MOLECULAR AND CELLULAR BIOLOGY, Oct. 1988, p. 3975–3987
0270-7306/88/103975-13$02.00/0

Vol. 8, No. 10

B-Cell Control Region at the 5′ End of a Major Histocompatibility Complex Class II Gene: Sequences and Factors

A. DORN, H. J. FEHLING, W. KOCH, M. LE MEUR, P. GERLINGER, C. BENOIST, AND D. MATHIS*

Laboratoire de Génétique Moléculaire des Eucaryotes du Centre National de la Recherche Scientifique, Unité 184 de Biologie Moléculaire et de Génie Génétique de l'Institut National de la Santé et de la Recherche Médicale, Faculté de Médecine, Institut de Chimie Biologique, 11, rue Humann, 67085 Strasbourg Cedex, France

Received 18 April 1988/Accepted 18 July 1988

Transcription of major histocompatibility complex class II genes is elaborately regulated. Mouse class II genes are transcribed primarily in B cells, peripheral macrophages and interdigitating cells, and thymic cortical and medullary cells. In this study, we began to identify the DNA sequences and protein factors that control expression of a class II gene in B cells, addressing in particular how closely they resemble those that regulate immunoglobulin gene expression. We describe a region upstream of the E_α gene that is crucial for its transcription in the B cells of transgenic mice but is less important in cultured B-cell lines. The sequence of this region reveals several familiar motifs, including a second X-Y pair reminiscent of that residing in the promoter-proximal region of all class II genes, a B motif strikingly homologous to that associated with the immunoglobulin kappa gene enhancer, several Ephrussi motifs, and a Pu box-like sequence very similar to that implicated in simian virus 40 and lymphotrophic papovavirus expression in B cells. Careful study of the proteins that bind specifically to these different motifs prompts us to suggest that major histocompatibility complex class II and immunoglobulin genes rely on quite different factors to achieve B-cell-specific expression.

Studies on tissue-specific regulation of transcription have usually focused on genes that operate in one cell lineage (for reviews, see references 17 and 44). The genes coding for albumin, insulin, the globins, and the immunoglobulins are good examples; each is expressed in a single differentiated cell type as the result of tight transcriptional control. In contrast, class II genes of the major histocompatibility complex (MHC) operate in a broader set of contexts; cells as diverse as macrophages, thymic epithelial cells, and B lymphocytes all express MHC class II genes. An important question to address is whether unique or overlapping sets of DNA sequences modulate expression in these disparate cells.

Another question to pursue is whether controls operating on class II gene expression in a certain cell type bear any relationship to controls that operate on other genes expressed specifically in the same type of cell. For example, MHC class II molecules and immunoglobulins are both prominent B-cell surface markers, and they appear at similar stages of differentiation. The precise patterns of their expression within the B-lymphocyte lineage do differ, however (for reviews, see references 20 and 70). Murine MHC class II molecules are displayed prominently on mature B cells but are generally absent from the surface of pre-B and immunoglobulin-secreting plasma cells. In addition, the MHC class II genes respond to effectors which seem not to influence immunoglobulin gene transcription and vice versa. Despite (perhaps because of) these differences, it would be of great interest to compare the mechanisms which control immunoglobulin and MHC class II gene expression in B cells. Is there more than one route to achieve specific expression in the B-lymphocyte lineage? What molecular events dictate the precise pattern of transcription through B-cell differentiation? What mechanism(s) permits fluctuation of expression under the influence of lymphokines and other effectors?

Much recent effort has been devoted to unravelling the mechanisms that control immunoglobulin gene expression. Evidence exists that the B-cell specificity of transcription is dictated by a tissue-specific enhancer (7, 25, 47, 51, 54, 59), a tissue-specific promoter (19, 21, 26, 45, 55), tissue-specific silencer elements (32, 33, 37, 73, 76), and perhaps sequences that reside in the body of the gene (27). Within these different elements, crucial 6- to 12-base-pair (bp) sequence motifs have been delineated and protein factors which bind to them have been identified. The most emphasis has been placed on the following three motifs. (i) Sequence comparisons have revealed a conserved octamer (ATTTGCAT) in the heavy- and light-chain promoters and in the heavy-chain enhancer (19, 52). This motif was shown by transfection studies to be of great functional importance (6, 16, 23, 42, 78). The octamer serves as the recognition site for lymphoid-cell-specific and ubiquitous protein factors (1, 8, 11, 31, 38, 39, 43, 49, 60, 62, 63, 67–69). (ii) Gel retardation assays have revealed a site in the κ gene enhancer that is recognized by a B-cell-specific factor (2, 63). Mutation of this so-called B sequence (GGGGACTTTCC) abolishes expression of a linked gene in B lymphocytes (42). Further study revealed that this factor is inducible in non-B cells by phorbol esters (64), perhaps a correlate of the finding of similar motifs in the simian virus 40 (SV40) enhancer, the human immunodeficiency virus long terminal repeat, and the *H-2K* gene (4, 34, 50). The B motif, when polymerized, can act as a lymphoid-cell-specific enhancer (56). (iii) An in vivo footprinting technique has revealed B lineage-specific protein contacts on the heavy-chain enhancer (10, 18). Examination of these contacts allowed the definition of a consensus Ephrussi sequence (GCCAGGTGGC) which has been observed at multiple locations in both heavy- and light-chain enhancers (63). Mutation of an Ephrussi sequence can affect expression of linked genes in B cells, but because of redundancy of this element, generation of a pronounced effect often requires simultaneous mutation of at least two Ephrussi motifs (23, 42). Ubiquitous nuclear proteins have been shown to bind to at least some of this set of sequences (63, 75).

* Corresponding author.

Much less is known about the regulation of MHC class II genes in B cells. Two important sequence elements, the X and Y boxes, have been identified in the −50 to −150 region, but their influence is felt in all cell types that express class II antigens (9, 15, 66). Both X and Y serve as recognition sites for apparently ubiquitous sequence-specific DNA-binding proteins (14, 15, 30, 37a, 48, 66); in fact, the latter is recognized by a CCAAT box-binding factor, NF-Y (14). The 5′-flanking region of the murine class II gene E_α has B-cell-specific enhancer activity (37a; W. Koch, C. Benoist, and D. Mathis, submitted for publication). This activity is due to the interplay of multiple elements. There is a strong nonspecific enhancer in the −2,172 to −1,148 region and a weaker but B-cell-specific enhancer in the −215 to +12 region. Other class II genes have been purported to possess lymphoid-cell- or B-cell-specific enhancers in the 5′-flanking region (24, 71) or within an intron (71, 74). In short, we are just beginning to understand how MHC class II genes function specifically in B cells.

In this report, we concentrate on a region (−1906 to −1180) that is required for transcription of the E_α gene in B cells, but not in other class II antigen-expressing cells, of transgenic mice. Surprisingly, this sequence is not crucial for expression of E_α in transfected B-cell lines. This stretch of DNA bears several interesting sequence motifs: a second X and Y box in reverse orientation, a motif very similar to the κ enhancer B sequence, Ephrussi sequences, etc. A careful analysis of the proteins that bind to these different motifs suggests that different factors mediate B-cell-specific expression of immunoglobulin and MHC class II genes.

MATERIALS AND METHODS

Creation and characterization of transgenic mice. The WE32-25 transgenic line carries a wild-type E^k_α gene and has been described previously (15). The Sma58 and Sma65 lines were derived from C57BL/6 × SJL F_2 hybrid eggs injected with an E^k_α fragment spanning positions −1180 to +6800. The lines were propagated by back-crossing to C57BL/6.

Total RNA was prepared from mouse tissues as previously described (41). For anti-immunoglobulin panning experiments, we followed the technique described in reference 79. RNA analysis by quantitative S1 nuclease mapping was performed as previously described (15, 41). Cytofluorimetric profiles were obtained after indirect staining with the biotinylated monoclonal antibody 14.4.4 (anti-E_α) and phycoerythrin-avidin. Electronic gating on immunoglobulin M (IgM)-positive B lymphocytes was set after staining with a fluorescein isothiocyanate-conjugated rabbit anti-mouse IgM reagent.

Transfections. Transfections into the human B-lymphoblastoid cell line Raji were performed by the DEAE-dextran technique as described in detail elsewhere (37a), and cytoplasmic RNA transcribed from the transfected DNA was quantitated by S1 nuclease mapping.

Nucleotide sequencing. The sequenced E^k_α fragments originate from clone λA/Jp34.13 (46), subcloned into M13 in both orientations. Nested sets of deletions were generated by controlled exonuclease III-mung bean nuclease digestions as described in reference 81. The sequences of these templates were determined by the dideoxynucleotide chain termination method (61). For a depiction of the direction and extent of the individual sequences obtained, see Fig. 4B. All of the sequence shown in Fig. 4A was confirmed in both orientations, except for the stretch downstream of the *Xmn*I sites, which we have previously determined (46).

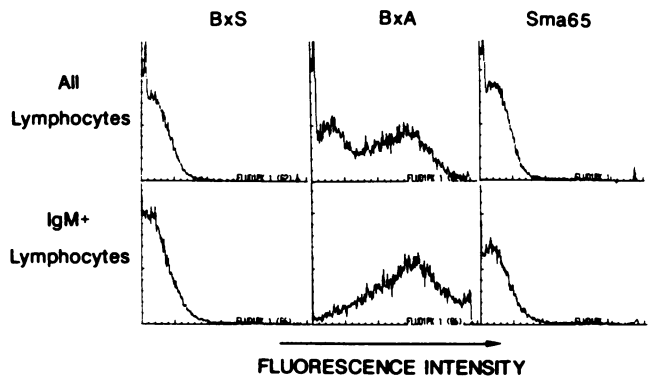

FIG. 1. Lack of E^k_α expression on B lymphocytes from Sma transgenic mice. Splenocytes from an Sma65 transgenic mouse or from E_α-negative and -positive controls (B × S and B × A, respectively) were stained with fluorescent anti-E_α monoclonal antibodies and analyzed by flow cytometry. Two profiles are shown for each animal, one corresponding to the whole splenocyte population (all lymphocytes) and the other corresponding to IgM-positive B lymphocytes selected by electronic gating.

In vitro DNA-protein interactions. DNA-binding proteins were extracted from nuclei with 1 M NaCl by the protocol of Dignam et al. (13), modified as described previously in detail (14). The gel retardation and methylation interference assays have also been described elsewhere (14, 15, 30).

RESULTS AND DISCUSSION

The −1,906 to −1,180 region is required for expression of the E_α gene in B cells of transgenic mice. Several mouse lines carrying an MHC class II transgene have been produced over the past few years (15, 41, 58, 72a, 77, 80). Most of these derive from injection of the E_α gene into embryos of C57BL/6 × SJL mice. The recipient strain does not transcribe its own E_α loci because of a large deletion that encompasses the promoter region and exon 1. If 1,906 bp of 5′-flanking DNA is included on the injected fragment, the E_α transgene is expressed in all of the usual compartments of the immune system: B cells, peripheral macrophages and interdigitating cells, and the thymic epithelium and cortex (15, 41, 72a, 77).

However, a different pattern of expression occurs when fewer 5′-flanking sequences are present on the injected DNA. The two transgenic lines presented here (Sma58 and Sma65) derive from C57BL/6 × SJL embryos injected with an E^k_α fragment that bears only 1,180 bp of 5′-flanking DNA. Both lines carry low transgene copy numbers (two to four), and the injected DNA is transmitted in a Mendelian fashion (data not shown). We noticed quickly that both Sma58 and Sma65 showed defective expression of the transgene in their B-cell compartments.

Figure 1 illustrates the dearth of E_α on B cells from one such line, Sma65. Splenocytes were doubly decorated with a fluorescein isothiocyanate-conjugated anti-IgM reagent and a phycoerythrin-avidin-biotin-labeled anti-*E* reagent. Gating on IgM-positive cells demonstrated that essentially all of the B cells displayed E_α in the C57BL/10 × A/J positive control (Fig. 1) and essentially none did so in the C57BL/6 × SJL recipient strain. Strikingly, the E_α transgene was not detectably expressed on B cells from Sma65 mice. The same result was obtained with Sma58 animals (data not shown).

Since lack of E_α expression on B cells ensued from a deletion of 5′-flanking DNA, it was difficult to envisage how

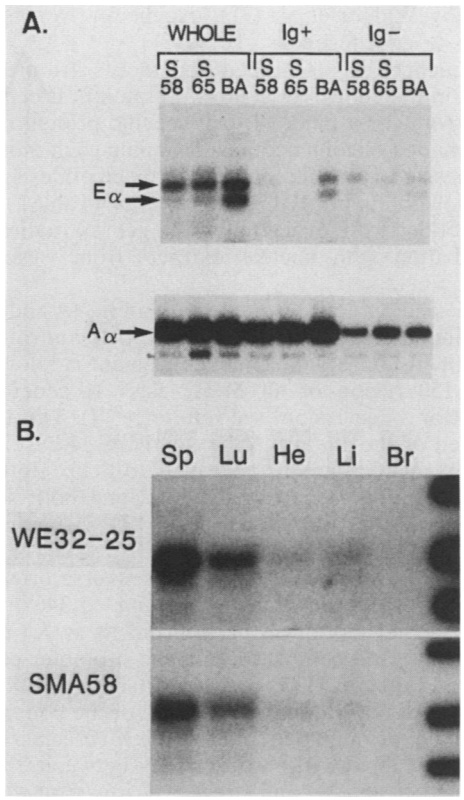

FIG. 2. Cell type-specific and tissue-specific expression of E_α RNA in Sma transgenic mice. (A) Spleen cells. Cytoplasmic RNA was prepared from splenocytes, with or without fractionation into nonadherent immunoglobulin-positive (Ig^+) and -negative (Ig^-) cells, by panning on anti-immunoglobulin serum-coated plates. The RNA was tested by S1 nuclease mapping with probes for the transgene-encoded E_α mRNA or for the host-encoded A_α mRNA as a control. Four micrograms of RNA was used in the whole-splenocyte (WHOLE) lanes, 1 μg was used in the Ig^+ lanes, and 2 μg was used in the Ig^- lanes. Lanes S58 and S65 contained cells from Sma58 or Sma65 transgenic mice, and lanes BA contained cells from a control (C57BL/10 × A/J) F_1 positive control mouse. (B) Other tissues. RNAs were extracted from various tissues of an Sma58 mouse or a wild-type transgenic (WE32-25) mouse which carries a full-length E^k_α gene. E_α mRNA was tested by S1 nuclease mapping as previously described (41). Abbreviations: Sp, spleen; Lu, lung; He, heart; Li, liver; Br, brain.

this defect could derive from a post-transcriptional event. Nevertheless, we sought to demonstrate a deficiency at the level of E_α mRNA. The organ which normally expresses the highest level of MHC class II mRNA is the spleen. This organ is composed of diverse cell types, including macrophages and interdigitating cells that strongly express class II molecules, B cells, and erythrocytes and T cells that do not express class II molecules. To enrich for B cells, we conducted a panning experiment. Splenocytes were incubated on plastic dishes to remove adherent cells, and then B cells were positively selected on anti-immunoglobulin antibody-coated petri dishes. RNA was prepared from these populations, and E_α mRNA was evaluated by quantitative S1 mapping. Figure 2A shows a clear diminution of E_α mRNA in the immunoglobulin-positive B-cell-enriched population from both the Sma58 and Sma65 transgenic lines relative to C57BL/10 × A/J controls. Such a decrease was not evident in the immunoglobulin-negative B-cell-depleted population. A_α mRNA, encoded by the host MHC, served as an internal control in these experiments.

The defect in E_α expression is specific to B cells of Sma58 and Sma65 transgenic mice. S1 nuclease analysis demonstrated that the tissue distribution of E_α mRNA is normal—essentially indistinguishable from that observed with a wild-type line, WE32.25 (Fig. 2B). There were high E_α mRNA levels in spleen and lung tissues, lower levels in liver and heart tissues, and no detectable signal in brain tissue. In these positive tissues, including the spleen, class II molecules occurred on interdigitating cells, certain epithelial cells, and macrophages. Expression of E_α on non-B cells in the spleens, lymph nodes, and thymuses of Sma transgenic mice has been confirmed by immunohistological studies reported elsewhere (72a). These analyses reveal clearly, for example, that thymic cortical and medullary cells stain normally for E_α. Finally, we also showed that gamma interferon strongly stimulates E_α RNA levels in peritoneal cells (mostly macrophages) from Sma transgenic mice (data not shown).

We conclude, then, that the −1,906 to −1,180 region is crucial for expression of the E_α gene in B cells of transgenic mice but that it has little influence on other class II positive cells or on the response to gamma interferon. This assertion is based on the two lines we produced and two similar lines recently described by Widera and co-workers (77). This stretch of DNA will be referred to as the B-cell control region.

Thus, different cell types rely on different DNA sequence blocs to control E_α transcription. This mode of regulation is similar to that used by a number of *Drosophila* gene systems (22, 29). There are fewer examples of mammalian genes displaying such a cell type-specific partitioning of regulatory elements. Noteworthy is the lysozyme gene, with its widely spaced macrophage-specific and steroid-responsive elements (72). There are, on the other hand, several examples of the same regulatory elements being implicated in expression of a given gene in two different tissues (28, 65).

The −1,906 to −1,180 region plays only a minor role in expression of E_α in transfected B-cell lines. A more precise delineation of the critical DNA sequences in the B-cell control region is clearly needed. This would be a monumental task if we continued to use transgenic mice as an assay system. Hence, we sought to evaluate the influence of the −1,906 to −1,180 region on expression of E_α genes transfected into cultured B-cell lines. Unfortunately, none of the existing lines originate from mice that are E_α nonexpressers, creating difficulties in distinguishing transcripts that originate from the transfected versus endogenous genes. We avoided the use of a reporter gene, because it was not known whether E_α introns and exons carry transcriptional regulatory elements and because it seemed important to test the same constructs that had been evaluated in transgenic mice.

To circumvent this problem, we chose to transfect E_α into a human B-lymphoblastoid line—Raji; we already knew that at least some E_α transcription signals function adeptly in this line (37a). Figure 3 shows an S1 nuclease analysis of RNA extracted from Raji cells 48 h after transfection with E_α constructs bearing deletions at the 5′ end. Only a minor (2.7-fold by densitometry) reduction in E_α transcript level occurred upon deletion to position −1,180, and this approximate level was maintained upon further deletion to position −215. These results were obtained twice with two independent plasmid preparations.

A similar lack of influence was observed with the murine B-cell line M12 (data not shown). In some experiments, we

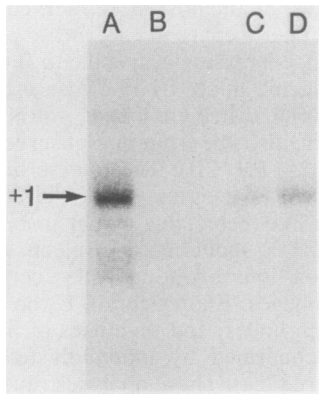

FIG. 3. E_α expression in transfected B-lymphoma cells. Human Raji cells were transfected with plasmids carrying different E^k_α gene fragments, and cytoplasmic RNA transcribed from the transfected DNA was quantitated by S1 nuclease analysis. Lanes: A, transfection with a full-length E^k_α plasmid (−2173 to +6800); B, mock transfection; C and D, transfections with plasmids carrying 5′-truncated E_α genes (to −215 and −1180, respectively). Numbering is from the major transcription start site (46).

attempted to exploit a 6-bp mismatch between the transfected (or injected) E_α gene and the endogenous E_α locus, a mismatch that had been created by transformation of a restriction enzyme site (15). Another E_α derivative was also tested, which carries another tag mutation in the 5′ untranslated region, allowing detection of the transfected E_α gene (S. Viville, unpublished data). Results of these transfections into M12 cells were consistent with the Raji cell experiment; densitometric analysis demonstrated only a 1.5- to 2-fold reduction in E_α transcripts resulting from a 5′ deletion to −1,180.

We conclude, then, that the −1,906 to −1,180 region exerts much less influence in directing expression in cultured B-lymphoma lines, as measured by a short-term transfection assay, than it does in B cells of transgenic mice. Although we cannot rule out a low level of expression in B cells of Sma58 mice, fluorescence-activated cell sorter analysis points to a more-than-10-fold reduction from wild-type levels of E_α mRNA in normal resting B cells. In contrast, deletion from −1,906 to −1,180 only resulted in a 1.5- to 2.7-fold reduction in RNA levels after transfection. Such a dichotomy between results in transfection versus transgenic experimental systems has been documented previously (12, 28, 57). It may reflect events that occur early in the B-cell differentiation pathway or the influence of lymphokines whose effect on E_α would be mediated through the upstream region—events that obviously would be bypassed in tissue culture cells. Precedents can be cited for this type of phenomenon (1, 35). The dichotomy may reflect, more trivially, the artificial nature of transfections into cultured cells (high DNA copy number, extrachromosomal chromatin packaging, etc.).

Sequence analysis revealed familiar motifs in the B-cell control region, some of which are recognized by sequence-specific DNA-binding proteins. An alternative strategy to define more precisely the DNA sequences that direct E_α expression in B cells might be to locate motifs on the −1,906 to −1,180 fragment that serve as recognition sites for B-cell-specific regulatory proteins. Thus, we sequenced the 5′-flanking region of the E^k_α gene from positions −2,173 to +12 by following the strategy illustrated in Fig. 4. The sequence obtained was very similar to the E^d_α sequence recently described by Widera et al. (77), except for 16 single-base differences (Fig. 4, legend).

The sequence was searched for motifs (i) previously implicated in the control of gene expression in B cells or (ii) previously shown to bind a B-cell-specific protein or both. Double-stranded oligonucleotides spanning each motif were then synthesized, as were control oligonucleotides spanning the immunoglobulin or MHC gene homolog (Table 1). These oligonucleotides were evaluated in the gel retardation assay after incubation with nuclear extracts from various cell types.

The E^k_α 5′-end sequence is presented in Fig. 4A and reveals the following interesting motifs in the B-cell control region.

X-Y motif. A conserved sequence element is found in the −50 to −150 region of all MHC class II genes so far examined (for a discussion, see reference 14). This element is composed of the 14-base Y box and the 14-base X box, separated by an invariant distance of 17 to 18 bp. Both X and Y are transcriptional control elements, and both serve as recognition sites for sequence-specific DNA-binding proteins.

Astonishingly, a second X-Y pair, in reverse orientation, was located between positions −1,392 and −1,346 on the E_α gene; this second motif will be referred to as X′-Y′. It is boxed in Fig. 4 and compared with the promoter-proximal X-Y pair in Fig. 5A. The comparison between X-Y and X′-Y′ illustrates several features typical of the conservation of X and Y boxes in human and murine MHC class II gene promoters: (i) more stringent homology between Y and Y′ than between X and X′ and (ii) nearly invariant distances between the motifs. Intrigued by this finding, we searched the 5′ ends of the two other MHC class II genes for which there is extensive sequence information. Indeed, an X′-Y′ motif was found for A_β at about positions −1590 and −1545 (40) and for E_β at about positions −990 to −940 (24). These upstream motifs are also in inverted orientation (Fig. 5A).

Previous gel retardation experiments have revealed specific binding of a nuclear protein(s) to the E_α X box (15; Koch et al., in press). A typical result is shown in the leftmost panel of Fig. 6A. A set of three bands labeled NF-X became increasingly prominent with higher concentrations of poly(dI-dC). That this triplet actually represents protein binding to the X box has been confirmed by methylation interference mapping (37a; V. Kouskoff, unpublished data). The other bands probably represent nonspecific protein-DNA interactions, because they are relatively sensitive to poly(dI-dC) and because they also occurred with a control oligonucleotide with a random sequence replacing the X box (center panel). As shown in the rightmost panel of Fig. 6A, the X′ oligonucleotide was also capable of binding the NF-X protein(s) as efficiently as the X oligonucleotide. The extraneous bands differed by the two oligonucleotides, probably because of differences in the X- and X′-flanking sequences.

Past experiments have also documented specific protein binding to the E_α Y box (14, 15, 30, 37a). The protein of interest, NF-Y, is known to be a ubiquitous CCAAT box-binding protein. As indicated by the gel retardation experiment illustrated in Fig. 6B, the Y and Y′ sequences appeared to bind the same protein. This conclusion has been confirmed by methylation interference mapping of contact residues (Fig. 6C) and by competition experiments (data not shown). Careful titration of binding efficiency with the two oligonucleotides indicated that the Y and Y′ boxes had approximately equal affinities for NF-Y (data not shown).

Their striking sequence homology with the promoter-proximal X-Y pair, their conservation in class II α- and

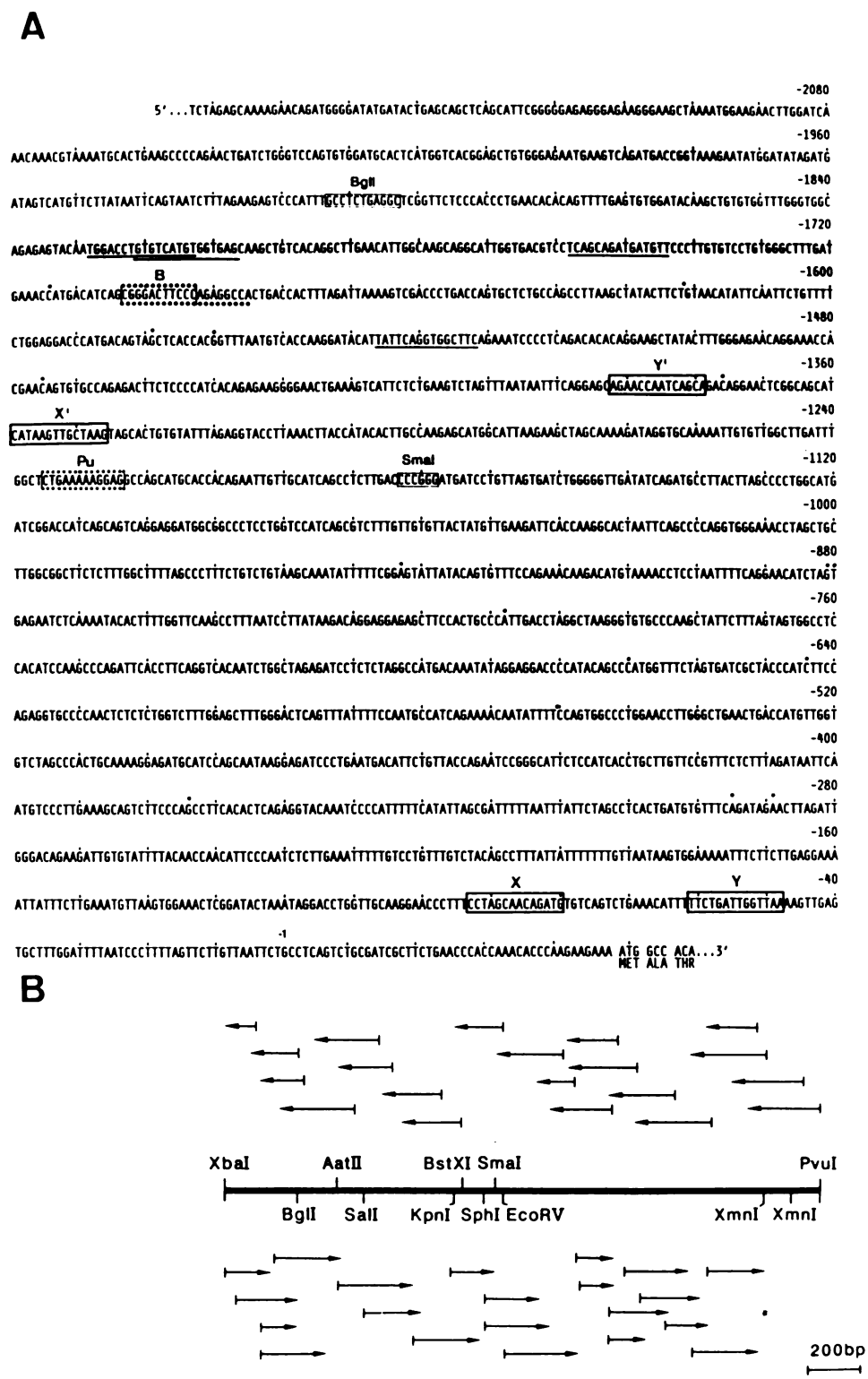

FIG. 4. Sequence of the E_α^k 5'-flanking region. (A) The sequence extends from the *Xba*I site at position −2172 to the beginning of the protein-coding sequence. Numbering begins at the major transcriptional start site (46). Thick boxes highlight the X and Y boxes—elements that are conserved in all MHC class II genes—and their X' and Y' duplicates. The B and Pu box discussed in the text are indicated by stippled boxes. The Ephrussi motif homology sequences are underlined. Thin boxes delineate the *Bgl*I and *Sma*I restriction enzyme sites used to prepare DNA fragments for injection in transgenic mouse experiments. The dots above the sequence indicate positions at which our E_α^k sequence differs from the E_α^d sequence recently published by Widera et al. (77). They are positioned at −1714, −1621, −1579, −1570, −1475, −1376, −943, −881, −880, −808, −670, −644, −560, −374, −295, and −289. (B) Arrows summarize the strategy used to determine the sequence. Several single-cut restriction enzyme sites are also indicated.

TABLE 1. Oligonucleotides used in this study

Motif	Sequence
Y	5′ GTCTGAAACATTTTTCTGATTGGTTAAAAGTTGAGTGCT 3′ 3′ CAGACTTTGTAAAAAGACTAACCAATTTTCAACTCACGA 5′
Y′	5′ AATTTCAGGAGCAGAACCAATCAGCAGACAGGAACTCGG 3′ 3′ TTAAAGTCCTCGTCTTGGTTAGTCGTCTGTCCTTGAGCC 5′
E_αBE40	5′ ATGACATCAGCGGGACTTCCCAGAGGCCACTGACCACTTT 3′ 3′ TACTGTAGTCGCCCTGAAGGGTCTCCGGTGACTGGTGAAA 5′
E_αB*E40	5′ ATGACATCAGCTTTACTTCCCAGAGGCCACTGACCACTTT 3′ 3′ TACTGTAGTCGAAATGAAGGGTCTCCGGTGACTGGTGAAA 5′
E_αBE*40	5′ ATGACATCAGCGGGACTTCCCAGATCAAACTGACCACTTT 3′ 3′ TACTGTAGTCGCCCTGAAGGGTCTAGTTTGACTGGTGAAA 5′
κBE40	5′ TCTCAACAGAGGGGACTTTCCGAGAGGCCATCTGGCAGTT 3′ 3′ AGAGTTGTCTCCCCTGAAAGGCTCTCCGGTAGACCGTCAA 5′
E_αE3	5′ CCTGTGTCATGTGGTGAGCAAG 3′ 3′ GGACACAGTACACCACTCGTTC 5′
E^*_αE3	5′ CCTGTGTCCTGGTGTGAGCAAG 3′ 3′ GGACACAGGACCACACTCGTTC 5′
μE3	5′ AGCAGGTCATGTGGCAAGGCTA 3′ 3′ TCGTCCAGTACACCGTTCCGAT 5′
X	5′ AAGGAACCCTTTCCTAGCAACAGATGTGTCAGTCTGAA 3′ 3′ TTCCTTGGGAAAGGATCGTTGTCTACACAGTCAGACTT 5′
X′	5′ TACACAGTGCTACTTAGCAACTTATGATGCTGCCGAGT 3′ 3′ ATGTGTCACGATGAATCGTTGAATACTACGACGGCTCA 5′
E_αB21	5′ ATCAGCGGGACTTCCCAGAGG 3′ 3′ TAGTCGCCCTGAAGGGTCTCC 5′
κB21	5′ ACAGAGGGGACTTTCCGAGAG 3′ 3′ TGTCTCCCCTGAAAGGCTCTC 5′
SVB	5′ GATCTGGAAAGTCCCCAG 3′ 3′ ACCTTTCAGGGGTCCTAG 5′
H2K^b	5′ CTGGGGATTCCCCAT 3′ 3′ CATGGACCCCTAAGGGGTAGAT 5′
PuE$_\alpha$	5′ TTGGCTCTGAAAAAGGAGGCCA 3′ 3′ AACCGAGACTTTTTCCTCCGGT 5′
PuSV	5′ TAACCTCTGAAAGAGGAACTTG 3′ 3′ ATTGGAGACTTTCTCCTTGAAC 5′
PuSV*	5′ TAACCTCGTCCAGAGGAACTTG 3′ 3′ ATTGGAGCAGGTCTCCTTGAAC 5′

β-chain genes, and their factor-binding properties all point to a functional importance for the X′ and Y′ boxes. Indeed, recent transfection studies demonstrate an involvement of X′-Y′ in a complex B-cell-specific enhancer associated with the E_α 5′ end (Koch et al., submitted). The head-to-head disposition of the promoter-proximal X-Y and promoter-distal X′-Y′ pairs hints that they could be involved in loop formation by dimerization of factors that bind X or Y or both. This proposition is currently being tested.

B motif. The sequence GGGGACTTTCC is a component of the κ gene enhancer, residing next to the Ephrussi sequence κ-E1. This B motif is recognized by a protein, NF-κB, that at first appeared to occur only in cells that express the κ gene (2, 63, 64). In particular, like κ mRNA

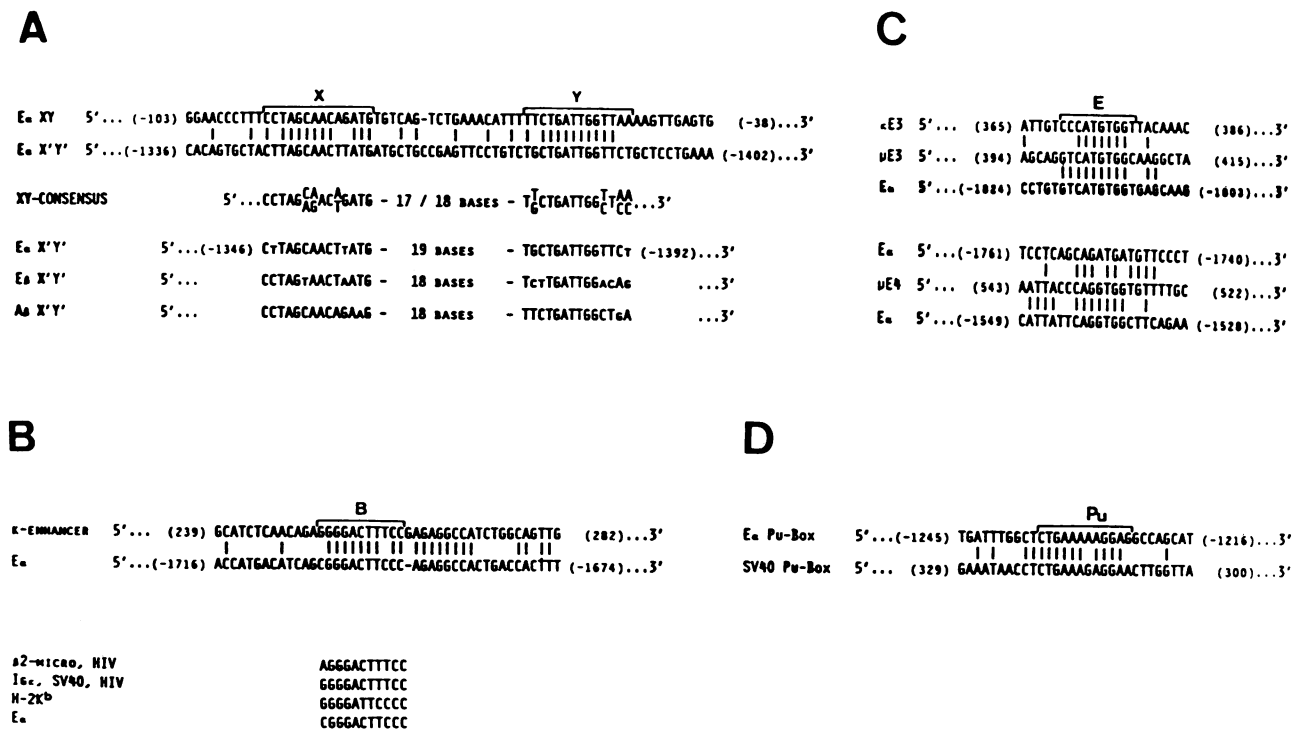

FIG. 5. Sequence homologies. (A) The promoter-distal X'-Y' pair from the E_α B-cell control region is aligned with its promoter-proximal X-Y homolog. Lines between the sequences indicate identical bases. Below is presented the X-Y consensus and an alignment of the E_α, A_β, and E_β X'-Y' motifs, all found on the antisense strand in a promoter-distal location. The A_β and E_β X'-Y' motifs are located at positions −1590 to −1545 and −990 to −940 relative to transcriptional start sites (approximated by homology to the E_α sequence). (B) Homology between the NF-κB binding site in the kappa enhancer and a similar stretch in the B-cell control region of E_α. Below, the B motifs from other genes are aligned; these are known to bind NF-κB or KBFI/H2TFI or both (see the text for references). β2-micro, β2-Microglobulin. (C) Sequence homologies between Ephrussi motifs in the E_α B-cell control region and the most closely related Ephrussi motif in the κ or μ immunoglobulin enhancer (63). (D) Homology between the SV40 Pu box and a similar sequence in E_α.

transcription, NF-κB could be induced in the pre-B-cell line 70Z/3 by lipopolysaccharide (LPS) treatment. More recently, it was established that many cell types, even non-κ gene expressers, can be induced to activate repressed NF-κB simply by treating them with the phorbol ester phorbol myristate acetate (3, 64).

A scan of the E_α 5' end revealed the very similar sequence CGGGACTTCCC at positions −1,703 to −1,693. In fact, the E_α-κ homology extends over a considerable distance, including the adjacent κ-E1 Ephrussi motif. Figure 5B presents a comparison of the relevant stretch. The immunoglobulin gene κ enhancer and E_α sequences actually seem to be members of a family of related motifs. As indicated in Fig. 5B, these motifs occur in the control regions of genes that are expressed primarily in lymphoid cells (the immunoglobulin gene κ enhancer, E_α, and human immunodeficiency virus) and the control regions of genes that are expressed quasiubiquitously (SV40 and *H-2K^b*).

To compare the proteins which bind to the E_α and κ sequences, we synthesized the appropriate 40-mers (E_αBE40 and κBE40) and evaluated their performance in the gel retardation assay. Figure 7A shows the patterns obtained after incubating the κBE40 and E_αBE40 oligonucleotides with a nuclear extract from 70Z/3 cells treated with LPS. Four bands were evident with the κBE40 oligonucleotide, but only two of these occurred with the E_αBE40 oligonucleotide. The bands were present with extracts from all of the cell types tested, whether lymphoid or not, except for band 3, which was detected only with B-lymphocyte extracts or extracts from LPS-activated 70Z/3 cells. To characterize the proteins represented by each band, we conducted a series of methylation interference, mutational sensitivity, and competition experiments (Fig. 7–9).

Band 1. Band 1 represents a protein that binds to the κ Ephrussi sequence. Methylation interference mapping unambiguously demonstrated protein contacts in this region (Fig. 8). Moreover, this band could be outcompeted by κBE40 but not by a 21-mer truncated just before the Ephrussi motif (κB21) (Fig. 9A and B). Interestingly, band 1 was not outcompeted by E_αBE40, which bears an Ephrussi motif-like sequence (Fig. 9E); this is consistent with the finding that gel retardation assays with this oligonucleotide showed no band 1 (Fig. 7A).

Band 2. Band 2 represents a protein that binds in the Ephrussi region of both κBE40 and E_αBE40. This band was unaffected by mutation of the E_α sequence in the 5' half of the B motif (E_αB*E40) but was reduced—actually converted to a doublet—by a mutation in the Ephrussi motif (E_αBE*40) (Fig. 7B). An extensive series of competition experiments also supported this conclusion in that band 2 was outcompeted by κBE40 but not κB21 (Fig. 9A and B); similarly, it was outcompeted by E_αBE40 but not E_αB21 (Fig. 9E and F); it was outcompeted by the B mutant E_αB*E40 but much more weakly by the Ephrussi mutant E_αBE*40 (Fig. 9I to L); and finally, it was not outcompeted by an oligonucleotide carrying the *H-2K^b* or SV40 B-related sequence (Fig. 9M to P). Repeated methylation interference experiments failed to convincingly identify contact bases for band 2 protein-DNA

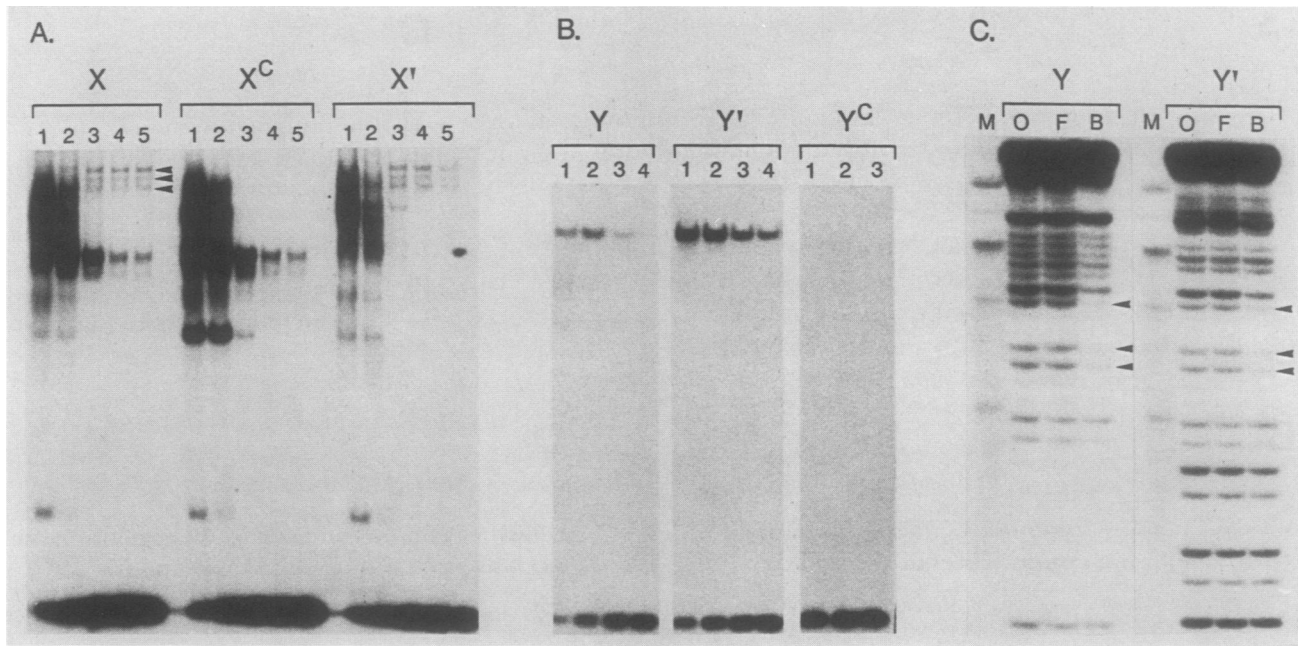

FIG. 6. Comparison of proteins binding to the promoter-proximal X and Y and promoter-distal X′ and Y′ boxes. (A and B) Gel retardation assays. A 1 M NaCl nuclear extract from M12 cells (1 to 3 μg of protein) was incubated with ^{32}P-labeled double-stranded oligonucleotides. Nucleoprotein complexes were resolved from free DNA by polyacrylamide gel electrophoresis and autoradiography. The oligonucleotides correspond to the sequences at and around the X, X′, Y, and Y′ boxes (Table 1). X^c and Y^c are control oligonucleotides which have the same 5′- and 3′-flanking sequences as the X and Y oligonucleotides, but the X and Y box segments have been replaced by random sequences. In panel A, lanes 1 to 5 correspond to increasing concentrations of nonspecific poly(dI-dC) competitor; i.e., 0, 50, 100, 200, and 500 ng/20-μl incubation reaction. The arrowheads indicate bands due to NF-X. In panel B, the poly(dI-dC) amounts were 50, 100, 200, and 500 ng in lanes 1 to 4, respectively. (C) Methylation interference mapping. Contact residues (arrowheads) were identified by methylation interference mapping, as previously described (15), after complex formation between NF-Y and Y or Y′ oligonucleotides. Lanes: O, starting material, partially methylated and then piperidine cleaved; F and B, DNA extracted from a retardation gel, either free (F) or bound to NF-Y (B) before piperidine cleavage. Arrowheads show positions where methylation interferes with protein-DNA complex formation.

complexes. There was variable evidence for contacts on the sense strand of the Ephrussi sequence and on the antisense strand of the 3′ half of the B sequence. The very similar behaviors of the E_α and κ bands 2 in the various competition experiments suggest that they represent the same or very closely related proteins.

Band 3. Band 3 most likely represents the protein termed NF-κB by Sen and Baltimore (63). This band did not occur when the κ oligonucleotide was incubated with a nuclear extract from uninduced 70Z/3 cells or from nonlymphoid cells (Fig. 7C); it did appear after incubation with a nuclear extract from 70Z/3 cells induced by LPS, cycloheximide, or LPS-cycloheximide (Fig. 7A and C; data not shown). Band 3 was outcompeted efficiently by both κBE40 and κB21, as well as by the SV40 and *H-2K^b* oligonucleotides (Fig. 9A, B, M, and N), indicating that it depends on interactions at the kappa B motif. NF-κB did not bind to the E_α B-related sequence, as evidenced by the absence of band 3 in gel retardation assays with a labeled E_α oligonucleotide (Fig. 7A). Some affinity did exist, however, as demonstrated by very weak competition for band 3 formation by cold E_αBE40 and E_αBE*40 but not by E_αB*E40 (Fig. 9E, J, and I).

Band 4. Band 4 represents a protein, probably different from NF-κB, that binds to the κ and E_α B motifs. This conclusion is most convincingly supported by the methylation interference map shown in Fig. 8. In addition, band 4 was clearly stronger in assays with the E_αBE40 and E_αBE*40 oligonucleotides than with the E_αB*E40 oligonucleotide, which bears a mutation in the B motif (Fig. 7). Finally, band 4 formation was outcompeted by κBE40, E_αBE40, and E_αBE*40 (Fig. 9A, C, E, G, J, and L). Interestingly, competition was much less pronounced with κB21 and E_αB21 (e.g., compare Fig. 9A and B or E and F) and was essentially nonexistent with the SV40 and *H-2K^b* B sequences (Fig. 9M to P), suggesting that the protein represented by band 4 also recognizes adjacent sequences and distinguishing the binding properties of the band 3 and 4 proteins.

One is led to question whether band 4 represents binding of a factor first identified as interacting with the *H-2K^b* enhancer, KBF1/H2TF1 (4, 34). The recognition sequence of this factor (GGGGATTCCCC) is very similar to that of NFκB (GGGGACTTTCC). Indeed, both proteins bind to both sequences, although KBF1/H2TF1 has a clear preference for the *H-2K^b* site (5). The proteins are also distinguishable on the basis of cell type distribution and contact sites on the recognition sequence. KBF1/H2TF1 is essentially ubiquitous and contacts all four Gs; NFκB activity is present constitutively only in B cells and contacts only three of the Gs (4, 5, 34, 63, 64). That band 4 could represent KBF1/H2TF1 binding was supported by its detection in assays with all of the cell types and by the observation that all four Gs are contacted in band 4 DNA. However, competition experiments suggest that the band 4 protein has less affinity for the *H-2K^b* sequence than for the immunoglobulin gene sequence (cf. Fig. 9A and M). This may be because the competing *H-2K^b* oligonucleotide is significantly shorter than the immunoglobulin gene oligonucleotide. Note also that κB21 competed less well than κBE40 (cf. Fig. 9A and B or C and D). Alternatively, we may be faced with a third protein which recognizes the B family of motifs.

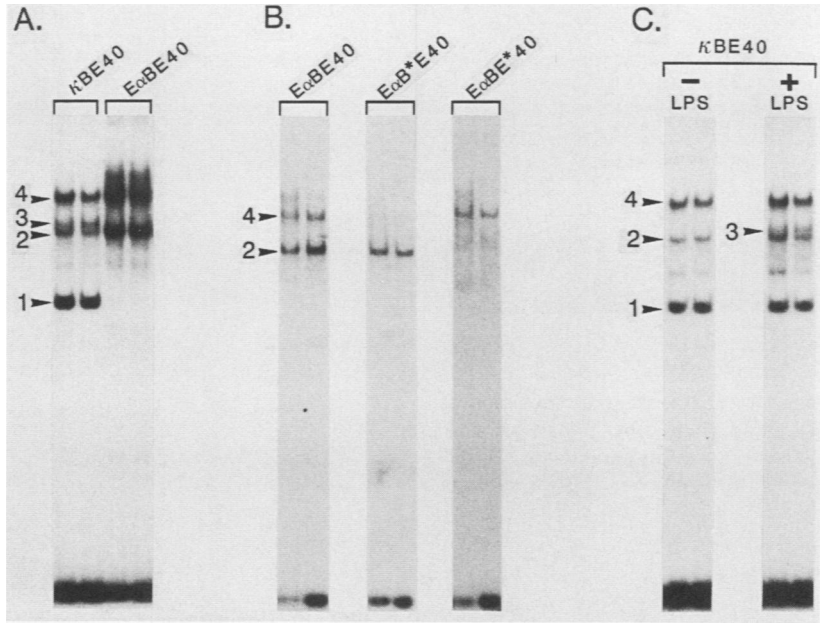

FIG. 7. Binding capacity of the B-Ephrussi motif complex from the E_α B-cell control region and the immunoglobulin gene kappa enhancer. Gel retardation assays were set up as described in the legend to Fig. 6. The 40-mer oligonucleotides carry sequences form the B-Ephrussi motif bloc of E_α or κ (E_αBE40 and κBE40; Table 1). E_αB*E40 and E_αBE*40 are mutations of E_αBE40, with several base changes in the B or Ephrussi motif, respectively. In panels A and B, the 1 M NaCl nuclear extract was from M12 cells. In panel C, the extract was from 70Z/3 cells grown with or without LPS for 16 h before harvest. In all panels, duplicate lanes correspond to incubation mixtures set up with 100 or 300 ng of poly(dI-dC).

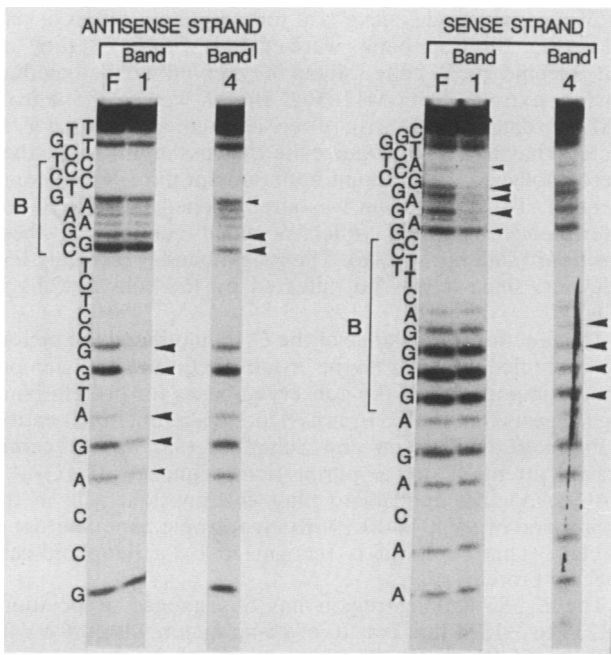

FIG. 8. Protein-DNA contacts in the B-Ephrussi motif bloc. Input DNA in the reaction was the κBE40 oligonucleotide, labeled and partially methylated on the sense or antisense strand. Material displayed in the different lanes was extracted from a retardation gel from either the free DNA band (F) or bands 1 and 4. Arrowheads point to contact bases.

It may be worth pointing out that the patterns we have shown with the immunoglobulin gene B oligonucleotide are significantly more complex than previously published patterns. This complexity is most likely due to the fact that we included the adjacent B and κE1 motifs on the same probe; we thus detected proteins that bind to B (bands 3 and 4), κE1 (band 1), and parts of both (probably band 2). Although it may complicate interpretation of the data, we feel that this strategy is preferable in that it more closely reflects the situation in vivo. B motifs have always been found closely juxtaposed to another element: an Ephrussi sequence in κ and E_α, an interferon response element in *H-2K*^{b} (36), several other motifs in SV40, and an Sp1-binding site in human immunodeficiency virus. Perhaps B proteins always act in concert with immediately adjacent factors.

To summarize, the B-Ephrussi sequence is capable of binding multiple proteins that recognize overlapping sites. Certain of these proteins are found by both the κ and E_α sequences. The functional relevance of the E_α element is not established. Strikingly, however, it does not bind the two proteins most convincingly implicated in immunoglobulin gene expression: NFκB and the κE1-binding protein.

Ephrussi motifs. An in vivo footprinting technique revealed several B-cell-specific protein contacts on the IgH enhancer (10, 18). These contacts clustered in four short stretches (later termed Ephrussi motifs) that had homology with the sequence CAGGTGGC. Three additional Ephrussi motifs were located along the κ gene enhancer, initially by sequence comparison and subsequently by functional studies (10, 42, 63). The seven Ephrussi motifs (μE1 to μE4 and κE1 to κE3) displayed a fairly loose sequence homology, implying that it would be difficult to identify such a segment by simply scanning a given sequence. However, it has previously been noted that Ephrussi motifs have certain

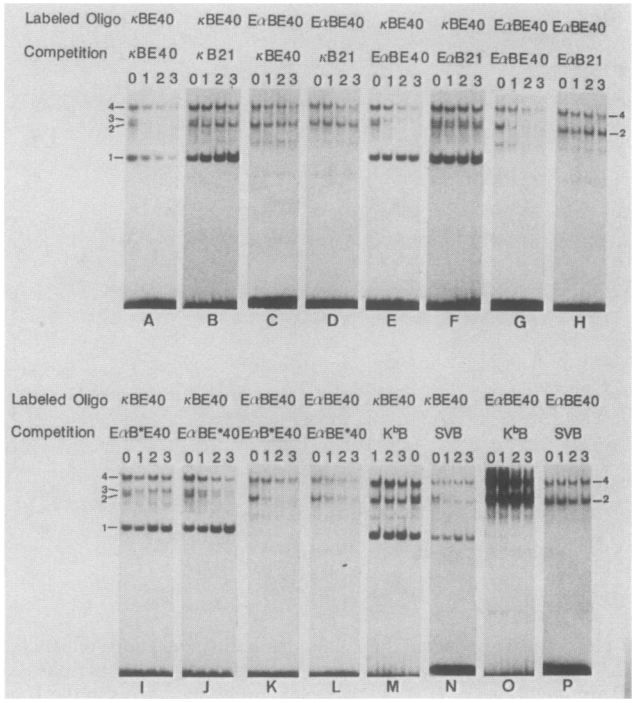

FIG. 9. Competition experiments with B-Ephrussi motif oligonucleotides. The experiments presented in each panel used fixed amounts of labeled oligonucleotide (Oligo), poly(dI-dC) and nuclear extract and increasing amounts of unlabeled double-stranded oligonucleotide as a competitor; i.e., 0, 50, 150, and 300 fmol in lanes 0 to 3, respectively. The sequences of all of the oligonucleotides are given in Table 1. κBE40, E_αBE40, E_αB*E40, and E_αBE*40 are described in the legend to Fig. 7. κB21 and E_αB21 are short versions of the former two and center on the B motif alone. K^bB and SVB are oligonucleotides with the B motif of *H2-K*^{b} or SV40 (kind gifts from M. Macchi and P. Chambon). Bands 1 to 4 are as defined in the legend to Fig. 7.

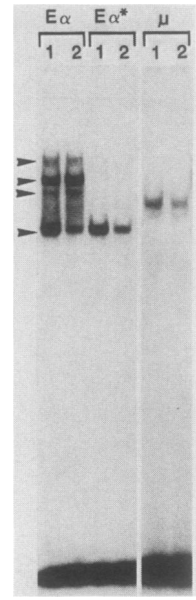

FIG. 10. Comparison of proteins binding to the E_α and immunoglobulin gene E3 motifs. Three oligonucleotides were tested in the gel retardation assay with 1 M NaCl nuclear extracts from M12 cells. E_αE3 carries the E3 sequence from E_α (Fig. 5 and Table 1), E^*_αE3 is a mutant thereof with three substitutions at conserved positions, and μE3 is an oligonucleotide centered around the corresponding motif in the IgH enhancer. Lanes 1 and 2 contained 100 and 300 ng, respectively, of poly(dI-dC). Arrowheads indicate the four bands discussed in the text.

invariant features. (i) All immunoglobulin gene copies have an A at position 2 and TG at positions 5 and 6 (CAGGTGGC) (10), and (ii) each has a minidyad axis of symmetry, beginning on the 3′ side with the TG at positions 5 and 6 (e.g., TACCCAGGTGGTGTT) (63).

Thus, we searched the 5′ end of the E_α gene in both orientations for Ephrussi motifs by using the following three sequential criteria: (i) at least a 10-of-15-base match to any of the seven immunoglobulin gene Ephrussi motifs; (ii) the presence of an A at position 2 and TG at positions 5 and 6; (iii) a minidyad beginning on the 3′ side, with the TG at positions 5 and 6. Considering these criteria, we identified the following three possible Ephrussi motifs on the −1,906 to −1,180 fragment: positions −1821 to −1807, homologous to both μE3 and κE3; positions −1758 to −1744, homologous to μE4; positions −1546 to −1532, also homologous to μE4. Figure 5C shows a comparison of the immunoglobulin gene and the putative E_α Ephrussi motifs. It may be worth noting that by using these three criteria we found only two other homologous segments along the sequenced 2,173 bp at the 5′ end of E_α.

We sought to determine whether the −1821 to −1807 sequence binds the same protein(s) as its immunoglobulin gene homolog. This particular E_α Ephrussi motif was chosen for study because it showed very good homology with two immunoglobulin gene elements that bind a single protein, μE3 and κE3. The following three oligonucleotides were evaluated in the gel retardation assay (Table 1): E_αE3; E^*_αE3, the same segment bearing substitutions at the conserved A and GT bases mentioned above; μE3, the corresponding sequence from the IgH enhancer (Table 1). The E_αE3 and μE3 oligonucleotides gave rise to completely different patterns (Fig. 10). Four bands were evident with E_αE3; they are not specific to B cells, since they could be detected in nuclear extracts from M12 (Fig. 10), as well as those from LMTK (data not shown); three of them must be due to factors that actually recognize the Ephrussi motif, since they were abolished by the point mutations in the E^*_αE3 oligonucleotide. This conclusion was strengthened by methylation interference mapping, which revealed contacts at these positions (data not shown). The fourth band is probably less relevant, since it was not affected by the substitutions in E^*_αE3.

The functional relevance of the E_αE3 motif and the factors we identified remains to be established, yet the demonstrated importance of the conserved bases for protein binding suggests that the E_α Ephrussi motif is a functional entity.

Pu motif. Pettersson and Schaffner (53) have recently drawn attention to a purine-rich sequence (CTGAAA GAGGAA) that appears to play an important role in the expression of SV40 and LPV (lymphotropic papovavirus) in B cells. This Pu motif is recognized by a lymphoid-cell-specific protein.

The E_α 5′-flanking region has a sequence at positions −1235 to −1224 that is a 10-of-12-bp match with the SV40-LPV motif (Fig. 5D). The two mismatches are reciprocal G→A and A→G substitutions. Although these changes seem rather mild, the first does occur at a major contact site for the Pu-binding protein.

Because of the similarity between the E_α and SV40-LPV sequences, we asked whether the E_α Pu-like sequence is

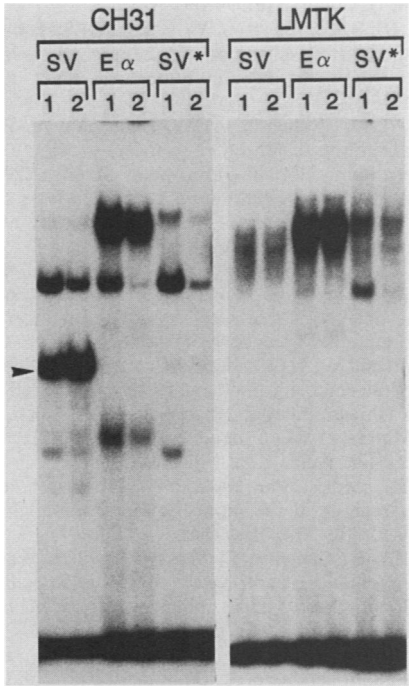

FIG. 11. Proteins binding to the E_α and SV40-LPV Pu box motifs. In this gel retardation experiment, nuclear extracts were from CH31 B-lymphoma cells or LMTK fibroblasts. The double-stranded oligonucleotides carry the SV40 Pu box (SV), its E_α homolog (E_α), or a mutation of the SV40 sequence (SV*) (Table 1). The arrowhead points to the lymphocyte-specific complex formed only with the SV40 Pu box. Lanes 1 and 2 contained 100 and 300 ng of poly(dI-dC), respectively.

recognized by the lymphocyte-specific Pu binding. Gel retardation assays with the following three double-stranded oligonucleotides were compared: PuSV, a 22-mer spanning the SV40 Pu box; PuE_α, the corresponding E_α 22-mer; and PuSV*, an oligonucleotide bearing mutations at critical positions of the SV40 sequence. Extracts from the B-lymphoma line CH31 contain a protein, absent in LMTK extracts, that binds to the SV40 Pu box (Fig. 11, arrow); methylation interference mapping confirmed that this protein corresponds to the one previously described (data not shown). As expected, binding was abolished by the PuSV* mutations; more unexpectedly, the E_α Pu-like sequence was also not recognized. The Pu-binding protein was detected in extracts from the B-lymphoma lines M12, CH27, A20, and WEHI 231, the pre-B-lymphoma line 70Z/3, and the macrophagelike line P388D1 (data not shown). In none of these cases did the protein bind detectably to the E_α oligonucleotide. These data imply that the Pu-like sequence is not important for controlling E_α transcription in B cells; at least the previously described lymphocyte-specific Pu-binding protein seems not to be involved.

Conclusion. The −1906 to −1180 region of the E_α gene is crucial for transcription in B cells of transgenic mice but is largely dispensable in other cell types. We sequenced this B-cell control region and searched for stretches homologous to motifs implicated in lymphocyte-specific expression of other genes. Glaringly absent was the immunoglobulin gene octamer ATTTGCAT, either as a perfect match or as a single-base mismatch. Provocatively present were a B-Ephrussi bloc, isolated Ephrussi motifs, and a Pu-like sequence. The E_α and immunoglobulin gene B-Ephrussi blocs bind some of the same proteins, but not those most convincingly implicated in immunoglobulin gene expression. The E_α Ephrussi motif at −1821 to −1807 binds a protein(s) that depends on conserved bases for effective attachment, but this protein does not seem to recognize the immunoglobulin gene μE3 counterpart. The E_α Pu-like sequence does not bind the lymphocyte-restricted protein that interacts with the SV40 and LPV homologs. On the basis of these findings, we are tempted to conclude that MHC class II and immunoglobulin genes rely on quite different factors for mediating B-cell-specific expression.

Consistent with this conclusion are the observations that two elements that control E_α expression in cultured B-cell lines reside in the −1906 to −1180 region and neither has a recognized counterpart in the immunoglobulin gene. The X′-Y′ pair, whose homology to the promoter-proximal X-Y pair is too striking to be fortuitous, is known to be one component of a complex B-cell-specific enhancer spaced along 2 kilobases of the 5′ end of E_α (37a; Koch et al., submitted). Another, even more crucial element of this enhancer is the W motif TGTTGCATC, located near the *Sma*I site. This motif was identified by scanning the −1906 and −1180 fragment for subfragments that bind proteins present in B-cell but not non-B-cell extracts (A. Dorn, C. Benoist, and D. Mathis, submitted for publication). We do not understand the relationship between the B-cell control region detected in transgenic mice and the B-cell-specific enhancer delineated in transfection experiments. Thus, we do not know what relevance X′-Y′ and W have for expression of E_α in B cells in mice.

Clearly, we must turn back to transgenic mice to assay the functional relevance of the various motifs identified in the B-cell control region. The value of this study is that it suggests priorities. We are encouraged to delete or mutate the X′-Y′ pair or the E_α Ephrussi motif at −1821 to −1807 but are discouraged from deleting the B-Ephrussi bloc or the Pu-like sequence. Such experiments are in progress. Meanwhile, on the basis of the DNA-protein-binding studies presented herein, we predict that the E_α story will not be just déjà vu.

ACKNOWLEDGMENTS

We are grateful to P. Gerber and G. Lang for skillful technical assistance, to C. Marfing for the injections, to P. Sourinphoumy and F. Mackay for maintaining the mice, and to C. Waltzinger for operating the cytofluorimeter. We thank M. Macchi for useful discussion and R. Accolla, G. Haughton, and C. Paige for cell lines.

A.D. received fellowships from the Deutsche Forschungsgemeinschaft and the Deutscher Akademischer Austauschdienst, H.J.F. received fellowships from the Carl Duisberg Foundation and the Boehringer-Ingelheim Foundation, and W.K. received fellowships from the Centre National de la Recherche Scientifique and the Deutscher Akademischer Austauschdienst. This work was supported by institutional grants from the Centre National de la Recherche Scientifique and the Institut National de la Santé et de la Recherche Médicale and by a grant from the Association pour la Recherche sur le Cancer to D.M. and C.B.

LITERATURE CITED

1. **Atchison, M. L., and R. P. Perry.** 1987. The role of the κ enhancer and its binding factor NF-κB in developmental regulation of κ gene transcription. Cell **48**:121–128.
2. **Augereau, P., and P. Chambon.** 1986. The mouse immunoglobulin heavy-chain enhancer: effect on transcription in vitro and binding of proteins present in HeLa and lymphoid B cell extracts. EMBO J. **5**:1791–1797.
3. **Baeuerle, P. A., and D. Baltimore.** 1988. Activation of DNA-binding activity in an apparently cytoplasmic precursor of the

NF-κB transcription factor. Cell **53**:211–217.
4. **Baldwin, A. S., and P. A. Sharp.** 1987. Binding of a nuclear factor to a regulatory sequence in the promoter of the mouse *H-2K*^{b} class I major histocompatibility gene. Mol. Cell. Biol. **7**: 305–313.
5. **Baldwin, A. S., and P. A. Sharp.** 1988. Two transcription factors, NF-κB and H2TF1, interact with a single regulatory sequence in the class I major histocompatibility complex promoter. Proc. Natl. Acad. Sci. USA **85**:723–727.
6. **Ballard, D. W., and A. Bothwell.** 1986. Mutational analysis of the immunoglobulin heavy chain promoter region. Proc. Natl. Acad. Sci. USA **83**:9626–9630.
7. **Banerji, J., L. Olson, and W. Schaffner.** 1983. A lymphocyte-specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes. Cell **33**:729–740.
8. **Bohmann, D., W. Keller, T. Dale, H. R. Schöler, G. Tebb, and I. W. Mattaj.** 1987. A transcription factor which binds to the enhancers of SV40, immunoglobulin heavy chain and U2 snRNA genes. Nature (London) **325**:268–272.
9. **Boss, J. M., and J. L. Strominger.** 1986. Regulation of a transfected human class II major histocompatibility complex gene in human fibroblasts. Proc. Natl. Acad. Sci. USA **83**:9139–9143.
10. **Church, G. M., A. Ephrussi, W. Gilbert, and S. Tonegawa.** 1985. Cell-type-specific contacts to immunoglobulin enhancers in nuclei. Nature (London) **313**:798–801.
11. **Davidson, I., C. Fromental, P. Augereau, A. Wildeman, M. Zenke, and P. Chambon.** 1986. Cell-type specific protein binding to the enhancer of simian virus 40 in nuclear extracts. Nature (London) **323**:544–548.
12. **Dente, L., U. Rüther, M. Tripodi, E. F. Wagner, and R. Cortese.** 1988. Expression of human α1-acid glycoprotein genes in cultured cells and in transgenic mice. Genes Dev. **2**:259–266.
13. **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.
14. **Dorn, A., J. Bollekens, A. Staub, C. Benoist, and D. Mathis.** 1987. A multiplicity of CCAAT box-binding proteins. Cell **50**: 863–872.
15. **Dorn, A., B. Durand, C. Marfing, M. Le Meur, C. Benoist, and D. Mathis.** 1987. The conserved MHC class II boxes—X and Y—are transcriptional control elements and specifically bind nuclear proteins. Proc. Natl. Acad. Sci. USA **84**:6249–6253.
16. **Dreyfus, M., N. Doyen, and F. Rougeon.** 1987. The conserved decanucleotide from the immunoglobulin heavy chain promoter induces a very high transcriptional activity in B-cells when introduced into an heterologous promoter. EMBO J. **6**:1685–1690.
17. **Dynan, W. S.** 1987. What mechanisms underlie tissue-specific gene transcription? Trends Genet. **3**:121–122.
18. **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.
19. **Falkner, F. G., and H. G. Zachau.** 1984. Correct transcription of an immunoglobulin κ gene requires an upstream fragment containing conserved sequence elements. Nature (London) **310**:71–74.
20. **Flavell, R. A., H. Allen, B. Huber, C. Wake, and G. Widera.** 1985. Organization and expression of the MHC of the C57 black/10 mouse. Immunol. Rev. **84**:29–50.
21. **Foster, J., J. Stafford, and C. Queen.** 1985. An immunoglobulin promoter displays cell-type specificity independently of the enhancer. Nature (London) **315**:423–425.
22. **Garabedian, M. J., M. C. Hung, and P. C. Wensink.** 1985. Independent control elements that determine yolk protein gene expression in alternative drosophila tissues. Proc. Natl. Acad. Sci. USA **82**:1396–1400.
23. **Gerster, T., P. Matthias, M. Thali, J. Jiricny, and W. Schaffner.** 1987. Cell type-specificity elements of the immunoglobulin heavy chain gene enhancer. EMBO J. **6**:1323–1330.
24. **Gillies, S. D., V. Folsom, and S. Tonegawa.** 1984. Cell type-specific enhancer element associated with a mouse MHC gene, E_β. Nature (London) **310**:594–597.
25. **Gillies, S. D., S. L. Morrison, V. T. Oi, and S. Tonegawa.** 1983. A tissue-specific transcription enhancer element is located in the major intron of a rearranged immunoglobulin heavy chain gene. Cell **33**:717–728.
26. **Gopal, V. T., T. Shimada, A. W. Baur, and A. W. Nienhuis.** 1985. Contribution of promoter to tissue-specific expression of the mouse immunoglobulin kappa gene. Science **229**:1102–1104.
27. **Grosschedl, R., and D. Baltimore.** 1985. Cell-type specificity of immunoglobulin gene expression is regulated by at least three DNA sequence elements. Cell **41**:885–897.
28. **Hammer, R. E., R. Krumlauf, S. A. Camper, R. L. Brinster, and S. M. Tilghman.** 1987. Diversity of alpha-fetoprotein gene expression in mice is generated by a combination of separate enhancer elements. Science **235**:53–58.
29. **Hiromi, Y., and W. J. Gehring.** 1987. Regulation and function of the drosophila segmentation gene fushi tarazu. Cell **50**:963–974.
30. **Hooft van Huijsduijnen, R., J. Bollekens, A. Dorn, C. Benoist, and D. Mathis.** 1987. Properties of a CCAAT box-binding protein. Nucleic Acids Res. **15**:7265–7281.
31. **Hromas, R., and B. Van Ness.** 1986. Nuclear factors bind to regulatory regions of the mouse kappa immunoglobulin gene. Nucleic Acids Res. **14**:4837–4848.
32. **Imler, J. L., C. Lemaire, C. Wasylyk, and B. Wasylyk.** 1987. Negative regulation contributes to tissue specificity of the immunoglobulin heavy-chain enhancer. Mol. Cell. Biol. **7**:2558–2567.
33. **Ishihara, T., A. Kudo, and T. Watanabe.** 1984. Induction of immunoglobulin gene expression in mouse fibroblasts by cycloheximide treatment. J. Exp. Med. **160**:1937–1942.
34. **Israël, A., A. Kimura, M. Kieran, O. Yano, J. Kanellopoulos, O. Le Bail, and P. Kourilsky.** 1987. A common positive trans-acting factor binds to enhancer sequences in the promoters of mouse H-2 and β_2-microglobulin genes. Proc. Natl. Acad. Sci. USA **84**: 2653–2657.
35. **Kelley, D. E., B. A. Pollok, M. L. Atchison, and R. P. Perry.** 1988. The coupling between enhancer activity and hypomethylation of κ immunoglobulin genes is developmentally regulated. Mol. Cell. Biol. **8**:930–937.
36. **Kimura, A., A. Israël, O. Le Bail, and P. Kourilsky.** 1986. Detailed analysis of the mouse H-2K^b promoter: enhancer-like sequences and their role in the regulation of class I gene expression. Cell **44**:261–272.
37. **Kitamura, D., H. Maeda, K. Araki, A. Kudo, and T. Watanabe.** 1987. Regulation of immunoglobulin gene transcription by labile repressor factor(s). Eur. J. Immunol. **17**:1249–1256.

37a. **Koch, W., S. Candeias, J. Guardiola, R. Accolla, C. Benoist, and D. Mathis.** 1988. An enhancer factor defect in a mutant Burkitt lymphoma cell line. J. Exp. Med. **167**:1781–1790.

38. **Landolfi, N. F., J. D. Capra, and P. W. Tucker.** 1986. Interaction of cell-type-specific nuclear proteins with immunoglobulin V_H promoter region sequences. Nature (London) **323**:548–551.
39. **Landolfi, N. F., J. D. Capra, and P. W. Tucker.** 1987. Protein-nucleotide contacts in the immunoglobulin heavy-chain promoter region. Proc. Natl. Acad. Sci. USA **84**:3851–3855.
40. **Larhammar, D., U. Hammerling, M. Denaro, T. Lund, R. A. Flavell, L. Rask, and P. A. Peterson.** 1983. Structure of the murine immune response I-A_β locus: sequence of the I-A_β gene and an adjacent β-chain second domain exon. Cell **34**:179–188.
41. **Le Meur, M., P. Gerlinger, C. Benoist, and D. Mathis.** 1985. Correcting an immune-response deficiency by creating E_α gene transgenic mice. Nature (London) **316**:38–42.
42. **Lenardo, M., J. W. Pierce, and D. Baltimore.** 1987. Protein-binding sites in Ig gene enhancers determine transcriptional activity and inducibility. Science **236**:1573–1577.
43. **Maeda, H., K. Araki, D. Kitamura, J. Wang, and J. Watanabe.** 1987. Nuclear factors binding to the human immunoglobulin heavy-chain gene enhancer. Nucleic Acids Res. **15**:2851–2869.
44. **Maniatis, T., S. Goodbourn, and J. A. Fischer.** 1987. Regulation of inducible and tissue-specific gene expression. Science **236**: 1237–1245.
45. **Mason, J. O., G. T. Williams, and M. S. Neuberger.** 1985. Transcription cell type specificity is conferred by an immuno-

globulin V_H gene promoter that includes a functional consensus sequence. Cell **41**:479–487.

46. **Mathis, D. J., C. O. Benoist, V. E. Williams II, M. R. Kanter, and H. O. McDevitt.** 1983. The murine Eα immune response gene. Cell **32**:745–754.
47. **Mercola, M., X. F. Wang, J. Olsen, and K. Calame.** 1983. Transcriptional enhancer elements in the mouse immunoglobulin heavy chain locus. Science **221**:663–665.
48. **Miwa, K., C. Doyle, and J. L. Strominger.** 1987. Sequence-specific interactions of nuclear factors with conserved sequences of human class II major histocompatibility complex genes. Proc. Natl. Acad. Sci. USA **84**:4939–4943.
49. **Mocikat, R., F. G. Falkner, R. Mertz, and H. G. Zachau.** 1986. Upstream regulatory sequences of immunoglobulin genes are recognized by nuclear proteins which also bind to other gene regions. Nucleic Acids Res. **14**:8829–8844.
50. **Nabel, G., and D. Baltimore.** 1987. An inducible transcription factor activates expression of human immunodeficiency virus in T cells. Nature (London) **326**:711–713.
51. **Neuberger, M. S.** 1983. Expression and regulation of immunoglobulin heavy chain gene transfected into lymphoid cells. EMBO J. **2**:1373–1378.
52. **Parslow, T. G., D. L. Blair, W. J. Murphy, and D. K. Granner.** 1984. Structure of the 5′ ends of immunoglobulin genes: a novel conserved sequence. Proc. Natl. Acad. Sci. USA **81**:2650–2654.
53. **Pettersson, M., and W. Schaffner.** 1987. A purine-rich DNA sequence motif present in SV40 and lymphotropic papovavirus binds a lymphoid-specific factor and contributes to enhancer activity in lymphoid cells. Genes Dev. **1**:962–972.
54. **Picard, D., and W. Schaffner.** 1984. A lymphocyte-specific enhancer in the mouse immunoglobulin κ gene. Nature (London) **307**:80–82.
55. **Picard, D., and W. Schaffner.** 1985. Cell type preference of immunoglobulin κ and λ gene promoters. EMBO J. **4**:2831–2838.
56. **Pierce, J. W., M. Lenardo, and D. Baltimore.** 1988. Oligonucleotide that binds nuclear factor NF-κB acts as a lymphoid-specific and inducible enhancer element. Proc. Natl. Acad. Sci. USA **85**:1482–1486.
57. **Pinkert, C. A., D. M. Ornitz, R. L. Brinster, and R. D. Palmiter.** 1987. An albumin enhancer located 10 kb upstream functions along with its promoter to direct efficient, liver-specific expression in transgenic mice. Genes Dev. **1**:268–276.
58. **Pinkert, C. A., G. Widera, C. Cowing, E. Heber-Katz, R. D. Palmiter, R. A. Flavell, and R. L. Brinster.** 1985. Tissue-specific, inducible and functional expression of the E_α^d MHC class II gene in transgenic mice. EMBO J. **4**:2225–2230.
59. **Queen, C., and D. Baltimore.** 1983. Immunoglobulin gene transcription is activated by downstream sequence elements. Cell **33**:741–748.
60. **Rosales, R., M. Vigneron, M. Macchi, I. Davidson, J. H. Xiao, and P. Chambon.** 1987. *In vitro* binding of cell-specific and ubiquitous nuclear proteins to the octamer motif of the SV40 enhancer and related motifs present in other promoters and enhancers. EMBO J. **6**:3015–3025.
61. **Sanger, F., A. R. Coulson, B. G. Barrell, A. J. H. Smith, and B. A. Roe.** 1980. Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. J. Mol. Biol. **143**:161–178.
62. **Schlokat, U., D. Bohmann, H. Schöler, and P. Gruss.** 1986. Nuclear factors binding specific sequences within the immunoglobulin enhancer interact differentially with other enhancer elements. EMBO J. **5**:3251–3258.
63. **Sen, R., and D. Baltimore.** 1986. Multiple nuclear factors interact with the immunoglobulin enhancer sequences. Cell **46**: 705–716.
64. **Sen, R., and D. Baltimore.** 1986. Inducibility of κ immunoglobulin enhancer-binding protein NF-κB by a posttranslational mechanism. Cell **47**:921–928.
65. **Shelley, C. S., and F. E. Baralle.** 1987. Dual tissue-specific expression of apo-AII is directed by an upstream enhancer. Nucleic Acids Res. **15**:3801–3821.
66. **Sherman, P. A., P. V. Basta, and J. P. Y. Ting.** 1987. Upstream DNA sequences required for tissue-specific expression of the HLA-DRα gene. Proc. Natl. Acad. Sci. USA **84**:4254–4258.
67. **Singh, H., R. Sen, D. Baltimore, and P. A. Sharp.** 1986. A nuclear factor that binds to a conserved sequence motif in transcriptional control elements of immunoglobulin genes. Nature (London) **319**:154–158.
68. **Sive, H. L., and R. G. Roeder.** 1986. Interaction of a common factor with conserved promoter and enhancer sequences in histone H2B, immunoglobulin, and U2 small nuclear RNA (snRNA) genes. Proc. Natl. Acad. Sci. USA **83**:6382–6386.
69. **Staudt, L. M., H. Singh, R. Sen, T. Wirth, P. A. Sharp, and D. Baltimore.** 1986. A lymphoid-specific protein binding to the octamer motif of immunoglobulin genes. Nature (London) **323**: 640–643.
70. **Sullivan, K. E., A. F. Calman, M. Nakanishi, S. Y. Tsang, Y. Wang, and B. M. Peterlin.** 1987. A model for the transcriptional regulation of MHC class II genes. Immunol. Today **8**:289–293.
71. **Sullivan, K. E., and B. M. Peterlin.** 1987. Transcriptional enhancers in the HLA-DQ subregion. Mol. Cell. Biol. **7**:3315–3319.
72. **Theisen, M., A. Stief, and A. E. Sippel.** 1986. The lysozyme enhancer: cell-specific activation of the chicken lysozyme gene by a far-upstream DNA element. EMBO J. **5**:719–724.

72a. **Van Ewijk, W., Y. Ron, J. Monaco, J. Kappter, P. Marrack, M. Le Meur, P. Gerlinger, B. Durand, C. Benoist, and D. Mathis.** 1988. Compartmentalization of MHC class II gene expression in transgenic mice. Cell **53**:357–370.

73. **Wall, R., M. Briskin, C. Carter, H. Govan, A. Taylor, and P. Kincade.** 1986. A labile inhibitor blocks immunoglobulin κ-light-chain-gene transcription in a pre-B leukemic cell line. Proc. Natl. Acad. Sci. USA **83**:295–298.
74. **Wang, Y., A. S. Larsen, and M. Peterlin.** 1987. A tissue-specific transcriptional enhancer is found in the body of the HLA-DRα gene. J. Exp. Med. **166**:625–636.
75. **Weinberger, J., D. Baltimore, and P. A. Sharp.** 1986. Distinct factors bind to apparently homologous sequences in the immunoglobulin heavy-chain enhancer. Nature (London) **322**:846–848.
76. **Weinberger, J., P. S. Jat, and P. A. Sharp.** 1988. Localization of a repressive sequence contributing to B-cell specificity in the immunoglobulin heavy-chain enhancer. Mol. Cell. Biol. **8**:988–992.
77. **Widera, G., L. C. Burkly, C. A. Pinkert, E. C. Böttger, C. Cowing, R. D. Palmiter, R. L. Brinster, and R. A. Flavell.** 1987. Transgenic mice selectively lacking MHC class II (I-E) antigen expression on B cells: an in vivo approach to investigate Ia gene function. Cell **51**:175–187.
78. **Wirth, T., L. Staudt, and D. Baltimore.** 1987. An octamer oligonucleotide upstream of a TATA motif is sufficient for lymphoid-specific promoter activity. Nature (London) **329**:174–178.
79. **Wysocki, L. J., and V. L. Sato.** 1978. "Panning" for lymphocytes: a method for cell selection. Proc. Natl. Acad. Sci. USA **75**:2844–2848.
80. **Yamamura, K., H. Kikutani, V. Folson, L. K. Clayton, M. Kimoto, S. Akira, S. Kashiwamura, S. Tonegawa, and T. Kishimoto.** 1985. Functional expression of a microinjected E_α^d gene in C57BL/6 transgenic mice. Nature (London) **316**:67–69.
81. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13 mp18 and pUC19 vectors. Gene **33**:103–119.