Functional Effects of a Natural Polymorphism in the Transcriptional Regulatory Sequence of *HLA-DQB1*

JON S. BEATY,* KENNETH A. WEST,† AND GERALD T. NEPOM

Virginia Mason Research Center, Seattle, Washington 98101, and Department of Immunology, University of Washington School of Medicine, Seattle, Washington 98195

Received 7 June 1995/Accepted 9 June 1995

DNA sequence polymorphism in the genes encoding HLA class II proteins accounts for allelic diversity in antigen recognition and presentation and, thus, in the role of these cell surface glycoproteins as determinants of the scope of the T-cell repertoire. In addition, sequence polymorphism in the promoter-proximal transcriptional regulatory regions of these genes has been described, particularly for the *HLA-DQB1* locus, where these differences may contribute to variation in locus- and allele-specific expression. In this study, we measured the effect of such regulatory sequence polymorphism on the expression of endogenous alleles of *DQB1* in heterozygous cells. Quantitative reverse transcriptase-mediated PCR analysis showed that expression of the *DQB1**0301 allele responded more rapidly to gamma interferon induction than that of *DQB1**0302. We have analyzed functional effects of a prominent allelic polymorphism that consists of a TG dinucleotide present between the W and X₁ consensus elements in the *DQB1**0302 allele but missing in the *DQB1**0301 allele. The dominant effect of this polymorphism was to introduce a variation in the spacing between the W and X₁ elements of these two alleles. A secondary compensatory effect was specific for the TG dinucleotide itself, which was essential for the binding of a nuclear protein complex to the *0302 regulatory region immediately 5' of the X₁ element. Derivatives of the *DQB1* 5' regulatory region were used to drive expression of the chloramphenicol acetyltransferase gene in transient transfections of human B-lymphoblastoid and gamma interferon-treated melanoma cell lines, demonstrating that the additional spacing between the W and X₁ elements caused by the presence of the TG dinucleotide in the *0302 allele resulted in reduced expression compared with that driven by the *0301 fragment; this difference overshadowed an up-regulating effect on expression which corresponded to the binding of the TG-dependent nuclear protein complex. The presence of this polymorphism in multiple *HLA-DQB1* alleles and in several species suggests selection for two alternative transcriptional regulatory mechanisms influencing expression of alleles of the same HLA locus.

Major histocompatibility complex class II genes encode cell membrane-associated heterodimeric glycoproteins which are highly polymorphic (24). Such polymorphism results in considerable flexibility for the recognition and binding of a broad variety of antigenic peptides (4), for the shaping of the T-cell repertoire in the thymus by positive and negative selection (16, 17), as well as for interallelic pairing of the α and β subunits to form mixed allotype heterodimers on the cell surface (18). Polymorphism in alleles of the *HLA-DQB1* locus, which encodes the β subunit of the HLA-DQ molecule, occurs both in the exons and in the 5' transcriptional control region (1). Similar polymorphism has been reported in the 5' regulatory regions of alleles of *HLA-DQA1*, which encodes the α subunit of the DQ heterodimer (28), and for *HLA-DRB* loci (22).

Within the 200 bp immediately 5' of the first exon are three sequence elements, referred to as the W (or Z), X, and Y boxes, which are found in the same relative position in all class II HLA genes. DNA-binding proteins which interact with these elements have been described, and several which influence expression of the *HLA-DRA* locus have been cloned and characterized (10, 35). The X box consists of two control elements, X₁ (gamma interferon [IFN- γ] responsiveness) and X₂, which in *DRA* contains sequences similar to those which confer cyclic

AMP and tetradecanoyl phorbol acetate responsiveness (20). RFX and TRAX1, nuclear proteins required for transcription of *DRA*, bind to the X₁ element of that gene (8, 14). The X₂ element of *DRA* is known to interact with XBP1, XBP2 and c-Jun; at least one of these factors (XBP2) acts cooperatively with RFX in order for binding to occur (32). The Y box contains an inverted CCAAT element, the binding to which by the activating protein, NF-Y, facilitates the binding of X-box-specific proteins (49); it is also bound by YB-1, which can repress IFN- γ induction of *DRA* (44). A protein required for expression of all HLA class II genes is CIITA, which probably functions by interacting with other HLA class II regulatory proteins rather than with DNA directly (40, 41).

Expression of *HLA-DQB1* is regulated largely at the transcriptional level by a similar set of sequence elements (3, 6, 10). In the *DQB1* locus, the W box contains the J element (IFN- γ responsiveness) and the 7-bp S element (constitutive and induced expression) (42). The *DQB1* Y element is identical to that in *DRA*, whereas the X box contains a number of locus-specific characteristics (43, 46). Additionally, in the region immediately 5' of the X₁ element, alleles of the *DQB1* locus contain sequences of two distinct types, differing by the presence or absence of the specific dinucleotide, TG. In previous studies, equivalent fragments of the 5' regions of two alleles which differ at this site, *DQB1**0301 and *DQB1**0302, were shown to drive expression of the chloramphenicol acetyltransferase (CAT) reporter gene with different efficiencies in transient transfections of human B-lymphoblastoid cells (1). This finding suggested that the polymorphism consisting of two nucleoside residues that occur in the *0302 allele at positions

* Corresponding author. Mailing address: Virginia Mason Research Center, 1000 Seneca St., Seattle, WA 98101. Phone: (206) 223-8813. Fax: (206) 223-7543. Electronic mail address: jbeaty@vmmc-im.vmmc.org.

† Present address: Department of Medicine, Dalhousie University, Halifax, Nova Scotia B3H 1V7, Canada.

FIG. 1. DNA sequence of 641 bp of the promoter and transcriptional regulatory regions of *DQB1**0301 and *DQB1**0302. Endonuclease cleavage sites are indicated, as are the W (J, S), X_1 , X_2 , and Y consensus elements. Polymorphism between the two alleles is signified by asterisks. Base numbering is indicated; the -1 base is that immediately 5' of the ATG start codon of the first exon. The *0302 sequence from positions -641 to -624 is conjectural, since a *0301-specific 5' PCR primer was used to generate the *0302 fragment used for cloning and subsequent sequencing.

-180 and -179, but which are lacking in *0301, may have functional consequences.

The functional significance of such transcriptional regulatory sequence polymorphism has been confirmed by results of steady-state *DQB1* mRNA analysis that we report here. Quantitative reverse transcriptase-mediated PCR (RT-PCR) of the allele-specific mRNAs derived from endogenous *0301 and *0302 in heterozygous primary skin fibroblasts showed that IFN- γ -induced transcription of *0301 occurred significantly sooner and more profoundly than did *0302 expression.

Also, we have investigated in more detail the effect of the -180/-179 polymorphism on the abilities of allelic promoter elements to mediate transcription in transient transfections of both B-lymphoblastoid and IFN- γ -treated melanoma cell lines. Here we describe two mechanisms by which this polymorphism affected regulation of transcription: by changing the spacing between the W and X_1 boxes and by changing the pattern of regulatory protein binding near the X_1 element. The increased W- X_1 spacing in *0302 due to the presence of the TG dinucleotide at positions -180 and -179 was the dominant effect, and it correlated with reduced expression compared with that of the *0301 allele. The presence of TG in *0302 is also associated with the binding of a novel nuclear protein or protein complex that is missing from the array of DNA-binding proteins interacting with the *0301 allele. Although the TG-associated protein complex was a strong positive effector of transcription, the effect of its binding to the *0302 regulatory region was superseded by the transcriptional down-regulating effect of the increased W- X_1 spacing. We suggest that tran-

scription is regulated differently between these two alleles of *HLA-DQB1*.

MATERIALS AND METHODS

Cell culture. BSM and PF97387 are Epstein-Barr virus-transformed human B-lymphoblastoid cell lines. They were grown in RPMI 1640 supplemented with 50 mM *N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid (HEPES), 10% fetal bovine serum, 1 mM sodium pyruvate, 0.2 mM L-glutamine, 50 U of penicillin G per ml, and 50 μ g of streptomycin sulfate per ml. Me9229/18 is a human melanoma cell line in which *DQB1* is inducible with IFN- γ (2); generously provided by J. L. Strominger). It was grown in Dulbecco's medium supplemented with 10% fetal bovine serum, 0.35% glucose, 1 mM sodium pyruvate, 0.2 mM L-glutamine, 50 U of penicillin G per ml, and 50 μ g of streptomycin sulfate per ml. Primary human skin fibroblasts were outgrown from a 6-mm-diameter skin punch biopsy. The fresh biopsy specimen was washed in sterile phosphate-buffered saline (PBS) containing 50 U of penicillin G and 50 μ g of streptomycin sulfate per ml and cut into 1- to 2-mm sections. These skin sections were then planted on the premoistened sides of 25-cm² culture flasks (four to five per flask) and incubated for several weeks with 1.0 ml of Waymouth medium supplemented with 20% fetal bovine serum, 0.2 mM L-glutamine, 50 U of penicillin G per ml, and 50 μ g of streptomycin sulfate per ml until the outgrowth of fibroblasts was confluent in the flasks. Subsequent passages of fibroblasts were grown in the same medium. All cell cultures were incubated at 37°C in a mixture of 5% CO₂ and 95% air.

Construction of *DQB1* regulatory region-CAT plasmids. Cloning of *DQB1**0301 and *DQB1**0302 5' regulatory regions has been described previously (1). The 641-bp 5' regulatory region fragments shown in Fig. 1 were generated by PCR (34) and ligated to *Kpn*I- and *Sma*I-treated pTCAT (Richard Harland, University of California, Berkeley). The resulting plasmids, pQ3.1P19CAT and pQ3.2P04CAT, were used as PCR templates to prepare the shorter wild-type and mutant 5' region fragments (positions -1 to -303) present in the plasmid DNAs diagrammed in Fig. 7. This PCR product DNA was treated with *Kpn*I and *Ppu*MI and ligated to either pQ3.1P19CAT or pQ3.2P04CAT which had been treated with the same endonucleases. These plasmids were then treated with *Kpn*I and

*Bsu*361, followed by Klenow DNA polymerase; the vector-containing fragment was purified by agarose gel electrophoresis and recircularized by treatment with T4 DNA ligase.

Oligonucleotide PCR primers (prepared with a model 391 DNA synthesizer; Applied Biosystems, Inc., Foster City, Calif.) included the following: the 5' primer, containing a *Kpn*I site, 5'-CACGGTACCTGAAACATGATCCA CATT-3'; the 3' primer for the 641-bp fragments, containing an *Hpa*I site, 5'-GTGGGTAAACAATTGAGACGAAG-3'; and several additional 3' primers for both alleles which extended from base positions -192 to -150 and spanned the *Pvu*MI cleavage site. These 3' primers contained various sequence changes either in the A tract or at positions -180 and -179 to introduce specific mutations in the PCR product DNA used to construct the 303-bp insert plasmids.

The cloned DNA inserts of all mutant plasmids were sequenced by the dideoxynucleotide chain termination method after alkaline denaturation of the plasmid template (12). Plasmid DNA for transient-transfection experiments was prepared by the alkaline lysis method and two consecutive isopycnic ultracentrifugation treatments in CsCl and ethidium bromide (23). The fidelity of each plasmid preparation was confirmed by DNA sequencing.

Transient transfections. Test CAT plasmid DNA was used to transfect B-lymphoblastoid cells by the DEAE-dextran technique (21) or melanoma cells by the calcium phosphate method (48). As a control for transfection efficiency, pEQ176, a plasmid bearing the cytomegalovirus immediate-early promoter-driven β -galactosidase gene (Adam Geballe, Fred Hutchinson Cancer Research Center, Seattle, Wash.), was used to cotransfect in each sample. Plasmid DNA concentrations were measured by fluorometry before each experiment. Actively growing B-lymphoblastoid cells (2×10^7) were treated with 200 μ g of DEAE-dextran per ml for 30 min in the presence of 10 μ g of CAT plasmid DNA and 1 μ g of pEQ176 DNA. The cells were then mixed with 0.1 volume of dimethyl sulfoxide for 3 min. After extensive washing, cells were suspended in 10 ml of growth medium. Melanoma cells (5×10^6 in a 100-mm-diameter culture dish inoculated 18 to 24 h earlier) were treated with calcium phosphate, 15 μ g of CAT plasmid DNA, and 1 μ g of pEQ176 DNA in 10 ml of medium. After 18 h, the cultures were washed thoroughly and 10 ml of fresh medium was added. When appropriate, IFN- γ (human recombinant IFN- γ ; R & D Systems, Minneapolis, Minn.) was added to the medium 24 h prior to harvesting at a concentration of 100 U/ml unless otherwise specified. Forty-eight hours after transfection, cells were lysed by three cycles of freeze-thawing, and aliquots of the lysates were used in both nonchromatographic CAT and β -galactosidase assays (27, 38). Results are expressed as index values calculated by dividing net CAT values by net β -galactosidase values for each sample (net values were calculated by subtracting from experimental CAT and β -galactosidase values those derived from the appropriate sham-transfected controls). Replicate samples from each experiment were averaged; independent transfection experiments were done with at least two independently purified plasmid DNA preparations for all experiments.

Quantitative RT-PCR. Primary human skin fibroblasts (2.5×10^6 passage 3 cells) were incubated for 24 h with 10 ml of growth medium in 100-mm-diameter culture dishes. IFN- γ (Genzyme Diagnostics, Cambridge, Mass.) was then added to a final concentration of 500 U/ml. After appropriate incubation periods, the cells were washed with 10 ml of PBS, harvested by scraping, and suspended in 10 ml of PBS. After centrifugation at $300 \times g$ for 5 min, the cell pellet was lysed in 0.4 ml of a guanidinium thiocyanate- and *N*-lauroyl sarcosine-containing extraction buffer from a QuickPrep Micro mRNA purification kit (Pharmacia Inc., Piscataway, N.J.). Poly(A) RNA was purified from each sample by using oligo(dT)-cellulose and other reagents supplied with this kit and stored in ethanol at -80°C .

mRNA was recovered by centrifugation from 200 μ l ($\sim 3 \times 10^5$ cell equivalents) of the ethanol-mRNA suspension, and cDNA was prepared by using reagents from a SuperScript II reverse transcriptase kit (Life Technologies, Gaithersburg, Md.). After treatment of the samples with 2 U of RNase H, 5 μ l of cDNA ($\sim 7 \times 10^4$ cell equivalents) was used as a template for PCR amplification of a 301-bp DNA fragment representing bases 825 to 1125 of glyceraldehyde-3-phosphate dehydrogenase (GAPDH) cDNA (5' primer [^{32}P end labeled], TGATGACATCAAGAAGGTGGT; 3' primer, CAGTGAGGGTCTCTCTCT TCC [corresponding to sites in exons 7 and 8, respectively]). PCR mixtures containing 50 mM Tris-HCl (pH 9.0), 1.5 mM MgCl_2 , 20 mM $(\text{NH}_4)_2\text{SO}_4$, 0.2 mM each dATP, dCTP, dGTP, and dTTP, and 25 pmol of each oligonucleotide primer, in a total volume of 99 μ l, were mixed in 0.5-ml tubes, overlaid with 100 μ l of mineral oil, and placed in a thermocycler for incubation. After 2 min at 94°C , 1 μ l (3 U) of Hot Tub DNA polymerase (Amersham Corp., Arlington Heights, Ill.) was added to each sample. Samples were then incubated for 30 cycles consisting of 1 min at each of the following temperatures: 94, 60, and 72°C . Incubation was ended after the samples were held at 72°C for 5 min. The PCR products from each sample were electrophoresed in a 6% nondenaturing acrylamide gel; after drying, GAPDH bands were quantified with a model GS-363 Molecular Imager (Bio-Rad Laboratories, Hercules, Calif.), and the respective values were used to adjust the volume of each cDNA sample to a normalized level for use in the *DQB1* quantitative PCR. Since GAPDH expression is unaffected by cell culture conditions and treatment with IFN- γ (7), the intensity variation of the 301-bp GAPDH bands reflects variation in the initial amounts of mRNA present in each sample and the efficiencies of the cDNA synthesis reactions.

The *DQB1* quantitative PCR mixtures, set up and incubated exactly as described above, contained the GAPDH-specific primer set and, in addition, a pair of oligonucleotide primers designed to amplify a 270-bp region of the third exon of *HLA-DQB1* (5' primer, GAGCCACAGTGACCATC; 3' primer [^{32}P labeled], TGATGGGGTCTCTGGAGG). These primers were designed to anneal to identical sites in the *0301 and *0302 alleles; sequence polymorphism within the amplified 270-bp region creates an allele-specific cleavage pattern when the PCR fragments are treated with *Hpa*II endonuclease, such that the *0301 and *0302 alleles generate 134- and 166-bp ^{32}P -labeled DNA fragments, respectively (the 301-bp GAPDH PCR-amplified fragment does not contain an *Hpa*II site, nor does the relevant intron). The *DQB1* quantitative PCR products were treated once with an equal volume of phenol-chloroform (1:1, [vol/vol]) and precipitated with ethanol. The DNA pellets were resuspended in 20 μ l of buffer containing 10 mM Tris-HCl (pH 7.4), 10 mM NaCl, and 0.1 mM EDTA. Three microliters of this DNA (representing $\sim 10^4$ cell equivalents) was treated with 9 U of *Hpa*II endonuclease in a volume of 15 μ l. A 6% nondenaturing polyacrylamide gel was loaded with 5 μ l of 1/30 dilutions of the *Hpa*II-treated samples (representing ~ 100 cell equivalents) and electrophoresed at 150 V (constant) for 2.5 h. The gel was dried and analyzed as described above.

As controls for genomic DNA contamination of the mRNA preparations, duplicate samples were treated exactly as described above except that reverse transcriptase was omitted from the cDNA synthesis reactions. In addition, a blank sample (containing no mRNA) was treated as described above to test for contamination of any of the reagents with PCR-amplifiable nucleic acid.

Nuclear protein extracts. Nuclear protein extracts were prepared by using a modified Shapiro-Sharp protocol (37). Actively growing cells (8×10^6) were harvested, suspended in hypotonic buffer (10 mM HEPES [pH 7.9], 0.75 mM spermidine, 0.15 mM spermine, 0.1 mM EDTA, 0.1 mM EGTA, 10 mM KCl, 1 mM dithiothreitol) and lysed in a Dounce homogenizer. Nuclei were recovered by centrifugation, and 0.1 volume of saturated $(\text{NH}_4)_2\text{SO}_4$ solution (~ 4 M at 4°C) was added. After centrifugation for 90 min at 40,000 rpm in a Beckman 70.1Ti rotor (4°C), the nuclear proteins were precipitated from the supernatant fraction by the addition of 0.3 g of solid $(\text{NH}_4)_2\text{SO}_4$ per ml. Phenylmethylsulfonyl fluoride (0.5 mM), 5 μ g of leupeptin per ml, and 1 mM dithiothreitol were present in all freshly prepared buffers. Concentrations of proteins in the extracts were determined by the Bio-Rad protein assay. Aliquotted samples of the nuclear extracts were stored in liquid nitrogen.

Electrophoretic mobility shift assays (EMSAs). Double-stranded DNA probes (Table 1) were prepared by annealing complementary oligonucleotides. Probes were radioactively labeled by filling in a 5' three-base single-stranded tail with Klenow DNA polymerase and the appropriate deoxynucleoside triphosphates, one of which was labeled at the α position with ^{32}P . To reduce nonspecific protein binding in the assay, 1.0 to 1.5 μ g of nuclear protein extract was preincubated with 0.05 to 1.0 μ g of poly(dI-dC)·poly(dI-dC) in 25 μ l of buffer containing 12 mM HEPES (pH 7.9), 60 mM KCl, 12% glycerol, 0.12 mM EDTA, and 0.3 mM dithiothreitol; incubation was for 15 min at ambient temperature. Labeled DNA probe (10^4 cpm, 20 to 40 pg of DNA) was then added, and incubation was continued for 15 min at ambient temperature (13). Protein-bound probe and free probe were separated by electrophoresis on a 4% polyacrylamide gel in 0.25 $\times$ TBE ($1 \times$ TBE is 89 mM Tris, 89 mM boric acid, and 2 mM EDTA); the gel was then dried and subjected to autoradiography. The specificities of protein-DNA complexes, seen as bands on the autoradiograms, were analyzed by competition experiments, in which unlabeled double-stranded oligonucleotide of a specific sequence (5- to 500-fold molar excess to probe) was preincubated with poly(dI-dC) and nuclear extract before addition of labeled DNA probe.

In vivo footprinting. BSM cells (2×10^7) were harvested and resuspended in 2 ml of medium containing 0.1% dimethyl sulfate. After 2 min of incubation at 37°C , the reaction was stopped by mixing with 50 ml of ice-cold PBS containing 1% 2-mercaptoethanol. The cells were centrifuged at $200 \times g$ for 5 min and washed three times in 50 ml of PBS. In vitro methylation of protein-free control DNA was done as described by Maxam and Gilbert (26). In vivo-methylated and control DNAs were purified by suspending 2×10^7 cells in 0.3 ml of PBS, adding 3 ml of a buffer containing 100 mM Tris-HCl (pH 7.9), 200 mM NaCl, 4 M urea, 0.5% *N*-lauroyl sarcosine, 10 mM 1,2-cyclohexanediamine tetraacetic acid, and 0.2 mg proteinase K per ml, and incubating the cells for 4 h at 37°C . The samples were extracted twice with phenol, twice with phenol-chloroform, and once with chloroform. DNA was precipitated in ethanol and resuspended in 10 mM Tris-HCl (pH 7.8) at a final concentration of 1 μ g/ μ l.

Guanine-specific piperidine cleavage of methylated DNA was performed as described by Maxam and Gilbert (26). Piperidine was removed by lyophilizing the samples three times in a Speed-Vac concentrator. DNA cleavage products were precipitated in ethanol and resuspended at a concentration of 1 μ g/ml in 10 mM Tris-HCl (pH 8.0).

Ligation-mediated PCR was performed as described by Mueller and Wold (29). Primer extension was carried out with 5 μ g of piperidine-cleaved DNA by using Sequenase version 2.0 T7 DNA polymerase (United States Biochemical Corporation). The ligation products were amplified for 17 cycles and then labeled in a two-cycle reaction using *Taq* DNA polymerase (Perkin-Elmer). After the addition of *Taq* stop buffer, the labeled products were extracted with phenol-chloroform, precipitated with ethanol, and resuspended in 20 μ l of formamide-dye loading buffer. Two to 5 μ l of sample was electrophoresed on 6% denaturing

TABLE 1. Oligonucleotide probes used for EMSAs

Description	Sequence ^a
*0301 WT XL.....	5' ACAAAAAAAAAA--TCTGCCAGAGACAGATGAGGTC 3' TGTTTTTTTTTT--AGACGGGTCTCTGTCTACTCCAG
*0302 WT XL.....	5' ACAAAAAAAAAATGTCTGCCTAGAGACAGATTAGGTC 3' TGTTTTTTTTTTACAGACGGATCTCTGTCTAATCCAG
*0301 + TG (–180/–179) XL.....	5' ACAAAAAAAAAATGTCTGCCAGAGACAGATGAGGTC 3' TGTTTTTTTTTTACAGACGGGTCTCTGTCTACTCCAG
*0302 – TG (–180/–179) XL.....	5' ACAAAAAAAAAA--TCTGCCAGAGACAGATGAGGTC 3' TGTTTTTTTTTT--AGACGGGTCTCTGTCTACTCCAG
*0302 CC to AA (–174/–173) XL.....	5' ACAAAAAAAAAATGTCTGAATAGAGACAGATTAGGTC 3' TGTTTTTTTTTTACAGACTTATCTCTGTCTAATCCAG
*0302 TG to AC (–180/–179) XL.....	5' ACAAAAAAAAAACTCTGCCTAGAGACAGATTAGGTC 3' TGTTTTTTTTTTTGAGACGGATCTCTGTCTAATCCAG
*0302 WT right medium.....	5' TGTCTGCCTAGAGACAGATTAGGTC 3' ACAGACGGATCTCTGTCTAATCCAG
*0302 WT left short.....	5' ACAAAAAAAAAATGTCTGCCTAG 3' TGTTTTTTTTTTACAGACGGATC
*0301 WT left short.....	5' ACAAAAAAAAAA--TCTGCCAG 3' TGTTTTTTTTTT--AGACGGGTCT
*0301 + AA XL.....	5' ACAAAAAAAAAAATCTGCCAGAGACAGATGAGGTC 3' TGTTTTTTTTTTTTAGACGGGTCTCTGTCTACTCCAG
*0302 – AA XL.....	5' ACAAAAAA--TGTCTGCCTAGAGACAGATTAGGTC 3' TGTTTTTTTT--ACAGACGGATCTCTGTCTAATCCAG

^a The XL probes represent sequence positions –191 to –156, inclusive (see Fig. 1). WT, wild type. Dashes indicate spacing for sequence alignment purposes. The 3'-most three bases in the upper strands were added by using the Klenow reaction and α -³²P-labeled nucleoside triphosphates (see Materials and Methods).

polyacrylamide gels. The gels were dried and used to expose X-AR X-ray film (Kodak) with an intensifying screen at –70°C for 5 to 20 h.

Oligonucleotides were synthesized and gel purified before use. Primers used for the primer extension (primer 1), amplification (primer 2), and labeling (primer 3) steps of the in vivo footprinting analyses are described in Table 2.

RESULTS

DNA sequence polymorphism in the 5' regulatory regions of alleles of *DQB1* affected their relative abilities to mediate transcription. Using PCR to amplify genomic DNA from *DQB1* homozygous cell lines, we generated 641-bp fragments representing the promoters and 5' transcriptional regulatory regions of the *DQB1* alleles, *0301 and *0302 (Fig. 1); these were cloned and sequenced in the promoter test plasmid vector, pTCAT (pQ3.1P19CAT and pQ3.2P04CAT, respectively).

TABLE 2. Primers used for in vivo footprinting analyses

Primer	Sequence	Melting temp (°C)
Plus strand		
1	ACATAATTGAGACGAAGGGAAAAAGT	54
2	AGTAGTGGTAGTCAACACAGCTTGG	61
3	GGTAGTCAACACAGCTTGGACCTGATGG	63
Minus strand		
1	TTGGCTTCTTAAATCCCTTAGG	54
2	CCCTTAGGCATTCAATCTTCAGCTC	59
3	CCCTTAGGCATTCAATCTTCAGCTCTTCT	61
Linker	ATGGTGACCCGGGAGATCTGAATTC	61

DNA sequence allelic polymorphism in this region is even more common than in the coding regions of the two alleles, and five such polymorphic residues occur in or near the W and X boxes, which are important transcriptional regulatory elements strongly conserved among all class II HLA genes.

These two CAT plasmid derivatives bearing the wild-type *0301 and *0302 5' regulatory sequences were used in transient-transfection experiments with an Epstein-Barr virus-transformed B-lymphoblastoid cell line, BSM. Figure 2A shows the results of one such experiment: the relative CAT value for *0302 was approximately 40% that of *0301. We tested these *DQB1*-CAT plasmids in a melanoma cell line, Me9229/18, in which *DQB1* is known to be IFN- γ inducible (2, 42). Melanoma cell transient transfectants were treated with various amounts of IFN- γ for 24 h before harvesting, and their lysates were tested for CAT activity. As can be seen in Fig. 2B, the *0301 regulatory fragment was quantitatively more responsive to IFN- γ in mediating expression of CAT than was that of *0302. While the time courses of IFN- γ induction were similar for the two alleles, the slope and plateau values for *0301 were greater than those for *0302. At 100 U of IFN- γ per ml, for example, the relative level of CAT expression in the *0302 transient transfectants was around 30% of that driven by *0301, similar to the levels seen with BSM cells.

Various amounts of IFN- γ were used to treat BSM cells in transient-transfection experiments, with no measurable effect on CAT levels (data not shown). HLA class II genes in Epstein-Barr virus-transformed lymphoblastoid cells appeared to be fully induced without exogenously supplied cytokines. Both

FIG. 2. Relative CAT activity mediated by *DQB1**0301 and *DQB1**0302 5' regulatory regions. The 641-bp *DQB1* 5' regulatory region-CAT plasmid DNAs were used to transiently transfect a B-lymphoblastoid cell line, BSM (A), and a melanoma cell line, Me9229/18 (B), treated with IFN- γ for 24 h. Each panel shows the averaged results of one representative experiment; the same results were obtained in three and two repetitions with BSM and Me9229/18 cells, respectively. Error bars signify the standard deviation for each averaged value.

BSM and Me9229/18 transient pQ3.1P19CAT or pQ3.2P04 CAT transfectants were treated with either interleukin-4 (100 and 500 U/ml; Genzyme) or granulocyte-macrophage colony-stimulating factor (100 and 500 pg/ml; generous gift of Immunex Corporation) no measurable change in relative CAT levels was noted after treatment with these cytokines (data not shown).

Endogenous *0301 and *0302 were expressed unequally in primary human skin fibroblasts. Steady-state *DQB1* mRNA levels in passage 3 primary human skin fibroblasts heterozygous for *HLA-DQB1* alleles (*DRB1**0401/*0407, *DQB1**0301/*0302) were measured by quantitative RT-PCR. Both *DQB1* alleles were amplified and subsequently separated by gel electrophoresis of allele-specific *HpaII* endonuclease fragments. This technique enabled us to compare the relative levels of *DQB1**0301- and *DQB1**0302-specific mRNAs in these cells after induction by IFN- γ for various lengths of time. Figure 3B shows a representative autoradiogram from one such experiment; *DQB1* mRNA was detectable after 12 h, and as in the *DQB1* promoter-driven CAT experiments, the *0301-specific component accumulated more rapidly than that of the *0302 allele. Band intensity values from phosphor imaging densitometry yielded *0301/*0302 mRNA ratios of 16.4:1, 7.0:1, 5.2:1, 4.9:1, and 4.6:1 at 12, 24, 48, 72, and 96 h, respectively; the ratios averaged 4.5:1 after 96 h.

These results were confirmed by using primary skin fibroblasts from two additional *DQB1* heterozygous donors (HLA class II genotypes; *DRB1**0401/*0401, *DQB1**0301/*0302 and *DRB1**0401/*0407, *DQB1**0301/*0302). RT-PCR results were identical to those presented in Fig. 3B; transcripts from the endogenous *DQB1**0301 allele were significantly increased relative to those from *DQB1**0302 in each case.

The allelic polymorphism at positions -180 and -179 contributed significantly to the expression differences between *0301 and *0302. The allelic hierarchy of transcription initiation strength in BSM, as shown in Fig. 2A, was the same as previously found in a study using shorter *DQB1* regulatory region DNA fragments, the 548-bp *HincII* fragments of these alleles (Fig. 1) (1). In that study, a polymorphism at positions -180 and -179, which consists of a TG dinucleotide present in *0302 and lacking in *0301, profoundly affects the protein

FIG. 3. Quantitative RT-PCR analysis of endogenous *DQB1**0301 and *DQB1**0302 expression in primary human skin fibroblasts. (A) Maps of the 270-bp PCR-generated DNA fragments from the third exon of the *0302 and *0301 alleles of *DQB1*. The positions of *HpaII* endonuclease cleavage sites and the sizes (in base pairs) of cleavage products are indicated. DNA sequence polymorphism between the two alleles results in a second *HpaII* site contained in the *0301 fragment which is not present in the *0302 fragment. The 3' PCR primer was end labeled with 32 P, enabling the detection and quantitation (after polyacrylamide gel electrophoresis) of the 166- and 134-bp *HpaII* fragments representing *DQB1**0302 and *DQB1**0301, respectively. (B) Autoradiogram of *HpaII*-treated GAPDH and *DQB1*-derived RT-PCR products electrophoresed in 6% nondenaturing polyacrylamide. The different primary fibroblast samples are identified by the length of IFN- γ treatment. The positions of the GAPDH (301 bp), *DQB1**0302 (166 bp), and *DQB1**0301 (134 bp) *HpaII*-PCR fragments are indicated. As controls, two B-lymphoblastoid cell lines were subjected to the same RT-PCR analysis; the lanes representing cell lines PF97387 (homozygous for *DQB1**0301) and BSM (homozygous for *DQB1**0302) are indicated. The faint band appearing above the 301-bp GAPDH band at a position equivalent to a DNA fragment of ~400 bp was also observed among the GAPDH-primer-only PCR products. Since its size is consistent with that expected for the 301-bp cDNA PCR product plus the 103-bp intron occurring between the two GAPDH primers and since reverse transcriptase-omitted controls contained no PCR product, the band was probably due to amplification of spliced GAPDH mRNA or of message resulting from low-level expression of a GAPDH pseudogene.

binding pattern at this site near the X₁ box. We therefore analyzed this polymorphic site to investigate its potential influence on expression differences between these two alleles.

Four *DQB1*-CAT test plasmids were constructed for use in transient-transfection experiments to evaluate the functional importance of the -180/-179 polymorphism. These CAT plasmids contained *DQB1* 5' regulatory fragments extending from bases -303 to -1, inclusive, for each allele (base -1 is the T residue immediately 5' of the ATG start codon of the *DQB1* gene). In addition to the plasmids containing wild-type sequence of each allele (pQ3.1P27CAT and pQ3.2P10CAT), a third contained *0301 sequence into which a TG dinucleotide was inserted at positions -180 and -179 and a fourth contained *0302 sequence from which the -180/-179 TG was deleted (pQ3.1P29CAT and pQ3.2P12CAT, respectively).

Table 3 shows the results of transient-transfection experiments for the cell lines BSM and Me9229/18. The relative CAT values obtained after transfection of BSM and IFN- γ -induced Me9229/18 cells with these mutant plasmids displayed virtually the same pattern. Insertion of TG at -180 and -179 reduced the efficiency of IFN- γ -induced CAT transcription mediated by the *0301 fragment by approximately 30 to 40%, whereas deletion of the TG at -180 and -179 from *0302 increased CAT

TABLE 3. The TG polymorphism influences efficiency of the *DQB1* promoter in BSM and Me9229/18 cells

Cell line	Transfecting DNA	Description ^a	IFN- γ ^b	Relative CAT activity ^c
BSM	pQ3.1P27CAT	*0301 wild type	NA ^d	10.0
	pQ3.1P29CAT	*0301 TG inserted	NA	6.1 $\pm$ 0.8
	pQ3.2P10CAT	*0302 wild type	NA	6.3 $\pm$ 0.5
	pQ3.2P12CAT	*0302 TG deleted	NA	10.8 $\pm$ 1.0
Me9229/18	pQ3.1P27CAT	*0301 wild type	+	10.0
			—	1.7 $\pm$ 0.7
	pQ3.1P29CAT	*0301 TG inserted	+	7.7 $\pm$ 1.1
			—	1.9 $\pm$ 0.4
	pQ3.2P10CAT	*0302 wild type	+	5.2 $\pm$ 2.2
			—	1.4 $\pm$ 0.6
	pQ3.2P12CAT	*0302 TG deleted	+	10.4 $\pm$ 3.9
			—	1.3 $\pm$ 0.7

^a *DQB1* regulatory region bases -303 to -1 , inclusive (see Fig. 1 and plasmid diagrams in Fig. 7) were placed in pTCAT; TG insertion or deletion mutations were directed to positions -180 and -179 .

^b IFN- γ was added to cultures of Me9229/18 transient transfectants at a final concentration of 100 U/ml 24 h before harvesting.

^c Mean results of three independent transient-transfection experiments with BSM and nine with Me9229/18, $\pm$ the calculated standard deviations. Relative CAT activities were calculated by normalizing the means of replicate CAT/ β -galactosidase ratios to that of pQ3.1P27CAT in the BSM experiments and pQ3.1P27CAT plus IFN- γ in the Me9229/18 experiments, both of which were assigned the value 10.0.

^d NA, not applicable.

expression to a level equivalent to that of wild-type *0301. These results suggested that the expression differences mediated by the TG polymorphism accounted for most, if not all, of the variation in expression corresponding to the promoter and 5' regulatory region fragments of the two alleles.

Allele-specific binding of nuclear protein to the *DQB1* transcriptional regulatory region. Figures 4A and C show the results of EMSAs using nuclear extracts from BSM and Me9229/18 cells, respectively, and four double-stranded oligonucleotide probes encompassing the $-180/-179$ site (Table 1). The bands indicated by arrows were present only with the *0302 wild-type probe and with the mutant *0301 probe containing the TG at positions -180 and -179 ; they did not occur with the other two probes which lacked the TG dinucleotide. A TG-specific band seen with the Me9229/18 extract (large arrow in Fig. 4C) comigrated with the TG-dependent band produced by the BSM extract. Two additional faster-migrating TG-specific bands, indicated by small arrows, were generated by the melanoma cell extract. The dependence of the appearance of these bands on the presence of the TG dinucleotide was confirmed by competition EMSA experiments in which only DNA fragments containing TG at -180 and -179 could compete for Me9229/18 or BSM protein complex binding with the wild-type *0302 or the TG insertion mutant *0301 probe (Fig. 4B and D).

Nuclear extracts were prepared from Me9229/18 cells which had been treated with IFN- γ for 24 h as well as from untreated melanoma cells. The binding pattern with the *0302 wild-type double-stranded oligonucleotide probe was identical when nuclear extracts from either IFN- γ -treated or untreated melanoma cells were used (data not shown).

The dependence of the complexes seen in EMSAs on the presence of the TG dinucleotide in the oligonucleotide probes implies the likelihood of a specific protein-DNA interaction around the $-180/-179$ site in the *DQB1* 5' regulatory region. This possibility was evaluated directly by in vivo footprinting analysis of the BSM cell line, which is homozygous for the *DQB1**0302 gene. As shown in Fig. 5, a number of hypomethylated G residues, as well as several hypersensitive purine residues, were seen in the X-box region, indicating footprint interactions with numerous residues throughout the -189 to -155 region. Five G residues on either the plus or minus

strand which show reduced intensity in the footprinted X-box region, at positions -179 , -174 , -173 , -170 , and -166 , are indicated. Although unable to discriminate between the different X-box-binding proteins, this technique confirms the apparent occupancy of the promoter region, including the -179 site in a cell which is expressing the *DQB1**0302 gene.

To evaluate the fine specificity of the interaction at positions -180 and -179 , and the contribution of surrounding nucleotides, we tested the ability of a variety of oligonucleotides with DNA sequence substitutions or truncations to compete for the formation of the protein-DNA complex shown in Fig. 4. Figures 6A and B show the results of these experiments for BSM and Me9229/18 nuclear extracts, respectively. A *DQB1**0302 double-stranded oligonucleotide containing a dinucleotide substitution at positions -174 and -173 was fully active in competing with the wild-type *0302 probe for the formation of the $-180/-179$ -associated complex (lanes 3 and 4), which indicates that although both of these areas were shown to be implicated in protein binding by the in vivo footprinting analyses (Fig. 5), they may represent adjacent but separate protein-DNA interactions. An *0302 double-stranded oligonucleotide in which the $-180/-179$ TG was substituted by AC was unable to compete for the complex formation with the wild-type probe (lanes 5 and 6). The 5' truncated, *0302 wild-type right medium oligonucleotide, which lacked residues to the left of the $-180/-179$ TG, did not compete against the full-length wild-type *0302 probe (lanes 7 and 8), while the 3' truncated *0302 wild-type left short oligonucleotide, which contained the poly(A) tract, did compete (lanes 9 and 10).

Functional duality of the $-180/-179$ polymorphism. In comparing the region immediately 5' of the X₁ boxes of *0301 and *0302, it is possible to view the $-180/-179$ polymorphism two different ways. One is that the TG dinucleotide provides the *0302 allele with a nuclear protein binding site that is absent in *0301. Another is that the presence or absence of the TG dinucleotide affects the spacing between the W and X₁ boxes in the transcriptional regulatory apparatus of *DQB1*. Both the X₁ element and the S and J elements (which overlap the W box; see Fig. 1) are required for IFN- γ induction of *DQB1* transcription (10, 42). W-X₁ box spacing in another, similarly regulated HLA class II locus, *DRA*, has been shown to affect transcriptional efficiency in an absolute manner; any

FIG. 4. EMSAs of the -180/-179 TG-specific protein complex. The arrows indicate bands which represent TG-specific probe binding activity. Probe sequences are shown in Table 1. (A) Direct binding assays with BSM nuclear extract. (B) Competition assays with BSM nuclear extract. Each sample contained 1.0 μ g of poly(dI-dC). (C) Direct binding assays with Me9229/18 nuclear extract. The sample in lane 1 is identical to that of lane 9 in panel A and is included on the gel for reference. (D) Competition assays with Me9229/18 nuclear extract. Each sample contained 0.2 μ g of poly(dI-dC).

number of base insertions or deletions from the optimal spacing between those two regulatory elements reduces initiation of transcription (47). Presumably, regulatory proteins that interact with the W and X₁ elements may also interact either with each other or with additional proteins, and these interactions are constrained if the spacing between the two elements is suboptimal.

To test the hypothesis that the W-X₁ box spacing variation component of the -180/-179 allelic polymorphism is functionally significant, additional mutant *DQB1*-CAT plasmids were designed to take advantage of the tract of nine contiguous A residues occurring between the W and X₁ boxes in these alleles. Insertion or deletion of A residues at that tract changed W-X₁ box spacing without altering adjacent sequences which might be involved in specific regulatory protein binding. The mutants were pQ3.1P32CAT, in which two A residues were inserted in the A tract of *DQB1**0301, and pQ3.2P14CAT, in which two A residues were deleted from *DQB1**0302.

Results of transient-transfection experiments with Me9229/18 cells are shown in Fig. 7. Insertion of two A residues in the A tract of the *0301 5' regulatory region

(pQ3.1P32CAT) (to mimic the W-X₁ box spacing of the *0302 allele) reduced IFN- γ -induced CAT expression by 70% compared with the level for wild-type *0301. Deletion of two A residues from *0302 (pQ3.2P14CAT) (to mimic the W-X₁ box spacing of the *0301 allele) significantly increased CAT expression to levels greater than those for wild-type *0302. Identical experiments with the B-lymphoblastoid BSM cell line produced similar results (data not shown), demonstrating that variation in W-X₁ box spacing of the alleles of *DQB1* clearly influenced CAT expression efficiency.

Two additional A-residue insertion mutants in which the W-X₁ spacing was four bases longer than the wild-type spacing in each allele were also evaluated, and the IFN- γ -induced CAT values obtained from each approximated those seen with wild-type *0302, that is, about 50% of those of *0301 (data not shown). Thus, within the range tested, the W-X₁ box spacing of the *0301 allele conferred optimal transcriptional efficiency.

To ensure that changing the size of the A tract in *DQB1**0302 was a specific test of the functional effects of W-X₁ box spacing variation and did not influence binding of the TG-specific protein complex, EMSA experiments were

FIG. 5. In vivo footprint of the plus and minus strands of the region encompassing the W, X, and Y boxes of the *DQB1*0302* transcriptional regulatory region in BSM cells. The lanes labeled C are the in vitro-treated control samples; the in vivo-treated, experimental samples are shown in the lanes marked E. Footprinted G residues are designated by arrows, and hypermethylated purine residues are designated by arrowheads. The sequences of both strands from base positions -189 through -155 are shown, with those that are either footprinted or hypermethylated in boldface letters.

done with probes in which the size of the A tract was altered. Figure 8 shows the results of EMSAs with the BSM nuclear extract in which the **0301* probe contained two extra A residues in the A tract (total of 11) and two A residues were removed from the A tract in the **0302* probe (total of 7). Additionally, the abilities of both mutant oligonucleotides and the wild-type **0302* oligonucleotide to compete with the **0302* -AA probe for TG-specific complex formation were tested (Fig. 8, lanes 9 to 14). These experiments demonstrated that a two-A-deletion mutation in the **0302* probe had no effect on the binding of the TG-specific nuclear protein complex: the protein complex was dependent on the presence of TG for binding, regardless of the length of the poly(A) tract.

Although W-X₁ box spacing variation was the most significant factor influencing CAT expression, the specific TG dinucleotide at -180 and -179 was also an influential component. For example, CAT values obtained from the **0301* plasmid into which TG was introduced at -180 and -179 (pQ3.1P29 CAT) were higher than those generated by the **0301* plasmid in which two A residues were inserted in the A tract (pQ3.1P32CAT) (Fig. 7), suggesting that binding of the TG-specific nuclear protein complex also contributed to the allelic variation in *DQB1* expression.

To clarify the relative importance of these two views of the -180/-179 polymorphism, additional mutant CAT plasmids were constructed: pQ3.1P36CAT, in which two A residues were replaced by a TG insertion into **0301* sequence at -180 and -179, and pQ3.2P20CAT, **0302* sequence with two A residues inserted and the -180/-179 TG deleted. In other words, despite the insertion or deletion of TG, the W-X₁ box spacing of the mutants remained identical to wild-type spacing for each allele.

The results of transient transfections of Me9229/18 cells with these mutants are also shown in Fig. 7. The binding of the TG-specific nuclear protein complex to the **0301* regulatory region (pQ3.1P36CAT) corresponded to CAT activity significantly increased over that of the wild type; on the other hand, lack of binding of the protein complex to the **0302* mutant (pQ3.2P20CAT) corresponded to reduced CAT expression

FIG. 6. EMSAs with various substitution and truncation mutant double-stranded oligonucleotides used as competitors. All samples contain a wild-type *DQB1*0302* double-stranded oligonucleotide probe representing nucleotide positions -191 to -156, inclusive (see Table 1). The arrows indicate bands which represent TG-specific probe binding activity. (A) BSM nuclear extract. Samples contain 1.0 μ g of poly(dI-dC). (B) Me9229/18 nuclear extract. Samples contain 0.2 μ g of poly(dI-dC). The large arrow indicates the band which comigrates with that produced by the BSM nuclear extract.

compared with wild-type **0302*. Thus, it was evident that both perspectives of the -180/-179 polymorphism were functionally significant: both the distance between the W and X₁ boxes and the TG present at that site acted as independent modulators of transcriptional efficiency.

DISCUSSION

The two principal observations from this study were that (i) transcriptional regulation of endogenous *HLA-DQB1* genes during IFN- γ induction occurs significantly differently in the **0301* allele than in the **0302* allele and (ii) conserved sequence polymorphism in the transcriptional regulatory region of *HLA-DQB1* results in two patterns of functional effects: both the spacing of the W and X₁ elements and allele-specific

FIG. 7. Relative CAT activities generated by wild-type and mutant *DQB1* 5' regulatory region-CAT plasmid DNAs in transient transfections of Me9229/18 cells. Regulatory region sequences from positions -208 to -161 are shown for each allele. Vertical lines in the sequence signify spacing made for alignment purposes only. Vector sequence is represented by heavy dark bars. Dashes signify insert sequence which is unchanged from the wild-type; // indicates a break in scale. Mutant base substitutions and insertions are indicated, as are deletions (Δ). CAT activity values obtained from pQ3.1P27CAT transfections were set to a relative value of 10.0, and all other results were normalized to that value; normalized values from multiple experiments were then averaged. The number of individual transfection experiment results contributing to the averaged values shown varied as follows: pQ3.1P27CAT, 12; pQ3.1P29CAT, 9; pQ3.1P32CAT, 8; pQ3.1P36CAT, 3; pQ3.2P10CAT, 11; pQ3.2P12CAT, 9; pQ3.2P14CAT, 8; and pQ3.2P20CAT, 3. The error bars represent the standard deviation for each value; outlying data were excluded only if they ranged greater than 2 standard deviations from the mean values.

DNA-protein interactions in the region influence gene expression.

Quantitative RT-PCR analysis of *DQB1* expression in *0301/*0302 heterozygous primary skin fibroblasts revealed that the *0302 allele responded much more slowly to induction by IFN- γ than did *0301. While the differences in mRNA levels diminished from ~16-fold at 12 h to ~5-fold at 72 h, the *0302 mRNA level never did reach that of *0301, even after 144 h of exposure to IFN- γ . If the subsequent appearance of the DQ3.2 protein on the cell surface is similarly delayed compared with that of the DQ3.1 protein, especially in response to cytokine induction in nonprofessional antigen-presenting cells, the predominant peptide-HLA complex displayed and, in turn, the predominant specificity of the T-cell response may be significantly skewed.

Transfection analysis of CAT plasmids revealed that the shorter spacing in the W-X₁ region of the *0301 allele mediated a higher level of expression than that of *0302. Insertion of the TG dinucleotide adjacent to the X₁ box, however, introduced a partially compensatory change, with the -180 and -179 base positions functioning as a discrete site for DNA-protein interactions in alleles which contain the TG. Spatial alignment of the W-box S element and the X₁ box is highly conserved in a broad range of class II major histocompatibility complex genes (10). Both of these elements are required for

IFN- γ responsiveness and cell type specificity of expression. Changes in the distance between the S and X₁ elements of *HLA-DRA*, for example, diminish IFN- γ -induced transcription (47), suggesting that regulatory proteins binding at these sites may interact with each other either directly or indirectly when the distance between the binding sites is optimum. A prominent feature of the region separating the W and X₁ elements in *HLA-DQB1* is a lengthy A tract which has the potential for influencing DNA bending and conformational rigidity (19). Native allelic length polymorphism and the mutations introduced in these experiments may generate variations in the torsional and lateral rigidity of the DNA (36). The basis for the observed changes in CAT activity in these mutants, as well as differences between native alleles, therefore, may be variation in the potential for bending of the A tract between the W and X₁ control elements.

Understanding the functional effects of these allelic differences, however, is complicated by our observation that the presence of TG at -180 and -179 in the wild-type *0302 allele corresponded to the binding of an additional nuclear protein or protein complex associated with regulatory activity. The protein binding to the site created by the presence of TG in *DQB1**0302 was highly specific; in EMSA experiments, it did not bind to the wild-type *0301 probe or to the *0302 probe from which the TG was deleted. Competition EMSA experi-

FIG. 8. EMSAs of the BSM nuclear extract with probes that varied in the size of the A tract between the W and X_1 boxes. Probe and competitor oligonucleotide sequences are shown in Table 1. Bands representing the TG-specific nuclear protein complex are indicated by the large arrow.

ments confirmed this specificity. Transient transfection assays with double-mutant CAT vectors helped discern the functional effects associated with the TG-specific protein complex: deleting the TG from the *0302 regulatory region while adding two A residues (to preserve the wild-type spacing of that allele) resulted in reduced CAT activity compared with that of the wild type; inserting TG and deleting two A residues from *0301 (again preserving *0301 wild-type spacing) indicated that binding of the TG-specific protein complex corresponded to an increase in transcriptional efficiency to 230% of that of wild-type *0301. Thus, the TG-specific binding protein complex behaved as a strong positive effector of transcription.

The presence of this TG dinucleotide in the 5' regulatory region of *DQB1* genes is not limited to the *0302 allele: *0101, *0102, *0303, and *0401 all have the TG motif. Of the alleles that we have sequenced, only *0301 and *0201 lack the dinucleotide (1). It is conserved in the *DQB* homologs of nonhuman primates such as chimpanzees (*Pan troglodytes*), owl monkeys (*Aotus trivirgatus*), and lemurs (*Lemur fulvus sanfordi* and *Lemur fulvus albifrons*) (9), as well as in cattle (11) and *I-Aβ* of mice (5). The number of bases separating the 3' base of the S element from the 5' base of the X_1 element ranges from 15 to 19 among these transcriptional control regions. None of the *DQB1* alleles and nonhuman homologs which have the TG motif at -180 and -179 share the 16-bp W- X_1 box spacing distance with *0301. It is possible that the loss of the transcriptional up-regulating TG motif in the 5' region of the *0301 allele has been compensated for by the shortening of the distance between the W and X_1 boxes and/or changing the degree of DNA bending at that site, which, as we have observed, increases the efficiency of *HLA-DQB1* expression.

Previous reports of X-box-binding proteins have predominantly defined interactions within the nonpolymorphic *DRA* gene (6, 31). The *DRA* X_1 element, however, lacks a site homologous to the -179/-180 TG dinucleotide in the *DQB1*

sequence and, indeed, fails to bind to the TG-protein complex prepared from a human B-cell extract (43).

The extent of the DNA sequence recognized by the TG-specific protein complex appeared to be limited to residues proximal to the X_1 element. Our data suggested that the in vivo footprint representing binding at the -174 and -173 positions is independent of the TG-specific interaction, since substitutions of residues -174 and -173 did not alter the in vitro TG complex binding. Furthermore, an oligonucleotide probe representing the TG-containing lemur homolog of *DQB1* in this region, which contains an A rather than a G at -175, produced the shifted TG complex in EMSAs and competed for the complex with the wild-type *0302 probe (9). It also appears that the TG-specific protein may require at least some of the residues of the A tract for binding to the *DQB1**0302 regulatory region, since oligonucleotides truncated at the A tract failed to compete for complex formation.

Other polymorphic sites in the transcriptional control region of *DQB1*, however, may influence the effect of TG-dependent protein binding. In transfection assays using plasmids pQ3.2P14CAT and pQ3.1P36CAT, for example, the up-regulating activity of the TG-binding protein or protein complex was modulated by other allelic polymorphism in the 5' regulatory region. Although both mutants contain the TG and have the same *0301-like W- X_1 box spacing, the CAT value corresponding to the *0302 mutant was approximately one-half of that of the *0301 mutant.

Whether the TG complex functions directly as a transcriptional activator, as a modulator of W-X box interactions and DNA bending, or simply as a steric inhibitor of other transcription factors is not known. It is of interest that a recently described protein, NF-X1, binds to the X_1 box of *DQB1* and functions as a down-regulator of its transcription (39). While the specific DNA binding sequence has not been reported for NF-X1, it may require nucleoside residues at the 5' end of the X_1 element as well as those adjacent to it. Possibly, binding of NF-X1 to *DQB1* is less efficient in those alleles containing the -180/-179 TG because of the interfering effect of the protein binding to that dinucleotide. Such abrogation of normal NF-X1 binding could produce the positive effect on *DQB1* transcription that we have observed to be associated with the TG-dependent protein binding.

Although it is not yet possible to state definitively whether the TG-specific EMSA band represented a single protein or a complex of proteins which bound to *DQB1**0302, some observations made in the course of these studies may be pertinent to that issue. Three TG-specific EMSA bands were seen with nuclear extracts from Me9229/18 cells; two of these had faster electrophoretic mobilities than the predominant TG band from BSM extracts. In addition, some BSM extracts often showed a minor faster-migrating band of the same mobility as that seen with Me9229/18 extracts (data not shown). A possible explanation for this finding is that the TG-specific complex in BSM extracts is a dimer or multimer, and some of the protein in Me9229/18 extracts may be present only in the smaller form seen in some BSM extracts. Alternatively, the smaller bands with TG binding activity may represent protein fragments encoded by the same gene with variable splicing, or they may result from partial degradation. Clarification of these issues and further evaluation of the in vivo significance of the TG-associated binding activity will require additional studies that are dependent on cloning of the relevant gene(s).

Transcriptional control for HLA class II genes is complex and must be sensitive to expression requirements of a given type of cell, its developmental stages, and the influence of cytokines. Quantitative differences in levels of class II expres-

sion influence antigen presentation to T cells (15, 25) and, in the extreme case, loss of transcription of these genes cause primary combined immunodeficiency (33).

Although models for immunologic mechanisms of HLA class II function have focused on the allelic polymorphism of the gene products at the amino acid level and effects on peptide binding and presentation (30, 45), the possible contribution of allelic expression differences has not yet been explored. Although the mechanisms of HLA class II-mediated antigen presentation and immunopathogenesis in autoimmunity are not clear, inappropriate expression of *DQB1* alleles or differential transcriptional response of specific alleles to cytokine stimuli may contribute to their role in disease associations and immune activation.

ACKNOWLEDGMENTS

We thank our colleagues Tom Cotner, Nicola Leech, Linda Shewey, and Teresa Sukiennicki for their helpful discussions and review of the manuscript and thank Kim Hegardt for technical assistance. We are grateful to Frederick A. Leaf for assistance with skin biopsies.

This work was supported by Public Health Service grant NIH DK41801, the Greenwall Foundation, a Juvenile Diabetes Foundation International Fellowship award (J.S.B.), and Medical Research Council of Canada Fellowship and Diabetes Research Council McAbee Fellowship awards (K.A.W.).

REFERENCES

- Andersen, L. C., J. S. Beaty, J. W. Nettles, C. E. Seyfried, G. T. Nepom, and B. S. Nepom. 1991. Allelic polymorphism in transcriptional regulatory regions of HLA-DQB genes. *J. Exp. Med.* **173**:181-192.
- Anichini, A., C. Castelli, G. Sozzi, G. Fossati, and G. Parmiani. 1988. Differential susceptibility to recombinant interferon-gamma-induced HLA-DQ antigen modulation among clones from a human metastatic melanoma. *J. Immunol.* **140**:183-191.
- 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**:681-715.
- Buus, S., A. Sette, S. M. Colon, C. Miles, and H. M. Grey. 1987. The relation between major histocompatibility complex (MHC) restriction and the capacity of Ia to bind immunogenic peptides. *Science* **235**:1353-1358.
- Celada, A., M. Shiga, M. Imagawa, J. Kop, and R. A. Maki. 1988. Identification of a nuclear factor that binds to a conserved sequence of the I-A β gene. *J. Immunol.* **140**:3995-4002.
- Cogswell, J. P., N. Zeleznik-Le, and J. P.-Y. Ting. 1991. Transcriptional regulation of the HLA-DRA gene. *Crit. Rev. Immunol.* **11**:87-112.
- Dukas, K., P. Sarfati, N. Vaysses, and L. Pradayrol. 1993. Quantitation of changes in the expression of multiple genes by simultaneous polymerase chain reaction. *Anal. Biochem.* **215**:66-72.
- Durand, B., M. Kobr, W. Reith, and B. Mach. 1994. Functional complementation of major histocompatibility complex class II regulatory mutants by the purified X-box-binding protein RFX. *Mol. Cell. Biol.* **14**:6839-6847.
- Gaur, L., L. Shewey, D. Sharkey-Mathis, and G. T. Nepom. 1994. Transcriptional regulatory elements in primate MHC-DQB genes. *Transgenics* **1**:267-276.
- Glimcher, L. H., and C. J. Kara. 1992. Sequences and factors: a guide to MHC class II transcription. *Annu. Rev. Immunol.* **10**:13-49.
- Groenen, M. A., J. J. Van der Poel, R. J. Dijkhof, and M. J. Giphart. 1990. The nucleotide sequence of bovine MHC class II DQB and DRB genes. *Immunogenet.* **31**:37-44.
- Hattori, M., and Y. Sakaki. 1986. Dideoxy DNA sequencing method using denatured plasmid templates. *Anal. Biochem.* **152**:232-238.
- Hennighausen, L., and H. Lubon. 1987. Interaction of protein with DNA in vitro. *Methods Enzymol.* **152**:721-735.
- Itoh-Lindstrom, Y., B. M. Peterlin, and J. P.-Y. Ting. 1995. Affinity enrichment and functional characterization of TRAX1, a novel transcription activator and X1-sequence-binding protein of HLA-DRA. *Mol. Cell. Biol.* **15**:282-289.
- Janeway, C. A., K. Bottomly, J. Babich, P. Conrad, S. Conzen, B. Jones, J. Kaye, M. Katz, L. McVay, D. B. Murphy, and J. Tite. 1984. Quantitative variation in Ia antigen expression plays a central role in immune regulation. *Immunol. Today* **5**:99-105.
- Kappler, J. W., N. Roehm, and P. Marrack. 1987. T cell tolerance by clonal elimination in the thymus. *Cell* **49**:273-280.
- Kaufman, J. F., C. Auffrey, A. J. Korman, D. A. Shackleford, and J. Strominger. 1984. The class II molecules of the human and murine major histocompatibility complex. *Cell* **36**:1-13.
- Kwok, W. W., D. Schwarz, B. S. Nepom, P. S. Thurtle, R. A. Hock, and G. T. Nepom. 1988. HLA-DQ molecules form α - β heterodimers of mixed allotype. *J. Immunol.* **141**:3123-3127.
- Leech, N., R. Sorrentino, D. K. McCulloch, and G. T. Nepom. Ultrastructural variation in allelic promoter elements of the DQB1 locus. *Hum. Immunol.*, in press.
- Liou, H.-C., M. R. Boothby, P. W. Finn, R. Davidson, N. Nabavi, N. J. Zeleznik-Le, J. P. Ting, and L. H. Glimcher. 1990. A new member of the leucine zipper class of proteins that binds to the HLA DR α promoter. *Science* **247**:1581-1584.
- Lopata, M. A., D. W. Cleveland, and B. Sollner-Webb. 1984. High-level transient expression of a chloramphenicol acetyltransferase gene by DEAE-dextran mediated DNA transfection coupled with a dimethyl sulfoxide or glycerol shock treatment. *Nucleic Acids Res.* **12**:5707-5717.
- Louis, P., R. Vincent, P. Cavadore, J. Clot, and J.-F. Eliaou. 1994. Differential transcriptional activities of HLA-DR genes in the various haplotypes. *J. Immunol.* **153**:5059-5067.
- Maniatis, T., E. F. Fritsch, and J. F. Sambrook. 1982. Molecular cloning: a laboratory manual, p. 13.11-13.13. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Marsh, S. G. E., and J. G. Bodmer. 1992. HLA class II nucleotide sequences, 1992. *Hum. Immunol.* **35**:1-17.
- Matris, L. A., L. H. Glimcher, W. E. Paul, and R. H. Schwartz. 1983. Magnitude of response of histocompatibility-restricted T cell clones is a function of the product of the concentrations of antigen and Ia molecules. *Proc. Natl. Acad. Sci. USA* **80**:6019-6023.
- Maxam, A. M., and W. Gilbert. 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol.* **65**:499-560.
- Miller, J. H. 1972. Experiments in molecular genetics, p. 403-404. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Morzycka-Wroblewska, E., J. I. Harwood, J. R. Smith, and M. F. Kagnoff. 1993. Structure and evolution of the promoter regions of the DQA genes. *Immunogenetics* **37**:364-372.
- Mueller, P. R., and A. Wold. 1989. In vivo footprinting of a muscle specific enhancer by ligation mediated PCR. *Science* **246**:780-786.
- Nepom, G. T. 1990. A unified hypothesis for the complex genetics of HLA associations with IDDM I diabetes. *Diabetes* **39**:1153-1157.
- Peterlin, B. M., G. Andersson, E. Löttscher, and S. Tsang. 1990. Transcriptional regulation of HLA class-II genes. *Immunol. Res.* **9**:164-177.
- Reith, W., M. Kobr, P. Emery, B. Durand, C.-A. Siegrist, and B. Mach. 1994. Cooperative binding between factors RFX and X2bp to the X and X2 boxes of MHC class II promoters. *J. Biol. Chem.* **269**:20020-20025.
- Reith, W., S. Satola, C. H. Sanchez, I. Amaldi, B. Lisowska-Grospierre, C. Griselli, 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**:897-906.
- Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Erlich. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**:487-491.
- Sanchez, C. H., W. Reith, P. Silacci, and B. Mach. 1992. The DNA-binding defect observed in major histocompatibility complex class II regulatory mutants concerns only one member of a family of complexes binding to the X boxes of class II promoters. *Mol. Cell. Biol.* **12**:4076-4083.
- Schleif, R. 1992. DNA looping. *Annu. Rev. Biochem.* **61**:199-223.
- Shapiro, D. J., P. A. Sharp, W. W. Wahli, and M. J. Keller. 1988. A high-efficiency HeLa cell nuclear transcription extract. *DNA* **7**:47-55.
- Sleigh, M. J. 1986. A nonchromatographic assay for expression of the chloramphenicol acetyltransferase gene in eukaryotic cells. *Anal. Biochem.* **156**:251-256.
- Song, Z., S. Krishna, D. Thanos, J. L. Strominger, and S. J. Ono. 1994. A novel cysteine-rich sequence-specific DNA-binding protein interacts with the conserved X-box motif of the human major histocompatibility complex class II genes via a repeated cys-his domain and functions as a transcriptional repressor. *J. Exp. Med.* **180**:1763-1774.
- Steimle, V., L. A. Otten, M. Zufferey, and B. Mach. 1993. Complementation cloning of an MHC class II transactivator mutated in hereditary MHC class II deficiency (or bare lymphocyte syndrome). *Cell* **75**:135-146.
- Steimle, V., C.-A. Siegrist, A. Mottet, B. Lisowska-Grospierre, and B. Mach. 1994. Regulation of MHC class II expression by interferon-gamma mediated by the transactivator gene CIITA. *Science* **265**:106-109.
- Sugawara, M., P. D. Ponath, J. Shin, Z. Yang, and J. L. Strominger. 1991. Delineation of a previously unrecognized cis-acting element required for HLA class II gene expression. *Proc. Natl. Acad. Sci. USA* **88**:10347-10351.
- Sukiennicki, T. L., L. M. Shewey, and G. T. Nepom. 1993. Locus and allele-specific DNA-protein interactions in the HLA-DQB1 X-box. *Immunol. Res.* **12**:317-329.
- Ting, J. P.-Y., A. Painter, N. J. Zeleznik-Le, G. MacDonald, T. M. Moore, A. Brown, and B. D. Schwartz. 1994. YB-1 DNA-binding protein represses interferon gamma activation of class II major histocompatibility complex

- genes. *J. Exp. Med.* **179**:1605–1611.
45. **Todd, J. A., J. I. Bell, and H. O. McDevitt.** 1987. HLA-DQ β gene contributes to susceptibility and resistance to insulin-dependent diabetes mellitus. *Nature (London)* **329**:599–604.
46. **Turco, E., B. J. Manfras, L. Ge, W. A. Rudert, and M. Trucco.** 1990. The X boxes from promoters of HLA class II B genes at different loci do not compete for nuclear protein-specific binding. *Immunogenetics* **32**:117–128.
47. **Vilen, B. J., J. F. Penta, and J. P.-Y. Ting.** 1992. Structural constraints within a trimeric transcriptional regulatory region: constitutive and interferon-gamma inducible expression of the HLA-DRA gene. *J. Biol. Chem.* **267**:23728–23734.
48. **Wigler, M., A. Pellicer, S. Silverstein, and R. Axel.** 1978. Biochemical transfer of single-copy eukaryotic genes using total cellular DNA as donor. *Cell* **14**:725–731.
49. **Wright, K. L., B. J. Vilen, Y. Itoh-Lindstrom, T. L. Moore, G. Li, M. Criscitiello, P. Cogswell, J. B. Clarke, and J. P.-Y. Ting.** 1994. CCAAT box binding protein NF-Y facilitates *in vivo* recruitment of upstream DNA binding transcription factors. *EMBO J.* **13**:4042–4053.