0022-1767/91/1469-3197$02.00/0
THE JOURNAL OF IMMUNOLOGY
Vol. 146, 3197-3204, No. 9, May 1, 1991
Printed in U.S.A.

NF-X, A TRANSCRIPTION FACTOR IMPLICATED IN MHC CLASS II GENE REGULATION[1]

VALÉRIE KOUSKOFF,* ROBERTO M. MANTOVANI,* SERGE M. CANDÉIAS,* ARNULF DORN,[2]* ADRIEN STAUB,* BARBARA LISOWSKA-GROSPIERRE,† CLAUDE GRISCELLI,† CHRISTOPHE O. BENOIST,[3]* AND DIANE J. MATHIS[3]*

*From the *Laboratoire de Génétique Moléculaire des Eucaryotes du C.N.R.S. Unité 184 de Biologie Moléculaire et de Génie Génétique de l'I.N.S.E.R.M. Institut de Chimie Biologique, Faculté de Médecine, 11, rue Humann, 67085 Strasbourg Cédex and †Hôpital Necker-Enfants Malades, Département de Pédiatrie, Unité d'Immunologie et d'Hématologie, 149, rue de Sèvres, 75743 Paris Cédex 15, France*

The X box has been shown in several assay systems to be a critical element of MHC class II gene promoters. Several X box-binding activities have been discovered in nuclear extracts from a variety of cell lines. The critical question is: which of these are responsible for mediating X box function? This report provides a further characterization of NF-X, a highly specific X box-binding activity we described previously. The cell-type distribution, structural features, and binding site characteristics of NF-X are analyzed in detail, to facilitate comparison with other reported activities. Most importantly, the functional relevance of NF-X is assessed by scanning mutagenesis, and the results indicate that this complex is indeed involved in regulating MHC class II gene expression. With these data in mind, the relationship between NF-X and RF-X, an X box-binding activity reported to be absent in patients with severe combined immunodeficiency, is discussed.

It is by now a truism to state that the primary level of gene regulation in higher eukaryotes is the initiation of mRNA synthesis (1, 2). The mechanisms controlling transcriptional initiation have been the subject of a massive research effort over the past decade. First came the definition of promoters and enhancers, critical DNA sequence elements usually located upstream of the mRNA start site; then, specific DNA-binding proteins were shown to interact selectively with short motifs within these elements. Although the precise molecular interactions remain largely undefined, it is thought that the combinatorial association of several such proteins is ultimately responsible for the efficiency and cell-type specificity of mRNA synthesis.

Class II genes of the MHC are no exception in their dependence on regulatory events at the level of transcriptional initiation. Much of the control of promoter activity and specificity can be attributed to two DNA sequence motifs, the X and Y boxes. These are 14 and 10 bp conserved segments, separated by a 19 to 20 bp spacer of fixed length but variable sequence. They are found in the −150 to −50 region of all MHC class II genes so far examined. Although the roles of X and Y are well established, the identity of the protein factors that mediate their effects is more controversial (3). A constellation of X box-binding proteins has been identified in various laboratories by gel retardation experiments or by binding-site screening of cDNA expression libraries (4–17). Several reports have described binding activities that seem to recognize the X box and bases immediately upstream (4–9, 12, 15–17). Although some of these probably represent the same protein(s), others appear different (3). This complexity is further compounded by the finding that the X region is actually composed of two overlapping motifs: the X box proper and the X2 box, which overlaps the 3′ end of the classical X sequence homology. X2 is related to the cAMP or TPA response elements (CRE, TRE), and the proteins that bind to it are members of the fos/jun family (10, 12, 14–18). The functional relevance of these various binding activities is the key question—it is clear that the sensitivity of techniques currently used for studying DNA-binding proteins far exceeds their discriminating power.

Among the very first X box-binding proteins to be described was NF-X, an activity initially discovered in this laboratory to bind to the X box of the Eα and DRα promoters (4, 8). In this report, to facilitate comparison with activities described by others, we have characterized more fully the cell-type distribution, structural properties, and binding characteristics of NF-X. Most important, we have addressed its functional relevance by creating panels of X box mutants and comparing the effects of alterations on NF-X binding and on transcriptional activity of an MHC class II gene promoter.

MATERIALS AND METHODS

Cell lines. Most of the cell lines were of murine origin: M12.4.1., TA3, CH27, WEHI-231 (B lymphoma lines that constitutively express MHC class II molecules); P388D1 (a macrophage-like line that can be induced to express MHC class II molecules by IFN-γ treatment); PD31, 70Z/3 (class II$^-$ preB lymphoma lines); LMTK (a fibroblast cell

Received for publication November 20, 1990.
Accepted for publication January 29, 1991.
The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
[1]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 a grant to D.M. and C.B. from the Association pour la Recherche sur le Cancer. V.K. and S.C. received bursaries from the Ministère de la Recherche et de la Technologie, R.M. a fellowship from the European Molecular Biology Organisation and A.D. a fellowship from the Deutscher Akademischer Austauschdienst.
[2]Present address F. Hoffmann la Roche Ltd, Department of Pharma Research/Infectious Diseases, CH-4002 Basel, Switzerland.
[3]Address correspondence and reprint requests to Drs. D. Mathis and C. Benoist, DOI/LGME, Faculté de Médecine, 11, rue Humann, 67085 Strasbourg Cédex, Paris.

TABLE I

Oligonucleotides used in this study

Eα X box 23 mer (X)
5′...ACCCTTTCCTAGCAACAGATGTG...3′
3′...TGGGAAAGGATCGTTGTCTACAC...5′
Eα X box 38 mer (X)
5′...AAGGAACCCTTTCCTAGCAACAGATGTGTCAGTCTGAA...3′
3′...TTCCTTGGGAAAGGATCGTTGTCTACACAGTCAGACTT...5′
Eα X box random sequence 38 mer (XC)
5′...AAGGAACCCTTTACGTTCTGGCTACATGTCAGTCTGGA...3′
3′...TTCCTTGGGAAATGCAAGACCGATGTACAGTCAGACTT...5′
DRα x box 61 mer
5′...ACCCTTTGCAAGAACCCTTCCCCTAGCAACAGATGCGTCATCTCAAAATATTTTTCTGATT...3′
3′...TGGGAAACGTTCTTGGGAAGGGGATCGTTGTCTACGCAGTAGAGTTTTATAAAAAGACTAA...5′
DRα X box random sequence 61 mer
5′...ACCCTTTGCAAGAACCCTTCCACGTTCTGGCTACACGTCATCTCAAAATATTTTTCTGATT...3′
3′...TGGGAAACGTTCTTGGGAAGGTGCAAGACCGATGTGCAGTAGAGTTTTATAAAAAGACTAA...5′
Pγ EFC box 23 mer (PY)
5′...AATTAGTTGCTAGGCAACTGGCC...3′
3′...TTAATCAACGATCCGGTTGACCGG...5′
Eβ X box 22 mer
5′...TCTCTAACTAGCAACTGATGAT...3′
3′...AGAGATTGATCGTTGACTACTA...5′
Aα X box 22 mer
5′...TCTGCAGCTGGCAACTGTGACG...3′
3′...AGACGTCGACCGTTGACACTGC...5′
Aβ box 22 mer
5′...AGTTTACCCAGAGACAGACGAC...3′
3′...TCAAATGGGTCTCTGTCTGCTG...5′

line). The undifferentiated teratocarcinoma T113 was a kind gift from Dr. D. Duboulle, EMBL, Heidelberg, Germany. Cells of human origin were EBV-transformed B lymphoblastoid lines derived from a normal individual (Debré) or from SCID[4] patients (Ramia, Nacera) (9).

Oligos. Oligos were synthesized on a Pharmacia Assembler (Pharmacia, Uppsala, Sweden), and purified by reverse-phase HPLC. The single-stranded oligos were 5′ end-labeled and annealed with their complement, and the resulting ^{32}P-labeled double-stranded fragments were purified by diffusion after electrophoresis on a neutral polyacrylamide gel. The sequences of the oligo pairs used in this study are listed in Table I.

Gel retardation assays. The 1 M NaCl extracts of nuclei were prepared according to the method of Dignam et al. (19), as detailed elsewhere (20).

Gel retardation assays were performed as previously described (20). Briefly, nuclear extract (1 to 5 μl), ^{32}P-labeled double-stranded oligo (5000 cpm) and poly(d(I-C)) (200 ng to 1 μg) were mixed in 25 μl final volume and incubated on ice for 30 min. Free and protein-bound DNA were then separated by electrophoresis on a 5% polyacrylamide gel, and the resulting bands revealed by autoradiography.

For the protease experiments, various amounts (see figure legend) of proteinase K or trypsin were added and incubated for 10 min either before or after the binding reaction. In these experiments, the DNA:protein complexes were separated on 7% polyacrylamide gels.

Methylation interference mapping was conducted as detailed elsewhere (20). In short, ^{32}P-labeled double-stranded oligos were partially methylated and then incubated with nuclear extract in the presence of poly[d(I-C)]. The free and protein-bound DNA molecules were separated by electrophoresis, extracted from the gel, cleaved with piperidine, and electrophoresed on a high resolution denaturing gel.

Mutants. A panel of mutant plasmids, each carrying a three-base mutation around the Eα X box, was constructed, beginning with the plasmid PX3. This plasmid contains the −215 to +12 stretch of the Eα promoter flanked by *Xho*I and *Bam*HI sites (21). Two rounds of PCR were used to construct the mutant fragments, using internal primers carrying the mutations and external primers at the *Xho*I or *Bam*HI sites and following a strategy essentially identical to the PCR-based mutagenesis procedure described (22). Amplifications were performed according to the Perkin-Elmer-Cetus (Emeryville, CA) protocol, adjusting the annealing temperature according to oligo length. The resulting DNA fragments were digested with *Xho*I and *Bam*HI, and introduced into PX3 to replace the wild-type Eα promoter segment.

From this set of plasmids we derived another panel carrying the same mutations on the pE "backbone," in which the −215 to +12 Eα fragment is placed immediately upstream of a reporter gene derived from rabbit β-globin. These plasmids were derived by simply removing the *Bam*HI-*Hin*dIII segment from the pX mutants (21).

In vitro transcription assays. In vitro transcription assays were performed as detailed elsewhere (22a) using transcriptionally active nuclear extracts from the MHC class II$^+$ B lymphoma CH27. RNA accurately initiated from the Eα promoter were measured by quantitative S1 mapping (21). The reported data are averages from at least three independent experiments with at least two different plasmid preparations.

[4] Abbreviations used in this paper: SCID, Severe Combined Immunodeficiency; oligo, oligonucleotide.

RESULTS

The focus of this report, NF-X, is a site-specific DNA-binding protein previously detected in nuclear extracts prepared from various cell lines (4, 8). NF-X bound specifically to X box oligonucleotides from the Eα and DRα promoters to form a complex that migrated as a characteristic doublet or triplet of bands not very far into the polyacrylamide gels used routinely in gel retardation assays (as compared with complexes formed by NF-Y, NF-κB, or octamer-binding proteins, for example). NF-X's specificity for the X box was established by methylation interference analysis and by gel retardation experiments using a control oligo carrying a highly mutated X sequence (see Table I for sequence). All other bands, which are also seen with the control oligo, thus represent nonspecific DNA binding proteins or factors recognizing flanking sequences, and are not considered further.

Cell-type distribution. Nuclear extracts prepared from a variety of cell-types were tested for NF-X activity by the gel retardation assay. NF-X was found in all extracts examined, as is partially illustrated in Figure 1, *A* and *B*. The NF-X bands, identifiable by their absence in the control oligo lanes, were clearly observed in extracts of nuclei from MHC class II$^+$ B lymphomas (WEHI-231, TA3) as well as from class II$^-$ preB cells and fibroblasts (70Z/3, LMTK). The induction of class II gene transcription by IFN-γ in macrophages was not accompanied by obvious changes in the abundance of NF-X (P388D1 + and − lanes). The ubiquitous nature of NF-X was also indicated by its occurrence in a completely undifferentiated teratocarcinoma (Fig. 1*B*). All extracts gave rise to at least two bands, referred to as NF-X_A and NF-X_B[5] (Fig. 1) (8).

[5] The two bands referred to as NF-X_A and NF-X_B were called NF-X1 and NF-X2 in a previous publication (8). This change in nomenclature was introduced to avoid confusion with the X2 motif and the factors that recognize it.

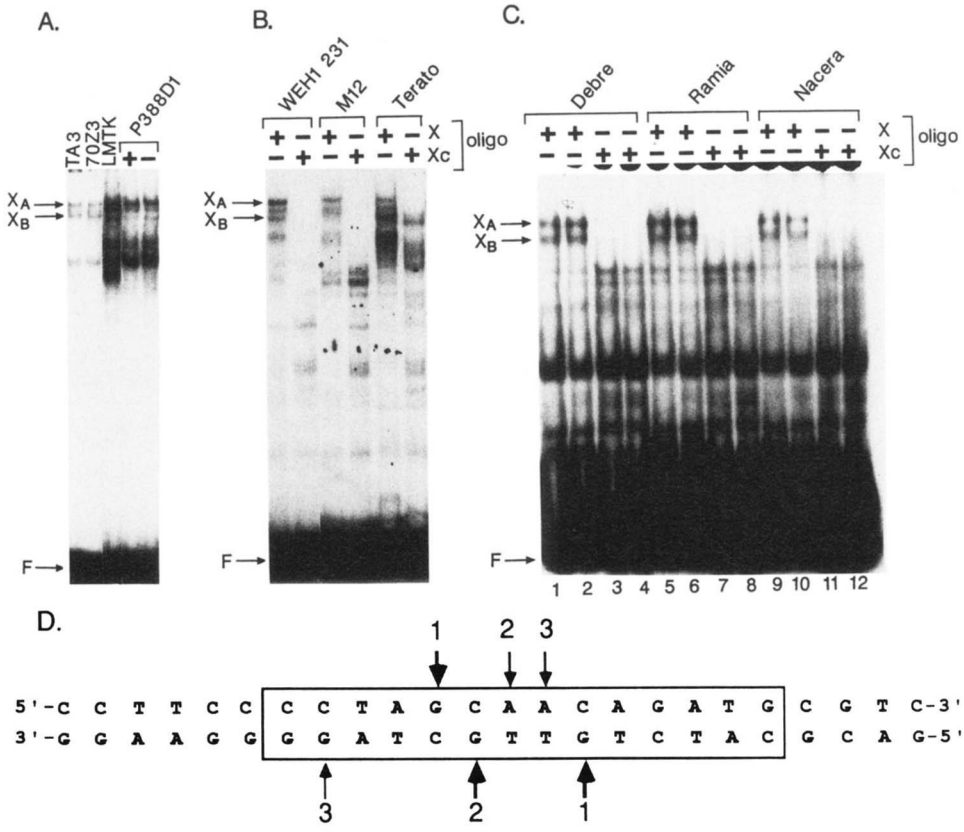

Figure 1. Cell-type distribution of NF-X. Nuclear extracts from various cell lines were incubated with ^{32}P-labeled oligos containing the X box from the Eα promoter region. The protein-DNA complexes were resolved by electrophoresis on a 5% polyacrylamide gel. *A*, Nuclear extracts originated from the class II$^+$ B lymphoma line TA3, the preB cell line 70Z/3, or the fibroblast line LMTK. For the macrophage-like cell line P388D1, extracts were prepared from cells grown with (+) or without (−) IFN-γ (36 h at 100 U/ml). The Eα 38 mer oligo was used (Table I). *B*, Extracts were prepared from class II$^+$ B lymphoma lines (WEHI 231, M12), or from class II$^-$ undifferentiated teratocarcinoma cells, and were incubated with the Eα X box oligo (X) or with the control oligo in which the X motif had been mutated (Xc; Table I). *C*, Nuclear extracts were prepared from lymphoblastoid cell lines derived from two SCID patients (Ramia, Nacera) or from a normal individual (Debré). For each extract/oligo combination, the incubation mixtures contained 1 μg of poly (d(I-C)) and 0 (*lanes 1, 3, 5, 7, 9,* and *11*) or 0.5 μg (*lanes 2, 4, 6, 8, 10,* and *12*) of *E. coli* DNA as non-specific competitor. The labeled oligo carried the X box and flanking sequences from the DRα gene (X), or a corresponding control (XC; Table I). *D*, Contact sites of NF-X on the DRα X box, as deduced from methylation interference experiments. The thickness of the *arrows* correlates with the degree of interference.

A third band was often, but not always, present between the other two; its presence and relative position seemed to vary somewhat with different extracts, even those prepared from the same cell line, and thus we suspect it may reflect partial proteolysis.

Another X box-binding activity, RF-X, appears to be absent from extracts prepared from B lymphoblastoid lines derived from SCID patients (9, 23). MHC class II genes are not expressed in these patients, a defect shown by genetic analyses to be due to the absence or dysfunction of a specific regulatory factor (24). Reith et al. (9, 23, 25) suggested that the absence of RF-X is responsible for the defective expression of class II genes in these patients. It was thus of interest to determine whether NF-X is similarly affected. Nuclear extracts were prepared from class II$^-$ lymphoblastoid lines derived from two SCID patients (Ramia and Nacera; incidentally two of the patients also studied by Reith et al. (9)). As shown in Figure 1*C*, there was no difference in appearance or abundance of NF-X between these extracts and a control extract made from a normal lymphoblastoid line (Debré). That these bands do correspond to human NF-X is substantiated by the use of control oligonucleotides (Fig. 1*C*), and by methylation interference footprints (Fig. 1*D*) very similar to those of NF-X on the Eα X box (see Fig. 3). This result was obtained with numerous extract preparations from these lines. In fact, two different sets of EBV-transformed isolates from the same patients were investigated, one of which corresponds to the very line used by Reith et al. (9). The retarded NF-X bands were also of similar aspect and abundance in normal and SCID lines when fragments of the DRα promoter (even as long as 61 bp) were used as templates (data not shown), and when the extract preparation and gel retardation protocols were slightly modified to follow those of Reference 9.

Structural features. The slow migration of NF-X in polyacrylamide gels suggested that it is a complex of rather large size. We have confirmed this point by Ferguson analysis (26), which showed the NF-X complexes to be in the range of several hundred kDa (data not shown). Yet it is unlikely that such a mass of protein is directly required for DNA binding. Indeed, most eukaryotic transcription factors have a modular organization, with a DNA-binding domain that is quite distinct from other functional regions (2). This also proved true of NF-X. As shown in Figure 2*A*, digestion of NF-X:oligo complexes with proteinase K gave rise to a very rapidly migrating retarded band, which was resistant to further protease attack. This small complex is specific of the X box, as it does not bind the control Xc oligo. The same band was obtained after trypsin and subtilisin A, but not pepsin or papain digestion (not shown). This band reflects a complex of very low m.w.: it was best seen on gels of higher polyacrylamide concentration than usual (Figs. 2, 4, and 5); its behavior when analyzed "à la Ferguson" indicated a molecular weight of less than 10 kDa (data not shown).

That this band indeed derived from NF-X was demonstrated by methylation interference experiments: the main DNA contacts were the same for the protease-resistant core and the intact NF-X complexes (Fig. 2*B*). There were some differences in the degree of interference at some positions, in particular at the last G on the antisense strand (*G4* in Fig. 3*B*) where methylation did not interfere with binding of the protease-resistant core. Interestingly, this is also the position where Reith et al. (25) find differences between monomeric and dimeric forms of RF-X. Somewhat surprisingly, the intensity of

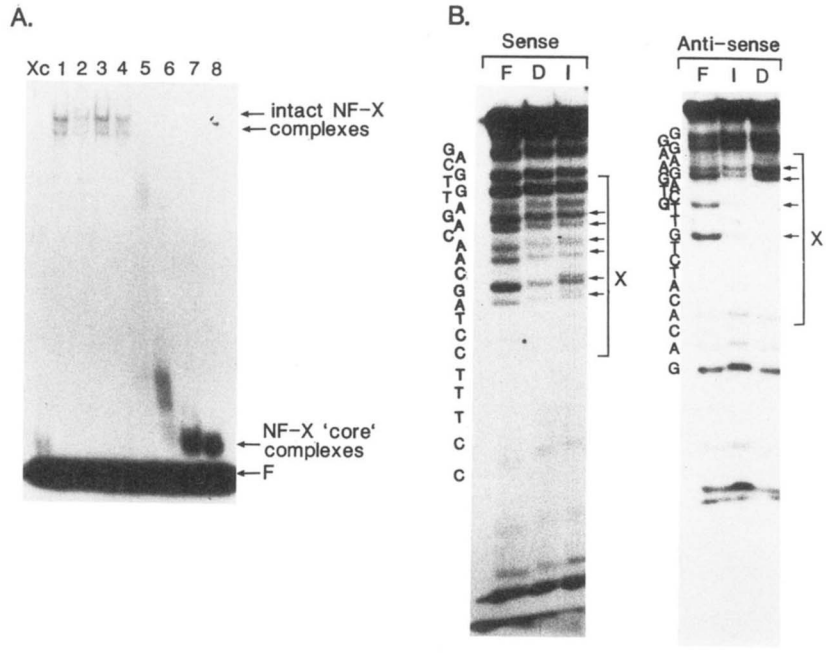

Figure 2. A protease-resistant DNA binding core. *A*, Gel retardation assay after proteinase K treatment. The retardation assay was performed as in Figure 1 except that the samples were treated for 10 min with various amounts of proteinase K before electrophoresis (*lane 1*, 0 ng; *2*, 0.5 ng; *3*, 1 ng; *4*, 5 ng; *5*, 10 ng; *6*, 50 ng; *7*, 100 ng; *8*, 300 ng). The samples were run on a 7% polyacrylamide gel to better visualize the resulting very low m.w. *B*, Comparison of the intact NF-X complex and the DNA-binding core by methylation interference mapping. Interference maps are visualized on the sense and antisense strands. *F*, free DNA; *D*, digested complex; *I*, intact complex. The *arrows* indicate bases whose methylation inhibits the binding of NF-X or the protease-resistant core.

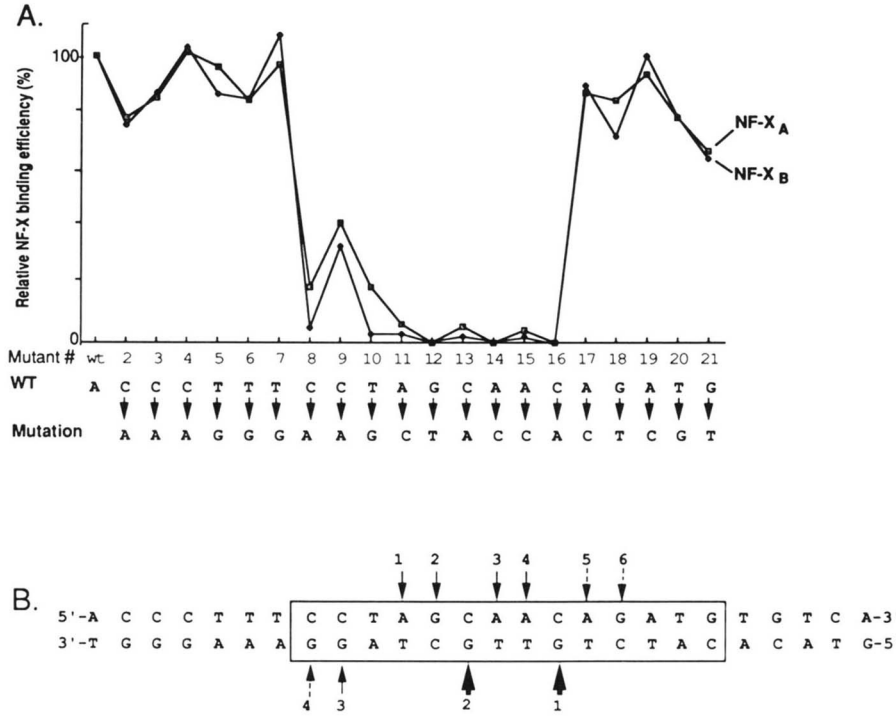

Figure 3. Definition of the critical nucleotides for NF-X binding. *A*, A set of 21 double-stranded oligos, each bearing a single site transversion from the Eα X 23 mer, was tested in the gel retardation assay. NF-X binding was quantitated by densitometry of the two bands corresponding to NF-X_A and NF-X_B, the value obtained for the wild-type oligo representing 100% binding efficiency. The *points* represent values averaged from at least three independent experiments, with a SD of less than 14% of the mean. *B*, The Eα X box sequence aligned with the diagram in *A*. *Arrows* point to contact residues identified by methylation interference, the thickness of the arrows correlating with the degree of interference (see Fig. 2*B*); the classical X homology is boxed.

the retarded bands was much greater for the core complex. This difference could reflect a higher stability for complexes formed with the small DNA-binding core—both before and during gel migration.

The intensity and positions on the gel of the protease-resistant DNA-binding core was the same in extracts from SCID or normal lymphoblastoid lines (data not shown).

Binding site characteristics. Previous methylation interference experiments (8) have established that NF-X contacts bases in the core of the X box and immediately upstream of it, much as RF-X does (9, 23). Further experiments established that NF-X_A and NF-X_B make identical contacts, and that these are the same whether extracts derive from class II^+ or class II^- cell lines (8) (data not shown). To further explore NF-X's binding characteristics, we synthesized a set of mutant double-stranded oligos, each consisting of the same 23-bp sequence except for a unique single-site transversion. Each oligo was evaluated in the gel retardation assay, and the relative binding efficiency was calculated by densitometry (Fig. 3). The most drastic mutations were those that touch the central part of the X box, over a length of 6 to 9 bp. Interestingly, this is the center of the sequence conservation seen when one compares X boxes in class II α and β genes from a variety of species (3). It is clear from the data in Figure 3 that the binding sites of NF-X_A and NF-X_B are exactly superimposable. Identical patterns were obtained using extracts of nuclei from class II^+ and class II^- cells (data not shown). Finally, the protease-resistant

DNA-binding core exhibited the same pattern of sensitivity to the binding-site mutants (data not shown).

NF-X was originally detected using Eα and DRα X box oligos (4, 8) (Fig. 1). It was important to determine whether NF-X also binds to the X boxes of other MHC class II genes, particularly because it has been proposed that distinct factors bind to and mediate the influence of the X box in different genes (12, 13). Gel retardation assays were performed with double-stranded oligos of identical length, spanning the X motifs from the Eα, Aα, Eβ, and Aβ genes. As illustrated in Figure 4, NF-X bands were observed with all oligos, both as intact complexes (*A*) and as protease-resistant cores (*B*). NF-X did seem to have a preference for the Eα oligo relative to the others (about 3- to 15-fold), which may explain why, in other reports, NF-X binding to Aα was obscured by the more intense signal from X2-binding proteins (12). This lower affinity of NF-X for Aα relative to Eα was also substantiated by competition experiments (Fig. 4*C*). A 10- to 15-fold greater amount of Aα X box competitor was required, compared to Eα X box competitor. Our data on this point agree with those of Dedrick and Jones (15).

To determine whether NF-X binding sites are restricted to MHC class II genes, we searched the literature for factors reported to have similar properties or target sequences. We did observe some similarity between the NF-X binding site, as defined in Figure 3, and the recognition sequence for a factor which interacts with a palindromic *GTTGC*NNN*GCAAC* motif in the polyomavirus and hepatitis virus enhancers (called "EF-C" or "EP"; 28–30). In addition, NF-X and EF-C/EP seemed to behave similarly in gel retardation experiments, both giving rise to a doublet or triplet of bands close to the gel origin. Methylation interference footprints of EF-C (29) are also very reminiscent of those we observe with NF-X on the Eα X box. To verify the possibility that these activities are identical, we made a direct comparison. Figure 5 shows that complexes formed with the Eα X box oligo and with an oligo carrying the polyoma palindrome (Py) comigrated on the gel. Cross-competition experiments confirmed that these complexes are the same; cold Eα and Py oligos compete against the formation of NF-X bands with either Py or Eα-labeled oligos. Further, cold Py oligo is a more efficient competitor than Eα, whichever labeled oligo is used. Interestingly, the NF-X/Py oligo complex gave rise to two bands after proteolytic cleavage (Fig. 5*B*), and the apparent sizes were such that one band could have represented a dimer complex. This interpretation is quite plausible because the Py oligo contains two palindromic NF-X binding sites, whereas the Eα oligo has only one.

Role in transcription. Although it is clear that the X box is a crucial promoter element of MHC class II genes, the identity of the factor or factors which mediate its influence remains uncertain (see *Introduction*). To investigate whether NF-X plays a role in X box function, we constructed a set of Eα promoters carrying short mutations along the X box and surrounding sequences. We wished to correlate the performance of these mutant promoters in functional assays with the ability of mutant fragments to bind NF-X (as detailed above).

The various mutations were introduced into the plasmid PE3, in which Eα promoter sequences (from position −215 to +12) drive transcription of a rabbit β-globin reporter gene (8, 21). We chose to introduce 3-bp replacements because we have found that single base changes, although capable of profoundly affecting factor binding in vitro, have only limited consequences on transcriptional activity (J. Bollekens and W. Koch, unpublished observations). This is probably due to the stabilizing effects that other factors have within an initiation complex. The replacements were introduced with synthetic oligonucleotides, using two rounds of PCR to assemble and amplify mutated fragments covering positions −215 to +12 of Eα (see *Materials and Methods*). These fragments were then inserted into PE3, replacing the corresponding wild-type fragment. The presence of each mutation and the integrity of neighboring sequences were verified by sequencing the whole −215 to +8 segment in each mutant plasmid. The position and sequence of the various replacements are diagrammed in Figure 6.

The effects of the mutations on promoter activity was evaluated by in vitro transcription. This assay, described in detail elsewhere (see footnote 5), relies on the use of transcriptionally active extracts from the class II$^+$ lym-

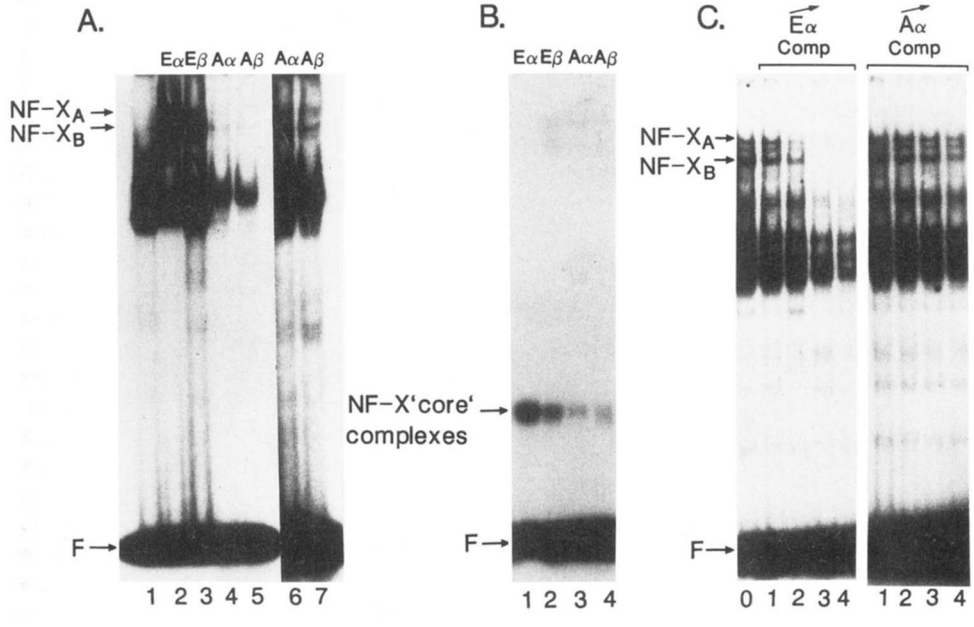

Figure 4. *A*, NF-X binding to diverse class II gene promoters. The gel retardation assay was conducted as described in Figure 1, with labeled oligos spanning the X motifs from several murine class II genes (sequences are listed in Table I). *Lane 1*, Eα X control oligo. *Lanes 6* and *7*, fivefold longer exposure of *lanes 4* and *5*. *B*, Binding of the NF-X protease-resistant core to different X boxes. The gel retardation assays was as in *A* except that the complexes were digested with proteinase K just before electrophoresis. *C*, Cold competition. Increasing amounts of unlabeled Eα or Aα X box oligo were used to compete against NF-X binding to labeled Eα X oligo. *Lane 0*, no competitor. *Lanes 1* to *4*, 100, 300, 900, and 2700 fmol of cold competitor, respectively.

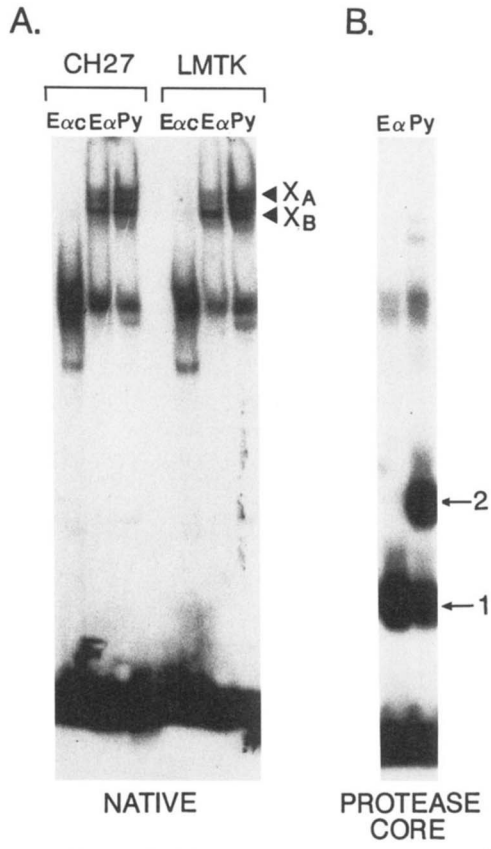

Figure 5. Binding of NF-X to the polyomavirus enhancer. *A*, Retardation assays were performed with Eα X box and control oligos or with an oligo containing the C motif of the polyoma enhancer (*Py*). The oligos were assayed with nuclear extracts from different cells: LMTK, a fibroblast cell line and CH27, a B lymphoma line. *B*, As *A*, but with proteinase K treatment just before electrophoresis of the complexes.

phoma line CH27. Transcription is initiated accurately and is critically dependent on the X and Y boxes (see footnote 5). Plasmid DNA were transcribed in this in vitro system, and the RNA produced were quantitated by S1 nuclease mapping. The results from several independent experiments are summarized in Figure 6. Confirming our previous observations (see footnote 5), sequences upstream of the X box were of little importance in this system: mutants 4 through 11 were essentially as active as the WT template. Within the X box, mutations 13 and 14 had the most drastic effect, reducing transcription 10-fold or more. Suggestively, the bases altered in these mutants correspond exactly to the NF-X binding core, as defined in Figure 3. The partial reduction seen with mutant 12 is also consistent with the binding data in Figure 3. We also noted a reasonably strong effect of mutations 16 and 17, which map in X2, immediately downstream from the classical X box. These effects suggest a role for X2 and X2-binding factors in Eα gene regulation.

DISCUSSION

Structure. We have defined NF-X as the complex formed upon incubation of nuclear extracts with a DNA fragment containing the X box from the Eα promoter. Its specificity was indicated by its absence when a control fragment carrying a highly mutated X box was used (4). We found NF-X to be quite large. Although several hundred kDa could conceivably reflect a single huge protein, it is more suggestive of a multimeric complex (homo- or heteromeric?). This notion is supported by our demonstration that NF-X is identical to EF-C, a factor known to bind as a dimer (28–30). We found that the polyoma enhancer oligo could bind one or two protease-resistant cores whereas the Eα oligo could bind only one (Fig. 5*B*). Inasmuch as the intact complexes looked exactly similar with both oligos (Fig. 5*A*), we believe that NF-X must be a dimer that can bind two linked motifs (as for the polyoma enhancer) or two independent oligonucleotides (as for Eα). Incidentally, a dimer with two X box-binding sites would provide a structural basis for the hypothesis we formulated some time ago that the inverted and duplicated X-Y motif in the far upstream region of several class II gene promoters (X′-Y′) could be juxtaposed to the X and Y boxes through dimerization of X or Y box-binding proteins (31): NF-X could fasten X-Y and X′-Y′ by binding to both the X and X′ boxes.

The relationship between NF-X_A and NF-X_B is not obvious at present. It is clear, though, that their binding properties are indistinguishable (Fig. 3) (8). It is conceivable that the two forms correspond to different gene products or that they result from modifications of products emanating from the same gene(s): alternative splicing, post-translational modifications, or just artefactual proteolysis.

The very small size (<10 kDa) of the DNA-binding core

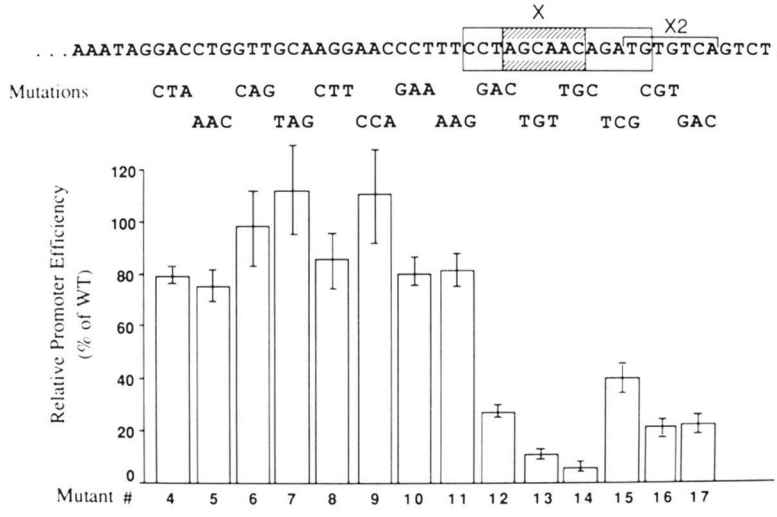

Figure 6. Functional evaluation of X box mutants. The X motif is indicated by a *solid box*, and the nucleotides defined as critical for NF-X binding by the saturation mutagenesis experiment are shown by a dashed box (Fig. 4). The position and exact sequence of each 3-bp mutation is shown below the sequence. The various mutant plasmids were tested in an in vitro transcription system (see footnote) with nuclear extracts from the class II^+ B lymphoma line CH27. Accurately initiated transcripts from the Eα promoter were quantitated by S1 mapping. The values displayed in the *bar graph* are averages from at least three independent experiments with at least two different plasmid preparations.

visualized after protease digestion contrasts with the very large apparent size of the intact complexes. Interestingly, the protease-resistant core bound more efficiently than intact NF-X in the gel retardation assay. As seen in Figure 2, there was a 5- to 10-fold increase in the amount of oligo retarded after proteolysis. The DNA-binding core also appeared less sensitive to small variations in the ligand sequence: the sharp preference for Eα over Aα which we observed with intact NF-X was less marked after proteolysis (Fig. 4). This difference in apparent affinities between the whole protein and the DNA-binding domain has several interpretations. Trivially, the smaller complex might simply be more stable during electrophoresis. More interestingly, the domains of NF-X that are removed by proteolysis might have a strong negative influence on the affinity of the DNA binding domain for its ligand. This type of interplay would be analogous to that already described for steroid hormone receptors, in which the hormone-binding domain modulates the action of the DNA-binding domain (32).

NF-X appears ubiquitous, because it can be detected in a large variety of cell-types irregardless of their class II phenotype (Fig. 1) (data not shown). There may be slight differences in the appearance of NF-X between lines, in particular when considering the presence or exact position of the third band often found between NF-X_A and NF-X_B. Yet this additional band tends to differ with independent extract preparations from the same cell line so that, given the high sensitivity of NF-X to proteolysis, we are reluctant to attribute much significance to these variations at this stage. We are thus faced, once again, with the paradox of the tissue-specific action of the X-Y motif, in spite of the ubiquity of NF-Y and NF-X. Conceivably, NF-X and NF-Y form a "supercomplex" with yet another factor with a restricted cell-type distribution. Alternatively, there may be subtle variation in NF-X between tissues, analogous to the tissue-specific differential splicing observed for NF-Y (X-Y. Li and R. Hooft van Huijsduijnen, unpublished observations).

Function. A key question to address for any DNA-binding activity discovered in vitro is: what is its functional relevance? Several experiments were performed to address this issue, and we believe there is now good evidence to indicate that NF-X is indeed involved in the regulation of MHC class II genes.

1) Data from the saturation mutagenesis experiment (Fig. 3) show that the core of the NF-X binding sequence is at the center of the X sequence homology. These results correlate well with sequence conservation amongst alpha and beta class II genes in various species (for compilation, see Ref. 3).

2) Strikingly, the mutations that affect transcription the most are exactly those that one would have predicted on the basis of these binding data: mutations 13 and 14 that are devastating for transcription (Fig. 6) fall exactly within the critical AGCAAC motif. Thus, there appears to be an excellent correlation between NF-X binding specificity and X box function.

These data, although correlative in nature, argue strongly for a functional role for NF-X in the transcriptional activity of the Eα promoter. Does this extend to other MHC class II genes or are there, as argued by others (12, 13), distinct X box-binding regulatory factors for each? We readily detected NF-X binding to the X boxes of three other murine class II genes. Although NF-X clearly had a preference for Eα, the difference was no more than 5- to 15-fold. Given the usual affinities of sequence-specific DNA-binding proteins for their ligand (on the order of 10^{-9}, 10^{-10} M), such a difference is unlikely to have very serious consequences for the formation of an initiation complex, particularly if one envisions the binding of a single factor as stabilized by interaction with other factors bound to adjacent sites. We thus consider it likely that NF-X plays a role in regulating the transcription of other class II genes. This view is in good agreement with the analysis of the Aα promoter by Dedrick and Jones (15) (L. Glimcher, personal communication).

NF-X and RF-X. If, as argued above, NF-X is indeed involved in class II gene regulation, what is the relationship between NF-X and RF-X, suggested by Reith et al. (9) to be the root of the genetic defect in class II-deficient SCID patients?

By certain criteria, the two activities are very similar: large size, very small DNA-binding core, possible dimer structure, and very similar binding requirements as judged from methylation interference experiments (compare Fig. 3B with Fig. 6 of Ref. 25). However, their appearance in gel retardation assays was clearly different and, most importantly, results with the SCID lines are conflicting: NF-X is not affected (this report) although RF-X is absent (9).

This contradiction may not be as absolute as it appears at first glance, considering new developments on RF-X. It now seems, on the basis of experiments with the cloned RF-X gene and its product, that SCID lines do contain RF-X mRNA and protein of normal abundance, size, and sequence (23, 25) (W. Reith and B. Mach, personal communication). It is quite possible, then, that RF-X and NF-X really represent the same entity, and that they are not affected by the genetic lesion in SCID patients. How, then, does one explain the initial observations on SCID lines? It is conceivable that the retarded band found by Reith et al. (9) to be missing in SCID extracts was really an altered form of NF-X, due either to post-translational modification or to association with another factor. Contrary to previous conclusions (9, 23), the SCID defect would then lie in another factor that binds to or modifies RF-X, not in RF-X proper.

Acknowledgments. We thank P. Gerber for technical assistance, and A. Ney, B. Boulay, and C. Werlé for help in preparing the manuscript.

REFERENCES

1. **Mitchell, P. J., and R. Tjian.** 1989. Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science 245:371.*
2. **Johnson, P. F., and S. L. McKnight.** 1989. Eukaryotic transcriptional regulatory proteins. *Annu. Rev. Biochem. 58:799.*
3. **Benoist, C., and D. Mathis.** 1990. Regulation of major histocompatibility complex class-II genes: X, Y and other letters of the alphabet. *Annu. Rev. Immunol. 8:681.*
4. **Dorn, A., B. Durand, C. Marfing, M. Le Meur, C. Benoist, and D. Mathis.** 1987. Conserved major histocompatibility complex class II boxes-X and Y-are transcriptional control elements and specifically bind nuclear proteins. *Proc. Natl. Acad. Sci. USA 84:6249.*
5. **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.*
6. **Sittisombut, N.** 1988. Two distinct nuclear factors bind the con-

served regulatory sequences of a rabbit major histocompatibility complex class II gene. *Mol. Cell. Biol. 8:2034.*

7. **Calman, A. F., and B. M. Peterlin.** 1988. Evidence for a trans-acting factor that regulates the transcription of class II major histocompatibility complex genes: genetic and functional analysis. *Proc. Natl. Acad. Sci. USA 85:8830.*
8. **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.*
9. **Reith, W., S. Satola, C. Herrero Sanchez, I. Amaldi, B. Lisowska-Grospierre, C. Griscelli, M. R. Hadam, and B. Mach.** 1988. Congenital immunodeficiency with a regulatory defect in MHC class II gene expression lacks a specific HLA-DR promoter binding protein, RF-X. *Cell 53:897.*
10. **Liou, H-C., M. R. Boothby, and L. H. Glimcher.** 1988. Distinct cloned class II MHC DNA binding proteins recognize the X box transcription element. *Science 242:69.*
11. **Sherman, P. A., P. V. Basta, T. L. Moore, A. M. Brown, and J. P. Y. Ting.** 1989. Class II box consensus sequences in the HLA-DRα gene: transcriptional function and interaction with nuclear proteins. *Mol. Cell. Biol. 9:50.*
12. **Boothby, M., H.-C. Liou, and L. H. Glimcher.** 1989. Differences in DNA sequence specificity among MHC class II X box binding proteins. *J. Immunol. 142:1005.*
13. **Celada, A., and R. Maki.** 1989. Evidence for multiple major histocompatibility class II X-box binding proteins. *Mol. Cell. Biol. 9:5219.*
14. **Hume, C. F., and J. S. Lee.** 1990. Congenital immunodeficiencies associated with absence of HLA class II antigens on lymphocytes result from distinct mutations in trans-acting factors. *Hum. Immunol. 26:288.*
15. **Dedrick, R. L., and P. P. Jones.** 1990. Sequence elements required for activity of a murine major histocompatibility complex class II promoter bind common and cell-type-specific nuclear factors. *Mol. Cell. Biol. 10:593.*
16. **Tsang, S. Y., M. Nakanishi, and B. M. Peterlin.** 1990. Mutational analysis of the DRA promoter: cis-acting sequences and trans-acting factors. *Mol. Cell. Biol. 10:711.*
17. **Liou, H -C., M. R. Boothby, P. W. Finn, R. Davidon, N. Nabavi, N. J. Zelenznik-Le, J. P. Y. Ting, and L. H. Glimcher.** 1990. A new member of the leucine zipper class of proteins that binds to the HLA DRα promoter. *Science 247:1581.*
18. **Lin Y-S., and M. R. Green.** 1989. Similarities between prokaryotic and eukaryotic cyclic AMP-responsive promoter elements. *Nature 340:656.*
19. **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.*
20. **Dorn, A., J. Bollekens, A. Staub, C. Benoist, and D. Mathis.** 1987. A multiplicity of CCAAT box-binding proteins. *Cell 50:863.*
21. **Koch, W., C. Benoist, and D. Mathis.** 1989. Anatomy of a new B-cell-specific enhancer. *Mol. Cell. Biol. 9:303.*
22. **Higuchi, R.** 1989. In *PCR technology.* H. Erlich, ed. Stockton Press, New York, pp. 61–70.

22a. **Viville, S., V. Jongeneel, W. Koch, R. Mantovani, F. Ruffenach, C. Benoist, and D. Mathis.** 1991. The Eα promoter: a linker-scanning analysis. *J. Immunol. 146:3211.*

23. **Reith, W., E. Barras, S. Satola, M. Kobr, D. Reinhart, C. Herrero Sanchez, and B. Mach.** 1989. Cloning of the major histocompatibility complex class II promoter binding protein affected in a hereditary defect in class II gene regulation. *Proc. Natl. Acad. Sci. USA 86:4200.*
24. **Griscelli, C., B. Lisowska-Grospierre, and B. Mach.** 1989. Combined immunodeficiency with defective expression in MHC class II genes. *Immunodeficiency Rev. 1:135.*
25. **Reith, W., C. Herrero-Sanchez, M. Kobr, P. Silacci, C. Berte, E. Barras, S. Fey, and B. Mach.** 1990. MHC class II regulatory factor RFX has a novel DNA-binding domain and a functionally independent dimerization domain. *Genes Dev. 4:1528.*
26. *Gel Electrophoresis of proteins: A practical approach.* 1981. B. D. Hames, D. Richwood, eds. IRL Press, p. 12.
27. **Kadonaga, J. T., and R. Tjian.** 1986. Affinity purification of sequence-specific DNA binding proteins. *Proc. Natl. Acad. Sci. USA 83:5889.*
28. **Ostapchuk, P., J. F. X. Diffley, J. T. Bruder, B. Stillman, A. J. Levine, and P. Hearing.** 1986. Interaction of a nuclear factor with the polyomavirus enhancer region. *Proc. Natl. Acad. Sci. USA 83:8550.*
29. **Ostapchuk, P., G. Scheirle, and P. Hearing.** 1989. Binding of nuclear factor EF-C to a functional domain of the hepatitis B virus enhancer region. *Mol. Cell. Biol. 9:2787.*
30. **Ben-Levy, R., O. Faktor, I. Berger, and Y. Shaul.** 1989. Cellular factors that interact with the hepatitis B virus enhancer. *Mol. Cell. Biol. 9:1804.*
31. **Dorn, A., H. J. Fehling, W. Koch, M. Le Meur, P. Gerlinger, C. Benoist, and D. Mathis.** 1988. B-cell control region at the 5′ end of a major histocompatibility complex class II gene: sequences and factors. *Mol. Cell. Biol. 8:3975.*
32. **Evans, R. M.** 1988. The steroid and thyroid hormone receptor super family. *Science 240:889.*