Proc. Natl. Acad. Sci. USA
Vol. 85, pp. 7322–7326, October 1988
Immunology

Characterization of the cDNA encoding a protein binding to the major histocompatibility complex class II Y box

(HLA/gene regulation)

DANIEL K. DIDIER, JOEL SCHIFFENBAUER, SUSAN L. WOULFE, MICHELLE ZACHEIS, AND BENJAMIN D. SCHWARTZ

Howard Hughes Medical Institute Laboratories and Department of Medicine, Division of Rheumatology, Washington University School of Medicine, Saint Louis, MO 63110

Communicated by Paul E. Lacy, June 13, 1988 (received for review April 14, 1988)

ABSTRACT **The expression of HLA class II genes is regulated by a series of cis-acting elements and trans-acting factors. Several cis-acting elements have been identified and have been termed the Z box, X box, Y box, octamer, and "TATA" box. The Y box contains an inverted CCAAT box. By probing a phage λgt11 library with double-stranded oligonucleotides, we have directly isolated a cDNA encoding a Y box-binding protein designated YB-1. YB-1 binding has an absolute requirement for the CCAAT box and relative specificity for the Y box. It has a M_r of 35,414, contains 18% basic residues, and contains putative nuclear localization signals. An inverse correlation of YB-1 and HLA-DR β chain mRNA levels suggests that YB-1 is a negative regulatory factor.**

The transcription of major histocompatibility complex class II genes is regulated by a series of cis- and trans-acting elements (1). For class II genes, the majority of the cis-acting elements has been localized to the 5′ flanking and intronic regions (2, 3). The promoter region within the 5′ flanking region contains a series of sequence motifs that are highly conserved among all class II genes and the invariant chain gene (1–4). These motifs have been termed the Z box, X box, Y box, octamer, and "TATA" box. The Y box contains an inverted CCAAT box. The functional role of these motifs has been explored by using transfection and transgenic systems. Both the X box and Y box have been shown to be essential for transcription, while γ interferon inducibility has been ascribed variably to the X box, Y box, and Z box (1, 5–8). Investigations of the trans-acting factors interacting with the cis-acting elements have recently begun and have largely exploited the ability of trans-acting elements within crude nuclear extracts to bind to double-stranded (ds) oligonucleotides corresponding to sequences within the 5′ flanking region (5, 7–11). We have utilized an alternative strategy in which we have probed a phage λgt11 library with ds oligonucleotides to directly isolate the cDNAs encoding the trans-acting factors. We have isolated and sequenced a cDNA that encodes the Y box-binding protein, YB-1.* YB-1 has specificity for the CCAAT box within the Y box, has a M_r of 35,414, contains 18% basic residues, and contains putative nuclear localization signals. In addition, YB-1 mRNA levels inversely correlate with HLA-DR β chain mRNA levels.

MATERIALS AND METHODS

λgt11 Library. A λgt11 library was constructed from Swei cells as described (12).

ds Oligonucleotide Synthesis. ds oligonucleotides were prepared as described (11). Either the complementary strands were first hybridized and then γ-^{32}P-end-labeled, or each strand was γ-^{32}P-end-labeled and then hybridized. The labeled ds oligonucleotides were purified by gel electrophoresis through a 15% polyacrylamide gel.

Filter Hybridization. The amplified library was plated, and the plates were overlaid with nitrocellulose filters saturated with 10 mM isopropyl thiogalactoside. Filters were prehybridized with 5% Carnation instant milk in 10 mM Hepes (pH 8.0) for 1 hr at room temperature and then washed twice for 10 min with 10 mM Hepes (pH 8.0) (13). Hybridization was done in 250 mM NaCl/5 mM $MgCl_2$/1 mM dithiothreitol/0.1 mM EDTA/10 mM Hepes, pH 8.0, containing 10 μg of salmon sperm DNA or 10 μg of poly(I-C) per ml. After 10 min, 10^8 cpm of end-labeled ds oligonucleotide was added, and the incubation was continued for 2 hr at room temperature. Filters were washed twice for 10 min with hybridization buffer at room temperature and were autoradiographed.

Molecular Weight Estimation of the Fusion Protein β-Galactosidase–YB-1. *Escherichia coli* strain Y1088 or Y1090 (10^8) was infected with 10^6 plaque-forming units of the phage λgt11-YB-1 and grown at 37°C until partial lysis was achieved. Isopropyl thiogalactoside was added to a final concentration of 10 mM, and the incubation was continued an additional 30 min at 37°C. Phenylmethylsulfonyl fluoride was added to a final concentration of 200 μg/ml, and the incubation was continued another 30 min at 37°C. The bacteria were centrifuged at 4550 × *g* for 5 min at 4°C. The supernatant was concentrated 20-fold in a Centricell ultrafilter (Polyscience, Warrington, PA) to 1 ml. Ten microliters of the supernatant was run in each of four lanes on a $NaDodSO_4$/4–20% polyacrylamide gradient gel (Novex, Encinitas, CA) and transferred to Nitroplus 2000 (Micron Separations, Westboro, MA) by using the Novex immunoblotting unit. For detection of DNA-binding proteins (14), the blots were prehybridized and hybridized as described above, and a γ-^{32}P-end-labeled class II box was used as the DNA probe. For immunoblotting (13), the blots were exposed to 1% bovine serum albumin in 150 mM NaCl/0.05% Tween-20/10 mM Tris·HCl, pH 8.0, for 30 min at room temperature (13), then exposed to rabbit anti-β-galactosidase (5 Prime to 3 Prime, Paoli, PA) in the same buffer for 30 min at room temperature, and washed three times in the same buffer. The antibody bound to the fusion protein was visualized by using a Protoblot immunoblotting system (Promega Biotec, Madison, WI) according to the directions supplied by the manufacturer.

DNA Sequencing. YB-1 cDNA was isolated from λgt11 by partial digestion with *Eco*RI. The 1.5-kb fragment was ligated into phage M13mp18 in both orientations. The M13mp18-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: ds, double stranded; *tk*, thymidine kinase gene.

*The sequence reported in this paper is being deposited in the EMBL/GenBank data base (IntelliGenetics, Mountain View, CA, and Eur. Mol. Biol. Lab., Heidelberg) (accession no. J03827).

YB-1 clones were digested with *Bam*HI and *Pst* I. The *Bam*HI ends were digested with *Exo* III and mung bean nuclease as recommended by the manufacturer (Stratagene Cloning Systems, La Jolla, CA) and religated. After transformation, single-stranded cDNA was prepared and sequenced by the dideoxy chain-termination method (15).

Slot Blots. Total RNA was extracted from 2×10^8 cells by the method of Chirgwin *et al.* (16). Total RNA (3 μg) was absorbed onto nitrocellulose by using a slot-blot system (Schleicher & Schuell), dried, and baked 2 hr at 80°C. Filters were prehybridized and hybridized as described by Maniatis *et al.* (17). The first 1100-base-pair (bp) *Eco*RI fragment from YB-1 cDNA and the entire HLA-DR5 β_1 chain cDNA (18) were ^{32}P-labeled by using the random hexamer technique (31) to a specific activity of $\approx 1 \times 10^9$ cpm/μg and were used to probe the slot blots. Filters were washed at 50°C and subjected to autoradiography.

RESULTS AND DISCUSSION

We initially designed our studies to isolate directly the cDNA encoding the regulatory proteins that interact with the X or Y box, or both, of the *HLA-DR* α chain gene. Therefore, we synthesized a 48-bp ds oligonucleotide corresponding to the 5′ region of the *HLA-DR* α chain gene, which spans the X box, spacer region, and Y box. We denoted this ds oligonucleotide the "class II box" (Table 1).

A cDNA library from the class II-positive lymphoblastoid B-cell line Swei was made by using λgt11, and the library was grown in *E. coli* strain Y1090 to produce a fusion protein between the products of the β-galactosidase gene and the cloned cDNA. The amplified library of 1×10^7 clones was plated. The plates were overlaid with nitrocellulose filters saturated with isopropyl thiogalactoside to induce expression of the fusion proteins. Previously, such λgt11 libraries have been screened with antibodies to the desired protein. However, we screened the library by hybridizing the filters with the ^{32}P-end-labeled class II box in the presence of poly(I-C) (to inhibit nonspecific binding) to identify proteins that bind to the X or Y box. Upon initial screening, 34 clones appeared positive. After three rounds of plaque purification, only 3 clones remained positive. One clone hybridized with the ds oligonucleotide corresponding to an HLA-DP α chain coding sequence (Table 1) and was therefore nonspecific; the other two clones were specific, but identical. This clone was expanded and characterized.

To assign the specificity of the protein encoded by this cDNA, filters were probed with the class II box or with ds oligonucleotides corresponding to the X box, the Y box, the spacer region, and an HLA-DP α chain coding sequence as a nonspecific control (Fig. 1). The entire class II box bound strongly to the protein produced, while the nonspecific HLA-DP α chain probe did not bind at all. The X box bound very weakly, the spacer region bound at an intermediate level, and the Y box bound most strongly. These results suggested that the cloned cDNA encodes a protein that has affinity for the Y box. This protein was designated YB-1.

To determine whether YB-1 was absolutely specific for the Y box or would bind to other known CCAAT box-containing promoters, we probed filters with CCAAT-box-containing ds oligonucleotides corresponding to a long-Y box containing additional base pairs 5′ to the Y box, the herpes simplex virus thymidine kinase gene (*tk*) promoter (19), and the rabbit β-globin gene promoter (20) (Table 1 and Fig. 2). Once again the long-Y box bound most strongly, the *tk* promoter oligonucleotide bound weakly, and the rabbit β-globin gene promoter oligonucleotide did not bind at all. Mutations of the CCAAT sequence within the long-Y box and the *tk* promoter oligonucleotide (Table 1) abrogated binding (Fig. 2), indicating that YB-1 protein was indeed a CCAAT box-binding protein with relative specificity for the Y box. Because it is a CCAAT-box binding protein, the binding of YB-1 to other CCAAT box-containing sequences is expected (11). The stronger binding of YB-1 to the *tk* promoter than to the other promoters containing CCAAT boxes can be explained by the similarity in the sequences surrounding the CCAAT boxes in the long-Y box and *tk*. However, the specificity of YB-1 for the long-Y box is demonstrated by the loss of binding of YB-1 to *tk* at 68°C and the retention of binding of YB-1 to the long-Y box at this temperature (data not shown).

The molecular weight of the fusion protein produced by the λgt11-infected *E. coli* was determined by protein blotting with a specific DNA probe and a specific antibody probe after $NaDodSO_4$/PAGE (Fig. 3). The class II box probe (lanes 1 and 2) and anti-β-galactosidase antibody (lanes 3 and 4) both bound to a band of $M_r \approx 150{,}000$. The molecular weight of the portion of the β-galactosidase protein contained within the fusion protein is ≈114,000; thus, the estimated molecular weight of YB-1 by this method is ≈36,000.

The cDNA encoding YB-1 was characterized and sequenced. The nucleotide and derived amino acid sequence is shown in Fig. 4. The size of the cDNA was ≈1.5 kb. It

Table 1. Oligonucleotides used in these studies

Class II box	5′–CCTAGCAACAGATGCGTCATCTCAAAATATTTTTCTGATTGGCCAAAG–3′
	3′–GGATCGTTGTCTACGCAGTAGAGTTTTATAAAAAGACTAACCGGTTTC–5′
Y box	5′–CTGATTGGCCAAAG–3′
	3′–GACTAACCGGTTTC–5′
Long Y (LY) box	5′–ATTTTTCTGATTGGCCAAAG–3′
	3′–TAAAAAGACTAACCGGTTTC–5′
Mutated LY box	5′–ATTTTTCTGCTGGGCCAAAG–3′
	3′–TAAAAAGACGACCCGGTTTC–5′
X box	5′–CCTAGCAACAGATG–3′
	3′–GGATCGTTGTCTAC–5′
Spacer	5′–CGTCATCTCAAAATATTTTT–3′
	3′–GCAGTAGAGTTTTATAAAAA–5′
HLA-DP α coding	5′–GGGGAGTTTATGTTTGAATTTG–3′
	3′–CCCCTCAAATACAAACTTAAAC–5′
tk	5′–GCGTCTTGTCATTGGCGAATTCG–3′
	3′–CGCAGAACAGTAACCGCTTAAGC–5′
Mutated *tk*	5′–GCGTCTTGTCAGGTGCGAATTCG–3′
	3′–CGCAGAACAGTCCACGCTTAAGC–5′
β-globin	5′–GTGTGAGCAGATTGGCCCTTACCAGGGTGT–3′
	3′–CACACTCGTCTAACCGGGAATGGTCCCACA–5′

CCAAT box and mutated CCAAT box sequences are underlined.

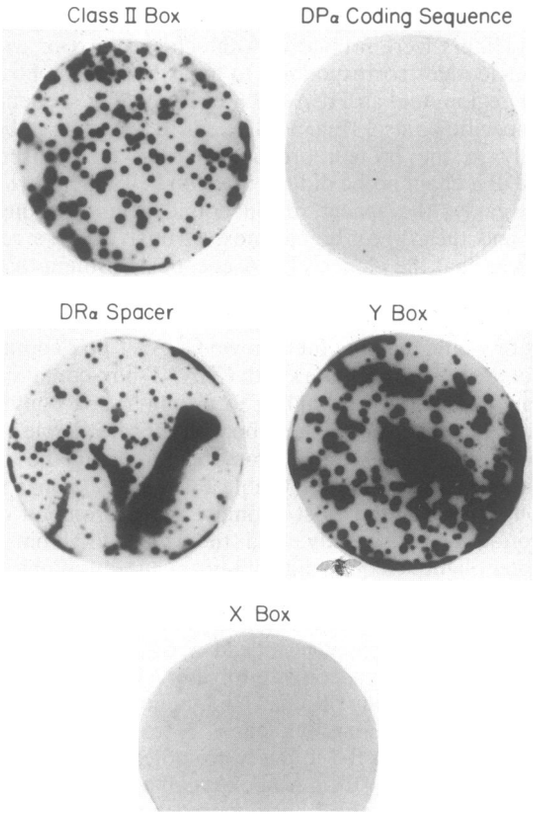

FIG. 1. Specificity of YB-1 for the class II Y box. The clone encoding YB-1 was plated and overlaid with nitrocellulose filters impregnated with 10 mM isopropyl thiogalactoside to induce fusion protein synthesis. Replicate filters were hybridized with ^{32}P-labeled ds oligonucleotides corresponding to the HLA-DR α chain class II box, the X box, the Y box, spacer region, and a HLA-DP α chain coding sequence (see Table 1). The class II box and Y box were bound most strongly, with less strong binding of the spacer and weak binding of the X box. No binding was seen with the HLA-DP α chain coding sequence.

contained an open reading frame of 1077 bases and a 3′-untranslated region of 435 bases. The nature of fusion proteins requires that the entire open reading frame be transcribed. The protein sequence derived from the entire open reading frame contains 359 amino acids and has a calculated M_r of 39,576 in excellent agreement with the estimates obtained by protein blotting with specific DNA and antibody probes. RNA blot-hybridization analysis revealed that the mRNA hybridizing to the 1.5-kb YB-1 cDNA is also 1.5 kb, indicating that the YB-1 cDNA is full length. The full-length nature of the cDNA was confirmed by sequence analysis of four independent YB-1 clones representing both orientations. Moreover, a translation initiation consensus sequence (ACC ATG) (21) is present beginning at base 124. These results indicate that the YB-1 cDNA has a 5′ untranslated region of 126 bases and an open reading frame of 951 bases and that the methionine encoded by the codon at base position 127 represents the beginning of the actual YB-1 when it is made as an independent (i.e., nonfusion) protein. The protein beginning with this consensus sequence contains 317 amino acids and has a calculated M_r of 35,414.

There are several features noteworthy about the YB-1 sequence. Eighteen percent of the residues of the protein are basic, and the protein is enriched for proline and glycine. Two potential N-linked oligosaccharide sites are present. When compared to sequences contained within GenBank (Beckman Microgenie Version 3.2), 67% homology at the protein level was observed between residues 278 and 292 of YB-1 and between residues 10 and 24 of Caspian sturgeon protamine.

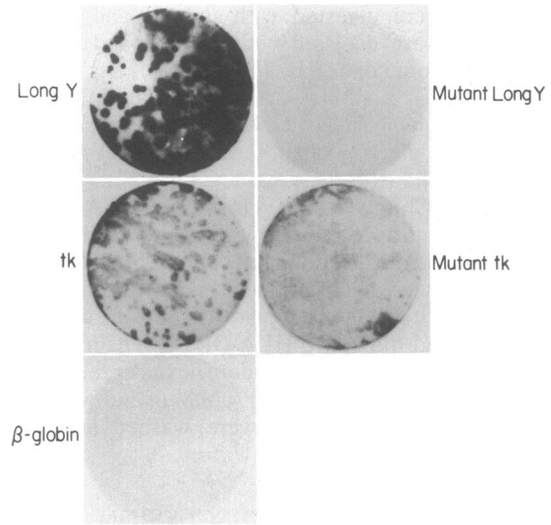

FIG. 2. Specificity of YB-1 for the CCAAT box of the Y box. Replicate filters prepared as in Fig. 1 were hybridized with ^{32}P-labeled ds oligonucleotides corresponding to the long-Y box, to the promoter regions of herpes simplex virus *tk* and rabbit β-globin gene, and to a mutated long-Y box and a mutated *tk* promoter, both with mutations within the CCAAT box (see Table 1). Strong binding was seen with the long-Y box, whereas weak binding was seen with *tk*. No binding was seen with the rabbit β globin gene oligonucleotide. Mutations within the CCAAT box of the long-Y box and *tk* abrogated binding to YB-1.

No "zinc finger" motifs (22–28) were found. Although a consensus sequence for translocation of proteins to the nucleus does not exist, nuclear localization signals that have been characterized typically contain a cluster of three to six basic residues in a short peptide of roughly four to nine amino acids (29). YB-1 contains at least six such sequences: Arg-150

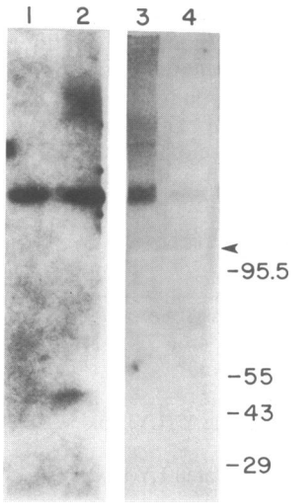

FIG. 3. Molecular weight estimation of the β-galactosidase–YB-1 fusion protein. *E. coli* Y1088 (lanes 1 and 3) or Y1090 (lanes 2 and 4) containing the λgt11-YB-1 clone were grown until partial lysis was achieved and were induced to secrete the fusion protein. After centrifugation, the supernatant containing the fusion protein was electrophoresed through a $NaDodSO_4$/4–20% gradient polyacrylamide gel, and transfer was made to nitrocellulose. The β-galactosidase–YB-1 fusion protein was detected either by protein blotting with ^{32}P-end-labeled class II box as the DNA probe (lanes 1 and 2) or with anti-β-galactosidase as the antibody probe (lanes 3 and 4). Molecular weight markers are shown on the right ($\times 10^{-3}$). The fusion protein had an estimated M_r of ≈150,000 by both detection modalities. The arrow indicates the position to which native β-galactosidase (M_r = 116,000) migrates.

```
YB-1

                                     30                                                        60
CCG GGA GCG GAG AGC GGA CCC CAG AGA GCC CTG AGC AGC CCC ACC GCC GCC GCC GGC CTA
Pro Gly Ala Glu Ser Gly Pro Gln Arg Ala Leu Ser Ser Pro Thr Ala Ala Ala Gly Leu

                                     90                                                       120
GTT ACC ATC ACA CCC CGG GAG GAG CCG CAG CTG CCG CAG CCG GCC CCA GTC ACC ATC ACC
Val Thr Ile Thr Pro Arg Glu Glu Pro Gln Leu Pro Gln Pro Ala Pro Val Thr Ile Thr

                                    150                                                       180
GCA ACC ATG AGC AGC GAG GCC GAG ACC CAG CAG CCG CCC GCC GCC CCC CCC GCC GCC CCC
Ala Thr Met Ser Ser Glu Ala Glu Thr Gln Gln Pro Pro Ala Ala Pro Pro Ala Ala Pro
         1
                                    210                                                       240
GCC CTC AGC GCC GCC GAC ACC AAG CCC GGC ACT ACG GGC AGC GGC GCA GGG AGC GGT GGC
Ala Leu Ser Ala Ala Asp Thr Lys Pro Gly Thr Thr Gly Ser Gly Ala Gly Ser Gly Gly
        21
                                    270                                                       300
CCG GGC GGC CTC ACA TCG GCG GCG CCT GCC GGC GGG GAC AAG AAG GTC ATC GCA ACG AAG
Pro Gly Gly Leu Thr Ser Ala Ala Pro Ala Gly Gly Asp Lys Lys Val Ile Ala Thr Lys
        41
                                    330                                                       360
GTT TTG GGA ACA GTA AAA TGG TTC AAT GTA AGG AAC GGA TAT GGT TTC ATC AAC AGG AAT
Val Leu Gly Thr Val Lys Trp Phe Asn Val Arg Asn Gly Tyr Gly Phe Ile Asn Arg Asn
        61
                                    390                                                       420
GAC ACC AAG GAA GAT GTA TTT GTA CAC CAG ACT GCC ATA AAG AAG AAT AAC CCC AGG AAG
Asp Thr Lys Glu Asp Val Phe Val His Gln Thr Ala Ile Lys Lys Asn Asn Pro Arg Lys
        81
                                    450                                                       480
TAC CTT CGC AGT GTA GGA GAT GGA GAG ACT GTG GAG TTT GAT GTT GTT GAA GGA GAA AAG
Tyr Leu Arg Ser Val Gly Asp Gly Glu Thr Val Glu Phe Asp Val Val Glu Gly Glu Lys
       101
                                    510                                                       540
GGT GAG GAG GCA GCA AAT GTT ACA GGT CCT GGT GGT GTT CCA GTT CAA GGC AGT AAA TAT
Gly Glu Glu Ala Ala Asn Val Thr Gly Pro Gly Gly Val Pro Val Gln Gly Ser Lys Tyr
       121
                                    570                                                       600
GCA GCA GAC CGT AAC CAT TAT AGA CGC TAT CCA CGT CGT AGG GGT CCT CCA CGC AAT TAC
Ala Ala Asp Arg Asn His Tyr Arg Arg Tyr Pro Arg Arg Arg Gly Pro Pro Arg Asn Tyr
       141
                                    630                                                       660
CAG CAA AAT TAC CAG AAT AGT GAG AGT GGG GAA AAG AAC GAG GGA TCG GAG AGT GCT CCC
Gln Gln Asn Tyr Gln Asn Ser Glu Ser Gly Glu Lys Asn Glu Gly Ser Glu Ser Ala Pro
       161
                                    690                                                       720
GAA GGC CAG GCC CAA CAA CGC CGG CCC TAC CGC AGG CGA AGG TTC CCA CCT TAC TAC ATG
Glu Gly Gln Ala Gln Gln Arg Arg Pro Tyr Arg Arg Arg Arg Phe Pro Pro Tyr Tyr Met
       181
                                    750                                                       780
CGG AGA CCC TAT GGG CGT CGA CCA CAG TAT TCC AAC CCT CCT GTG CAG GGA GAA GTG ATG
Arg Arg Pro Tyr Gly Arg Arg Pro Gln Tyr Ser Asn Pro Pro Val Gln Gly Glu Val Met
       201
                                    810                                                       840
GAG GGT GCT GAC AAC CAG GGT GCA GGA GAA CAA GGT AGA CCA GTG AGG CAG AAT ATG TAT
Glu Gly Ala Asp Asn Gln Gly Ala Gly Glu Gln Gly Arg Pro Val Arg Gln Asn Met Tyr
       221
                                    870                                                       900
CGG GGA TAT AGA CCA CGA TTC CGC AGG GGC CCT CCT CGC CAA AGA CAG CCT AGA GAG GAC
Arg Gly Tyr Arg Pro Arg Phe Arg Arg Gly Pro Pro Arg Gln Arg Gln Pro Arg Glu Asp
       241
                                    930                                                       960
GGC AAT GAA GAA GAT AAA GAA AAT CAA GGA GAT GAG ACC CAA GGT CAG CAG CCA CCT CAA
Gly Asn Glu Glu Asp Lys Glu Asn Gln Gly Asp Glu Thr Gln Gly Gln Gln Pro Pro Gln
       261
                                    990                                                      1020
CGT CGG TAC CGC CGC AAC TTC AAT TAC CGA CGC AGA CGC CCA GAA AAC CCT AAA CCA CAA
Arg Arg Tyr Arg Arg Asn Phe Asn Tyr Arg Arg Arg Arg Pro Glu Asn Pro Lys Pro Gln
       281
                                   1050                                                      1080
GAT GGC AAA GAG ACA AAA GCA GCC GAT CCA CCA GCT GAG AAT TCC CGC TCC CGA GGC TGA
Asp Gly Lys Glu Thr Lys Ala Ala Asp Pro Pro Ala Glu Asn Ser Arg Ser Arg Gly End
       301                                                                     317
                                   1110                                                      1140
GCA GGG CGG GGC TGA GTA AAT GCC GGC TTA CCA TCT CTA CCA TCA TCC GGT TTA GTC ATC

                                   1170                                                      1200
CAA CAA GAA GAA ATA TGA AAT TCC AGC AAT AAG AAA TGA ACA AAA GAT TGG AGC TGA AGA

                                   1230                                                      1260
CCT AAA GTA CTT GCT TTT TGC CGT TTG CAA CCA GAT AAA TAG AAC TAT CTG CAT TAT CTA

                                   1290                                                      1320
TGC AGC ATG GGG TTT ATA TTT TAC TAA GAC GCT CTT TGG TAT ACA ACG GTT TTA AAA GCC

                                   1350                                                      1380
TGG TTT TCT CAA TAC GCC TTA AAG GTT TTA AAT TGT TTC ATA TCT GGT CAA GTT GAG ATT

                                   1410                                                      1440
TTT AAG AAC TTC ATT TTT AAT TTG TAA TAA AAG TTT ACA ACT TGA TTT TTT CAA AAA AGT

                                   1470                                                      1500
CAA CAA ACT GCA AGC ACC TGT TAA TAA AGG TCT TAA ATA ATA AAA AAA AAA AAA AAA AAA

AAA AAA AAA AAA
```

FIG. 4. The nucleotide and derived amino acid sequence of YB-1 cDNA. The cDNA is 1512 bases in length and has an open reading frame of 1077 bases and a 3′ untranslated region of 435 bases. The most 5′ 126 bases represent the 5′ untranslated region, but because YB-1 is made as a fusion protein, this portion of the mRNA is translated; therefore, the derived 41 amino acid sequence is given. When made as an independent protein, the protein begins with the methionine, which is at position 1. The translation initiation site is circled, two potential N-linked glycosylation sites are boxed, and six putative nuclear localization signals are underlined.

to Arg-152; Arg-185 to Arg-192; Arg-199 to Arg-205; Arg-242 to Arg-247; Arg-279 to Arg-283; and Arg-288 to Arg-291.

A possible regulatory function of YB-1 was elucidated by correlating levels of YB-1 mRNA with levels of class II expression. YB-1 mRNA and HLA-DR β chain mRNA levels were examined by slot-blot analysis in a highly positive class II-expressing lymphoblastoid B-cell line, Swei; a class II-negative T-cell line, CCRF/CEM; a class II-positive T-cell line, HUT 78; a class II-negative monocytoid line, U937; and γ interferon- and phorbol 12-myristate 13-acetate (PMA)-stimulated U937, which is class II-positive (Fig. 5). YB-1 mRNA is found in both class II-negative and class II-positive cells, and there appears to be an inverse correlation with YB-1 mRNA levels and levels of HLA-DR β chain mRNA. Thus, for example, the lowest levels of YB-1 mRNA are observed concurrently with the highest levels of HLA-DR β chain mRNA in Swei. At the other extreme, the highest YB-1 mRNA levels are seen when no HLA-DR β chain mRNA is detectable in the class II-negative CCRF/CEM line. Phorbol 12-myristate 12-acetate and γ interferon stimulation of U937 cells results in a decrease in YB-1 mRNA levels and an increase in HLA-DR β chain mRNA to detectable levels. These results have been confirmed by RNA blot-hybridization analysis. This inverse relationship of YB-1 and class II mRNAs suggests that YB-1 is a negative regulatory factor.

Dorn *et al.* (11) and Van Huijsduignen *et al.* (10) have characterized a factor from nuclear extracts of a murine B-cell lymphoma line that has specificity for the Y box; this factor they have termed nuclear factor Y or NF-Y. This factor binds to the inverted CCAAT box within the Y box and shows weaker binding to CCAAT sequences in other promoters, particularly the herpes simplex virus *tk* promoter. It is distinct from other CCAAT binding factors such as the CCAAT binding protein (CBP) (32) and the CCAAT binding transcription factor/nuclear factor 1 (CTF/NF-1) (33). Therefore, it is instructive to compare YB-1 and NF-Y. Molecular weight estimates, based on glycerol gradient centrifugation, suggest a M_r of 50,000–70,000 for NF-Y compared with a M_r of 35,414 for YB-1. These molecular weights were derived by different methods; thus, these differences may not be real. The binding of both YB-1 and NF-Y to the Y box is resistant to at least 69°C, and NF-Y binding is resistant to protease. NF-Y binding to the Y box is sensitive to *o*-phenanthroline and is totally lost at 8 mM *o*-phenanthroline, suggesting that NF-Y may be a metalloprotein. In contrast, YB-1 binding to the Y box is totally resistant to even 10 mM *o*-phenanthroline (data not shown). The binding of both YB-1 and NF-Y to the Y box is resistant to EDTA and high salt, and at least NF-Y is also resistant to

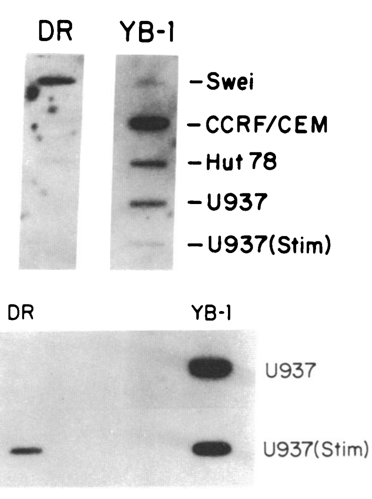

FIG. 5. (*Upper*) Slot blots of YB-1 mRNA and HLA-DR β chain mRNA. mRNA (3 μg) from a class II-positive B-cell line, Swei; from a class II-negative T-cell line, CCRF/CEM; from a class II-positive T-cell line, HUT 78; from a class II-negative monocytoid line, U937; and from U937 stimulated with γ interferon and phorbol 12-myristate 13-acetate to induce class II expression were assayed for levels of HLA-DR β chain mRNA and YB-1 mRNA by hybridization with cDNA probes. The two mRNAs appeared to correlate inversely (see text). (*Lower*) Repeat of the experiment on U937 and stimulated U937 cells using 10 μg of mRNA.

urea. YB-1 mRNA is found in both class II-positive and -negative cells, and YB-1 mRNA levels appear to inversely correlate with *HLA-DR* expression. Levels of NF-Y mRNA have not been quantitated, but NF-Y is also found in both class II-negative and -positive cells. The data at this point, though certainly not conclusive, suggest that YB-1 and NF-Y are most likely different proteins.

Chodosh *et al.* (30) have characterized three distinct CCAAT binding proteins, CP1, CP2, and NF-1, present in HeLa cell extracts and found that each one appears to be composed of a heterodimer. If all, or most, CCAAT-binding proteins are indeed heterodimers, it is conceivable that the nuclear factor NF-Y may also be a heterodimer and that YB-1 may be one of its subunits.

The authors thank Mrs. Lorraine Whiteley and Mrs. Pat Parvin for their expert secretarial assistance. These investigations were supported in part by National Institutes of Health Grant AI-18925.

1. Sullivan, K. E., Calman, A. F., Nakanishi, M., Tsang, S. Y., Wang, Y. & Peterlin, B. M. (1987) *Immunol. Today* **8,** 289–293.
2. Mathis, D. J., Benoist, C. D., Williams, Y. E., Kanter, M. R. & McDevitt, H. O. (1983) *Cell* **32,** 745–754.
3. Kelly, A. & Trowsdale, J. (1985) *Nucleic Acids Res.* **13,** 1607–1621.
4. O'Sullivan, D. M., Larhammar, D., Wilson, M. C., Peterson, P. A. & Quaranta, V. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 4484–4488.
5. Dorn, A., Durand, B., Marfing, C., Lemeur, M., Benoist, C. & Mathis, D. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 6249–6253.
6. Boss, J. M. & Strominger, J. L. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 9139–9143.
7. Sherman, P. A., Basta, P. V. & Ting, J. P.-Y. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 4254–4258.
8. Miwa, K., Doyle, C. & Strominger, J. L. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 4939–4943.
9. Miwa, K. & Strominger, J. L. (1987) *Nucleic Acids Res.* **15,** 8057–8067.
10. Van Huijsduijnen, R. A. M. H., Bollekens, J., Dorn, A., Benoist, C. & Mathis, D. (1987) *Nucleic Acids Res.* **15,** 7265–7282.
11. Dorn, A., Bollekens, J., Staub, A., Benoist, C. & Mathis, D. (1987) *Cell* **50,** 863–872.
12. Schiffenbauer, J., Didier, D. K., Klearman, M., Rice, K., Shuman, S., Tieber, V. L., Kittlesen, D. J. & Schwartz, B. D. (1987) *J. Immunol.* **139,** 228–233.
13. Towbin, H., Staehelin, T. & Gordon, J. (1979) *Proc. Natl. Acad. Sci. USA* **76,** 4350–4354.
14. Bowen, B., Steinberg, J., Laemmli, U. K. & Weintraub, H. (1980) *Nucleic Acids Res.* **8,** 1–20.
15. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
16. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) *Biochemistry* **24,** 5294–5299.
17. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY), pp. 324–328.
18. Tieber, V. L., Abruzzini, L. F., Didier, D. K., Schwartz, B. D. & Rotwein, P. (1986) *J. Biol. Chem.* **261,** 2738–2742.
19. Graves, B. J., Johnson, P. F. & McKnight, S. L. (1986) *Cell* **44,** 565–576.
20. Dierks, P., Van Ooyen, A., Cochran, M. D., Dobkin, C., Reiser, J. & Weissman, C. (1983) *Cell* **32,** 695–706.
21. Kozak, M. (1986) *Cell* **44,** 283–292.
22. Kadonaga, J. T., Carner, K. R., Masiarz, F. R. & Tijan, R. (1987) *Cell* **51,** 1079–1090.
23. Miller, J., McLachlan, A. D. & Klug, A. (1985) *EMBO J.* **4,** 1609–1614.
24. Page, D. C., Mosher, R., Simpson, E. M., Fisher, E. M. C., Mardon, G., Pollack, J., McGillvray, B., de La Chapelle, A. & Brown, L. G. (1987) *Cell* **51,** 1091–1104.
25. Hollenberg, S. M., Weinberger, C., Ong, E. S., Cerelli, G., Oso, A., Lebo, R., Thompson, E. B., Rosenfeld, M. G. & Evans, R. M. (1985) *Nature (London)* **318,** 635–641.
26. Green, S., Walter, P., Kumar, V., Krust, A., Bornert, J. M., Argos, P. & Chanton, P. (1986) *Nature (London)* **320,** 134–139.
27. Arriza, J. L., Weinberger, C., Cerelli, G., Glaser, T. M., Handelin, B. L., Houseman, D. E. & Evans, R. M. (1987) *Science* **237,** 268–275.
28. Petkovich, M., Brand, N. J., Krust, A. & Chambon, P. (1987) *Nature (London)* **330,** 444–450.
29. Dingwall, C. & Laskey, R. A. (1986) *Annu. Rev. Cell. Biol.* **2,** 367–390.
30. Chodosh, L. A., Baldwin, A. S., Carthew, R. W. & Sharp, P. A. (1988) *Cell* **53,** 11–24.
31. Feinberg, A. P. & Vogelstein, B. (1983) *Anal. Biochem.* **132,** 6–13.
32. Graves, B. J., Johnson, P. F. & McKnight, S. L. (1986) *Cell* **44,** 565–576.
33. Jones, K. A., Yamamoto, K. R. & Tijan, R. (1985) *Cell* **42,** 559–572.