

Enhancer Complexes Located Downstream of Both Human Immunoglobulin C α Genes

By Frederick C. Mills, Nagaradona Harindranath, Mary Mitchell, and Edward E. Max

From the Laboratory of Cell and Viral Regulation, Division of Hematologic Products, Center for Biologics Evaluation and Research, Food and Drug Administration, Bethesda, Maryland 20892

Summary

To investigate regulation of human immunoglobulin heavy chain expression, we have cloned DNA downstream from the two human C α genes, corresponding to the position in the mouse IgH cluster of a locus control region (LCR) that includes an enhancer which regulates isotype switching. Within 25 kb downstream of both the human immunoglobulin C α 1 and C α 2 genes we identified several segments of DNA which display B lymphoid-specific DNase I hypersensitivity as well as enhancer activity in transient transfections. The corresponding sequences downstream from each of the two human C α genes are nearly identical to each other. These enhancers are also homologous to three regions which lie in similar positions downstream from the murine C α gene and form the murine LCR. The strongest enhancers in both mouse and human have been designated HS12. Within a 135-bp core homology region, the human HS12 enhancers are ~90% identical to the murine homolog and include several motifs previously demonstrated to be important for function of the murine enhancer; additional segments of high sequence conservation suggest the possibility of previously unrecognized functional motifs. On the other hand, certain functional elements in the murine enhancer, including a B cell-specific activator protein site, do not appear to be conserved in human HS12. The human homologs of the murine enhancers designated HS3 and HS4 show lower overall sequence conservation, but for at least two of the functional motifs in the murine HS4 (a κ B site and an octamer motif) the human HS4 homologs are exactly conserved. An additional hypersensitivity site between human HS3 and HS12 in each human locus displays no enhancer activity on its own, but includes a region of high sequence conservation with mouse, suggesting the possibility of another novel functional element.

The regulation of human immunoglobulin heavy chain gene expression is incompletely understood, despite clinically significant conditions in which specific isotypes are inappropriately up- or downregulated, e.g., allergies due to inappropriate IgE response, and various forms of immunodeficiency associated with low IgA expression. Clearly, cytokines and interactions between B and T cells play a role in regulating isotype switching, and *cis* elements in the IgH gene locus which mediate these effects have been documented in the murine and human promoters of the sterile transcripts associated with each heavy chain constant region gene (1, 2). However, in the mouse an additional control region which contributes to regulation of isotype switching has been reported to lie downstream from C α , and the corresponding region of the human heavy chain locus has not yet been investigated.

The existence of a regulatory region downstream from murine C α was originally inferred when it was found that plasmacytomas which had undergone spontaneous deletions of the only heavy chain enhancer then known, which lies in the intron between JH and C μ , nevertheless remained capable of high level immunoglobulin secretion (3–6). Conversely, a myeloma subclone which retained the intronic enhancer but lost a segment of DNA downstream from the murine C α gene was found to have markedly reduced its heavy chain gene expression (7). A systematic search in the homologous region of the rat heavy chain locus revealed an enhancer (8), and a homologous mouse enhancer designated 3' α E was found soon after (9, 10) positioned ~16 kb downstream from C α . The mouse and rat 3' α E segments lie in opposite orientations and are flanked by inverted repeats (9). In addition to the 3' α E, Matthias and Baltimore also reported a weak enhancer in mouse lying only 4 kb downstream from C α (Fig. 1 and reference 11).

F.C. Mills and N. Harindranath both made substantial contributions to this work.

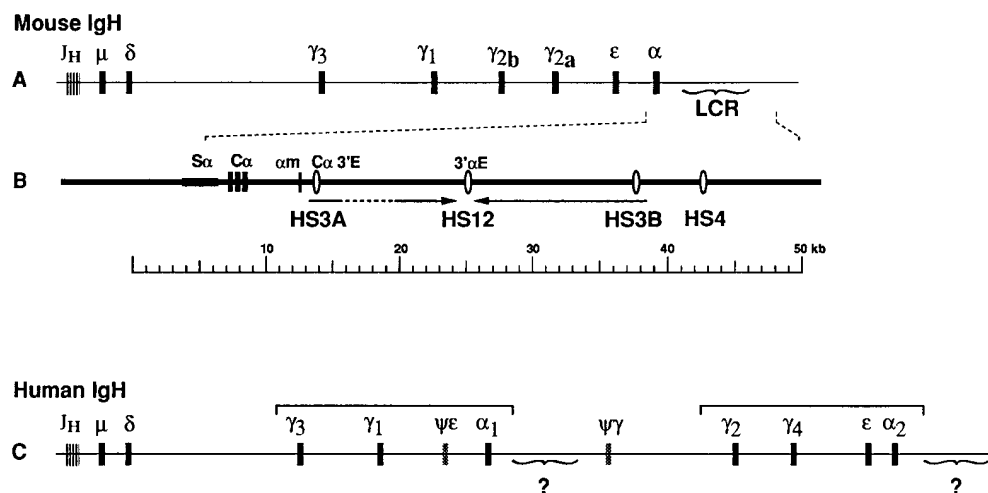


Figure 1. Comparison of IgH loci of mouse and human. Line A shows a map of the murine IgH locus, from which the region downstream from C α is expanded in line B. The murine enhancers designated C α 3'E (11) and 3' α E (9) are shown as vertical ovals, along with the DNase I hypersensitivity site designations (12). We have distinguished the two copies of HS3 sequence as HS3A and HS3B; these are included in a large palindrome (arrows) that flanks HS12, according to the sequence analysis of Chauveau and Cogné (13). Line C shows the human IgH locus, illustrating the γ - γ - ϵ - α duplication units (brackets) and the possibility of two regions homologous to the murine LCR.

More recently, Madisen and Groudine (12) analyzed B cell-specific DNase I hypersensitivity downstream from C α and identified four hypersensitive sites. HS1 and HS2 fall in the previously described 3' α E, whereas HS3 and HS4 lie further downstream and identify two new regions with somewhat weaker enhancer activity in transient transfection assays. The HS3 sequence is almost identical to that of the enhancer described by Matthias and Baltimore but has an inverted orientation. This reflects the fact that the sequence surrounding the HS12-3' α E is present in the mouse in a long inverted repeat which includes HS3 sequences at both ends (Fig. 1 and reference 13). When constructs containing HS3, HS12, and HS4 linked to a reporter gene were transfected into a B cell line, subsequently isolated stable transfectants were found to express the reporter gene in a position-independent manner. This suggested that the three enhancer sequences (HS12, HS3, and HS4) acted together as a locus control region (LCR)¹. LCRs, first defined in the globin locus (14), activate large domains of chromatin in vivo (100 kb in the human β globin locus), and, as components of DNA constructs in transgenic mice, support gene expression proportional to the number of integrated copies. In contrast, integrated gene constructs lacking LCR sequences are variably expressed, depending on the positions of integration. LCRs typically contain several DNase I hypersensitive sites, which often represent DNA with enhancer activity. In addition to the LCRs found in the β globin and mouse IgH loci, LCRs have also been described as associated with macrophage-specific lysozyme, CD2, and α/β TCR loci (15).

Analyses of the regulatory regions downstream from murine C α have identified several motifs which bind specific transcription factors to mediate different aspects of regulation of enhancer function. The 3' α E has been found to ac-

tivate transcription strongly in plasmacytomas, but only weakly in earlier B lymphoid cells. Part of this developmental change is attributable to a motif known as E5, which matches the "E-box" consensus binding site (CANNTG) characteristic for members of the basic helix-loop-helix family of transcription factors. The contribution of the E5 site to enhancer activity is inhibited in early stages of development by the dominant negative nuclear regulator Id3, which is expressed in early B lineage cells but downregulated in plasma cells (16). At least four other motifs in the 3' α E have been reported to contribute to enhancer activity specifically in plasmacytomas, motifs whose contribution in B cells is inhibited by BSAP (the B cell-specific activator protein), which disappears as B cells mature to plasma cells. These sites include α P (17), the octamer motif (ATG-CAAAT; reference 18), a κ B-like site (16), and a G-rich sequence (19). In B cells, BSAP prevents the binding of the transcriptional activator NF- α P to the α P site, and causes the octamer, G-rich, and κ B-like motifs to exert an active repressive influence on transcription (17, 19–21).

Apart from the motifs mediating upregulation of the 3' α E during maturation to plasma cells, a response element in the enhancer for activation induced by B cell receptor cross-linking has been traced to partially overlapping sites for the ETS family member Elf-1 and for members of the AP-1 transcription factor family (22). Two other motifs in the enhancer have been proposed to contribute to its regulation, but are less well documented: the μ E1 and the μ B motifs, which were first noted in the rat 3' α E and which are partially conserved in mouse. The HS3 and HS4 enhancer regions of mouse have been studied in less detail, but the HS4 enhancer apparently contains functional Oct-1 and BSAP binding sites (23).

A role for the 3' α E in isotype switching was revealed by experiments in which this region was replaced by a neomycin resistance gene through homologous recombination in embryonic stem cells which were then used to reconstitute the B cell population in RAG-2 knockout mice. The re-

¹Abbreviations used in this paper: BAC, bacterial artificial chromosome; BSAP, B cell-specific activator protein; HSE, heat shock element; HSTF, heat shock transcription factors; LCR, locus control region.

sulting B cells showed normal V(D)J recombination but marked deficiencies in switching to IgG2a, IgG2b, IgG3, and IgE in vitro, whereas expression of IgM and IgG1 was normal (24). This observation suggests that the enhancer exerts isotype-specific effects on switch recombination, possibly through its regulation of germline transcription of the different isotypes before switch recombination.

Because the enhancer regions lying downstream from the mouse $C\alpha$ gene have been found to be important for heavy chain gene expression and isotype switching, there has been considerable interest in determining how homologous regions might regulate immunoglobulin gene expression in humans. The human heavy chain locus includes two γ - γ - ϵ - α segments (25, 26), apparently the product of a large duplication in the primate lineage (27). Isotypes from the upstream duplication, comprising the γ 3- γ 1- $\psi\epsilon$ - α 1 constant region genes, are generally expressed at a much higher level than those of the downstream γ 2- γ 4- ϵ - α 2 duplication. The existence of two $C\alpha$ genes in humans suggests the possibility that two 3' enhancer complexes may regulate the locus, one downstream from each $C\alpha$ gene; differences in these complexes could contribute to the differential regulation of the two duplications. Moreover, individuals who have a $\psi\epsilon$ - α 1- $\psi\gamma$ deletion on one chromosome show reduced expression of the downstream γ 2 and γ 4 genes on that chromosome, indicating that this deletion may have removed a region which exerts distal control over at least some of the human heavy chain genes (28). Finally, the possibility that there are two 3' enhancer complexes makes the human IgH locus an attractive candidate for study because there could be interactions between two adjacent LCRs, a situation which has not been described in any other system.

Several laboratories have attempted to characterize sequences lying downstream from the two human $C\alpha$ genes, but technical difficulties have impeded this work. Gene walking downstream from the $C\alpha$ genes has been difficult, apparently because of a segment of 20-bp tandem repeats which lies almost immediately downstream of the most 3' exon of both $C\alpha$ genes. These tandem repeats, described independently by three laboratories (29–31), include the sequence GATC recognized by the isoschizomer restriction enzymes *Sau*3A and *Mbo*I. Since commercial human DNA libraries in λ phage have been constructed using genomic DNA fragments generated by partial *Mbo*I/*Sau*3A digestion, the repeated *Sau*3A sites downstream of the $C\alpha$ genes make it unlikely that library clones isolated by hybridization to $C\alpha$ probes will contain DNA downstream from the repeats.

As an initial step toward defining the role of global control regions in the activation of the human IgH genes, our laboratory has sought to determine: (a) are enhancer complexes located downstream from the human $C\alpha$ genes; and (b) how do any such human enhancers correspond to the regulatory sequences described downstream from the murine $C\alpha$ gene? Towards this end we have successfully applied several strategies other than gene walking from $C\alpha$ to obtain DNA clones extending downstream from the human

$C\alpha$ genes, enabling us, by functional analysis and sequencing, to characterize enhancer regions 3' of each $C\alpha$ gene homologous to those of the mouse HS12, HS3, and HS4.

Materials and Methods

Cloning of 3' α Regions from Human Genomic DNA. To obtain DNA between $C\alpha$ 1 and the previously reported $\psi\gamma$ pseudogene which lacks associated $S\gamma$ sequences, we initially sought clones containing the pseudogene, which should hybridize to a $C\gamma$ probe but not to an $S\gamma$ probe. We screened a commercial library of partial *Mbo*I-digested human placental DNA in the λ FixII phage vector (Stratagene, La Jolla, CA) with a $C\gamma$ probe: probe f (Fig. 2), a 7-kb *Hind*III fragment, was isolated from a pBR322 plasmid clone originally derived from λ phage CH-Ig-H-g-11(32) (Health Science Research Resources Bank, Osaka, Japan; e-mail: hsrhb@nihs.go.jp). $C\gamma$ + plaques were replated and duplicate plaque-lift filters were hybridized with the $C\gamma$ probe and an $S\gamma$ probe (1,100-bp *Kpn*I to *Pst*I fragment containing human $S\gamma$ 2; reference 33). Southern blots of *Bam*HI-digested DNA from eight candidate clones ($C\gamma$ +, $S\gamma$ -) were hybridized with a 32 P-labeled oligonucleotide Pseudogam-1 (see Table 1 for sequences of oligonucleotides) specific for the $\psi\gamma$ hinge region at 48°C in hybridization buffer (1 M NaCl, 0.1 M sodium phosphate, pH 7.0, 10% dextran sulfate, 10 mM EDTA, 1% SDS), followed by washing at 48°C in 1× SSC, 10 mM EDTA, 0.1% SDS. This analysis identified three overlapping $\psi\gamma$ + clones ($\psi\gamma$ -25, $\psi\gamma$ -2, and $\psi\gamma$ -38) extending over an ~30-kb region containing the $\psi\gamma$ gene (Fig. 2).

To obtain clones for the expected homologous region downstream of the $C\alpha$ 2 region, two probes derived from the farthest upstream $\psi\gamma$ clone ($\psi\gamma$ -25) were used to rescreen two duplicate plaque lifts of the λ FIXII library. Probe b (Fig. 2) was a 4.5-kb fragment extending from the 5' end of $\psi\gamma$ -25 to the first internal *Eco*RI site. Probe e was a 1.3-kb PCR fragment which lies ~2.5 kb upstream from $\psi\gamma$ and which was obtained by amplification from a $\psi\gamma$ -25 subclone containing a 10.5-kb *Eco*RI-*Xba*I fragment which extends from the farthest downstream *Eco*RI site in $\psi\gamma$ -25 to an artificial *Xba*I site at the 3' end of the phage insert. Primers used for this amplification were a reverse sequencing primer (Table 1) and *elkb*2, a primer based on sequence data from a $\psi\gamma$ -25 subclone. This probe contains a sequence homologous to a gene encoding the transcription factor *elk-1*, which is known to lie on the X chromosome (34). Clones EG3-6 and EG3-5 were selected for the following properties: they hybridized to probe e but not to a $C\gamma$ probe (as expected for DNA downstream of $C\alpha$ 2), and they had restriction maps distinct from that of the α 1 locus and that of the *elk-1* gene on the X chromosome, as established by genomic Southern blots of DNA from a mouse/human somatic hybrid cell line (NIGMS repository No. GM06318B; Coriell Institute for Medical Research, Camden, NJ) carrying the human X chromosome as its only human DNA. Clone H3E-1 was isolated on the basis of hybridization with probe b but not with a $C\gamma$ probe. By restriction maps and Southern blots, clones H3E-1, EG3-6, and EG3-5 were found to overlap with each other, forming a contig of ~30 kb.

From genomic Southern blots, the phage contigs downstream of the two $C\alpha$ genes were found to lie ~20 kb away from clones containing the respective $C\alpha$ genes. To bridge these gaps, a bacterial artificial chromosome (BAC) library of human genomic DNA (Genome Systems, Inc., St. Louis, MO) was screened using both a $C\alpha$ membrane exon probe (Fig. 2, probe a) made by amplification from the $C\alpha$ 1 plasmid GMA5 using primer 5'TM-A1

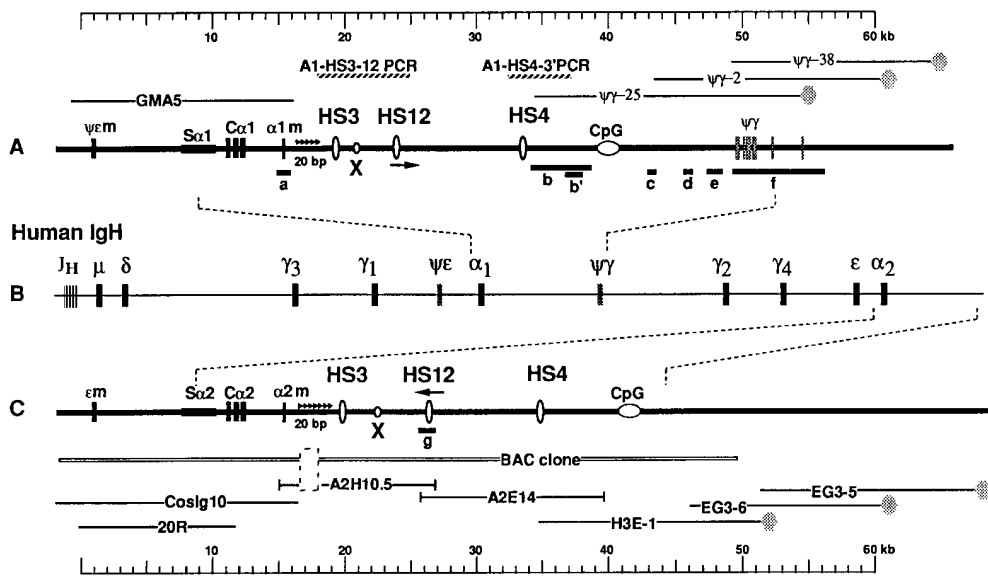


Figure 2. Regulatory loci downstream of human C α 1 and C α 2. Lines A and C, based on this study, show an expanded map of the region downstream of C α 1 and C α 2, respectively, as well as available DNA clones, which are shown above (α 1) or below (α 2) the line: phage clones are marked with diagrammatic phage heads, while the subclones of PCR-amplified segments A1-HS3-12 and A1-HS4-3' are drawn with hatched lines; and a BAC clone is drawn as a double line, containing a deletion (*dashed box*). Vertical ovals mark DNase I sites demonstrating enhancer activity and named according to the homologous murine HS sites. A series of small triangles identifies the 20-bp repeats located downstream from human C α genes. X marks the

position of a DNase I site which shows human/mouse sequence conservation. The position of a CpG island previously identified by Southern blotting is also shown (*oval*). The arrow under HS12 in line A indicates the orientation of this sequence, which is the same as that of the homologous mouse HS site, but opposite from the orientation of HS12 in the α 2 locus (line C). The thick black lines under the maps of lines A and C (*single lower case letters*) represent hybridization probes used in this study.

in combination with primer 3'TM-B (35) and a probe extending from the most 5' HindIII site in $\psi\gamma$ -25 to a position \sim 1,300 bp upstream (Fig. 2, probe b'). A single BAC clone was obtained, spanning \sim 120 kb, from \sim 20 kb upstream of C γ 2 to 35 kb downstream from the C α 2 membrane exon. From this clone, designated BAC 11771, two subclones were prepared: a 10.5-kb HindIII fragment (A2H10.5) and an overlapping 14-kb EcoRI fragment (A2E14). Fragments of these subclones were found to have sequence homology to murine HS12, HS3, and HS4, and enhancer activity in a transient transfection assay (see Results).

Because several library screens were unsuccessful in identifying phage or BAC clones covering the gap between C α 1 and our downstream phage contig, selected DNA segments in this region were amplified by PCR. Southern blots of human genomic DNA indicated that the HS12 regions near α 1 and α 2 lie in opposite orientation. This inversion made it possible to selectively amplify α 1-derived sequence from genomic DNA using two primers which both corresponded to the sense strand of the α 2 locus. Amplification with the upstream primer SA2.5-A2 and the downstream primer SA2.1-A2 was performed using Taq polymerase XL (PE Applied Biosystems, Foster City, CA) for 32 cycles (94°C for 1 min, 61°C for 2 min, and 72°C for 10 min), and yielded the 5.5-kb A1-HS3-12 PCR fragment (Fig. 2). From this fragment a 982-bp α 1 HS-3 product was amplified using primers SA2.5A and SA2.6A; and 964 bp α 1HS12T and 892 bp α 1HS12B fragments were amplified using primers SA2.1A and SA2.2B. For comparisons between enhancer activities of corresponding α 1 and α 2 fragments, the homologous 1070-bp A2HS12 PCR product was generated with the same primer pair but using plasmid pEH1.3 as template. In experiments seeking the α 1 homolog of HS4, amplifications exploiting the inversion were unsuccessful. As an alternative strategy, we prepared an α 1-specific genomic template for PCR: a 23-kb HindIII fragment from the human myeloma HS Sultan (American Type Culture Collection, Rockville, MD) which Southern blots indicated extended from the α 1 membrane exon to a HindIII site in our $\psi\gamma$ -25 clone, including DNA homologous to HS4. The corresponding regions

from α 2 fall in 12-kb and 14-kb HindIII fragments. A 23-kb preparative electrophoresis fraction of HindIII-digested DNA from HS Sultan was used as template for amplification with primers SA8.A and SA11.B based on sequence from the α 2 locus. The resulting 4.2-kb fragment (A1-HS4-3'PCR; Fig. 2) was cloned into a λ ZAP Express phage vector (Stratagene), yielding the clone α 1HS4-4.5; the cloned insert showed a restriction map consistent with the α 1 locus and distinct from the α 2 locus. From the 4.2 A1-HS4-3'PCR fragment a 468-bp segment containing the α 1 HS4 region was amplified using primers SA.8A and SA.9B; this segment was used for sequence analysis and enhancer studies. With the same primer pair, the corresponding α 2 fragment A2-HS4 was generated using plasmid pA2E14 as template in order to enable meaningful α 1 versus α 2 comparisons of HS4 enhancer activity.

DNase I Hypersensitive Site Analysis. Nuclei were prepared and digested with DNase I according to a previously described protocol (36). K562 and HS Sultan cells (both obtained from American Type Culture Collection) were grown to densities of $5\text{--}8 \times 10^5$ cells/ml. For each experiment, $3\text{--}6 \times 10^8$ cells were harvested, lysed by addition of NP-40, centrifuged through a 1.7 M sucrose cushion, and resuspended in 5 ml; 450- μ l aliquots of suspended nuclei were treated with serially diluted DNase I (Boehringer-Mannheim, Indianapolis, IN) to give final DNase I concentrations of 0–8 μ g/ml. Nuclei were digested with DNase I for 3 min at 25°C. For one experiment HS Sultan nuclei were digested with the restriction endonuclease SspI (New England Biolabs, Beverly, MA) for 15 min at 37°C. DNase I (or SspI) digestion was terminated by adding 50 μ l 1% SDS, 100 mM EDTA. DNA samples were deproteinized for 5–48 h at 37°C using proteinase K (Boehringer-Mannheim) at a final concentration of 100 μ g/ml. DNA was purified by phenol/chloroform extraction and ethanol precipitation, resuspended in 50–100 μ l deionized water, and DNA concentrations were measured using a Fluorometer (TKO 100; Hoeffer, San Francisco, CA). 5- μ g DNA samples were digested for 5–24 h in 50 μ l of appropriate restriction enzyme buffer with BglII, EcoRI, or HindIII (New England Biolabs). To

Table 1. Sequences of Oligonucleotides Used in This Study

Oligonucleotide name	Sequence
Pseudogam-1	AGATGCCACCATGTCAAGT
Reverse sequencing primer	AACAGCTATGACCATG
elkb2	TAAGCTGTCTGAGAGAAAGTTGGGGGAGG
5'TM-A1	CTGTTACACGAGTCTGGGCCTGG
3'TM-B	TCCAAGAGGTTCCCTCCACACTTCC
SA2.5-A2	<u>ggccggtacc</u> GGATCCCGGTTCCCTGATCACTG
SA2.1-A2	<u>ggccggtacc</u> CTTCCTGCCAACCTGGGGGCTG
SA2.5A	<u>ggccacgcgt</u> GGATCCCGGTTCCCTGATCACTG
SA2.6A	<u>ggccacgcgt</u> CCACAGTCACTGCCAGATGCTC
SA2.1A	<u>ggccacgcgt</u> CTTCCTGCCAACCTGGGGGCTG
SA2.2B	<u>ggccacgcgt</u> GGCTTTTGCCAGTCCCTCCTAC
SA11.B	CAGTGCCCCAACCCAGGACGCCAGCC
SA.8A	<u>ggccacgcgt</u> CGCTCGCTGCCCCACTCAGGAGG
SA.9B	<u>ggccacgcgt</u> CTCCTAGCAGGGTCTCCTCCCTGG
β5PR-A	GAGGAGAAGTCTGCCGTTACTGCC
β3PR-B	GGCACAATCCAGATGCTCAAGGCC

All primer sequences are presented 5' to 3'. For certain primers derived from human genomic DNA sequence, additional nucleotides were added at the 5' end to provide restriction sites; these nongenic nucleotides are written in lowercase letters, with the restriction site underlined. All oligonucleotides were prepared in the Facility for Biotechnology Research (Center for Biologics Evaluation and Research, Bethesda, MD).

assess the SspI sensitivity of the β globin locus in HS Sultan nuclei, a 1,511-bp human β globin probe was amplified from total human genomic DNA using primers β5PR-A and β3PR-B. Restriction-digested samples, together with ³²P-labeled size markers, were electrophoresed, blotted, and hybridized with the probes indicated in the figure legends. After washing the membranes, radioactive images were obtained with a PhosphorImager (Molecular Dynamics, Sunnyvale, CA).

Enhancer Assays. To analyze DNA fragments for enhancer activity, we used the luciferase reporter plasmid pGL3 (Promega Corp., Madison, WI), modified so that the SV40 promoter between BglIII and HindIII sites was replaced by a 76-bp Vκ promoter (37) containing an octamer motif and TATA box. This plasmid, named pGL3-Vκ, served as an enhancerless, promoter-only control. Fragments to be assayed for enhancer activity were blunt-ended with Klenow fragments of DNA polymerase and ligated with MluI linkers, or amplified with primers incorporating an MluI (or KpnI) site; the fragments were then cloned into the MluI (or KpnI) site in the polylinker upstream of the promoter in pGL3-Vκ. Plasmid DNAs for transfection were twice purified on a CsCl gradient.

Approximately 5×10^6 cells in mid-log phase were washed with PBS and electroporated with either 5 μg of promoter-only control construct or equimolar amounts of enhancer test constructs along with 5 μg of CMV-β-galactosidase plasmid (pCMVβ; Clontech, Palo Alto, CA) as an internal control. Each plasmid was electroporated in triplicate in 400 μl PBS using an Electro Cell Manipulator (model 600; BTX Inc., San Diego, CA) at 200 V, 900 μF capacitance, and 13 Ω resistance in a 0.2-cm electroporation cuvette. Immediately after transfection cells were returned to culture medium. 24 h after transfection, the cells were harvested, washed with PBS, and lysed in 50 μl reporter lysis buffer (Promega Corp.) at room temperature for 20 min. After centrifugation at

15,000 rpm for 5 min, 15 μl of supernatant extracts were assayed for β-galactosidase using Galacto-Light Plus Chemiluminescent Reporter Assay kit (Tropix Inc., Bedford, MA) according to the manufacturer's instructions. To assay for luciferase activity, 35 μl of cell extracts was incubated with 100 μl of luciferase substrate [470 μM luciferin in 20 mM tricine, 1.07 mM (MgCO₃) × Mg(OH)₂ × H₂O, 2.67 mM MgSO₄, 0.1 mM EDTA, 33.33 mM DTT, 270 μM coenzyme A, and 540 μM ATP]. For both β-galactosidase and luciferase assays light output was detected in a luminometer (Dynatech Instruments, Chantilly, VA).

Sequence Analysis. A PCR-based methodology employing P³³-labeled ddNTPs was used for sequencing reactions (ThermoSequenase kit; Amersham Life Science, Arlington Heights, IL). All samples were amplified for 50 cycles (95°C for 30 s; 60°C for 30 s; 72°C for 60 s). Sequencing reactions were electrophoresed on 6% gels, dried, autoradiographed, and read manually. All reported sequences were read on both A and B strands, except for the sequences of the HS12 59-bp repeats, which, despite several attempts employing various strategies, could be read only on one strand.

Results

Cloning DNA Downstream of Cα1 and Cα2. Because attempts to clone the human 3' Cα regions using strategies based on gene walking from Cα or cross-species hybridization had failed in other laboratories, we used an alternative strategy based on the fact that a Cγ-like pseudogene (ψγ) has been described downstream of human Cα1 (32, 38). We reasoned that we could clone DNA downstream of Cα1 by walking upstream from ψγ. The similarity in the restriction maps reported downstream of the two Cα genes

based on Southern blotting (26) furthermore suggested that these regions would be highly similar to each other. Therefore, probes derived by walking upstream from $\psi\gamma$ should also hybridize to clones containing DNA from downstream of $C\alpha 2$.

By screening a phage library for clones that hybridized to $C\gamma$ but not $S\gamma$, and then screening those clones with a $\psi\gamma$ -specific oligonucleotide (see Materials and Methods), we obtained three overlapping clones spanning ~ 30 kb (Fig. 2). To obtain clones covering the corresponding region from the $\alpha 2$ locus, the library was then rescreened with probes b and e (Fig. 2). Clones deriving from downstream of $C\alpha 2$ were selected on the basis of restriction site differences between the $\alpha 1$ and $\alpha 2$ which had been established by analysis of genomic Southern blots and comparisons with our cloned DNA from the $\alpha 1$ duplication. Three clones deriving from the $\alpha 2$ locus and spanning ~ 30 kb were obtained (Fig. 2).

To estimate the distance between our two phage contigs and the corresponding $C\alpha$ genes, we performed genomic Southern blot experiments using a panel of restriction enzymes with known sites in the $C\alpha$ loci and in our contigs. For example, co-migrating BglII bands of ~ 35 kb were found to hybridize to both probe a (from the $C\alpha$ membrane exon αm) and probe b (Fig. 2); these experiments suggested a gap of ~ 20 kb between each contig and the corresponding membrane exon of $C\alpha$ (data not shown). For the $\alpha 2$ locus, the gap was bridged by a BAC clone (see Materials and Methods). The corresponding regions from the $\alpha 1$ locus have resisted direct cloning and have been obtained by PCR using primers designed from the sequence of the $\alpha 2$ locus, as described in Materials and Methods.

DNase I Hypersensitive Site Analysis. To map potential regulatory sequences downstream of the $C\alpha$ genes, we used fragments from the cloned DNA downstream of the $C\alpha 2$ gene to search for DNase I hypersensitivity sites; in cells where such enhancers or promoters are active, they generally are hypersensitive to endonucleases, apparently because binding of transcription factors disrupts the nuclease protection afforded by nucleosomes. Intact nuclei from the human myeloma HS Sultan were incubated briefly with various concentrations of DNase I; DNA samples purified from these treated nuclei were then analyzed using several Southern blot strategies in order to localize the positions of DNase I cleavage. The analyses were complicated by the fact that all of the probes we used hybridized to both the $\alpha 1$ and $\alpha 2$ loci, but various restriction map differences between the two loci allowed us to position all the hypersensitivity sites with respect to restriction sites mapped from our clones.

Fig. 3 A demonstrates analyses of DNA from the promyelocyte K562 line, representative of a cell not expressing immunoglobulin genes, and the myeloma HS Sultan line. The DNase I-treated DNA was digested with BglII, which cuts ~ 1 kb upstream of the membrane exon of both $C\alpha$ genes; the blots were hybridized with a probe corresponding to this exon. Since the next downstream BglII site is >20 kb away, this strategy can display hypersensitivity sites

over this wide distance from the α membrane exon. The blot demonstrates at least seven hypersensitivity sites, which were subsequently assigned to the $\alpha 1$ or $\alpha 2$ locus by other blotting experiments and were named according to sequence similarity to the homologous murine regions as described below. The sites we have tentatively designated $\alpha 1X$ and $\alpha 2X$ do not correspond to any reported murine enhancer sequence. None of these sites were visible in the DNA from K562, in which the enhancer region is expected to be inactive.

The BglII blot of HS Sultan DNA fails to resolve the two HS4 sites because they are too far away from the BglII sites. The two HS4 sites were resolved by an alternative Southern blot strategy employing EcoRI digests of the DNA from DNase I-digested nuclei (Fig 3 B). To determine which of the two resulting bands represented HS4 from the $\alpha 1$ versus $\alpha 2$ loci, we exploited the observation from our sequence analysis that recognition sites for the restriction enzyme SspI lie in each HS4 site. Since many regulatory regions accessible to DNase I are also accessible to restriction endonucleases, we digested HS Sultan nuclei with SspI and localized the cleavage sites by isolating the DNA, digesting with EcoRI, and hybridizing Southern blots with probe b' as shown in Fig. 3 B. From the map positions of the $\alpha 1$ and $\alpha 2$ SspI sites we assigned the SspI hypersensitivity bands as shown in Fig. 3. A control experiment on the same DNA isolated from SspI-digested HS Sultan nuclei demonstrated that SspI recognition sites in the β globin locus were not cut, indicating that the SspI sites associated with HS4 were indeed hypersensitive in HS Sultan cells (data not shown; see Materials and Methods).

Fig. 3 C illustrates an experiment which allowed assignment of several HS sites to the $\alpha 2$ locus rather than $\alpha 1$. DNA samples from DNase I-treated HS Sultan nuclei were digested with HindIII and hybridized with a probe for the HS12 site derived from the $\alpha 2$ locus. Although this probe hybridizes to both the $\alpha 1$ and $\alpha 2$ loci, in this DNA the $\alpha 1$ band is ~ 23 kb, so large that any fragments from this locus that were generated by HindIII and DNase I and which hybridized to the probe would be larger than the 12-kb band representing the $\alpha 2$ locus. Thus all HS bands <12 kb derive from $\alpha 2$. This blot therefore defines the position of the HS12, X, and HS3 sites in the $\alpha 2$ locus. By implication, the other HS sites in the BglII blot of Fig. 3 A must derive from $\alpha 1$. (It should be noted that a common allele in other DNA samples shows an additional polymorphic HindIII site which cuts within the 23 kb corresponding to the HS Sultan band; this allele would have confounded the above strategy, but was absent in HS Sultan.)

Enhancer Function. To analyze enhancer activity in the cloned DNA downstream of the $C\alpha 2$ locus, fragments containing one or more DNase I hypersensitive sites were subcloned into a luciferase reporter gene driven by a $V\kappa$ promoter as described in Materials and Methods. Luciferase activity in each sample was normalized to the β -galactosidase activity of the promoter-only control plasmid, and expressed as fold-increase over luciferase activity of that control plasmid.

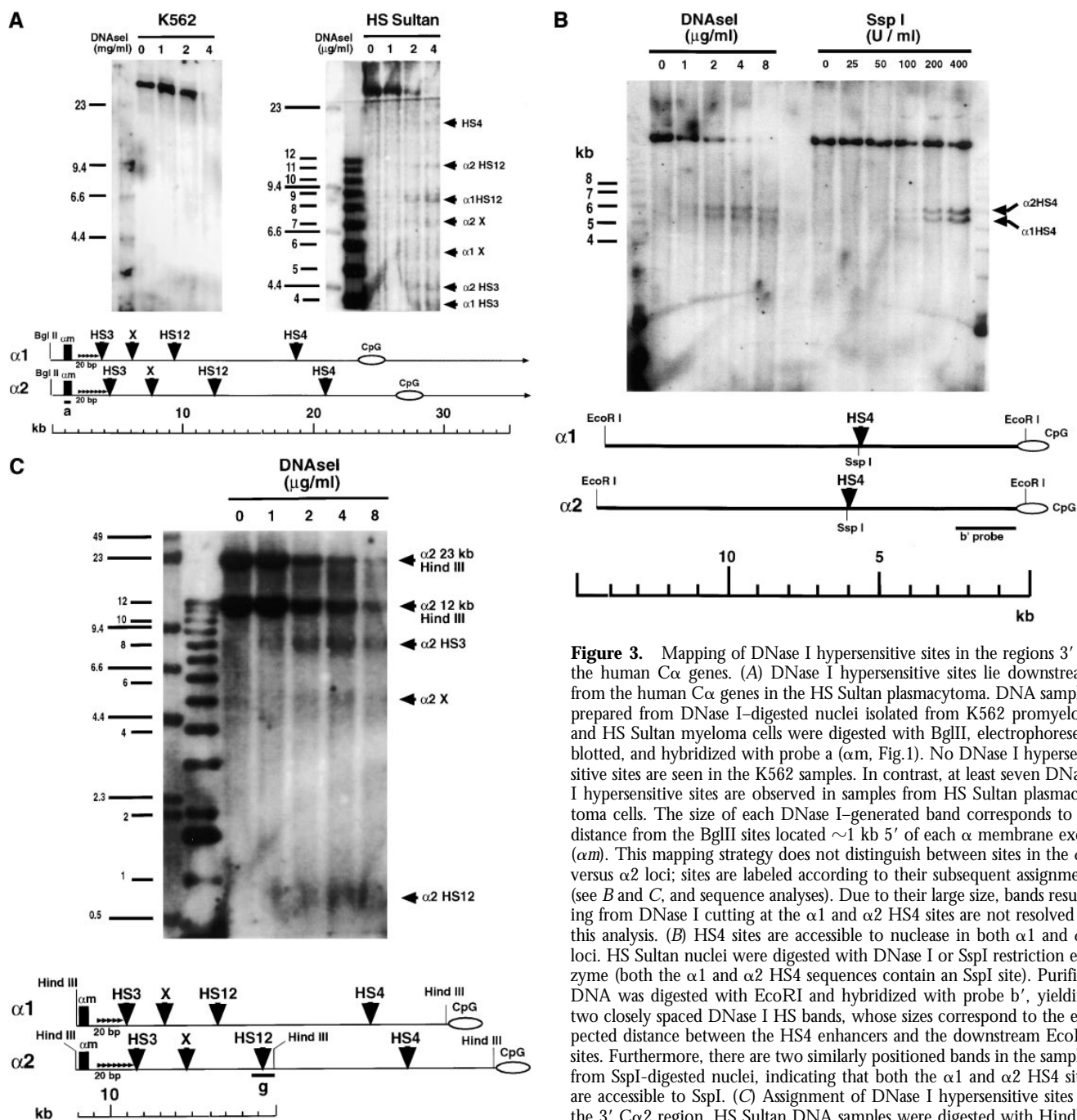


Figure 3. Mapping of DNase I hypersensitive sites in the regions 3' of the human $C\alpha$ genes. (A) DNase I hypersensitive sites lie downstream from the human $C\alpha$ genes in the HS Sultan plasmacytoma. DNA samples prepared from DNase I-digested nuclei isolated from K562 promyeloid and HS Sultan myeloma cells were digested with BglII, electrophoresed, blotted, and hybridized with probe a (αm , Fig. 1). No DNase I hypersensitive sites are seen in the K562 samples. In contrast, at least seven DNase I hypersensitive sites are observed in samples from HS Sultan plasmacytoma cells. The size of each DNase I-generated band corresponds to its distance from the BglII sites located ~ 1 kb 5' of each α membrane exon (αm). This mapping strategy does not distinguish between sites in the $\alpha 1$ versus $\alpha 2$ loci; sites are labeled according to their subsequent assignment (see B and C, and sequence analyses). Due to their large size, bands resulting from DNase I cutting at the $\alpha 1$ and $\alpha 2$ HS4 sites are not resolved in this analysis. (B) HS4 sites are accessible to nuclease in both $\alpha 1$ and $\alpha 2$ loci. HS Sultan nuclei were digested with DNase I or SspI restriction enzyme (both the $\alpha 1$ and $\alpha 2$ HS4 sequences contain an SspI site). Purified DNA was digested with EcoRI and hybridized with probe b', yielding two closely spaced DNase I HS bands, whose sizes correspond to the expected distance between the HS4 enhancers and the downstream EcoRI sites. Furthermore, there are two similarly positioned bands in the samples from SspI-digested nuclei, indicating that both the $\alpha 1$ and $\alpha 2$ HS4 sites are accessible to SspI. (C) Assignment of DNase I hypersensitive sites to the 3' $C\alpha 2$ region. HS Sultan DNA samples were digested with HindIII and hybridized with probe g ($\alpha 2$ HS12, Fig. 1). Because DNase I-generated bands from the $\alpha 1$ region which hybridize to this probe are expected to be larger than the 12-kb $\alpha 2$ HindIII fragment, all bands < 12 kb must result from DNase I cutting in the 3' $\alpha 2$ region, with the size of these bands corresponding to their distance from the 3' end of the 12-kb $\alpha 2$ HindIII fragment. This analysis allows assignment of three DNase I sites to the $\alpha 2$ locus, thus making it possible to assign the other hypersensitive sites seen in the BglII analysis of Fig. 2 A to the $\alpha 1$ locus.

The luciferase assays (Fig. 4) revealed strong enhancer activity in the 5-kb SmaI-HindIII fragment (SM5) which was found to contain sequence homologous to murine HS12 (see below). Within this fragment, enhancer activity seemed to be confined to the 1.3-kb EcoRI-HindIII fragment (EH1.3) containing the HS12 site. This segment was further cut into an upstream 0.3-kb EcoRI-PstI fragment

(EP300), a 0.3-kb PstI-PstI fragment (P300), and a 0.6-kb PstI-HindIII fragment (PH600). Of these, only the P300 fragment, which contained HS12, showed enhancer activity. However, it should be noted that the enhancer activity of P300 was less than that measured for a larger PCR-generated fragment A2HS12. Thus it is possible that additional elements that do not show intrinsic enhancer activity when

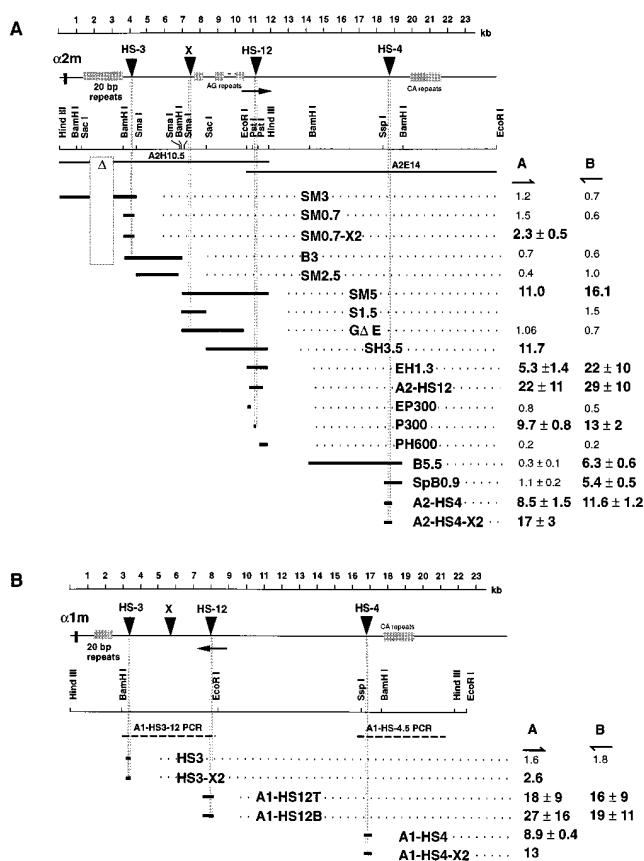


Figure 4. Enhancer activity of selected regions downstream of human C α 1 and C α 2 genes. (A) Analysis of the locus downstream of α 2, which was studied in detail. The map shows the position of DNase I sites; below this are diagrammed the restriction sites defining the boundaries of each fragment tested for enhancer activity by insertion into pGL3-V κ , transfection into the human myeloma HS Sultan, and assay of resulting luciferase activity, as described in the text. The enhancer activities are given for constructs in the A orientation (the same orientation with respect to transcribed strands of immunoglobulin and luciferase) or the opposite B orientation, where examined. The luciferase activities were normalized to β -galactosidase activity encoded by a cotransfected plasmid, and expressed as fold-increase over the activity of an enhancerless control plasmid. For fragments showing enhancer activity, assays were performed at least in triplicate, and standard deviations are given. (B) Comparable analysis of selected fragments amplified from the homologous locus downstream from C α 1.

isolated in constructs can nevertheless augment the activity of the core HS12 enhancer lying in the P300 fragment. Furthermore, the EH1.3 fragment, but not the slightly shorter A2HS12 fragment, showed significantly less enhancer activity in the A orientation, suggesting the possibility of inhibitory sequences located near one of the ends of EH1.3.

Two fragments containing the α 2 HS3 site showed no significant enhancer activity in the luciferase construct, but a plasmid in which the HS3-containing 0.7-kb BamHI-SmaI fragment was dimerized showed consistent low activity (Fig. 4, SM0.7-X2). The homologous murine enhancer was reported to show greater activity with a c-myc promoter than with an immunoglobulin V λ promoter (12), but we found no significant increase in enhancer activity

when the V κ promoter was replaced by a human c-myc promoter (data not shown; construct described in Materials and Methods).

Two restriction fragments, B5.5 and SpB0.9, containing the α 2 HS4 site demonstrated significant stimulation of luciferase activity, but only when cloned in the B (inverted) orientation (Fig. 4 A). When these fragments were oriented so that the DNA strand continuous with the transcribed strand in the luciferase gene corresponded to the transcribed strand of the immunoglobulin gene in genomic DNA (the A orientation), luciferase activity was no greater than that from the promoter-only plasmid. Unexpectedly, a smaller fragment (the 468-bp PCR-generated fragment designated A2-HS4, designed to span the sequence showing the strongest homology to the murine HS4 site) was found to give a strong activation of luciferase activity. The stronger activation of luciferase conferred by the short A2-HS4 PCR fragment compared with that of the longer SpB0.9 fragment (in the same A orientation) suggests that the latter may contain an inhibitory sequence that is either position- or orientation-dependent. This possibility is currently being explored.

Because of the lack of genomic clones spanning the HS3, HS12, and HS4 regions of α 1 locus, fragments corresponding to each of these sites were obtained by PCR and analyzed; and, to facilitate α 1 versus α 2 locus comparisons, the same primer pairs were used to generate corresponding fragments from the α 2 locus. As shown in Fig. 4, the activity of each HS site fragment from α 1 was similar to that from the α 2 locus (compare SM0.7-X2 α 2 versus HS3-X2 α 1; A2HS12 α 2 versus A1HS12T α 1; and A2-HS4 α 2 versus A1-HS4 α 1). The PCR amplification of the HS12 site yielded two fragments, designated A1HS12T (top) and A1HS12B (bottom), which differ in size by <0.1 kb and may represent alleles (see below). Both fragments showed substantial enhancer activity (Fig. 4 B).

Since the mouse 3' α enhancers show different activation patterns during B cell differentiation, with HS4 being activated at the pre-B cell stage, while HS12 and HS3 are only active in mature B cells, we examined the activity of the human 3' α 2 enhancers in a range of human and mouse B cell lines (Table 2). The pattern of activation for the human HS12 enhancer is similar to that of mouse HS12; i.e., human HS12 is inactive in the human pro-B cell line FLEB-14 and the mouse pre-B cell line 18-81, but functions in the human mature B cell line Raji, as well as three plasmacytomas (HS Sultan, human; S194, mouse; and MOPC 315, mouse). HS3 is also inactive in mouse 18-81 pre-B cell line, but shows modest activity in most of the more mature lines tested. HS3 shows surprisingly strong activity in the mouse S194 myeloma, indicating that unknown factors varying between cell lines at similar stages of differentiation can modulate the activity of this enhancer. Finally, the HS4 enhancer shows strong activity in the human pro-B cell line FLEB14, and is also variably active in all of the more mature cell lines (except 18-81) in which this enhancer was assayed in the B orientation.

Sequence Analysis. The nucleotide sequence of all DNase I hypersensitivity sites was determined, revealing ~99% se-

Table 2. Cell Specificity of Human $\alpha 2$ Enhancer Elements

Cell line	HS12 (EH1.3)		HS3 (SM0.7)			HS4 (SpB0.9)	
	A	B	A	B	X2	A	B
Human							
FLEB 14 (pro B)	1.9 \pm 0.1	1.9 \pm 0.1	ND	ND	ND	2.8 \pm 0.2	18.4 \pm 2.4
Raji (Mature B)	3.1 \pm 2.2	3.2 \pm 1.8	0.7 \pm 0.2	1.0 \pm 0.3	2.8 \pm 0.5	1.2 \pm 0.1	2.7 \pm 0.5
HS Sultan (Plasmacytoma)	5.3 \pm 1.4	22 \pm 10	1.54	0.59	2.3 \pm 0.5	1.1 \pm 0.2	5.4 \pm 0.5
Mouse							
18-81 (Pre-B)	0.9 \pm 0.2	1.1 \pm 0.3	1.2 \pm 0.1	0.8 \pm 0.2	1.6 \pm 0.5	0.5 \pm 0.1	0.9 \pm 0.2
S 194 (Plasmacytoma)	5.3 \pm 0.2	14.4 \pm 1.0	8.5 \pm 0.8	2.6 \pm 0.4	16.3 \pm 0.3	7.2 \pm 1.4	28.4 \pm 2.3
MOPC 315 (Plasmacytoma)	5.9 \pm 1.1	16.9 \pm 1.9	1.4 \pm 0.2	0.7 \pm 0.2	2.9 \pm 1.9	1.3 \pm 0.6	5.8 \pm 0.9

The numbers given represent the fold-increase of luciferase activity seen with a promoter-only control plasmid, with standard deviation. Numbers in bold were judged to represent significant enhancer activity.

quence identity between the human $\alpha 1$ and $\alpha 2$ elements, and similarity between the human and mouse enhancers ranging from 74 to 90%.

HS12, the strongest enhancer, showed 90% sequence identity to the homologous murine enhancer over a 135-bp core homology (Fig. 5 A). In the $\alpha 2$ locus, four tandem repeats with a 59-bp consensus sequence lie immediately upstream of the HS12 core. However, this sequence has been inverted in Fig. 5 A, to facilitate comparison with the homologous $\alpha 1$ sequence in opposite orientation. In the corresponding region of $\alpha 1$ (which, due to the inversion, lies downstream of the core homology region) a 115-bp deletion removes the second and third repeats; the $\alpha 1$ HS12B region shows an additional deletion of 70 bp. The core homology region includes several of the functional motifs identified in the murine enhancer: the AP1-Ets site reported to confer responsiveness to B cell receptor cross-linking (22, 39); an exact octamer sequence (ATGCAAAT); and a μ E5 site (except in the $\alpha 1$ HS12B sequence, in which the μ E5 is missing owing to the 70-bp deletion). The sequences of these three motifs from the human HS12 are identical to their murine homologs except for a single base change in the AP1 site which causes the human sequence to exactly match the consensus AP1 site, where the murine motif has one mismatch. The murine element designated μ E1 (39), which has never been thoroughly documented even in the murine enhancer, is poorly conserved in the human homologs. Although the murine binding site for NF- κ B lies outside the 135-bp region of strongest sequence similarity, a reasonable match to consensus for this element is found in a position roughly homologous to the murine κ B site in the $\alpha 1$ HS12T and $\alpha 2$ HS12 sequences, but is part of the 70 bp deleted in $\alpha 1$ HS12B. One of the mouse BSAP sites (BSAP2) is not conserved, but most residues in a second mouse BSAP binding site (BSAP1) are

maintained in the human $\alpha 2$ HS12 enhancer. The murine α P site, which binds to an ETS-related transcription factor which augments enhancer activity (17), is not conserved in the human sequences.

The DNase I hypersensitive sites lying <3 kb downstream from the membrane exons of $C\alpha 1$ and $C\alpha 2$, roughly in the position of the weak enhancer reported by Matthias and Baltimore (11) and here designated HS3A, were found to contain sequences which are 74% identical to the murine HS3 over a 200 bp core segment. The two human HS3 segments are identical in the 326 bp shown (Fig. 5 B), and lie in the same orientation as HS3A. We have assumed that the correct orientation of the murine HS3A sequence is that described by Chauveau and Cogné (13). This orientation is opposite to that of the murine HS3B, which lies downstream from HS12 in the mouse, as described by Madisen and Groudine (12). Of the enhancer/HS sites downstream of murine $C\alpha$, HS3 is the least well investigated for functional motifs, in part because of its weak enhancer activity. Independent sequence analysis of the murine HS3A and HS3B regions detected several similarities to octamer motifs, AP1 sites, and consensus E box motifs (CANNTG). The AP1 site identified in Fig. 4 is a precise match to the consensus AP1 binding motif TGANTCA (40) in the two human and two mouse HS3 sequences, and the murine sequence has been shown to bind to c-jun and c-fos in vitro (Neurath, M., personal communication). Similarly, several of the E box consensus motifs in the murine sequence have been shown to bind in vitro to proteins of the HLH family (Neurath, M., personal communication); some of these motifs are conserved in the two human HS3 sequences. The significance of the conserved motifs remains uncertain in the absence of a functional analysis of HS3 sequences.

The DNase I hypersensitivity sites furthest downstream

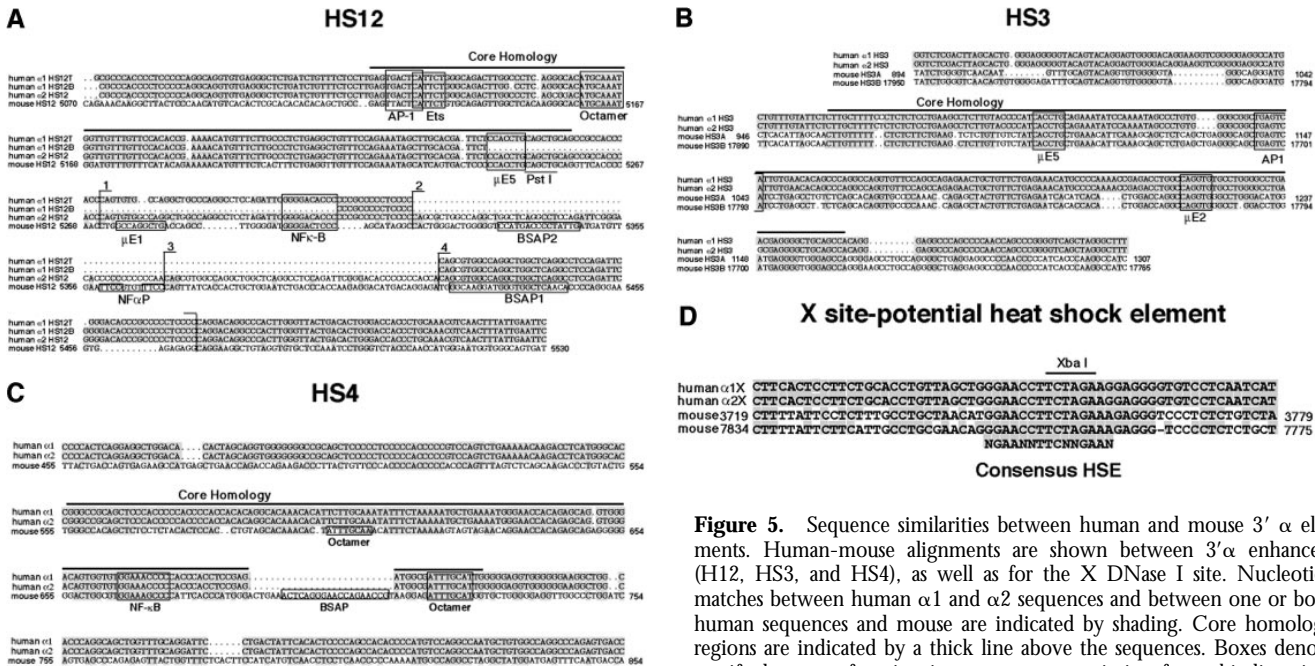


Figure 5. Sequence similarities between human and mouse 3' α elements. Human-mouse alignments are shown between 3' α enhancers (H12, HS3, and HS4), as well as for the X DNase I site. Nucleotide matches between human $\alpha 1$ and $\alpha 2$ sequences and between one or both human sequences and mouse are indicated by shading. Core homology regions are indicated by a thick line above the sequences. Boxes denote motifs shown to function in mouse as transcription factor binding sites. For HS12, HS3, and HS4, 50–100 bp of sequence flanking the core homology to the coding strand of the mouse heavy chain locus. Numbering for mouse HS12, HS3A, HS3B, and X segments is according to reference 13 (EMBL/GenBank/DBJ accession numbers X96607 and X96608), while numbering for mouse HS4 is according to reference 12 (EMBL/DBJ/GenBank accession number S74166). (A) HS12 sequences ($\alpha 2$ sequence inverted). Overlining highlights the striking 135-bp core segment which is 90% homologous between human and mouse. The sequence alignment has been extended downstream from the core to include additional transcription factor motifs which are functional in mouse. Vertical lines indicate the boundaries of the GC-rich 59-bp repeat units. (B) HS3. Comparison of the nearly identical human $\alpha 1$ and $\alpha 2$ HS3 sequences with mouse HS3A and HS3B sequences, which are also nearly identical, shows that there is a 200-bp core segment which is 74% homologous between the mouse and human sequences. (C) HS4. Excluding the 25-bp gap containing the mouse HS4 BSAP site, the 145 core HS4 region is 76% homologous between human and mouse. (D) X site. Near the center of a 61-bp segment which has 70% human-mouse homology, there is a 20-bp sequence which matches at 19 positions between humans and mice. In both mice and humans this segment contains a consensus HSE (41, 42). The sequences of the human enhancers and X sites are available from EMBL/GenBank/DBJ under accession numbers AF013718 ($\alpha 1$ HS3), AF013719 ($\alpha 2$ HS3), AF013720 ($\alpha 1$ X), AF013721 ($\alpha 2$ X), AF013722 ($\alpha 1$ HS12T), AF013723 ($\alpha 1$ HS12B), AF013724 ($\alpha 2$ HS12), AF013725 ($\alpha 1$ HS4), and AF013726 ($\alpha 2$ HS4).

from human $C\alpha 1$ and $C\alpha 2$ in Fig. 3, which we have designated HS4, are 76% similar to the murine HS4 site over a core 145-bp sequence which spans the three functional motifs demonstrated in murine HS4 (23); see Fig. 5 C. The NF- κ B motif and the downstream octamer motif in Fig. 5 C are both precisely conserved; in the murine HS4 these motifs both contribute to functional enhancer activity (23). In contrast, the BSAP site which upregulates murine HS4 enhancer activity in B cells but downregulates it in pre-B cells is completely absent from the human HS4 sequences.

Lying between HS3 and HS12 in both the $\alpha 1$ and $\alpha 2$ loci are DNase I hypersensitive sites which are not associated with any of the known enhancer elements, but that do map to the position of a 61-bp segment of 70% mouse-human homology. These conserved regions are provisionally designated X sites in part because of their unknown function, and in part because both the human and mouse segments contain XbaI restriction enzyme sites. In mouse this sequence is duplicated as part of the large inverted repeat centered on HS12, so that one copy lies between HS3A and HS12 while a second copy lies between HS12 and HS3B. A segment of (GA)_n repeats is found near the X

site in the direction of HS12 at an interval of 80 bp for both mouse X sites (13) and an interval of ~70 bp for the human (data not shown). Within the 61-bp conserved segment, the most highly conserved sequence is a consensus heat shock element (HSE; 41, 42). An HSE could potentially bind heat shock transcription factors (HSTFs), which are known to activate several heat shock response genes (HSP70, HSP90) in response to cellular stress such as heating (43–45). Fragments containing an X site do not appear to dramatically affect enhancer activity in HS Sultan, but may contribute to regulation through mechanisms not captured in our transient transfection assays.

Discussion

Structure of Human 3' α Regions: Evolutionary Implications. The enhancers clustered 3' of $C\alpha$ in the mouse IgH locus can activate the upstream genes, functioning as an LCR. In human IgH locus, arrays of enhancers homologous to those 3' of mouse $C\alpha$ are located at two positions within the human IgH locus, 3' of each $C\alpha$ gene. On the basis of sequence homology and conservation of restriction

Downloaded from http://rupress.org/jem/article-pdf/186/6/845/11367497-0796.pdf by guest on 20 September 2021

sites between the two human enhancer arrays, it is apparent that these enhancers lie near the 3' ends of the two duplication units which encompass the $\gamma 3$ - $\gamma 1$ - $\psi \epsilon$ - $\alpha 1$ and $\gamma 2$ - $\gamma 4$ - ϵ - $\alpha 2$ gene clusters (25–27), indicating that the 3' α enhancer arrays were present in an approximation of the human arrangement preceding the duplication event that gave rise to the present human IgH locus structure. Moreover, the arrangement downstream of both human C α genes is 5'-HS3-HS12-HS4-3', in contrast to the large palindromic structure downstream from mouse C α that contains a 5'-HS3A-HS12-HS3B-HS4-3' arrangement (13). Therefore, although an arrangement containing an HS3 enhancer proximal to the C α membrane exon and an HS4 enhancer farther downstream would seem to have been present in the common ancestor of rodents and primates, the mouse HS3A-HS12-HS3B palindrome probably arose after the primate-rodent divergence. Finally, there is a major structural difference between the 3' $\alpha 1$ and 3' $\alpha 2$ enhancer arrays; namely, that a DNA segment containing HS12 is inverted between the two loci. Using probes containing the 135-bp human HS12 core, it should now be possible to examine DNA from a number of primates for inversion of 3' $\alpha 2$ HS12 relative to 3' $\alpha 1$ HS12; such data may indicate when in evolution the inversion event occurred, and which orientation was present initially in the locus. What caused this inversion? Interestingly, the single mouse HS12 lies in opposite orientation from the rat HS12, and both are flanked by inverted repeats (13) which are known to mediate inversions in other genomic contexts, e.g., in the iduronate-2-sulfatase gene causing Hunter syndrome (46) and in the factor VIII gene (47). Limited Southern blot experiments have not provided evidence for inverted repeats flanking the human HS12 sequences (data not shown). Some hints about the mechanism of the inversion may be found when the inversion breakpoints are identified and sequenced, work currently in progress in our laboratory.

HS12 Structure and Function. The 135-bp HS12 core homology sequence is likely to contain essential motifs important for the strong, late, B cell-specific enhancer activity characteristic of HS12 in mice and humans. Although the function of transcription factor binding sites within the human HS12 core has not yet been demonstrated experimentally, this segment contains sequences nearly identical to the murine AP1, ETS, Oct, and, in $\alpha 1$ HS12T and $\alpha 2$ HS12, μ E5 motifs, all of which are functional in the mouse HS12 enhancer. However, the high degree of sequence conservation in the HS12 core homology extends beyond the transcription factor-binding sites identified in the mouse enhancer, indicating that there may be additional conserved motifs that have not been characterized in either mice or humans.

Despite the fact that a number of other transcription factor motifs whose function has been demonstrated in the mouse HS12 lie outside the HS12 135-bp core homology and are absent in one or more of the human $\alpha 1$ HS12T, $\alpha 1$ HS12B, and $\alpha 2$ HS12 enhancers that we have studied, these enhancers all show roughly equivalent activities. This result suggests that elements missing from these enhancers

are not essential for enhancer function in HS Sultan. These inconsistently conserved elements include human sequences corresponding to μ E5 and NF- κ B sites (absent from $\alpha 1$ HS12B) and the BSAP2 site (absent from both $\alpha 1$ alleles). On the other hand, a 1.3-kb $\alpha 2$ EcoRI-HindIII fragment containing the $\alpha 2$ HS12 core plus considerable flanking sequence shows a dependence of enhancer activity on orientation of the fragment (Table 2), suggesting that uncharacterized elements beyond the HS12 core may have some inhibitory function.

Outside the human HS12 core are GC-rich 59-bp repeat units which by themselves do not have enhancer activity in the HS Sultan myeloma (EP300, Fig. 4 A), and are not conserved between mice and humans (Fig. 5 A). However, it is possible that these repeats contribute to enhancer activity because they are present in the A2HS12 PCR-generated fragment, which gives significantly higher enhancer activity than we observe in the p300 fragment containing the core homology. Deletions of the 59-bp repeats have given rise to apparent allelic polymorphisms, as evidenced by $\alpha 1$ HS12T (deletion of the second and third repeats found in $\alpha 2$ HS12), $\alpha 1$ HS12 (deletion extending from 28 bp 5' of the first repeat through the third repeat), and other alleles (Harindranath, N., unpublished results).

HS3 and HS4. The other two enhancer components of the mouse 3' α LCR, HS3 and HS4, are weaker enhancers than HS12, but nonetheless are essential for locus control activity (12). Although these elements are less well characterized than the HS12 enhancer, the existing data indicate general human-mouse similarity of the HS3 and HS4 elements, with some notable differences.

In the mouse system, the HS3A element assayed in CAT reporter gene constructs driven by *c-fos* or thymidine kinase promoters (11) showed weak enhancer activity, although the nearly identical HS3B enhancer showed substantial activity in certain constructs with other promoters tested by another laboratory (12). These disparate results resemble our data on the human $\alpha 1$ and $\alpha 2$ HS3 elements in that single copies and dimers of human HS3 generally gave very low enhancer activity, except in the mouse S194 myeloma in which the same constructs gave substantial enhancer activity comparable to that of HS12 (Table 1). Taken together, these data suggest that HS3, though typically the weakest of the 3' α enhancers, contains uncharacterized motifs that in some cells and/or in combination with certain promoters, can mediate a strong enhancer function.

HS4 is the most downstream 3' α enhancer in both mice and humans, and shows activity intermediate between that of HS3 and HS12. The HS4 enhancer data in the mouse (12, 48), as well as our data on the human $\alpha 1$ and $\alpha 2$ HS4 elements, demonstrate that HS4 is active from the early stages of the B cell lineage onward (Table 2), and thus is qualitatively different from HS3 and HS12. In mouse HS4, there is a binding site for the BSAP, which is expressed in the early B cell lineage. However, in the human $\alpha 1$ and $\alpha 2$ HS4 enhancers, the BSAP site is deleted, indicating that BSAP binding is not an essential feature for HS4 activation

in human pre-B cells. The human HS4 is inactive in the 18-81 mouse pre-B cell line, which was reported to support the activity of mouse HS4 (12). The significance of this difference is not clear; it could be related to the BSAP site deletion in the human HS4, or perhaps to other differences between the mouse and human HS4 sequences.

X Site. The DNase I X sites may represent novel control elements that function together with the HS3, HS12, and HS4 enhancers to activate the IgH locus. Although the significance of the conserved HSE motif is unclear, binding of HSTF protein to an HSE has been shown to be critical for maintaining the DNase I hypersensitivity of the yeast HSC82 gene promoter (49). Heat shock activation of the *Drosophila* HSP70 gene promoter results from binding of HSTF to HSE sites after accessibility of the HSEs has been established by binding of the GAGA protein to adjacent (GA)_n repeats (50). This demonstrated interaction between HSEs and (GA)_n motifs suggests that the location of a (GA)_n repeat region 70–80 bp away from the X site HSE may be of some significance. Furthermore, in the context of IgH gene regulation, it is of interest to note that HSE motifs have been shown to respond to IL-2 and IL-4 (51).

Potential Locus Control Region. In the mouse, it has been demonstrated that when HS3, HS12, and HS4 are linked together in a construct containing the *c-myc* gene and stably transfected into the Raji human B cell line, the *c-myc* gene is transcribed independent of integration site (12). This observation suggests that the HS1234 combination confers LCR activity, although LCRs have more typically been described based on position-independent transcription of mouse transgenes rather than genes transfected into a cell line. Because the regions 3' of the human C α 1 and C α 2 genes contain similar HS3, HS12, and HS4 elements that function as enhancers, it is reasonable to hypothesize that these elements also function together in the human system as LCRs. The different arrangement of 3' α enhancers in mice and humans (HS3A-HS12-HS3B-HS4 versus HS3-HS12-HS4) may cause some functional differences in these control regions. Moreover the distance between 3' α enhancers also differs between mice and humans, with the mouse

enhancer complex spanning a 30-kb region (13), whereas both the human 3' α 1 and 3' α 2 enhancers span \sim 15 kb.

Our finding that arrays of enhancers homologous to those in the mouse 3' α LCR lie downstream of both human C α genes raises the possibility that differences in the activation of each human γ - γ - ϵ - α duplication unit result from differences between the putative 3' α 1 and 3' α 2 LCRs. Even though sequence comparison shows that there is near identity between homologous enhancer elements in the α 1 versus α 2 locus (Fig. 5), transcription and expression of the upstream heavy chain duplication unit (γ 3- γ 1- ψ ϵ - α 1) is greatly elevated relative to the downstream unit (γ 2- γ 4- ϵ - α 2; reference 2). This difference could result from the fact that the 3' α 2 HS12 element is inverted relative to the 3' α 1 HS12, and is also at a greater distance from HS3 than in the α 1 locus, possibly reducing synergistic interactions between HS3 and HS12. Alternatively, it may be that only the 3' α 1 enhancers are activated early in B cell development, possibly falling under the influence of the upstream E μ enhancer, which itself can function as an LCR (52, 53). In this model, the E μ and 3' α 1 enhancers together would activate a large domain encompassing the first duplication unit, whereas the second duplication unit and the 3' α 2 enhancers would fall outside this combined domain. Thus activation of the second duplication unit would depend solely on the 3' α 2 enhancers, and expression of genes in this unit might therefore be reduced. Clarification of the basis for this difference will have to await experiments that involve specific deletion of either the 3' α 1 or 3' α 2 enhancers, as well as studies identifying matrix attachment sites and chromatin insulator elements that define domains within the human IgH region (52–54).

In the work presented here, we have laid the foundation for experimental studies on the activation of the human IgH gene transcription, as well as the regulation of isotype switching. In addition, knowledge of the action of these enhancers on distant constant region genes should contribute to a general understanding of the mechanisms underlying activation of large gene clusters.

We thank Drs. Rose Mage, Louis Staudt, and Steven Bauer for discussions and useful comments on this manuscript. We also thank Dr. Markus Neurath for communicating unpublished results on functional elements in the mouse HS3 enhancers.

Address correspondence to Dr. Frederick C. Mills, Division of Hematologic Products, FDA/CBER/HFM-541, Bldg. 29A, RM 2B09, 29 Lincoln Dr., MSC 4555, Bethesda, MD 20892-4555. Phone: 301-827-1808; FAX: 301-480-3256; E-mail: millsf@fdacb.cber.fda.gov

Received for publication 12 May 1997 and in revised form 16 July 1997.

Note added in proof. While this manuscript was under review, related investigations by two other laboratories came to our attention. Chen, C., and B.K. Birshtein (1997. *J. Immunol.* 159:1310–1318.) have described the HS12 enhancers from the α 1 and α 2 loci; and recently others have characterized the HS3 and HS12 enhancers from the α 1 locus (M. Cogné, personal communication).

References

1. Coffman, R.L., D.A. Leberman, and P. Rothman. 1993. The mechanism and regulation of immunoglobulin isotype switching. *Adv. Immunol.* 54:229–270.
2. Sideras, P., L. Nilsson, K.B. Islam, I.Z. Quintana, L. Freihof, G.J. Rosen, L. Hammarstrom, and C.I. Smith. 1992. Transcription of unrearranged Ig H chain genes in human B cell malignancies. Biased expression of genes encoded within the first duplication unit of the Ig H chain locus. *J. Immunol.* 149: 244.
3. Klein, S., F. Sablitzky, and A. Radbruch. 1984. Deletion of the IgH enhancer does not reduce immunoglobulin production of a hybridoma IgD class switch variant. *EMBO (Eur. Mol. Biol. Organ.) J.* 3:2473–2476.
4. Eckhardt, L.A., and B.K. Birshtein. 1985. Independent immunoglobulin class-switch events occurring in a single myeloma cell line. *Mol. Cell. Biol.* 5:856–868.
5. Wabl, M., and P.D. Burrows. 1984. Expression of immunoglobulin heavy chain at a high level in the absence of a proposed immunoglobulin enhancer in *cis*. *Proc. Natl. Acad. Sci. USA.* 81:2452–2455.
6. Aguilera, R.J., T.J. Hope, and H. Sakano. 1985. Characterization of immunoglobulin enhancer deletions in murine plasmacytomas. *EMBO (Eur. Mol. Biol. Organ.) J.* 4:3689–3693.
7. Gregor, P.D., and S.L. Morrison. 1986. Myeloma mutant with a novel 3' flanking region: loss of normal sequence and insertion of repetitive elements leads to decreased transcription but normal processing of alpha heavy-chain gene products. *Mol. Cell. Biol.* 6:1903.
8. Pettersson, S., G.P. Cook, M. Bruggemann, G.T. Williams, and M.S. Neuberger. 1990. A second B cell-specific enhancer 3' of the immunoglobulin heavy-chain locus. *Nature (Lond.)* 344:165–168.
9. Dariavach, P., G.T. Williams, K. Campbell, S. Pettersson, and M.S. Neuberger. 1991. The mouse IgH 3'-enhancer. *Eur. J. Immunol.* 21:1499–1504.
10. Lieberson, R., S.L. Giannini, B.K. Birshtein, and L.A. Eckhardt. 1991. An enhancer at the 3' end of the mouse immunoglobulin heavy chain locus. *Nucleic Acids Res.* 19:933–937.
11. Matthias, P., and D. Baltimore. 1993. The immunoglobulin heavy chain locus contains another B-cell-specific 3' enhancer close to the constant region. *Mol. Cell. Biol.* 13:1547–1553.
12. Madisen, L., and M. Groudine. 1994. Identification of a locus control region in the immunoglobulin heavy-chain locus that deregulates *c-myc* expression in plasmacytoma and Burkitt's lymphoma cells. *Genes Dev.* 8:2212–2226.
13. Chauveau, C., and M. Cogné. 1996. Palindromic structure of the IgH 3' locus control region. *Nat. Genet.* 14:15–16.
14. Talbot, D., P. Collis, M. Antoniou, M. Vidal, F. Grosveld, and D.R. Greaves. 1989. A dominant control region from the human β -globin locus conferring integration site-independent gene expression. *Nature (Lond.)* 338:352–355.
15. Ernst, P., and S.T. Smale. 1995. Combinatorial regulation of transcription. I. General aspects of transcriptional control. *Immunity.* 2:311–319.
16. Meyer, K.B., M. Skoberg, C. Margenfeld, J. Ireland, and S. Pettersson. 1995. Repression of the immunoglobulin heavy chain 3' enhancer by helix-loop helix protein Id3 via a functionally important E47/E12 binding site: implications for developmental control of enhancer function. *Eur. J. Immunol.* 25:1770–1777.
17. Neurath, M.F., E.E. Max, and W. Strober. 1995. Pax5 (BSAP) regulates the murine immunoglobulin 3' α enhancer by suppressing binding of NF- α P, a protein that controls heavy chain transcription. *Proc. Natl. Acad. Sci. USA.* 92: 5336–5340.
18. Ernst, P., and S.T. Smale. 1995. Combinatorial regulation of transcription II: the immunoglobulin μ heavy chain gene. *Immunity.* 2:427–438.
19. Singh, M., and B.K. Birshtein. 1993. NF-HB (BSAP) is a repressor of the murine immunoglobulin heavy-chain 3' α enhancer at early stages of B cell differentiation. *Mol. Cell. Biol.* 16:3611–3622.
20. Neurath, M.F., W. Strober, and Y. Wakatsuki. 1994. The murine Ig3' α enhancer is a target site with repressor function for the B cell lineage-specific transcription factor BSAP (NF-HB, S α -BP). *J. Immunol.* 153:730–742.
21. Singh, M., and B.K. Birshtein. 1996. Concerted repression of an immunoglobulin heavy-chain enhancer, 3' α -E(12). *Proc. Natl. Acad. Sci. USA.* 93:4392–4397.
22. Grant, P.A., C.B. Thompson, and S. Pettersson. 1995. IgM receptor-mediated transactivation of the IgH 3' enhancer couples a novel Elf-1-AP-1 protein complex to the developmental control of enhancer function. *EMBO (Eur. Mol. Biol. Organ.) J.* 14:4501–4513.
23. Michaelson, J.S., M. Singh, C.M. Snapper, W.C. Sha, D. Baltimore, and B.K. Birshtein. 1996. Regulation of 3' IgH enhancers by a common set of factors, including κ B-binding proteins. *J. Immunol.* 156:2828–2839.
24. Cogné, M., R. Lansford, A. Bottaro, J. Zhang, J. Gorman, F. Young, H.-L. Cheng, and F.W. Alt. 1994. A class switch control region at the 3' end of the immunoglobulin heavy chain locus. *Cell.* 77:737–747.
25. Flanagan, J.G., and T.H. Rabbitts. 1982. Arrangement of human immunoglobulin heavy chain constant region genes implies evolutionary duplication of a segment containing γ , ϵ and α genes. *Nature (Lond.)* 300:709–713.
26. Hofker, M.H., M.A. Walter, and D.W. Cox. 1989. Complete physical map of the human immunoglobulin heavy chain constant region gene complex. *Proc. Natl. Acad. Sci. USA.* 86:5567–5571.
27. Kawamura, S., and S. Ueda. 1992. Immunoglobulin C_H gene family in hominoids and its evolutionary history. *Genomics.* 13:194–200.
28. Hammarstrom, L., A.O. Carbonara, M. DeMarchi, G. Lefranc, M.-P. Lefranc, and C.I.E. Smith. 1987. Generation of the antibody repertoire in individuals with multiple immunoglobulin heavy chain constant region gene deletions. *Scand. J. Immunol.* 25:189–194.
29. Gualandi, G., D. Frezza, A. Scotto d'Abusco, E. Bianchi, S. Gargano, S. Giorgi, A. Fruscalzo, and E. Calef. 1995. Integration of an Epstein-Barr virus episome 3' into the gene encoding immunoglobulin heavy-chain α 1 in a lymphoblastoid line. *Gene.* 166:221–226.
30. Chen, C., and B.K. Birshtein. 1995. A region of 20-bp repeats lies 3' of human Ig Ca1 and Ca2 genes. *Int. Immunol.* 8: 115–122.
31. Kang, H.K., and D.W. Cox. 1996. Tandem repeats 3' of the IGHA genes in the human immunoglobulin heavy chain cluster. *Genomics.* 35:189–195.
32. Takahasi, N., S. Ueda, M. Obata, T. Nikaido, S. Nakai, and

- T. Honjo. 1982. Structure of human immunoglobulin gamma genes: implications for evolution of a gene family. *Cell*. 29: 671–679.
33. Mills, F.C., M.P. Mitchell, N. Harindranath, and E.E. Max. 1995. Human Ig S γ regions and their participation in sequential switching to IgE. *J. Immunol.* 155:3021–3036.
 34. Rao, V.N., K. Huebner, M. Isobe, A. Ar-Rushdi, C.M. Croce, and E.S.P. Reddy. 1989. *elk*, tissue-specific ets-related genes on chromosomes X and 14 near translocation breakpoints. *Science (Wash. DC)*. 244:66–70.
 35. Bensmana, M., and M.-P. Lefranc. 1990. Gene segments encoding membrane domains of the human immunoglobulin gamma 3 and alpha chains. *Immunogenetics*. 32:321–330.
 36. Siebenlist, U., L. Hennighausen, J. Battey, and P. Leder. 1984. Chromatin structure and protein binding in the putative regulatory region of the c-myc gene in Burkitt lymphoma. *Cell*. 37:381–391.
 37. Bergman, Y., D. Rice, R. Grosschedl, and D. Baltimore. 1984. Two regulatory elements for immunoglobulin κ light chain gene expression. *Proc. Natl. Acad. Sci. USA*. 81:7041–7045.
 38. Bensmana, M., H.S. Huck, G. Lefranc, and M.-P. Lefranc. 1988. The human immunoglobulin pseudo-gamma IGHGP gene shows no major structural defect. *Nucleic Acids Res.* 16: 3108.
 39. Grant, P.A., V. Arulampalam, L. Arhlund-Richter, and S. Pettersson. 1992. Identification of Ets-like lymphoid specific elements within the immunoglobulin heavy chain 3' enhancer. *Nucleic Acids Res.* 20:4401–4408.
 40. Janknecht, R., and A. Nordheim. 1993. Gene regulation by Ets proteins. *Biochim. Biophys. Acta*. 1155:346–356.
 41. Amin, J., J. Ananthan, and R. Voellmy. 1988. Key features of heat shock regulatory elements. *Mol. Cell. Biol.* 8:3761–3769.
 42. Xiao, H., and J.T. Lis. 1988. Germline transformation used to define key features of heat-shock response elements. *Science (Wash. DC)*. 239:1139–1142.
 43. Zimarino, V., C. Tsai, and C. Wu. 1990. Complex modes of heat shock factor activation. *Mol. Cell. Biol.* 10:752–759.
 44. Schuetz, T.J., G. Gallo, J.L. Sheldon, P. Tempst, and R.E. Kingston. 1991. Isolation of α cDNA for HSF2: evidence for two heat shock factor genes in humans. *Proc. Natl. Acad. Sci. USA*. 88:6911–6915.
 45. Rabindran, S.K., G. Giorgi, J. Clos, and C. Wu. 1991. Molecular cloning and expression of a human heat shock factor, HSF1. *Proc. Natl. Acad. Sci. USA*. 88:6906–6910.
 46. Bondeson, M.L., N. Dahl, H. Malmgren, W.J. Kleijer, T. Tonnesen, B.M. Carlberg, and U. Pettersson. 1995. Inversion of the IDS gene resulting from recombination with IDS-related sequences is a common cause of the Hunter syndrome. *Hum. Mol. Genet.* 4:615–621.
 47. Naylor, J.A., D. Buck, P. Green, H. Williamson, D. Bentley, and F. Giannelli. 1995. Investigation of the factor VIII intron 22 repeated region (int22h) and the associated inversion junctions. *Hum. Mol. Genet.* 4:1217–1224.
 48. Michaelson, J.S., S.L. Giannini, and B.K. Birshstein. 1995. Identification of 3' α -hs4, a novel Ig heavy chain enhancer element regulated at multiple states of B cell differentiation. *Nucleic Acids Res.* 23:975–981.
 49. Erkine, A.M., C.C. Adams, T. Dikaen, and D.S. Gross. 1996. Heat shock factor gains access to the yeast HSC82 promoter independently of other sequence-specific factors and antagonizes nucleosomal repression of basal and induced transcription. *Mol. Cell. Biol.* 16:7004–7017.
 50. Tsukiyama, T., P.B. Becker, and C. Wu. 1994. ATP-dependent nucleosome disruption at a heat-shock promoter mediated by binding of GAGA transcription factor. *Nature (Lond.)*. 367:525–532.
 51. Metz, K., J. Ezernieks, W. Sebald, and A. Duschl. 1996. Interleukin-4 upregulates the heat shock protein Hsp90a and enhances transcription of a reporter gene coupled to a single heat shock element. *FEBS (Fed. Eur. Biochem. Soc.) Lett.* 385: 25–28.
 52. Jenuwein, T., W.C. Forrester, R.G. Qiu, and R. Grosschedl. 1993. The immunoglobulin mu enhancer core establishes local factor access in nuclear chromatin independent of transcriptional stimulation. *Genes Dev.* 7:2016–2032.
 53. Forrester, W.C., C. van Genderen, T. Jenuwein, and R. Grosschedl. 1994. Dependence of enhancer-mediated transcription of the immunoglobulin mu gene on nuclear matrix attachment regions. *Science (Wash. DC)*. 265:1221–1225.
 54. Chung, J.H., M. Whitely, and G. Felsenfeld. 1993. A 5' element of the chicken β -globin domain serves as an insulator in human erythroid cells and protects against position effect in drosophila. *Cell*. 74:505–514.