome this limitation by polymerization of the core sequences. Three different wild-type sequence oligonucleotides were synthesized (Fig. 6). APRC includes the AP-1 site and the ADA-NF1 and ADA-NF2 sites; RCE includes the ADA-NF1, ADA-NF2, and TCF-1α sites (44); and RC contains only the ADA-NF1 and ADA-NF2 sites. The resulting plasmids were transfected into the high-expressing cell line MOLT 4, and CAT activity was normalized relative to the level of expression of the parental plasmid. As expected from the BAL-31 data, a single copy of the core sequence had low enhancer activity. However, multiple copies of the oligomers had potent enhancer activity in MOLT 4 cells (up to 1,000-fold stimulation of the parental construct for tetramers). The addition of the AP-1 site made no difference compared with the RC oligonucleotide. The addition of the TCF-1α site gave a consistent (six of seven constructions) but low (maximally only twofold) increase in the apparent activation. In light of the DNase I footprint observed over this site, TCF-1α may play a role in the function of the enhancer. However, since the 28-bp RC oligonucleotide was sufficient to be strongly recognized by the enhancer-binding factors when polymerized, neither the AP-1 nor TCF-1α motif is likely to be the primary HS III enhancer-activating factor.

Interestingly, the enhancer strength of the RC oligonucleotide is relatively constant per copy, acting as an approximately sevenfold multiplier. This is essentially identical to the strength of single copies of the 100-bp C fragment (Fig. 4) alone or when combined with larger HS III fragments, providing further indication that the 28-bp RC oligonucleotide represents the essential core of the HS III enhancer.

To map the critical bases within the RC oligonucleotide, a series of oligonucleotides (M1 to M7; Fig. 6) that contained adjacent two-base alterations of the RC sequence were synthesized. Each of the seven mutations of the core region prevented the ability of the polymerized oligonucleotide to act as an enhancer. Clearly, bases within the entire RC region that encompass binding sites for both ADA-NF1 and ADA-NF2 are essential for formation of the active enhancer.

To evaluate the cell type specificity of the RC enhancer core elements, the resulting plasmids were transfected into MOLT 4, CEM, and Raji cell lines (Fig. 7). The enhancer effect was much lower in CEM cells than in MOLT 4 cells and was not detectable in Raji cells. Sequence analysis of the plasmids indicated a variety of absolute and relative orien-

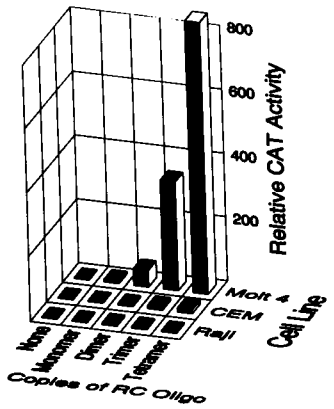


FIG. 7. Cell type specificity of the enhancer core. Each of the plasmids containing polymerized RC enhancer core elements was transfected into the human lymphoid cell lines MOLT 4, CEM, and Raji and assayed for transient CAT activity expression. CAT activities are averages of two independent transfections for each plasmid and are the fold increases of expression compared with the oligonucleotide-lacking parent vector as determined in each cell line.

tations of the polymerized core sequences, but this had little or no effect on the relative enhancement of expression by either the monomers or the multimers.

Transgenic analysis of the ADA thymic regulatory region. To delineate the essential segments of the ADA CAT transgene, we made deletions of both the ADA gene promoter and enhancer and evaluated their expression in transgenic mice (Fig. 8). All results were obtained from F_1 offspring of established transgenic mouse lines, and usually three or four independent mouse lines were obtained from each construct and analyzed. CAT activity was consistently highest in the thymus, marrow, and spleen. The inclusion of a sufficiently large segment of the first-intron regulatory region, as contained in mouse lines 22 to 38, resulted in significant CAT expression in all tissues tested, but expression was always the highest in the thymus by usually 10- to 100-fold. In the thymus, the range of CAT specific activity divided by gene copy number was 16,000 to 63,000. Expression in the marrow and spleen was lower and much more variable than expression in the thymus. CAT expression in other tissues was low, as previously described in detail (3).

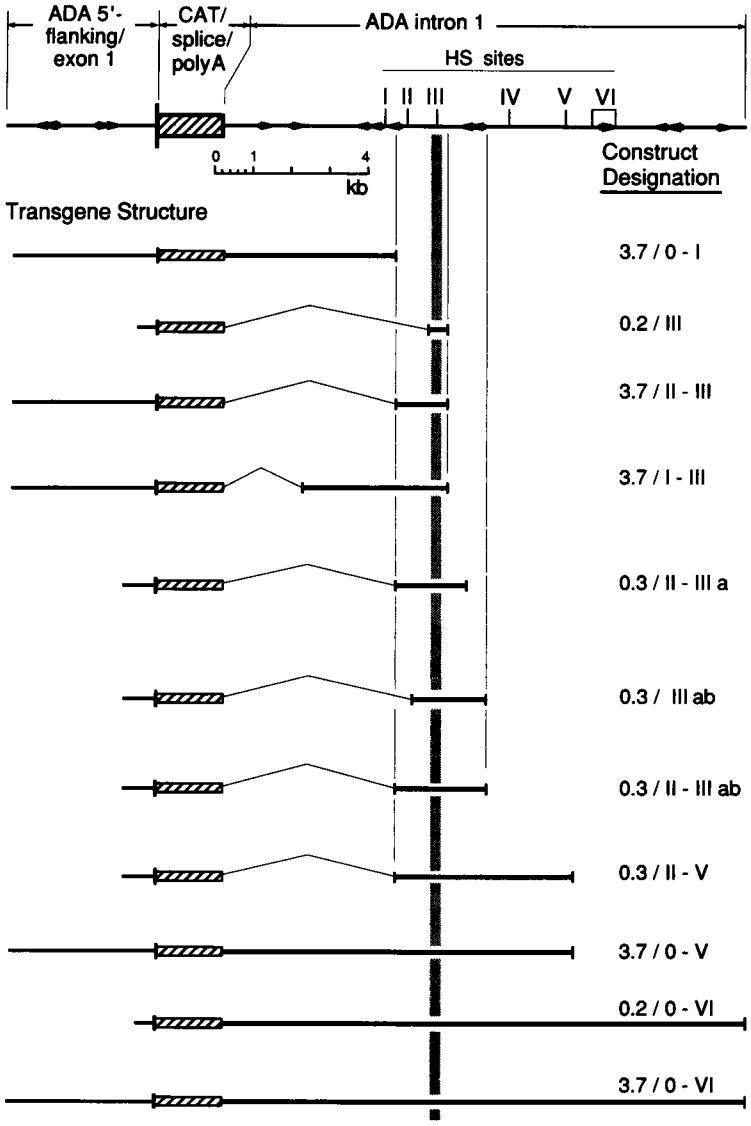
(i) Lack of a role for distal 5′-flanking/promoter region DNA. Results for mouse lines 31 to 33 indicated that only 232 bp of 5′-flanking DNA was fully capable of responding to the intronic enhancer region (results for subsequent lines 22 to 29 confirmed this finding). This result is consistent with our previous transient transfection expression studies in MOLT 4 cells, which demonstrated that the short promoter was strongly activated by the intronic enhancer. However, the lack of any discernible effect by the additional 5′-flanking DNA in transgenic mice is in marked contrast to the strong negative effects of these sequences in transient transfection assays. We previously observed that expression from the core promoter was markedly inhibited by additional 5′-flanking DNA sequences, but that the extent of the inhibition was most prominent in the absence of the intronic enhancer segment. Thus, we had hypothesized that CAT expression in tissues outside of the thymus, such as mature lymphoid cell types of the spleen, might be higher in the absence of the 5′-flanking negative regulatory sequences. This was not observed in spleen or bone marrow or in tissues that were low expressers in the previously reported lines 30 and 34 to 38 (3; additional tissue data not shown). Once the segment of the promoter region that was essential for high-level thymic expression was shown to be quite small, we proceeded to delimit the enhancer region.

(ii) Copy-consistent thymic expression and variable spleen and bone marrow expression. Constructs used to make transgenic mouse lines 22 to 38 generated consistently high thymic CAT expression (31,700 ± 13,000 U per gene copy, mean ± standard deviation). The same mice, however, exhibited much more variability of bone marrow and spleen expression (930 ± 780 and 400 ± 490 U per gene copy, respectively). Thus, in mice made from the larger transgenes, expression in the thymus was more consistently controlled than expression in the spleen and bone marrow. When adjusted for transgene copy number, thymic CAT mRNA was present at levels similar to those of endogenous ADA mRNA in mouse lines 35 and 38 (3). Also, the intrathymic distributions of CAT mRNA in line 38 and ADA enzyme in human thymus are similar (3). Thus, we conclude that the larger constructs closely mimic the intrathymic regulation of the human ADA gene but not in its consistency of expression in spleen and bone marrow.

(iii) Loss of the 3′ *Alu*-containing segment leads to line variability of thymic (but not bone marrow) expression. The smallest segment of the intronic enhancer region that gave consistent thymic expression was a 2.3-kb segment that included about 1 kb on each side of the HS III enhancer region. Each end of the 2.3-kb segment appears to play a role. The most characteristic aspect of mice made transgenic with constructs that lacked the downstream nonhypersensitive segment (lines 6 to 18) is the wide variability in their copy-adjusted expression of CAT in thymus (6,917 ± 11,540 U per gene copy). Lines 13 to 18 in particular lacked only the terminal 655 bp of the segment, 530 bp of which is *Alu* repetitive DNA. Lines 6 to 12 lacked the terminal 1,048 bp. Careful examination of these data suggests that the deviation from the expected range is greatest among the mouse lines with higher copy number. These results suggest that the terminal nonhypersensitive portion of the active 2.3-kb fragment is likely to be responsible for the ability of the enhancer region to acquire insertion site independence of expression in the thymus. Interestingly, comparison of lines 9 to 12 versus the remainder of lines 6 to 18 suggests that the ability to overcome position effects in bone marrow did not predict that position effects would be overcome in developing thymocytes. However, the level of bone marrow expression in lines 9 to 12 was consistently higher (5,000 ± 1,995 U per gene copy) than in any of the other groups of mice, and this occurred in the absence of the segment of DNA downstream of HS III.

(iv) Deletion of the 5′ HS II region from the active regulatory domain leads to line variability and possibly low expression. Since virtually the entire 3′ end of the 2.3-kb intronic fragment appeared critical for insertion site independence in thymus, we tested the significance of the 5′ HS II region by deleting it from the otherwise complete 2.3-kb region. In the three lines generated from this construct, 19 to 21, thymic CAT activity was low and markedly variable (1,400 ± 1996 U per gene copy). Minimal enhancer sequences (HS III only) did not activate transgene expression in two lines (4 and 5), but more of these lines will need to be analyzed in order to make a firm conclusion in light of the variability of expression observed in lines 6 to 18. Thus, sequences both upstream and downstream of the HS III enhancer region are critical for establishing a strong and consistently active enhancer domain. Furthermore, considerably more *cis* information is necessary for enhancement of transgene expression than is required for enhancement in transient transfection expression assays.

DNase I hypersensitivity. We examined DNase hypersensitivity at HS III to address the question of the potential mechanism by which different chromosomal insertion sites could strongly influence thymic expression of transgene constructs used to make lines 6 to 18 while having little effect on those used to generate lines 22 to 38. Previously, we observed complete conversion of thymic nuclear chromatin at the HS III site to cleaved DNA in mouse lines that included the complete intron segment (Fig. 8, lines 37 and 38; see Fig. 5 in reference 3). This finding implied that all 60 and 100 copies of the transgene contained HS III in a hypersensitive configuration. We examined DNase I hypersensitivity in lines 6 and 8 (Fig. 8, transgene construction ADA CAT 3.7/II-III) as being representative of high and low gene expression per copy. Both lines of mice were able to form some degree of hypersensitivity at HS III. This finding implies that within the region there is *cis* information that can specify where hypersensitivity should occur. However, as shown in Fig. 9, HS III sequences were significantly more hypersensitive to DNase I in mouse line 8 than in line 6. We also examined DNase I hypersensitivity in lines 10, 12, 23, and 32 (not shown). As observed in lines 6 and 8, there was excellent correlation of level of expression with degree of hypersensitivity. In all,



Construct Designation	Mouse Line	copy no.	CAT activity divided by transgene copy number: thymus	marrow	spleen
3.7 / 0 - I	1	*12*	0.0	0.0	0.0
	2	*5*	0.1	0.0	0.0
	3	*6*	0.2	0.0	0.1
0.2 / III	4	*2*	0	0	0
	5	*2*	0	0	0
3.7 / II - III	6	*20*	800	23	9
	7	*60*	680	130	9
	8	*3*	19 000	2 700	390
3.7 / I - III	9	*12*	480	4 900	250
	10	*30*	1 300	7 000	400
	11	*60*	2 200	5 800	140
	12	*3*	9 300	2 300	70
0.3 / II - III a	13	*1*	43 000	260	77
	14	*24*	0	0	0
	15	*7*	11 000	110	54
	16	*150*	360	290	150
	17	*2*	18 000	27	24
	18	*150*	360	160	73
0.3 / III ab	19	*12*	3 700	64	18
	20	*2*	120	11	1
	21	*120*	380	660	210
0.3 / II - III ab	22	*3*	26 000	440	37
	23	*8*	63 000	1 700	200
	24	*6*	55 000	1 000	260
	25	*2*	52 000	1 000	1 400
0.3 / II - V	26	*8*	32 000	700	180
	27	*5*	23 000	360	84
	28	*20*	23 000	3 100	1 700
	29	*100*	34 000	890	450
3.7 / 0 - V	30	*10*	31 000	210	67
0.2 / 0 - VI	31	*100*	30 000	350	590
	32	*60*	20 000	280	110
	33	*100*	34 000	720	330
3.7 / 0 - VI	34	*1*	28 000	210	23
	35	*10*	22 000	nd	66
	36	*15*	16 000	nd	19
	37	*60*	30 000	1 700	720
	38	*100*	20 000	1 300	520

FIG. 8. Evidence that HS III enhancer domain must be combined with upstream and downstream segments to obtain high-level and gene-copy-proportional reporter expression in transgenic mouse thymus. DNA fragments from various ADA CAT plasmids were digested with restriction enzymes to yield DNA fragments used to generate transgenic mice. A scale diagram of the largest ADA CAT transgene isolated as a fragment from plasmid pADA CAT 4/12 is shown at the top. In each construction, CAT coding, splicing, and polyadenylation sequences of pSV0-CAT replace the translated portion of the ADA gene first exon and the 5′ splice site of the first intron. Arrowheads mark *Alu*-type repetitive DNA sequences. HS sites are indicated by roman numerals. The relevant segments of the ADA gene promoter and first intron fused to the CAT cartridge are depicted. The vertical gray segment centered at HS III is the enhancer region as defined by BAL-31 deletion analysis. The vertical fine lines represent the minimal sequence that exhibits position-independent transgene reporter expression. Constructs are designated by the amount of 5′-flanking DNA, HS sites included, and additional nonhypersensitive flanking DNA. Copy number is the number of transgene copies per diploid genome equivalent (assume 6 × 10E9 bp per diploid mouse and human genome). Copy number is Southern blot determined by comparison of various amounts of transgenic mouse liver DNA to similarly cut human DNA. Band intensities were analyzed by visual inspection and by surface emission scanning (Materials and Methods). We consider the listed values to be valid within a factor of 2 or better of the actual number. CAT activity is shown as picomoles of acetylated product per hour per 100 μg of protein per transgene copy number. No transgenic lines obtained were excluded from the data unless a rearrangement of the transgene was observed. All analyzed mice were F_1 or F_2 offspring except line 2, which was F_0.

three of three low-range expressors formed poor hypersensitivity, and five of five high-range expressors formed good hypersensitivity. Therefore, it appears that DNase I hypersensitivity must be acquired at HS III to allow the enhancer domain to direct high-level transgene expression.

DISCUSSION

To assess mechanisms involved in controlling expression of the human ADA gene in developing T cells of the thymus, we examined reporter gene expression in transgenic mice

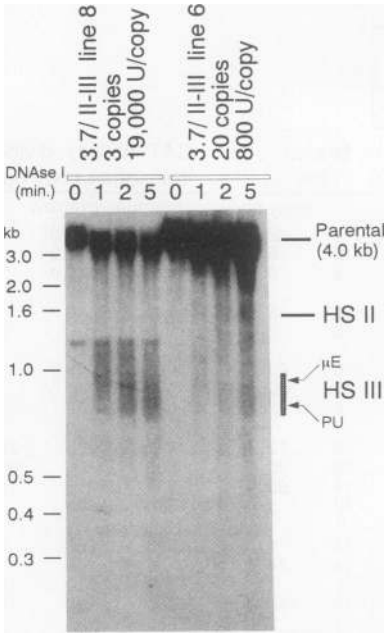


FIG. 9. Evidence that formation of DNase I hypersensitivity is variable in the absence of the facilitator segment but correlates with transgene expression and can occur precisely at HS III. Mouse lines (Fig. 8) are derived from transgene 3.7/II-III, which demonstrates variable thymic expression in independent mouse lines. Thymic nuclei were exposed to DNase I; DNA was purified, cut with *Xho*I, electrophoresed on a 1.8% agarose–1× Tris-borate-EDTA gel, transferred to a membrane, and hybridized to a fragment from the 5′-flanking DNA of the construction. In effect, this fragment looks upstream at the enhancer region from the adjacent transgene. The locations of hypersensitivity and the positions of μE and PU are determined from the migration of adjacent molecular size markers. The stippled box represents the BAL-31-delimited enhancer region. Line 8 exhibited an anomalous band of less than single-copy intensity. The autoradiogram was exposed at −70°C for 90 h with an intensifying screen.

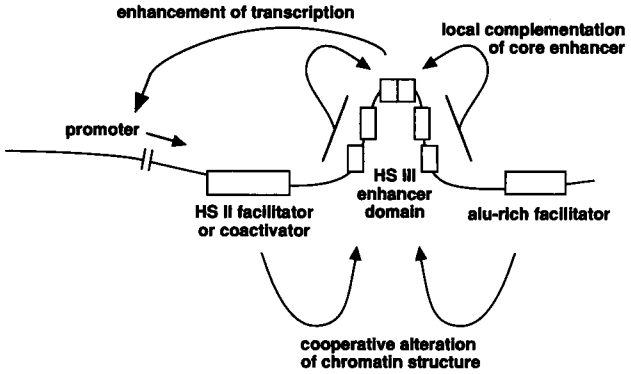


FIG. 10. Model of the ADA gene first-intron thymic regulatory region. A series of interactions permit the HS III enhancer domain to function efficiently in chromatin of developing thymocytes. At the center of HS III, a 28-bp enhancer core is bound by at least two factors capable of distal promoter activation. The enhancer strengths of the core elements are, in turn, augmented by a series of flanking elements that lie within an approximately 200-bp domain that corresponds to the limits of HS III hypersensitivity. In addition, segments that flank HS III outside the hypersensitive region are essential for its function in transgenic mouse chromatin. Both an upstream segment and a downstream segment are required to prevent variation in gene expression among independent transgenic lines. We suggest that the segments act as facilitators because occasional transgene insertion sites can compensate for their absence. Preliminary evidence indicates that the downstream facilitator ensures that HS III can become DNase I hypersensitive. We hypothesize that the process of facilitation leads to additional or more-stable interactions of *trans* activators with the enhancer domain.

and transfected lymphoid cells. By comparing the effects of progressive deletions upon enhancer activity with the locations of DNase I hypersensitivity, nuclear factor binding, and the ability of the region to become DNase I hypersensitive, we have defined and mapped determinants of a thymic regulatory region contained in a 2.3-kb segment of the human ADA gene first intron. The most remarkable feature of the overall control region is the extent to which its function is dependent on an organized set of local and regional interactions among its elements and segments. This is the case within the critical 28-bp enhancer core, between the core and bilaterally flanking elements encompassing 200 to 300 bp corresponding to HS III, and between the HS III enhancer and regionally flanking segments that encompass 2.3 kb. The effects of the latter segments are apparent only in transgenic mice. The upstream segment is HS II, and the downstream segment is nonhypersensitive and composed mainly of *Alu* sequences. The flanking segments of the thymic regulatory region appear to ensure the appropriate presentation, within the context of nuclear chromatin, of *cis*-regulatory information contained within the enhancer domain. A model indicating the arrangement of the HS II coactivator segment, the HS III enhancer domain itself composed of core and coactivating elements, and the downstream nonhypersensitive facilitator segment is shown in Fig. 10.

Core elements of the HS III enhancer domain. A 28-bp cell-type-specific enhancer core sequence occurs directly at the center of HS III. From the results of DNase I footprints, gel shift methylation interference analysis, and functional analysis of polymerized and mutated core motifs, we believe that at least two factors present in MOLT 4 and thymic nuclear extracts, ADA-NF1 and ADA-NF2, are required for generating the activating complex able to be formed at the 28-bp enhancer core. ADA-NF1 was detected in MOLT 4 nuclear extracts and was selectively prevented from contacting an extended fragment of the enhancer by the 28-mer core enhancer RC oligonucleotide without affecting binding to adjacent TCF-1α and PU segments (Fig. 3B, lane 5). Within the ADA-NF1 site is a consensus for the E12 and E47 products of the E2A locus (55). Since the E2A gene products are expressed in most tissues, this factor may be necessary but is unlikely to be sufficient or specific for cell-type-specific transactivation. E2A gene products are basic helix-loop-helix proteins able to form homo- and heterodimeric complexes with other proteins via their helix-loop-helix domains and to interact with DNA via their basic regions. Depending on the specific complex formed, active binding or inactive nonbinding transcription factors are generated (5, 6, 55, 62). Regardless, neither ADA-NF1 nor ADA-NF2 can on its own form an effective enhancer core transactivator.

ADA-NF2 was detectable in whole cell thymic extracts and required for its binding eight G residues within a 10-base stretch at the 3′ end of the RC oligonucleotide (Fig. 5). ADA-NF2 was a rapidly migrating complex in the gel shift analysis, and this finding raises concerns about whether it is

specific and whether it is a proteolytic fragment of a larger binding factor. The evidence for specificity, as is inherent with interference methods, is that single G methylations within the binding site nearly completely prevented the complex from forming. DNase I footprint analysis suggested that the residues of the ADA-NF2 interaction were unprotected or even hypersensitive to DNase I. However, specifically mutated oligonucleotides M1 to M4 altered bases implicated in the binding of ADA-NF2. Analyzed by transfection, each mutation caused nearly complete loss of enhancer activity by the polymerized mutant oligomers (Fig. 6). This finding provides strong evidence that a factor that requires these bases for its binding is likely to be involved in the active functional enhancer complex. No obvious homology to known elements was found within the ADA-NF2 footprint.

Elements within HS III that complement the core enhancer. Since the highest activity of the enhancer domain required that the core be complemented by flanking sequences, it was possible to discern the presence of elements within these sequences as breakpoints for enhancer activity in sufficiently deleted enhancer fragments. Several sequence homologies to known factors were observed at these breakpoints (Fig. 3C). ADA-NF1 is flanked just upstream by a consensus AP-1 site (2) and downstream by a TCF-1α site (69). The TCF-1α site is squarely centered in the footprint given the same designation in Fig. 3 and 6. The TCF-1α sequence is recognized by related T-cell-specific factors that have been implicated in the function of the T-cell receptor α (TCRα) enhancer (42, 69, 73) and the CD3-ε enhancer (67). The mechanism for the function of this element appears to involve the ability to strongly bend DNA (30) without itself acting as a transactivator, apparently thus cooperating strongly with other cell-type-specific *trans* activators in some, but not all, sequence contexts (70). In the ADA HS III enhancer, TCF-1α also does not function as a simple positively acting core or core-complementing element when adjoined to ADA-NF1 and ADA-NF2 because its inclusion or exclusion from polymerized oligonucleotides had little effect on reporter expression upon transfection in MOLT 4 cells. A distinct possibility, however, is that the TCF-1α site is essential for the region to function at a high level from within chromatin in transgenic constructions. In light of its ability to bend DNA (30), we are especially interested in whether its presence is essential for the HS III domain to adopt the hypersensitive configuration. Similar questions can be asked of factors that recognize the AP-1 consensus immediately 5′ of the ADA-NF1 core element. It is also noteworthy that another TCF-1α site occurs in the HS III enhancer immediately 5′ of the upstream μE element discussed below.

The downstream PU footprint (which is very purine rich on the upper strand) is bounded by matches for the recognition sequence (GAGGAA) for the PU-box-binding factor PU.1 (47). PU.1 is a member of the c-*ets* multigene family. One member of this family, *ets*-1, is expressed as the protein Ets-1 in a lymphoid cell-specific fashion (7). The downstream PU box in the ADA enhancer (which in BAL-31 analysis correlates most closely with the location of a downstream element that coactivates the core enhancer) is contained within DNA segments that are highly homologous to segments in the Moloney murine sarcoma virus LTR and human TCRα enhancer that bind Ets-1 (39, 41). The Moloney murine sarcoma virus sequence (TCGGAGAGCGG AAGCG) and TCRα sequence (AGAGGATGTGG) are 81% (13 of 16 bases) and 91% (10 of 11 bases) homologous to the overlapping ADA sequences TCAGAGAGAGGAAGTG and AGAGGAAGTGG, respectively. Recently, Pongubala et al. (57) showed that PU.1 can form an active enhancer complex in plasmacytoma B cells by interacting through its 43 amino acid PEST (proline-glutamic-serine-threonine) domain with an additional factor (NF-EM5). The two factors are bound to adjacent sites in the immunoglobulin κ-chain 3′ enhancer, but NF-EM5 is unable to bind on its own. Presumably, similar interactions might occur within the ADA gene HS III enhancer with other *ets* family members.

The coactivating region upstream of the enhancer core also contains several elements known to bind positively acting *trans* activators, including the sequence designated μE in Fig. 3C. μE was found to be a perfect match for the consensus E-box sequence initially observed as four related regions within the immunoglobulin μ heavy-chain enhancer (25). Subsequently, these sequences have been found to be distinct in their factor binding (60). Three homologous sequences have also been observed in the immunoglobulin κ light-chain enhancer (54). These sequences are reported to bind ubiquitous factors (60), such as the E12 and E47 products of the E2A locus.

To summarize, there is an essential core region of the HS III enhancer, and we have detected two factors that make contact with the sequences that are present in the core. The nucleotide sequence of the core enhancer is not homologous to sequences of previously reported elements. Expression of factors in their active state that bind to these elements occurs in a specific stage of T-cell differentiation, as exemplified by the MOLT 4 cell line. Specificity is evident from the lack of significant functional enhancer activity by the polymerized core in CEM and Raji lymphoid cells, previously shown to transcribe the ADA gene at much lower rates than do MOLT 4 cells (49). At least three additional factors interact with sequences that bilaterally flank the core to generate a strong enhancer domain capable of strongly activating the ADA promoter. Our current hypothesis to explain the greater than 100-fold variation in ADA specific activity in different lymphoid cells is that the expression of factors interacting with the flanking coactivating elements occurs in a variety of stages of differentiation of lymphoid cells, but the expression of the active forms of ADA-NF1 and/or ADA-NF2 that bind to the central enhancer core is more restricted to specific stages of T-cell differentiation in which ADA is abundant.

Consistent enhancer function in transgenic mouse thymus. Results obtained for transgenic mice implicated two additional segments whose effects were not otherwise observable in transient transfection expression assays. In addition, we found that strong negative regulatory effects observed in transient transfection expression assays had no effect on the level or pattern of expression in transgenic mice. Our understanding of how these extra-enhancer segments function is presently quite limited, but their importance in transgenic mice is quite clear, as discussed below.

HS II region. The HS II region contains several *cis*-regulatory element consensus sequences that are also present in HS III, including μE, AP-1, and PU.1, and an NF-κB site, but there are no sites similar to ADA-NF1, ADA-NF2, or TCF-1α. Deletion of the HS II region from the active regulatory domain led to variable and low expression among independent lines of transgenic mice (Fig. 8, lines 19 to 21). This finding indicates that in the absence of the HS II region, HS III flanked with the downstream facilitator region is not consistently expressed. What is not clear, however, is whether in a larger group of transgenic mice, occasional

insertion sites might allow high-level expression. If in a larger series, some transgenic animals exhibited occasional high-level expression, the conclusion would be first that the HS II region facilitates the establishment of enhancer function at HS III, but without a high degree of specificity. If no transgenic lines are expressed at as high a level, the conclusion would be that HS II provides a specific coenhancer function when combined with the HS III enhancer. If this were the case, HS II could either be making direct interactions with a complex at the promoter region or be interacting only with HS III. In contrast, in the three cell lines that we have transfected, HS II does not have detectable independent enhancer activity (Fig. 1), nor do we see evidence of an enhancer-complementing activity of HS II when BAL-31 deletants were assayed by transfection (Fig. 2). On this basis, we suggest that HS II functions as a coactivator of HS III in transgenic mice and is also necessary to ensure consistent high level expression.

A downstream DNase I hypersensitivity facilitator segment. Transgenic mice that lacked the nonhypersensitive segment of DNA downstream of HS III exhibited substantial variability of copy-adjusted CAT expression (Fig. 8, lines 6 to 18). In contrast, presence of the downstream nonhypersensitive segment largely prevented position-dependent variation of expression. Some of the transgenic mouse lines made without this segment exhibited higher-level reporter expression. This finding suggests that given facilitation by their insertion site, the *cis* elements of HS II and HS III are sufficient to direct high-level expression. All of the higher-expressing lines occurred among mice with low numbers of transgene copies. However, low copy number did not ensure high-level expression, and several of the lines with higher copy numbers exhibited expression up to 10% of the expected level. Thus, *cis* information can be supplied by the insertion site that makes up for the lack of the downstream segment. This facilitative effect of some insertion sites appears able to be transmitted to only a limited number of transgene copies. Regardless of how the effect is transmitted, from the percentage of low-copy-number mice that exhibited high-level expression, it is clear that the additional *cis* information must be present in a significant number of genomic insertion sites.

We sought to analyze DNase I hypersensitivity at HS III in order to evaluate what impediment prevented function of the enhancer region when the downstream facilitator segment was absent. DNase I hypersensitivity was able to develop directly over the enhancer domain in relatively high expressing lines that lacked the downstream segment. However, it did not form as well, if at all, in low-expressing lines made from the same constructs. This finding implies that *cis* information necessary to specify where DNase I hypersensitivity is to be formed is associated with the hypersensitive site itself. It also implies that the downstream nonhypersensitive segment functions to ensure that hypersensitivity can be formed at HS III, and it is this effect that permits consistent gene-copy-proportional expression. When the formation of hypersensitivity at the HS III enhancer site is facilitated by a chromosomal insertion site, high-level expression from the shorter constructs is allowed.

Sequences responsible for the facilitative activity of the downstream nonhypersensitive segment correspond to a 630-bp sequence (Fig. 8, lines 13 to 18 versus lines 22 to 25) of which 550 bp is *Alu*-type repetitive DNA. The facilitator activity, however, is unlikely to be conferred by any random *Alu*-containing segment, because the incorporation of *Alu*-containing DNA from the 5′-flanking segment of the ADA gene into multicopy concatameric tandem arrays of transgenes (Fig. 8, lines 6 to 12) failed to confer this effect. We do not know whether the downstream facilitator can have any function of its own. Since specificity of the thymic regulatory region appears to be dictated by the HS III domain, the facilitator may act rather nonspecifically. Therefore, we speculate that the downstream segment may be facilitative of other enhancer specificities when combined in chimeric regulatory constructions. In this regard, it may be important to also note that our data indicate that the downstream facilitator acts on HS II and HS III rather than on HS III alone.

HS site formation does not dictate high-level enhancer function. It is unclear whether the *trans*-acting factors that interact with the enhancer domain and cause hypersensitivity to be formed at HS III are the same as those that determine enhancer activation. Given the need for downstream facilitator sequences, it is clear that the presence of HS III-binding factors alone is not sufficient to ensure that the hypersensitive configuration is adopted by HS III. However, because several chromosomal insertion sites allowed HS III to become hypersensitive in the absence of the facilitator region, the primary determinants of hypersensitive site formation are most likely associated with HS III itself. Nevertheless, the factors that cause high-level transactivation are not likely to be the same as those necessary for the development of hypersensitivity. The basis for this conclusion is that MOLT 4 cells and CEM cells exhibit virtually identical degrees of DNase I hypersensitivity at HS III (see Fig. 5 and 6 in reference 3). This finding suggests that the HS III is just as accessible in CEM cells as in MOLT 4 cells. In transfection assays, the HS III enhancer function was much weaker in CEM cells than in MOLT 4 cells. This was the case with both the intact domain (Fig. 1) and the polymerized enhancer core segment (Fig. 7). Similarly, MOLT 4 cells transcribe the ADA gene at a 10-fold-higher rate than do CEM cells (49). Thus, the degree of hypersensitivity of the domain is not determined by the level of enhancer function. Rather, hypersensitivity is permissive and required for the ability of the HS III domain to function as an enhancer. The formation of hypersensitivity is likely to be determined by a subset of factors different from those that determine the highest level of cell-type-specific expression.

The extent of *trans*-acting factor occupancy of HS sites is generally not known. Elgin has described an "open for business" model for the role of hypersensitivity (23), but it is unclear whether enhancer site hypersensitivity is necessary only for genes that are to be expressed or induced at a relatively high level. Thus, it might be considered surprising that the human ADA gene requires a high degree of chromatin structure for the proper control of its expression, despite producing an mRNA that comprises only 0.1% of total mRNA in human cortical thymocytes (1, 3). However, many examples of DNase I-hypersensitive regulatory regions have been identified in a variety of genes that are expressed at lower levels than β-globin genes, such as the genes for mouse and human Thy-1 (31, 68), human CD2 (37), chicken lysozyme (9), mouse CD3-δ (29), and brown fat mitochondrial uncoupling protein (12). It is thus likely that many control regions associated with comparatively low levels of gene transcription must nevertheless adopt a DNase I-hypersensitive configuration for efficient access to *trans* activators. Conversely, the β-globin LCR has been shown to become hypersensitive developmentally prior to acquisition of strong enhancer capability (27, 64, 66).

Does the facilitator segment act as a boundary element? The downstream facilitator segment may represent a new class of

genetic boundary elements. Two observations make us suspect that a boundary function is involved: (i) the region is non-DNase I hypersensitive and therefore is unlikely to bind conventional transactivators, and (ii) the segment facilitates position-independent enhancer function and thus seems to insulate the enhancer from the effects of the chromosome insertion site. The existence of boundary elements that are external to transcribed regions of eukaryotic genes has been inferred from direct observations of chromatin structure and of the effects of perturbed boundaries on gene expression (for a review, see reference 22). Kellum and Schedl (46) have demonstrated chromatin boundary function by nonhypersensitive DNA segments termed specialized chromatin structures (SCS and SCS') that flank the *Drosophila* heat shock locus 87A7. When bilaterally flanking a reporter gene that contained regulatory elements, SCS largely prevented the suppressive effects (position effect variegation) of some insertion sites. The potential complexity of factors that interact with boundary segments is suggested by identification of at least 20 distinct complementation groups in *Drosophila* cells that increase or decrease position effect variegation (75). Another class of boundary element that appears to generate segregation of regulatory domains has been implicated in the 5'-flanking region of the P2 promoter of the *Antennapedia* gene (11; reviewed in reference 22). Deletion of these domains causes regulatory segments of the P2 distal 5'-flanking DNA to function inappropriately. The alteration may be a failure to cordon off a distal enhancer from a proximal enhancer at specific stages of segmentation, suggesting an inducible type of boundary element. The alpha-fetoprotein silencer could be thought of similarly (14, 65). A third class of boundary function has been demonstrated for a segment of DNA (A element) that binds to nuclear scaffolds and flanks the chicken lysozyme gene immediately upstream of an interval of general DNase I hypersensitivity. Stief et al. (61) have shown that the A element confers a high degree of position-independent and copy-proportional activity to an enhancer/promoter reporter construct in stable transformants of a cell line. This finding suggests that anchoring to structural components of the nucleus is associated with function of regulatory boundary domains.

The human β-globin LCRs may function by a somewhat distinct mechanism (for a review, see reference 26). Three of the four HS in the cluster are not enhancers in transient expression assays, but all confer position-independent copy-proportional expression in stable transformants of erythroleukemic cells and in transgenic mice (28, 50, 56, 58, 63). This finding suggests that their activity requires a chromatin-associated structure or activity. At variance from the ADA gene thymic regulatory region, small (200- to 300-bp) fragments corresponding to the limits of the HS engender these activities. Like each of the locus-flanking boundary elements discussed above, the β-globin LCRs function outside of the transcriptional units. In contrast to the ADA gene thymic regulatory region, it is not known whether the β-globin LCR could function from within an intron. Similarly, it is also not known whether transcription can cross an A element or an SCS.

Conclusion. Figure 10 summarizes in a hypothetical scheme an organized set of local and regional interactions among elements and segments of the ADA gene thymic regulatory region. Within the critical 28-bp enhancer core, at least two factors bind to adjacent sequences, and point mutations of either site destroy the ability of the polymerized core to act as an enhancer. The core is active in a highly cell-type-specific fashion, and its activity is strongly complemented by bilaterally flanking elements encompassing 200 to 300 bp corresponding to the limits of HS III. We do not yet know how the core and core-flanking sequences synergize within HS III. This may occur in a localized fashion as indicated, or there may be distinct interactions of core and core-flanking binding factors with a factor(s) bound to the promoter region. A similar question can be posed for the function of HS II: does it interact with the promoter directly, or does it act at HS III? As discussed above, it is likely but unproved to act at HS III. Finally, our most provocative finding is that there is a facilitator segment downstream of the two hypersensitive sites that prevents line to line variation in gene expression per transgene copy. Its presence seems to ensure that HS III can become hypersensitive. To our knowledge, the failure of a regulatory region to function in transgenes has never been associated with its inability to become DNase I hypersensitive. The facilitator segment is highly *Alu* sequence rich, but it could not be substituted for by an *Alu*-containing segment from the distal 5'-flanking segment of the promoter region. Flanking facilitators may be difficult to identify, but they appear critical for the ability of some regulatory regions to gain control of a chromatin domain. We postulate that the ADA gene first-intron facilitator interacts with a novel class of nuclear components that allows the enhancer domain to in turn be more easily restructured by cell-type-specific *trans* activators. In providing a chromosomal context, the facilitator assembly could be interacting with the HS III enhancer domain or it could be insulating the domain from negative effects of some chromosomal insertion sites. In either event, a facilitative mechanism may be a key to understanding how nucleosomes can be replaced with *trans* activators at a regulatory region as a function of developmental gene activation, even in the case of a gene expressed at low levels such as the ADA gene in thymocytes.

ACKNOWLEDGMENTS

This work was supported by grant 1-1205 from the March of Dimes (B.J.A.) and by NIH grants HD 19919 (J.J.H.) and GM 42969 (D.A.W.). R.N.S. was supported by training grant HL 07527.

We are grateful for technical assistance by John Gustofson, Jennifer Ruschulte, and Leah Sparks.

REFERENCES

1. **Adrian, G., and J. J. Hutton.** 1983. Adenosine deaminase mRNAs in lymphoblast cell lines derived from leukemic patients and patients with hereditary adenosine deaminase deficiency. J. Clin. Invest. **71**:1649–1660.
2. **Angel, P., M. Imagawa, R. Chiu, B. Stein, R. J. Imbra, H. J. Rahmsdorf, C. Jonat, P. Herrlich, and M. Karin.** 1987. Phorbol ester-inducible genes contain a common cis element recognized by a TPA-modulated trans-acting enhancer. Cell **49**:729–739.
3. **Aronow, B., D. Lattier, R. Silbiger, M. Dusing, J. Hutton, G. Jones, J. Stock, J. McNeish, S. Potter, D. Witte, and D. Wiginton.** 1989. Evidence for a complex regulatory array in the first intron of the human adenosine deaminase gene. Genes Dev. **3**:1384–1400.
4. **Baldwin, A. S.** 1990. Mobility shift DNA-binding assay using gel electrophoresis, p. 12.2.1. *In* F. M. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.), Current protocols in molecular biology. Wiley Interscience, New York.
5. **Beckmann, H., L.-K. Su, and T. Kadesch.** 1990. TFE3: a helix-loop-helix protein that activates transcription through the immunoglobulin enhancer μE3 motif. Genes Dev. **4**:167–179.
6. **Benezra, R., R. L. Davis, D. Lockshon, D. L. Turner, and H. Weintraub.** 1990. The protein Id: a negative regulator of helix-loop-helix proteins. Cell **61**:49–59.

7. **Bhat, N. K., K. L. Komischlies, S. Fujiwara, R. J. Fisher, B. J. Mathieson, T. A. Gregorio, H. A. Young, J. W. Kasik, K. Ozato, and T. S. Papas.** 1989. Expression of *ets* genes in thymocyte subsets. J. Immunol. **142**:672–678.
8. **Blom vanAssendelft, G., O. Hanscombe, F. Grosveld, and D. R. Greaves.** 1989. The β-globin dominant control region activates homologous and heterologous promoters in a tissue-specific manner. Cell **56**:969–977.
9. **Bonifer, C., M. Vidal, F. Grosveld, and A. E. Sippel.** 1990. Tissue specific and position independent expression of the complete gene domain for chicken lysozyme in transgenic mice. EMBO J. **9**:2843–2848.
10. **Borzy, M. S., H. Schulte-Wissermann, E. Gilbert, S. D. Horowitz, J. Pellet, and R. Hong.** 1979. Thymic morphology in immunodeficiency diseases: results of thymic biopsies. Clin. Immunol. Immunopathol. **12**:31–51.
11. **Boulet, A. M., and M. P. Scott.** 1988. Control elements of the P2 promoter of the Antennapedia gene. Genes Dev. **2**:1600–1614.
12. **Boyer, B. B., and L. P. Kozak.** 1991. The mitochondrial uncoupling protein gene in brown fat: correlation between DNase I hypersensitivity and expression in transgenic mice. Mol. Cell. Biol. **11**:4147–4156.
13. **Brinster, R. L., J. M. Allen, R. R. Behringer, R. E. Gelinas, and R. D. Palmiter.** 1988. Introns increase transcriptional efficiency in transgenic mice. Proc. Natl. Acad. Sci. USA **85**:836–840.
14. **Camper, S. A., and S. M. Tilghman.** 1989. Postnatal repression of the α-fetoprotein is enhancer independent. Genes Dev. **3**:537–546.
15. **Chechik, B., W. P. Schrader, and J. Minowada.** 1981. An immunomorphologic study of the distribution of adenosine deaminase distribution in human thymus tissue, normal thymocytes, and hematopoietic cell lines. J. Immunol. **126**:1003–1007.
16. **Chinsky, J. M., V. Ramamurthy, W. C. Fanslow, D. E. Ingolia, M. R. Blackburn, K. T. Shaffer, H. R. Higley, J. J. Trentin, F. B. Rudolph, T. B. Knudsen, and R. E. Kellems.** 1990. Developmental expression of adenosine deaminase in the upper alimentary tract of mice. Differentiation **42**:172–183.
17. **Chodosh, L. S.** 1990. Methylation interference assay for analysis of DNA-protein interactions, p. 12.3.1. *In* F. M. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.), Current protocols in molecular biology. Wiley Interscience, New York.
18. **Choi, T., M. Huang, C. Gorman, and R. Jaenisch.** 1991. A generic intron increases gene expression in transgenic mice. Mol. Cell. Biol. **11**:3070–3074.
19. **Chung, S., and R. P. Perry.** 1989. Importance of introns for expression of mouse ribosomal protein gene rpL32. Mol. Cell. Biol **9**:2075–2082.
20. **Collis, P., M. Antoniou, and F. Grosveld.** 1990. Definition of the minimal requirements within the human β-globin gene and the dominant control region for high level expression. EMBO J. **9**:233–240.
21. **Dignam, J. D., R. M. Lebovitz, and R. G. Roeder.** 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res. **11**:1475–1489.
22. **Eissenberg, J. C., and S. C. R. Elgin.** 1991. Boundary functions in the control of gene expression. Trends Genet. **7**:335–340.
23. **Elgin, S. C. R.** 1988. The formation and function of DNase I hypersensitive sites in the process of gene activation. J. Biol. Chem. **263**:19259–19262.
24. **Emerson, B. M., C. D. Lewis, and G. Felsenfeld.** 1985. Interaction of specific nuclear factors with the nuclease-hypersensitive region of the chicken adult beta-globin gene: nature of the binding domain. Cell **41**:21.
25. **Ephrussi, A., G. M. Church, S. Tonegawa, and W. Gilbert.** 1985. B-lineage-specific interactions of an immunoglobulin enhancer with cellular factors in vivo. Science **227**:134–140.
26. **Felsenfeld, G.** 1992. Chromatin as an essential part of the transcriptional machinery. Nature (London) **355**:219–224.
27. **Forrester, W. C., S. Takegawa, T. Papayannopoulou, G. Stamatoyannoupoulos, and M. Groudine.** 1987. Evidence for a locus activation region: the formation of developmentally stable hypersensitive sites in globin-expressing hybrids. Nucleic Acids Res. **15**:10159–10177.
28. **Fraser, P., J. Hurst, P. Collis, and F. Grosveld.** 1990. DNase I hypersensitive sites 1, 2, and 3 of the human β-globin dominant control region direct position-independent expression. Nucleic Acids Res. **18**:3503–3507.
29. **Georgopoulos, K., P. vandenElsen, E. Bier, A. Maxam, and C. Terhorst.** 1988. A T cell-specific enhancer is located in a DNase I-hypersensitive area at the 3′ end of the CD3-delta gene. EMBO J. **7**:2401–2407.
30. **Giese, K., J. Cox, and R. Grosschedl.** 1992. The HMG domain of lymphoid enhancer factor 1 bends DNA and facilitates assembly of functional nucleoprotein structures. Cell **69**:185–195.
31. **Gordon, J. W., P. G. Chesa, H. Nishimura, W. J. Rettig, J. E. Maccari, T. Endo, E. Seravalli, T. Seki, and J. Silver.** 1987. Regulation of thy-1 gene expression in transgenic mice. Cell **50**:445–452.
32. **Gorman, C.** 1985. High efficiency gene transfer into mammalian cells, p. 143–190. *In* D. M. Glover (ed.), DNA cloning, vol. 2. IRL Press, Washington, D.C.
33. **Gorman, C., L. F. Moffat, and B. H. Howard.** 1982. Recombinant genes which express chloramphenicol acetyltransferase in mammalian cells. Mol. Cell. Biol. **2**:1044–1051.
34. **Gorman, C. M., G. T. Merlino, M. C. Willingham, I. Pastan, and B. H. Howard.** 1982. The Rous sarcoma virus long terminal repeat is a strong promoter when introduced into a variety of eukaryotic cells by DNA-mediated transfection. Proc. Natl. Acad. Sci. USA **79**:6777.
35. **Gosseye, S., N. Diebold, C. Griscelli, and C. Nezelof.** 1983. Severe combined immunodeficiency disease: a pathological analysis of 26 cases. Clin. Immunol. Immunopathol. **29**:58–77.
36. **Gottesdiener, K. M., B. A. Karpinski, T. Lindsten, J. L. Strominger, N. H. Jones, C. B. Thompson, and J. M. Leiden.** 1988. Isolation and structural characterization of the human 4F2 heavy-chain gene, an inducible gene involved in T-lymphocyte activation. Mol. Cell. Biol. **8**:3809–3819.
37. **Greaves, D. R., F. D. Wilson, G. Lang, and D. Kioussis.** 1989. Human CD2 3′-flanking sequences confer high-level, T cell-specific, position-independent gene expression in transgenic mice. Cell **56**:979–986.
38. **Grosveld, F., G. B. van Assendelft, D. R. Greaves, and G. Kollias.** 1987. Position-independent, high-level expression of the human β-globin gene in transgenic mice. Cell **51**:975–985.
39. **Gunther, C. V., J. A. Nye, R. S. Bryner, and B. J. Graves.** 1990. Sequence specific binding of the proto-oncogene ets-1 defines a transcriptional activator sequence within the long terminal repeat of the Moloney murine sarcoma virus. Genes Dev. **4**:667–679.
40. **Hirschhorn, R., V. Roegner-Maniscalco, L. Kuritsky, and F. S. Rosen.** 1981. Bone marrow transplantation only partially restores purine metabolites to normal in adenosine deaminase-deficient patients. J. Clin. Invest. **68**:1387.
41. **Ho, I. C., N. K. Bhat, L. R. Gottschalk, T. Lindsten, C. B. Thompson, T. S. Papas, and J. M. Leiden.** 1990. Sequence-specific binding of human ets-1 to the T-cell receptor α locus. Science **250**:814–818.
42. **Ho, I. C., L. H. Yang, G. Morle, and J. M. Leiden.** 1989. A T-cell-specific transcriptional enhancer element 3′ of Ca in the human T-cell receptor a locus. Proc. Natl. Acad. Sci. USA **86**:6714–6718.
43. **Hogan, B., F. Constantini, and E. Lacy.** 1986. Manipulating the mouse embryo. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
44. **Jones, N. C., P. W. J. Rigby, and E. B. Ziff.** 1988. Trans-acting protein factors and the regulation of eukaryotic transcription: lessions from studies on DNA tumor viruses. Genes Dev. **2**:267–281.
45. **Karpinski, B. A., L. Yang, P. Cacheris, G. D. Morle, and J. M. Leiden.** 1989. The first intron of the 4F2 heavy chain gene contains a transcriptional enhancer element that binds multiple nuclear proteins. Mol. Cell. Biol. **9**:2588–2597.
46. **Kellum, R., and P. Schedl.** 1991. A position-effect assay for boundaries of higher order chromosomal domains. Cell **64**:941–950.

47. **Klemsz, M. J., S. R. McKercher, A. Celada, C. VanBeveren, and R. A. Maki.** 1990. The macrophage and B cell-specific transcription factor PU.1 is related to the ets oncogene. Cell **61**:113–124.
48. **Lang, G., D. Wotton, M. J. Owen, W. A. Sewell, M. H. Brown, D. Y. Mason, M. J. Crumpton, and D. Kioussis.** 1988. The structure of the human CD2 gene and its expression transgenic mice. EMBO J. **7**:1675–1682.
49. **Lattier, D. L., J. C. States, J. J. Hutton, and D. A. Wiginton.** 1989. Cell type specific transcriptional regulation of the human adenosine deaminase gene. Nucleic Acids Res. **17**:1061–1076.
50. **Lowrey, C. H., D. M. Bodine, and A. W. Nienhuis.** 1992. Mechanism of DNase I hypersensitive site formation within the human globin locus control region. Proc. Natl. Acad. Sci. USA **89**:1143–1147.
51. **Ma, D. D. F., T. A. Sylwestrowicz, S. Granger, M. Massaia, R. Franks, G. Janossy, and A. V. Hoffbrand.** 1982. Distribution of terminal deoxynucleotidyl transferase and purine degradative and synthetic enzymes in subpopulations of human thymocytes. J. Immunol. **129**:1430–1435.
52. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring, N.Y.
53. **McKnight, S. L., E. R. Gavis, and R. Kingsbury.** 1981. Analysis of transcriptional regulatory signals of the HSV thymidine kinase gene: identification of an upstream control region. Cell **25**:385.
54. **Meyer, K. B., M. J. Sharpe, M. A. Surani, and M. S. Neuberger.** 1990. The importance of the 3′-enhancer region in immunoglobulin κ gene expression. Nucleic Acids Res. **18**:5609–5615.
55. **Murre, C., P. S. McCaw, H. Vaessin, M. Caudy, L. Y. Jan, Y. N. Jan, C. V. Cabrera, J. N. Buskin, S. D. Hauschka, A. B. Lassar, H. Weintraub, and D. Baltimore.** 1989. Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. Cell **58**:537–544.
56. **Philipsen, S., D. Talbot, P. Fraser, and F. Grosveld.** 1990. The β-globin dominant control region: hypersensitive site 2. EMBO J. **9**:2159–2167.
57. **Pongubala, J. M. R., S. Nagulapalli, M. J. Klemsz, S. R. McKercher, R. A. Maki, and M. L. Atchison.** 1992. PU.1 recruits a second nuclear factor to a site important for immunoglobulin κ 3′ enhancer activity. Mol. Cell. Biol. **12**:368–378.
58. **Pruzina, S., O. Hanscombe, D. Whyatt, F. Grosveld, and S. Philipsen.** 1991. Hypersensitive site 4 of the human β-globin locus control region. Nucleic Acids Res. **19**:1413–1419.
59. **Ryan, T. M., R. R. Behringer, N. C. Martin, T. M. Townes, R. D. Palmiter, and R. L. Brinster.** 1989. A single erythroid-specific DNase I super-hypersensitive site activates high levels of human β-globin gene expression in transgenic mice. Genes Dev. **3**:314–324.
60. **Sen, R., and D. Baltimore.** 1987. Nuclear factors interacting with the immunoglobulin heavy and κ enhancers, p. 83–101. *In* D. Granner, M. G. Rosenfeld, and S. Chang (ed.), Transcriptional control mechanisms. Alan R. Liss, Inc., New York.
61. **Stief, A., D. M. Winter, W. H. Sratling, and A. E. Sippel.** 1989. A nuclear attachment element mediates elevated and position-independent gene activity. Nature (London) **341**:343–345.
62. **Sun, X.-H., N. G. Copeland, N. A. Jenkins, and D. Baltimore.** 1991. Id proteins Id1 and Id2 selectively inhibit DNA binding by one class of helix-loop-helix proteins. Mol. Cell. Biol. **11**:5603–5611.
63. **Talbot, D., S. Philipsen, P. Fraser, and F. Grosveld.** 1990. Detailed analysis of the site 3 region of the human β-globin dominant control region. EMBO J. **9**:2169–2178.
64. **Tuan, D., W. Solomon, Q. Li, and I. M. London.** 1985. The "β-like globin" gene domain in human erythroid cells. Proc. Natl. Acad. Sci. USA **82**:6384–6388.
65. **Vacher, J., and S. M. Tilghman.** 1990. Dominant negative regulation of the mouse α-fetoprotein gene in adult liver. Science **250**:1732–1735.
66. **van Assendelft, G. B., O. Hanscombe, F. Grosveld, and D. R. Greaves.** 1989. The β-globin dominant control region activates homologous and heterologous promoters in a tissue-specific manner. Cell **56**:969–977.
67. **van de Wetering, M. M., M. Oosterwegel, D. Dooijes, and H. Clevers.** 1991. Identification and cloning of TCF-1, a T-lymphocyte-specific transcription factor containing a sequence specific HMG box. EMBO J. **10**:123–132.
68. **Vidal, M., R. Morris, F. Grosveld, and E. Spanopoulou.** 1990. Tissue-specific control elements of the Thy-1 gene. EMBO J. **9**:833–840.
69. **Waterman, M. L., W. H. Fischer, and K. A. Jones.** 1990. A thymus-specific member of the HMG protein family regulates the human T-cell receptor Cα enhancer. Genes Dev. **5**:656–669.
70. **Waterman, M. L., and K. A. Jones.** 1990. Purification of TCF-1α, a T-cell specific transcription factor that activates the T-cell receptor Cα gene enhancer in a context-dependent manner. New Biol. **2**:621–636.
71. **Wiginton, D. A., D. J. Kaplan, J. C. States, A. L. Akeson, C. M. Perme, I. J. Bilyk, A. J. Vaughn, D. J. Lattier, and J. J. Hutton.** 1986. Complete sequence and structure of the gene for human adenosine deaminase. Biochemistry **25**:8234–8244.
72. **Wilson, R. B., M. Kiledjian, R. Benezra, P. Zwollo, S. M. Dymecki, S. V. Desiderio, and T. Kadesch.** 1991. Repression of immunoglobulin enhancers by the helix-loop-helix protein Id: implications for B-cell development. Mol. Cell. Biol. **11**:6185–6191.
73. **Winoto, A., and D. Baltimore.** 1989. A novel, inducible and T cell-specific enhancer located at the 3′ end of the T cell receptor a locus. EMBO J. **8**:729–733.
74. **Witte, D. P., D. A. Wiginton, J. J. Hutton, and B. J. Aronow.** 1991. Coordinate developmental regulation of purine catabolic enzyme expression in gastrointestinal and post-implantation reproductive tracts. J. Cell Biol. **115**:179–190.
75. **Wustmann, G., J. Szidonya, H. Taubert, and G. Reuter.** 1989. The genetics of position-effect variegation modifying loci in Drosophila melanogaster. Mol. Gen. Genet. **217**:520–527.