# DO $\beta$ : a new $\beta$ chain gene in HLA-D with a distinct regulation of expression

### Cécile Tonnelle, Robert DeMars<sup>1</sup> and Eric O.Long

Laboratory of Immunogenetics, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Building 5, Room B1-11, Bethesda, MD 20205, and <sup>1</sup>Laboratory of Genetics and Department of Human Oncology, University of Wisconsin, Madison, WI 53706, USA

Communicated by B.Dobberstein

The HLA-D region of the human major histocompatibility complex encodes the genes for the  $\alpha$  and  $\beta$  chains of the DP, DQ and DR class II antigens. A cDNA clone encoding a new class II  $\beta$  chain (designated DO) was isolated from a library constructed from mRNA of a mutant B-cell line having a single HLA haplotype. Complete cDNA clones encoding the four isotypic  $\beta$  chains of the DR1, DQw1, DPw2 and putative DO antigens were sequenced. The DO $\beta$  gene was mapped in the D region by hybridization with DNA of HLA-deletion mutants. DO $\beta$  mRNA expression is low in B-cell lines but remains in mutant lines which have lost expression of other class II genes. Unlike other class II genes  $DO\beta$  is not induced by  $\gamma$ -interferon in fibroblast lines. The DO $\beta$  gene is distinct from the DP $\beta$ , DQ $\beta$  and DR $\beta$  genes in its pattern of nucleotide divergence. The independent evolution and expression of  $DO\beta$ suggest that it may be part of a functionally distinct class II molecule.

*Key words:* cDNA clone/histocompatibility/HLA-D/immune response/nucleotide sequence

## Introduction

The major histocompatibility complex (MHC) encodes antigens which control the initial steps of immune responses. MHC antigens are required in the recognition of foreign antigens by T lymphocytes. The class I antigens of the MHC, which are expressed on virtually all nucleated cells, are primarily involved in antigen presentation to cytotoxic T cells, while class II MHC antigens are generally required in antigen presentation to T cells of the helper/inducer subset (Swain, 1983). Class II antigens are expressed in antigen-presenting cells and B lymphocytes. Their expression can also be induced in human T lymphocytes upon antigenic or mitogenic stimulation and in a variety of cells by treatment with  $\gamma$ -interferon. MHC antigens display a very high degree of allelic polymorphism and T cells usually recognize foreign antigens only in the context of a self MHC antigen (Zinkernagel and Doherty, 1979). In fact, MHC antigens themselves elicit strong immune responses in allogeneic hosts as observed in graft rejections and mixed lymphocyte reactions.

In addition to the allelic polymorphism there is another level of complexity: multiple isotypic forms of class I and class II genes are present in the MHC. In man there are at least three class II antigens, HLA-DP, -DQ and -DR, each consisting of distinct  $\alpha$  and  $\beta$  chains, which are encoded by genes in the HLA-D region of the MHC (Auffray and Strominger, 1985). The  $\alpha$  and  $\beta$  chains associate non-covalently and are both transmembrane glycoproteins with a membrane-proximal immunoglobulin-like domain. There is evidence that the three class II antigens in man can be involved in antigen presentation to T cells (Qvigstad *et al.*, 1984). The existence of the DR, DQ and DP antigens may serve to diversify the range of immune responses that each individual can mount. An alternative but not exclusive view is that each of these class II antigens plays a distinct role in the generation of immune responses.

The molecular analysis of the HLA-D region has revealed multiple genes within each subregion: one  $\alpha$  and three  $\beta$  genes in DR, and two  $\alpha$  and two  $\beta$  genes in both DQ and DP (Auffray and Strominger, 1985). An additional  $\alpha$  gene named DT (Inoko *et al.*, 1984) or DZ (Trowsdale *et al.*, 1985) has been described. Correlations have been made between human and murine class II genes based on their structural similarities. The  $\alpha$  and  $\beta$  chain genes of DR and DQ are the respective counterparts of the I-E and I-A genes of the murine I region. Two additional  $\beta$  genes named A $\beta$ 2 and A $\beta$ 3 have been found in the I region (Larhammar *et al.*, 1983a; Wake *et al.*, 1985). A $\beta$ 3 appears to be the counterpart of the human DP $\beta$  gene but no human equivalent for A $\beta$ 2 has been reported yet. It remains to be established how many of these class II genes in man and mouse are functional and whether other undetected genes exist.

We have approached these questions in man by constructing a cDNA library from mRNA of a B-cell line which has lost a complete HLA-haplotype (Kavathas et al., 1980). Such a hemizygous mutant containing only non-allelic genes presents an advantage over homozygous cell lines. Even when homozygosity results from consanguinous pairings it is possible that a recombination event, undetectable by serological means, introduces heterozygosity at certain loci (Termijtelen et al., 1983; Robinson et al., 1984). The library was made in a eukaryotic expression vector such that a functional dissection of truly isotypic class II antigens will be possible. Cell surface expression of DR antigens in transfected murine and human cells was obtained from these cDNA clones (Sekaly et al., 1985). We describe here complete cDNA clones for the  $\beta$  chains of the DR, DQ and DP antigens. By searching for new  $\beta$  chain clones at a low stringency of hybridization, we isolated an additional cDNA clone, designated DO $\beta$ , which shows good structural similarity with the murine A $\beta$ 2 gene. Several properties of DO $\beta$  distinguish it from the DP $\beta$ , DQ $\beta$  and DR $\beta$  genes.

## **Results and Discussion**

The HLA haplotype-loss mutant 45.1 was derived from the heterozygous B-cell line 721 by  $\gamma$ -irradiation and immunoselection. The remaining haplotype carries the DR1, DQw1 and DPw2 class II specificities. Cytoplasmic poly(A)<sup>+</sup> RNA from mutant 45.1 was used to construct a cDNA library by the procedure of Okayama and Berg (1983). 360 000 independent clones were generated and amplified in culture. A plasmid preparation of the entire library was linearized and size-fractionated by agarose gel electrophoresis. Clones with long cDNA inserts of the DR $\beta$ , DQ $\beta$  and DP $\beta$  chains were localized by blotting part

DO DP DQ DR	AACTCATTCTGAAGGAGGCTGACGATTTTACTGTCICAT TTT TTT CCT IIC ICC GGA AACTCATTCTGAAGGAGGCTGACGATTTTACTGTCICAT TTT TTT CCT IIC ICC GGA TTTTC-GACTGCAGCT-TTTCG-CACTT- <u>-CC</u> AGC T-C [A-G]AG AAG TCT C CAG TTTTATTCTTCTGCCAGGTACATCAGATCCATCAGGTCTGAGGTG-GT-GACT-CC-CTACTTTTCCCCG-C-AAT [A-G]-C- IGG AAG AAG TCT I C-G TAGTTCTCCCCTGAGTG-GA-T-G-CTGCTTCTG-CCC-GGT-CTGTC-IG -CC AGC [A-G]GTG T-T C AAG	-25 62 63 105 77
DO DP DQ DR	S G W V P W V V A L L V N L T R L D S S M T Q G T D S P TCT GGG TGG GTC CCC TGG GTG GCT CTG CTA GTG AGT CTG ACC CGA CTG GAT TCC TCC ATG ACT CAA GGC ACA GAC TCT CCA GT TCT GCC C ACA ACG -C- TTATG GTG CTC A-A G G	4 146 147 189 16 1
DO DP DQ DR	E D F V I Q A K A D C Y F T N G T E K V Q F V V R F I F GAA GAT TTT GTG ATT CAG GCA AAG GCT GAC TGT TAC TTC ACC GAC GGA ACA GGA AAG GTG CAG TTT GTG GTC AGA TTC ATC TTT G AAC C-T T-C0- CG- CAGAC GCG TTT C-G CGC T-C CAGAC G	32 230 225 273 245
DO DP DQ DR	N L E E Y V R F D S D V G M F V A L T K L G Q P D A E Q AAC TTG GAG GAG TAT GTA CGT TTC GAC AGT GAT GTG GGG ATG TTT GTG GCA TTG ACC AAG CTG GGG CAG CCA GAT GCT GAG CAG CG	60 314 309 357 329
DO DP DQ DR	W       N       S       R       L       E       R       S       R       Q       A       V       D       G       V       C       R       H       N       Y       R       L       G       A       P         TGG       AAC       AGC       CGG       GTG       GAG       CGG       GTG       GAG       CAC       CAC       AAG       CTG       CTG       CAC       AAG       CTG       CTG       CAC       AAG       CAC       AAG       CTG       CTG       CAC       AGG       CTG       CCC       CTG       CAC       CAC       AAG       CAC       AAG       CTG       CAC       AGG       CCC       CTG       CCC       CTG       AGG       CTG       CTG       CTG       AGG       CTG       CTG       CTG       CTG       AGG       CTG       CTG	88 398 393 441 413
DO DP DQ DR	F T V G R K V Q P E V T V Y P E R T P L L H Q H N L L H TTC ACT GTG GGG AGA AAA GTG CAA CCA GAG GTG ACA GTG TAC CCA GAG AGG ACC CCA CTC CTG CAC CAG CAT AAT CTG CTG CAC A-GC C CA- C-C CGCGT AG ATT -CC TCC -AAG GGG -C TGCCC T GT- GGG -TC C CAG -G GC CACC CC TCC AG GGG -CC A CCC AT GA CA- C-G CGT G-GT AT T TCA -A AG G-C GCC GT- G AT CA- C-G CG T G-G C GT	116 482 477 525 497
DO DP DQ DR	C S V T G F Y P G D I K I K W F L N G Q E E R A G V M S TGC TCT GTG ACA GGC TTC TAT CCA GGG GAT ATC AAG ATC AAG TGG TTC CTG AAT GGG CAG GAG GAG AGA GCT GGG GTC ATG TCC CACG -ATC AGCT C-A G CGA	144 566 561 609 581
DO DP DQ DR	T G P I R N G D W T F Q T V V M L E M T P E L G H V Y T ACT GGC CCT ATC AGG AAT GGA GAC TGG ACC TTT CAG ACT GTG GTG ATG CTA GAA ATG ACT CCT GAA CTT GGA CAT GTC TAC ACC C AA- TG C-T	172 650 645 693 665
DO DP DQ DR	C L V D H S S L L S P V S V E H R A Q S E Y S H R K M L TGC CTT GTC GAT CAC TCC AGC CTG CTG AGC CCT GTT TCT GTG GAG TGG AGA GCT CAG TCT GAA TAT TCT TGG AGA AAG ATG CTG AAGG A GATTC A-CAGA	200 734 729 777 749
DO DP DQ DR	S G I A A F L L G L I F L L V G I V I Q L R A Q K G Y V AGT GGC ATT GCA GCC TTC CTA CTT GGG CTA ATC TTC CTT CTG GTG GGA ATC GTC ATC CAG CTA AGG GCT CAG AAA GGA TAT GTG -CGA GCGG -GC G-GG C A TG- GGA C T C T CC AGG AGC AG- G G G G-G G G GG- C-T C C-T A AG- AG- AGC A TT C-A CGA A G-C -GG -G G-GGC GC	228 8 18 8 13 8 53 8 33
DO DP DQ DR	R T 9 M S G N E V S R A V L L P 9 S C AGG ACG CAG ATG TCT GGT AAT GAG GTC TCA AGA GCT GTT CTG CTC CCT CAG TCA TGC TAAGGTCCTCACTAAGCTTGCTCTCTCTGGAGCC G-A T-T GCA (TAA) A-AG TIC CT- AG CTG AAA AGAG TGC -TT AGG AAA AGCAT-IGCTGTGTTT-G-TAG-AICTG-CTC-A CTC CT- CTC CTC AGA T-T TAA -TA AG- CTG GTTATCATICTGTGACTG-TGTCCCTG G-A CTT CCA A-A TTC CT- AG- G A GAC -A- AIA GG- A-A ACCTICTGC-AGCTT-GCAGAATGAAAAGCTT	247 909 904 928 924
DO DP DQ DR	TGAAGTAGTGÁTGAGTAGTCTGGGCCCTGGĠTGAGGTAAAĠGACATTCATĠAGGTCAATGŤTCTGGGAATÁACTCTCTTCĊCTGATCCTTĠGAGGAGGCCCĠAACTGATTC GCAG-CCTTCA-C-TCCAAATTGGA-AGTGCT-CCGAAGTGGCTCAG-T-CTATCATTCIGGATTCA-AG-ACT-TTTCT-T-ACTGG-CC CCCA-T-CCCAGC-GCTGT-T-AGCTT-CCCCCGTG-AT-AAGTGCCCTGACGTCACCGTGCAC-C-ACCAC-AGG-ATCTCCTTC-CATCCCCACC-AGGGC-G -CCT-CTTG-CA-TTATTCT-CCA-AAGA-AG-GCT-TCTCAGG-CCTGGTT-C-ACTGG-CA-C-GCAGAAAA-GT-CTCCCT-GTG-CCTAGCTCCTG	019 014 038 034
DO DP DQ DR	İGGAGCTCTGİGTTCTGAGAİCATGCATCTĊCCACCCATCİGCCTTCTCĊCTTCTACGIĞTACATCATTÀATCCCCATTĠCCAAGGGCAİTGTCCAGAAÀCTCCCCTGA CAACCA-G-TCCC-TCTTCT-AGCA-CA-A <u>AAT-AI</u> -AAACAACA-GAGT-TGT-T-C-T-T GCTGTGACC-TCCT-CA-TGA-CCAGAG-CA-TGCTA-A-GG-AGCTGCGTCTACTCAGGCCCCAAGGGGGATT-TGTTTCTGCTCCTC-GA-TGC- TTGCTGAAC-CAGCTGATG-CAG-G-CT-ATCT-CAAT-GTGCCC-T-TGC-TAA-CCGTATGG-C-CC-GTGCAT-TG-ACTCCCTG-A-GACA- 1	129 081 148 144
DO DQ DR	ĠACCTTACTCĊTTCCAGCCCĊAAATCATTTÁCTTTTCTGTĠGTCCAGCCCTACTCCTATAÁGTCATGATCTCCAAAGCTTÍCTGTCTTCCÁACTGCAGTCŤCCACAGTCT AGAGAAG-A-A-GAA-AA-ATT-CCTGAC-T-AGAG-T-T-TT-ATA <u>ATAA</u> -A-GATCC-G-GTT ACA-A-TACAT- <u>ATT-AA</u> TGTTTCAAG-TGGAGT-	239 267 183
DO	ΤCAGAAGACAAATGCTCAGGTAGTCACTGTTTCCTTTTCACTGTTTTTAAAAACCTTTTAITGTCA <u>AATAA</u> ATGGAGATACA	322

**Fig. 1.** Nucleotide sequences of four isotypic  $\beta$  chain cDNA clones. The clones were derived from mRNA of the B-cell line 45.1 (Kavathas *et al.*, 1980) which has a single HLA haplotype with the DR1, DQw1 and DPw2 class II specificities. The sequences have been aligned with respect to their reading frame and identity with DO $\beta$  is indicated by hyphens. The presentation in triplets starts with the most upstream initiation codon (in DQ $\beta$ ) and ends with the last termination codon (in DO $\beta$ ). The translation of DO $\beta$  is given with the single-letter amino acid code. Initiation and termination codons are boxed and poly(A) addition signals are underlined. Triangles indicate the positions of intron-exon boundaries as defined in DR $\beta$ , DQ $\beta$  and DP $\beta$  genes. The intron-exon organization of the DO $\beta$  gene is unknown. The putative exon 6 of DO $\beta$  could be a short intron because it starts with the sequence GT which is a potential splice donor. The length in bp of the GC-tails and of the AT-tails in the four clones were respectively as follows: 10 and 66 in DO $\beta$  #163, 18 and 73 in DP $\beta$  #003, 6 and 69 in DQ $\beta$  #021 and 14 and 126 in DR $\beta$  #008.

of the agarose gel to a nitrocellulose filter and hybridization with previously cloned cDNA fragments. The longest cDNA clones for DR $\beta$ , DQ $\beta$  and DP $\beta$  were extracted from slices of the agarose gel by electroelution, introduced into *Escherichia coli* after circularization with ligase and screened by colony hybridization of replica filters. The same set of filters was subsequently hybridized at a very low stringency and a clone was isolated which had not

been positive with DR $\beta$ , DQ $\beta$  and DP $\beta$  probes at high stringency. This new clone, designated DO $\beta$ , will be described after the DR $\beta$ , DQ $\beta$  and DP $\beta$  clones.

## The $\beta$ chains of HLA-DR1, -DQw1 and DPw2

Over 45 clones for the DR $\beta$  chain were isolated. One of the longest, DR $\beta$  #008, was chosen for sequence analysis (Figure



Fig. 2. Amino acid sequences deduced from four isotypic  $\beta$  chain cDNA clones. Residues identical to DO $\beta$  are indicated by hyphens. The sequences are separated into exons. The first exon encodes the signal sequence and the first four amino acids of the mature  $\beta$  chains. Exons 2 and 3 encode respectively the first and second extracellular domains. Exon 4 encodes the connecting peptide, the hydrophobic transmembrane region and the first cytoplasmic residues. Exons 5 and 6 encode the rest of the cytoplasmic tail. The intron-exon organization of the DO $\beta$  gene is not known. The putative exon 6 of DO $\beta$  could be a short intron because it starts with the sequence GT which is a potential splice donor. The coding sequence for DP $\beta$  stops in exon 5. The DQ $\beta$  genes have lost the use of exon 5 because of a splicing defect (Larhammar *et al.*, 1983b).

1). With 62 bp of 5'-untranslated sequence it is longer than other DR $\beta$  cDNA sequences published to date (Long et al., 1983; Gustafsson et al., 1984; Bell et al., 1985). The first nucleotide of the cDNA insert is 26 bp downstream from the putative TATA box of a sequenced DR $\beta$  gene (Larhammar *et al.*, 1985). The translation of this cDNA clone (Figure 2) defines the complete sequence of an HLA-DR1  $\beta$  chain. Alignment with DR $\beta$  sequences of other specificities (Götz et al., 1983; Long et al., 1983; Gustafsson et al., 1984) confirms that most of the variable residues are clustered in three regions of the first domain, at positions 9-13, 26-38 and 70-74 (not shown). The sequence of a DR $\beta$  cDNA clone from the DR1 B-cell line LG2 was reported (Bell et al., 1985). It differs from our clone #008 by seven amino acids. However, most of these differences can be accounted for by sequencing or editing errors in the LG2 clone (J.Bell, personal communication).

The nucleotide sequence of the longest DQ $\beta$  cDNA clone, DQ $\beta$ #021 (Figure 1), defines the complete sequence of a DQw1  $\beta$ chain (Figure 2). This clone has a very long 5'-untranslated region which starts 19 bp upstream of the first mRNA start mapped by primer-extension (Boss and Strominger, 1984). The 5' end of cDNA clone #021 is within the putative TATA box of the DQ $\beta$ gene (Boss and Strominger, 1984). It cannot be due to a cDNAcloning artifact because the additional 19 bp match perfectly the genomic sequence. The serological specificities that define the DQ antigens are in very strong linkage disequilibrium with DR specificities. For instance, DQw1 is almost always associated with the DR1, DR2 and DRw6 specificities. Biochemical studies have shown that some variability in the  $\beta$  chains, but not the  $\alpha$  chains, existed in the DQw1 antigens associated with DR1, DR2 and DRw6 (Shackelford et al., 1983; Bono and Strominger, 1983; Giles et al., 1985). A logical nomenclature would be to call these different DQw1 antigens respectively DQw1.1, DQw1.2 and DQw1.6. The DQw1.1  $\beta$  chain sequence deduced from clone #021 is indeed guite different from the amino acid sequence determined for the  $\beta$  chain of DQw1.2 (Götz et al. 1983). Most of the variability is in the first domain where the two sequences share 88% homology. Two DQ $\beta$  cDNA clones, pII- $\beta$ -1 and pII- $\beta$ -2, from a DR3,w6 – DQw1,w2 cell line have been sequenced (Schenning et al., 1984). Their respective allelic specificities have been deduced as DQw2 for pII- $\beta$ -1 by comparison with a DQw2 genomic sequence (Boss and Strominger, 1984) and as DQw1.6 for pII- $\beta$ -2 because it probably represents the other allele. The amino acid sequence homology of pII- $\beta$ -1 with our DQw1.1  $\beta$ sequence is only 77% in the first domain. On the other hand, pII- $\beta$ -2 is identical to our DOw1.1  $\beta$  sequence, except for a single silent base change. The difficulty in assigning the proper specificity to clones isolated from heterozygous cell lines emphasizes the usefulness of haplotype-loss mutants such as 45.1.

The nucleotide sequence of clone  $DP\beta \#003$  (Figure 1) defines the complete sequence of a  $DPw2\beta$  chain (Figure 2). The amino acid sequence is identical to a sequence deduced from two overlapping cDNA clones obtained from different cell lines (Kappes *et al.*, 1984). The nucleotide sequences differ by a single silent change.

## A new class II $\beta$ chain gene in the HLA-D region

After identifying DR $\beta$ , DQ $\beta$  and DP $\beta$  clones in the cDNA library by high stringency hybridizations, we hybridized the same set of filters at a very low stringency. A new clone,  $\beta \# 163$ , was isolated. The sequence (Figure 1) defines a new  $\beta$  chain which was given the designation DO. The open reading frame contains 273 codons, flanked by 56 bp of 5'-untranslated region and 447 bp of 3'-untranslated region. The 5'- and 3'-untranslated regions show no significant similarities with their counterparts in the DP, DQ and DR cDNA sequences. The translation of the

Table I. Comparison of nucleotide and amino acid sequences of class II  $\beta$  chain genes

	$\beta_1 \exp(270 \text{ bp})$						$\beta^2$ exon (282 bp)					Trans	Transmembrane exon (111 bp)						
	DO	DP	DQ	DR	Αβ	Εβ	DO	DP	DQ	DR	Αβ2	Αβ	Εβ	DO	DP	DQ	DR	Αβ	Εβ
		63	67	66	64	63		71	73	72	83	67	70		63	76	68	70	64
DP	53		76	76	72	67	67		78	77	71	76	77	43		72	68	67	68
DO	56	60		76	79	68	66	76		78	71	83	74	56	59		78	<u>82</u>	70
	57	65	61		72	79	66	73	71		73	71	83	54	54	78		73	87
AB2		•					79	60	64	64		66	70						_
Aβ	54	63	68	63		74	65	72	86	64	59		70	54	59	81	73		71
Εβ	51	58	59	<u>64</u>	59		54	69	<del>69</del>	83	64	64		51	56	73	<u>89</u>	67	

Numbers represent the percentage of identical positions. Nucleotide comparisons are in the upper right triangle and amino acid comparisons in the lower left triangle. Amino acids encoded by split codons between two exons have been included in the exon which carries two of the three nucleotides (see Figure 2). The murine sequences are from Saito et al. (1983) for E $\beta$  and Larhammar et al. (1983a) for A $\beta$  and A $\beta$ 2. Only the  $\beta$ 2 exon of A $\beta$ 2 has been sequenced. The numbers for the comparisons of related human and murine genes are underlined.



Fig. 3. Mapping of the DO $\beta$  gene within the MHC. 15  $\mu$ g of DNA from the B-cell line 721 and several of its HLA-deletion mutant derivatives were digested with the restriction endonuclease BamHI, electrophoresed in a 0.6% agarose gel and transferred to a nitrocellulose filter. The filter was first hybridized with the cDNA insert of DO $\beta$  (panel a) and was subsequently re-hybridized with an HLA-A3 gene fragment (panel b). The position and size in kb of  $\lambda$ -HindIII fragments are indicated on the left. 721 is an HLAheterozygous B-cell line. The following mutants have been derived from 721 by irradiation and immunoselection (described in Materials and methods): two haplotype-loss mutants which have retained either the DR3 haplotype (127) or the DR1 haplotype (45.1). Further mutations were induced which caused the complete loss of DR antigens (81.3, 102.2, 107, 175), of DQ and DR antigens (82.4, 84.5) and of all DP, DQ and DR antigens (174) Hybridization with the class I gene probe (b) shows that the DNA of mutant 174 is present on the filter and that it has the molecular genotype of the parental haplotype (45.1).

open reading frame reveals a sequence with characteristics of class II  $\beta$  chains (Figure 2). A hydrophobic signal sequence is followed by two disulfide-bonded domains. The cysteine residues, as well as an asparagine which could be a site for glycosylation, are all at conserved positions. Furthermore, several stretches of amino acids which are conserved between DP, DQ and DR  $\beta$ chains are also present in this new sequence. Hydrophobic amino



Fig. 4. Correlations between murine (H-2) and human (HLA) class II genes. The centromere is to the left and the telomere is to the right. The complement genes and the class I genes are telomeric to the HLA-D region. The murine I region has been divided into the I-A and I-E subregions because of a recombination hotspot. A rough molecular map of the class II genes identified so far is shown. Subregions of the HLA-D region have been aligned based on established structural relationships between human and murine genes. Except for the DP subregion (Gorski et al., 1984; Trowsdale et al., 1984) the arrangement of class II genes in HLA-D is not known. The DP and DQ subregions each contain two  $\alpha$  and two  $\beta$  genes. The DR subregion contains one  $\alpha$  and three  $\beta$  genes.

acids which could serve as a transmembrane region are followed by hydrophilic residues and a cytoplasmic tail which is longer than in other class II  $\beta$  chains. The possibility that a short intron is still present in the 3' part of clone #163 cannot be ruled out, even though the cDNA was derived from cytoplasmic RNA. In particular, the putative exon 6 starts with the sequence GT which is a potential splice donor.

The DP $\beta$ , DQ $\beta$  and DR $\beta$  sequences have diverged from each other to the same extent, both at the amino acid and at the nucleotide level (Long *et al.*, 1984a). The DO $\beta$  sequence is more diverged, particularly in the first domain (Table I). Comparisons of the DO $\beta$  sequence with all known  $\beta$  sequences of man and mouse reveal that the highest similarity is with the A $\beta$ 2 gene of the murine I region. In fact the nucleotide sequence identity of 83% in the second domain is the same as that for the DR $\beta$ /E $\beta$ and the DQ $\beta/A\beta$  pairs. It is reasonable to conclude that common ancestral genes for DR $\beta$  and E $\beta$ , for DQ $\beta$  and A $\beta$ , and for DO $\beta$  and A $\beta$ 2 existed before the divergence of man and mouse. Conservation in the non-coding regions of the DO $\beta$  and A $\beta$  genes (Larhammar *et al.*, 1983a) reinforce this conclusion.

To test whether the DO $\beta$  gene is in the human MHC, the DO $\beta$ cDNA clone was hybridized with DNA from several HLAdeletion mutants (Figure 3). Hybridizations with DNA digested with different restriction enzymes suggest that there is a single gene for DO $\beta$  (not shown). Mutant 174 which carries a homozygous deletion of the DP, DQ and DR genes, but retains the centromeric DP pseudogenes and the telomeric complement genes, has lost the DO $\beta$  gene. Therefore DO $\beta$  is most likely encoded in the HLA-D region. Several mutants which have lost



Fig. 5. Expression of the DO $\beta$  gene. RNA from the indicated cell lines was size-fractionated by gel electrophoresis and transferred to DBM paper. Panel a: 30  $\mu$ g poly(A)<sup>+</sup> RNA from each cell line was used. Hybridization was with a 900-bp NaeI-PstI fragment of the DOß cDNA insert nick-translated to >2 × 10<sup>9</sup> c.p.m./µg. Panel b: 4 µg poly(A)<sup>+</sup> RNA from each cell line was used. Hybridization was with the 520-bp and 380-bp PstI fragments from the cDNA insert of the clone DRBII (Long et al., 1983), nicktranslated to ~1 × 10<sup>9</sup> c.p.m./µg. CEM and Molt4 are two T-ALL cell lines. 45.1 and DR4,6 are two EBV-transformed B-cell lines. Raji is a Burkitt's lymphoma line. D1.2 and 2.2.5 are two mutants derived from Raji by immunoselection which have lost the expression of DR and DQ antigen genes. 4312 and 4429 are two SV40-transformed human Xeroderma pigmentosum fibroblast cell lines. 5887, also called 143B, is an osteosarcoma line. Fibroblast lines which were treated (+) or not treated (-) with  $\gamma$ -interferon (IF) are indicated. The Raji RNA sample in panel a suffered some degradation.

expression of the DR antigens or which have homozygous deletions of the DR and DQ genes retain the DO $\beta$  gene. Therefore,  $DO\beta$  maps either between the DP genes and the DQ-DR genes, or between the DQ-DR genes and the complement genes. The first possibility is more likely based on the alignment of structurally related human and murine genes (Figure 4). The  $DO\beta$ gene defines a fourth subregion of the HLA-D region. The sequence conservation between DO $\beta$  and A $\beta$ 2 also suggests that the overall genetic organization of the murine I-region and the HLA-D region may be quite similar. The division of the I-region into only two subregions, I-A and I-E, is somewhat artifactual due to a recombination hotspot (Steinmetz et al., 1982). The Iregion contains four  $\beta$  genes that are as distinct from each other as are the four DO, DP, DQ and DR  $\beta$  genes in man. It is of great interest to search for additional  $\alpha$  genes in the I-region that could be partners for the A $\beta$ 2 and A $\beta$ 3 genes. A potential partner for the DO $\beta$  gene exists. An  $\alpha$  chain gene distinct from DP $\alpha$  $DQ\alpha$  and  $DR\alpha$  genes was identified by Trowsdale *et al.* (1985) on a cosmid clone (DZ $\alpha$ ) and by Inoko et al. (1984) as a short cDNA clone (DT $\alpha$ ). There is, however, no evidence that the DO $\beta$  and the DZ $\alpha$  genes are closely linked or that their products are subunits of the same class II molecule.

## Distinct regulation of the $DO\beta$ gene expression

To test whether the DO $\beta$  gene was expressed in other B-cell lines besides 45.1 and whether its pattern of expression was similar

Table	II.Rate	of	divergence	between	β	chain ge	enes
-------	---------	----	------------	---------	---	----------	------

	Percent corrected divergence								
	Replaceme	nt	Silent						
Comparison	β1 exon	β2 exon	$\beta$ 1 exon	β2 exon					
DO/DP	39	23	109	98					
DO/DQ	34	22	81	80					
DO/DR	33	22	99	105					
DP/DQ	30	16	34	60					
DP/DR	26	17	40	76					
DQ/DR	29	18	34	88					
<b>Α</b> β2/Αβ	_	35	_	84					
Αβ2/Εβ	_	26	_	86					
Αβ/Εβ	32	26	32	99					
DO/Aβ2	_	12	_	58					
DQ/Aβ	23	11	34	64					
DR/Eβ	23	10	44	84					

The percent corrected divergence has been calculated for silent and replacement sites according to Perler *et al.* (1980). A number of 100 means that every potential silent site, or potential replacement site, has changed on average once. The published sequence of  $A\beta 2$  includes only the  $\beta 2$  exon (Larhammar *et al.*, 1983a). Other exons are not included because they are too short to yield meaningful numbers. Numbers for silent divergence which differ between the  $\beta 1$  and  $\beta 2$  exons have been boxed.

to other class II  $\beta$  genes, we hybridized the DO $\beta$  cDNA probe to RNA from various cells (Figure 5a). For comparison, the same RNA samples were hybridized with a DR $\beta$  cDNA probe (Figure 5b). In three different B-cell lines a 1.4-kb mRNA for DO $\beta$  was found but it is ~30 times less abundant than DR $\beta$  mRNA. Two T-ALL DR-negative cell lines (CEM and Molt4) were also negative for DO $\beta$  mRNA. Two mutants (D1.2 and 2.2.5) derived from the B-cell line Raji by mutagenesis and immunoselection against expression of class II antigens (Accolla, 1983) were also tested. These mutants have retained apparently intact class II genes but they lack expression of mRNA for the  $\alpha$  and  $\beta$  chains of class II antigens (Long et al., 1984b). Interestingly, mRNA for DO $\beta$  is still present in the two mutants showing that this gene is not subject to the same regulatory mechanism(s) as are the known class II genes. Induction of class II antigen expression by  $\gamma$ -interferon occurs in a number of cells which are normally class II-negative. The induction is at the level of transcription and *de novo* synthesis of mRNA for both  $\alpha$  and  $\beta$  chains of the DP, DQ and DR antigens takes place (Collins et al., 1984). We tested the effect of  $\gamma$ -interferon on three human fibroblast lines. After 4 days in the presence of  $\gamma$ -interferon, cell surface expression of HLA-DR antigens could be detected in most of the cells by flow cytometry (not shown) and mRNA for DR $\beta$  was present, although at a lower level than in B-cell lines (Figure 5b). mRNA for DO $\beta$  remained undetectable in the  $\gamma$ -interferon-treated cells (Figure 5a). Levels 100-fold lower than that of DR $\beta$  mRNA could have been detected in this experiment. We have also found that two transcripts from the DZ $\alpha$  gene of ~1.4 kb and ~3.5 kb were clearly induced in the same fibroblasts after  $\gamma$ -interferon treatment (not shown, see Materials and methods). It therefore seems unlikely that the DZ $\alpha$  and the DO $\beta$  chains constitute a class II heterodimer. The expression of  $DO\beta$  in class II-negative mutants and the lack of DO $\beta$  inducibility by  $\gamma$ -interferon suggest that DO $\beta$  encodes an atypical class II antigen  $\beta$  chain and that its role could be different.

Table III. CpG dinucleotide frequency in class II genes

Gene	First domain ex	on		Second domain	exon	
	(CpG)obs	(CpG)exp	(CpG)obs (CpG)exp	(CpG)obs	(CpG)exp	(CpG)obs (CpG)exp
DPα	0.029	0.058	0.500	0.021	0.085	0.247
DQα	0.028	0.063	0.444	0.004	0.063	0.063
DRα	0.029	0.053	0.547	0.028	0.074	0.378
DOβ	0.022	0.062	0.355	0.004	0.072	0.056
DPβ	0.076	0.090	0.844	0.021	0.081	0.259
DQβ	0.089	0.094	0.947	0.032	0.085	0.376
DRB	0.074	0.085	0.871	0.014	0.078	0.179

Observed (obs) and expected (exp) frequencies were calculated with the computer program of Queen and Korn (1980). The sequence of DR $\alpha$  is from Lee *et al.* (1982) and of DQ $\alpha$  and DP $\alpha$  from Auffray *et al.* (1984). The  $\beta$  sequences are from Figure 1.

## Evolution of four isotypic class II $\beta$ chain genes

We calculated the rates of replacement nucleotide changes (leading to changes in amino acids) and of silent nucleotide changes (leading to amino acid conservation) between the four isotypic  $\beta$  genes (Table II) according to the method of Perler *et* al. (1980). The replacement changes reflect, as expected, the divergence in amino acid sequences. Silent changes are generally assumed to occur independently of selection for protein structure and the rate of silent divergence is expected to be homogeneous in exons of the same gene. The accumulation of silent changes observed between DO $\beta$  and the other three  $\beta$  genes is reasonable for genes which duplicated before mammalian radiation (Perler et al., 1980) and is similar in the  $\beta 1$  and  $\beta 2$  exons. On the other hand, the silent changes between DP $\beta$ , DQ $\beta$  and DR $\beta$  genes were much lower in the  $\beta$ 1 exons than in the  $\beta$ 2 exons. We found the same result when comparing the murine  $A\beta$ and E $\beta$  genes: the  $\beta$ 1 domains have a 3-fold lower rate of silent changes than the  $\beta 2$  domains.

What mechanism(s) could account for the reduced observed silent divergence in the  $\beta$ 1 domains of some class II genes? Could exchanges of DNA segments between genes explain these results? In several cases of class I MHC genes and in one case of a mutant class II MHC gene evidence has been obtained that through some copy mechanism (double cross-over, gene conversion or other) a DNA segment was transferred from one gene to another (Weiss et al., 1983; McIntyre and Seidman, 1984; Widera and Flavell, 1984; Denaro et al., 1984). If such a copy mechanism operated between the  $\beta$ 1 domain exons of the DP, DQ and DR genes during evolution it would have led to an homogenization of their sequences. As a result the rate of silent changes could appear reduced. Two pieces of evidence argue against this interpretation. (i) The good structural conservation between class II genes of man and mouse (Table I) is incompatible with frequent intra-species genetic exchanges between genes of different subregions. This cross-species structural conservation is in contrast to class I genes which have not maintained structural counterparts between man and mouse or even between strains of mice (Kimball and Coligan, 1983). This loss of interspecies conservation is presumably due to fairly active genetic exchanges between class I genes within a species. (ii) The rate of silent changes between related human and murine genes,  $DR\beta/E\beta$  and  $DQ\beta/A\beta$ , is also reduced as observed for genes within species (Table II). Genetic exchanges cannot occur between man and mouse and obviously cannot account for the observed reduced rates of silent changes. Therefore, it is more likely that the low number of silent changes in the  $\beta$ 1 exons of DP $\beta$ , DQ $\beta$  and DR $\beta$  genes reflects a slower rate of divergence.

Our observation of a low rate of silent divergence in some  $\beta I$ exons could be related to another unusual feature of class II MHC genes. The  $\beta 1$  exons of class II genes in man and mouse are unusual in their content of CpG dinucleotides (Tykocinski and Max, 1984). DNA of higher vertebrates contains fewer CpG dinucleotides than expected on the basis of its C+G content. The ratio of (CpG) observed to (CpG) expected is  $\sim 0.25$  for all sequences examined. However, this CpG suppression was not found in the 5' regions of certain genes including the first two domain exons of MHC class I genes and the first domain exons of class II  $\beta$  chain genes (Tykocinski and Max, 1984). In order to further assess the correlation between these two observations (i.e., low silent divergence and lack of CpG suppression in  $\beta$ 1 exons) we measured the CpG dinucleotide frequency for human class II genes, including DO $\beta$  (Table III). The  $\alpha$  genes, which have a similar rate of silent divergence in the  $\alpha 1$  and  $\alpha 2$  exons (108)  $\pm$  35.6 for  $\alpha$ 1 and 92.7  $\pm$  10.0 for  $\alpha$ 2), show a 2-fold CpG suppression in the  $\alpha 1$  exons and the expected CpG suppression in the  $\alpha^2$  exons. However, unlike the DP $\beta$ , DQ $\beta$  and DR $\beta$  genes the DO $\beta$  gene shows CpG suppression in both exons  $\beta 1$  and  $\beta 2$ . Therefore only those class II gene exons with a low rate of silent divergence have clearly escaped CpG suppression. The significance of this correlation is unclear. Either the high C+G content per se causes a reduced mutation rate (e.g., due to a higher DNA helix stability) or a need to conserve C+G-rich sequences has selected against silent changes.

A reduced rate of silent changes between class I genes of man and mouse, which was ~60% of the rate observed between other genes of man and mouse, had been noted (Hayashida and Miyata, 1983) but it was less pronounced than the 2- to 3-fold difference described here. Gustafsson *et al.* (1984) analyzed the divergence between alleles of class II  $\alpha$  and  $\beta$  genes. Although the number of silent changes between alleles is small they concluded from an analysis of 13 sequences that the first and the second domains of alleles accumulated silent changes at the same rate. The reduced observed rate of silent changes is thus unique to isotypic forms of DP, DQ and DR  $\beta$  genes. Whatever mechanism accounts for this phenomenon and whatever the reason for its existence it is significant that the DO $\beta$  gene was not involved and has evolved independently from the DP $\beta$ , DQ $\beta$  and DR $\beta$  genes in this respect.

## Conclusions

The four complete cDNA sequences encoding isotypic  $\beta$  chains from a single HLA haplotype define a new subregion of HLA-D, in addition to the known DP, DQ and DR subregions. The new DO $\beta$  cDNA encodes an apparently expressible class II  $\beta$ chain. By sequence analysis DO $\beta$  was found to be the structural counterpart of the murine  $A\beta 2$  gene and to be more distantly related to other human class II  $\beta$  genes. An important objective is to define DO at the protein level and to determine which  $\alpha$ chain, if any, is associated with the  $\beta$  chain. The  $\alpha$  chain encoded in the DZ $\alpha$  gene is not a likely candidate because DZ $\alpha$  RNA is not coordinately expressed with DO $\beta$  mRNA. A molecular linkage map of class II genes in HLA-D would facilitate the assignment of the proper  $\alpha$  and  $\beta$  gene pairs. The evolution of DO $\beta$  was distinct from other class II  $\beta$  genes with respect to the rate of silent changes and the frequency of CpG dinucleotides in the first domain exon. Furthermore  $DO\beta$  expression is regulated independently of other class II genes. These two observations suggest strongly that it may be involved in functions that are distinct from those of the typical 'immune response' genes. Biochemical and cellular analyses are necessary to establish the existence and the function of this putative antigen. The complete sequence of the DO $\beta$  chain and the availability of an expressible cDNA clone provide means to approach these questions.

### Materials and methods

## General methods

Most manipulations of nucleic acids followed standard techniques (Maniatis *et al.*, 1982). Small plasmid preparations were obtained by the boiling method (Holmes and Quigley, 1981). Large amounts of plasmid were prepared by the boiling lysis followed by two consecutive bandings in CsCl with ethidium bromide. Restriction endonucleases were purchased from New England Biolabs, Boehringer Mannheim and International Biotechnologies Inc., and were used according to the supplier's instructions. Isotopes were purchased from Amersham and from New England Nuclear.

#### Construction of the cDNA library

The procedure of Okayama and Berg (1983) was used. This procedure yields cDNA clones which are expressible in eukaryotic cells because the vector pcD provides the SV40 early promoter and a splice site upstream of the cDNA inserts, and termination sequences downstream. We have removed the splice site from the original vector because it contained two alternative acceptor sites, one of which, if used, results in an AUG initiation codon upstream of the cDNA insert. The possibility of producing fused proteins with an abnormal N-terminal sequence was undesirable in view of future functional studies. The splice site was replaced by a polylinker as follows. pL1 was linearized with XhoI, treated with nuclease S1 (Bethesda Research Laboratories), digested with EcoRI and the large fragment was purified by agarose gel electrophoresis, electroelution and DEAE-cellulose chromatography (Clarkson et al., 1978). The plasmid pUC12 (P.L. Biochemicals) was linearized with HindIII, filled with Klenow DNA polymerase (Boehringer Mannheim), digested with EcoRI and the small polylinker fragment was purified by polyacrylamide gel electrophoresis, diffusion and DEAEcellulose chromatography. The pL1- and pUC12-derived fragments were ligated with T4 DNA ligase (New England Biolabs) to produce the plasmid pL2. The structure of resulting cDNA clones is shown in Figure 6. The lack of splice site in the transcription unit is apparently not a problem because cell surface expression of antigen was detected in cells transfected with these cDNA clones (Sekaly et al., 1985).

Cytoplasmic RNA from the HLA-hemizygous mutant B cell line 45.1 (Kavathas *et al.*, 1980) was prepared by centrifugation through CsCl as described (Long *et al.*, 1982a). Poly(A)<sup>+</sup> RNA was enriched by chromatography on oligo(dT)-cellulose (Type 3, Collaborative Research). The cDNA library was constructed exactly as described by Okayama and Berg (1982, 1983) except for the use of plasmid pL2 and the following modifications.

(i) The length of the T-tails added on *Kpn*I-digested pcDV1 was controlled by limiting the concentration of dTTP in the reaction. 80 pmol of dTTP were added per pmol of 3' end of pcDV1 vector (at ~0.1 mg/ml) and the incubation was for 30 min at 37°C in 0.11 M potassium cacodylate, 30 mM Tris-cacodylate, 0.1 mM dithiothreitol, 0.1 mM EDTA, 10 mM CoCl<sub>2</sub>, pH 7.4 and 200 units/ml of terminal deoxynucleotidyltransferase (P.L. Biochemicals). The exact length of T-tails obtained in preparative tailing reactions was measured. Aliquots of the reactions were labeled with [ $\alpha$ -<sup>32</sup>P]dideoxy-ATP and terminal deoxynucleotidyltransferase, digested with *Xho*I and electrophoresed in 8% polyacrylamide sequencing gels. The *KpnI-XhoI* fragment (35 bp) without tail was run in parallel. Only vector-primers with T-tails between 60 and 80 residues were used and purified as described (Okayama and Berg, 1982).

(ii) The plasmid pL2 was linearized with *Sacl* before tailing with dGTP. (iii) 60  $\mu$ g of poly(A)<sup>+</sup>-enriched RNA was annealed with 8  $\mu$ g of T-tailed vector-primer.



**Fig. 6.** Map of the cDNA cloning and expression vector. The vector-primer used to construct the cDNA library is pcDV1 from Okayama and Berg (1983). The linker fragment derived from pL1 (Okayama and Berg, 1983) was modified as described in Materials and methods. Briefly, the splice site was replaced by a polylinker fragment from the plasmid pUC12.



Fig. 7. Sequencing strategy for the four isotypic  $\beta$  chain clones. Bars and open arrows represent sequences determined by the chain termination method, crosses and closed arrows those determined by chemical cleavage. The closed and open bars at the ends represent the GC-tails and the AT-tails, respectively. Except for the AT-tails, all the sequences rely on at least two independent determinations and every restriction site has been read through. **A**, clone 45.1 DQ $\beta$  #163; **B**, clone 45.1 DP $\beta$  #003; **C**, clone 45.1 DQ $\beta$  #021; **D**, clone 45.1 DR $\beta$  #008.

(iv) The digestion with *Hind*III was an important parameter. We calibrated it with aliquots of the preparation which were then carried through the last steps of the procedure. The length of cDNA inserts in a random sample of plasmids was measured. The amount of *Hind*III which resulted in the highest number of *E. coli* transformants which carried plasmids with long inserts was chosen to prepare the rest of the library.

(v) T4 DNA ligase (New England Biolabs) was used, instead of *E. coli* DNA ligase, in 50 mM Tris-HCl (pH 7.5), 10 mM MgCl<sub>2</sub>, 10 mM DTT and 1 mM ATP.

(vi) Transformation of E. coli was with frozen competent DH1 cells exactly

#### C.Tonnelle, R.DeMars and E.O.Long

as described by Hanahan (1983). The material used in the final transformation corresponded to an initial amount of 2.3  $\mu$ g of vector-primer and yielded 360 000 independent clones ( $\sim 1.6 \times 10^5$  clones/ $\mu$ g vector-primer).

(vii) Total plasmid DNA was prepared from 2 l of the amplified library and was linearized with *ClaI*. Size selection and localization of long class II cDNA clones was carried out as described (Gunning *et al.*, 1983). Two size fractions containing the longest cDNA clones for the  $\alpha$  and  $\beta$  chains of DP, DQ and DR antigens were circularized with T4 DNA ligase, introduced into *E. coli* strain HB101 and spread over nitrocellulose filters for screening and storage (Hanahan and Meselson, 1980).

## Hybridizations

Bacterial colonies on nitrocellulose filters were screened by the procedure of Grunstein and Hogness (1975). After a wash in  $4 \times SSC$  (SSC = 0.15 M NaCl, 15 mM sodium citrate) and 5  $\times$  Denhardt (Denhardt = 0.02% bovine serum albumin, 0.02% polyvinylpyrrolidone, 0.02% Ficol) (Denhardt, 1966) at 55°C, filters were hybridized overnight at 65°C in sealed plastic bags with 50  $\mu$ g/ml denatured herring DNA, 25 mM sodium phosphate (pH 7.0), 0.6 M NaCl, 5 mM EDTA,  $1 \times \text{Denhardt}, 0.1\%$  SDS and  $4 \times 10^5$  c.p.m./ml of denatured <sup>32</sup>P-labeled DNA probe. Filters were washed at 65°C for 30 min each consecutively in  $4 \times 2 \times 2$  $1 \times 0.5 \times \text{and } 0.1 \times \text{SSC}$ . The DNA probes used were: a 520-bp and a 380-bp PstI fragment from the cDNA clone DRBII (Long et al., 1983), a full-length PstI insert of a DCß cDNA clone (Long et al., 1983) and a 850-bp PstI fragment of cDNA clone SB $\beta$  (Long et al., 1984a). The fragments were purified by two consecutive agarose gels. For low stringency screening, filters were hybridized at 55°C as described above except that NaCl was raised to 0.9 M and that poly(A) and poly(G) (Boehringer Mannheim) were included at 50 µg/ml each. The probe was a mixture of the 520-bp PstI DR $\beta$  and the 850-bp PstI SB $\beta$ fragments. Washes were at 55°C in  $6 \times$  SSC and at room temperature in  $2 \times SSC$ 

Southern blots (Southern, 1975) were performed as described (Wake *et al.*, 1982). Probes were a full-length *Bam*HI insert of the cDNA clone DO $\beta$  # 163 or a 1.8-kb *Hind*III-*BgI*II fragment containing the first three exons of the HLA-A3 class I gene (Cowan *et al.*, 1985).

For Northern blots (Alwine *et al.*, 1979), glyoxal-treated RNA was sizefractionated in 1.5% agarose gels (McMaster and Carmichael, 1977) and transferred to diazotized paper (Schleicher and Schuell). Hybridization and washes were as described (Long *et al.*, 1982b).

Two synthetic oligonucleotides overlapping by 10 bp (5' ATTGGGAGCTC-CAAGTGCCTATTCCACCAC and 5' TAGGGTCTCCATGGCATCTGGTGG-TGGAAT) were derived from the sequence of the fourth exon of the DZ $\alpha$  gene (Inoko et al., 1984) and were labeled as follows. 10 ng of each were mixed in 7.5 µl containing 2 µl of a 10-fold concentrated Klenow polymerase buffer (New England Nuclear), incubated at 90°C for 5 min and allowed to cool slowly to room temperature. In a final volume of 20  $\mu$ l, dithiothreitol was added to 2.5 mM, dGTP and dTTP to 50  $\mu$ M, [ $\alpha$ -<sup>32</sup>P]dCTP and [ $\alpha$ -<sup>32</sup>P]dATP to 1  $\mu$ M (100  $\mu$ Ci each). Three units of Klenow DNA polymerase (Bethesda Research Laboratories) were added and the sample was incubated at 23°C for 20 min and at 30°C for 30 min. A chase reaction was performed with an addition of all four dNTPs at 33 µM for 20 min at 30°C. After a phenol extraction the 50-bp double-stranded fragment was purified by chromatography over Sephadex G-50. The specific activity was ~1  $\times$  10<sup>10</sup> d.p.m./µg. Hybridization to Northern blots was as described above. The final wash was at 65°C for 30 min in 37.5 mM NaCl, 1 mM sodium phosphate (pH 7.0), 1 mM EDTA and 0.1% sodium lauryl sulfate.

#### DNA sequencing, analysis and synthesis

Nucleotide sequences were determined by the M13 chain termination method (Sanger *et al.*, 1980) and by the chemical degradation method (Maxam and Gilbert, 1980). Strategies are shown in Figure 7. A computer program that performs the calculations of silent and replacement changes according to Perler *et al.* (1980) was kindly provided by Dr. G.Gutman. Oligonucleotides were synthesized with the Model 380A DNA synthesizer (Applied Biosystems).

#### Cell lines

B cell lines were maintained in RPMI-1640 supplemented with 2 mM glutamine and 20% fetal calf serum (Hazleton Dutchland, Inc., Denver, PA 17517). Human fibroblasts from the Human Genetic Mutant Repository (Camden, NJ 08103) were maintained in Eagle's modified essential medium supplemented with 2 mM glutamine and 15% fetal calf serum. For  $\gamma$ -interferon induction, cells were incubated during 4 days with 500 units/ml of recombinant  $\gamma$ -interferon (a kind gift of Dr. N.Sarvar, Meloy Laboratories, Springfield, VA 22151). Cells in exponential growth were washed twice in phosphate-buffered saline and frozen at  $-70^{\circ}$ C prior to RNA extractions. The Epstein-Barr virus-transformed B lymphoblastoid cell line (LCL) 721 expresses the class II haplotypes DPw2, DQw1, DR1/DPw4, DQw2, DR3. Irradiation of LCL 721 with  $\gamma$ -rays followed by negative immunoselection with complement and appropriate antibodies resulted in the isolation of mutant 45.1 which expresses only the DR1-containing haplotype (Kavathas *et al.*, 1980) and mutant 127 which expresses only the DR3-containing haplotype (DeMars *et al.*, 1983). Mutant 45.1 has a large deletion involving the short arm of one chromosome 6 (Kavathas *et al.*, 1980). The presence of only one copy of the remaining MHC haplotype in 45.1 is further attested to by Southern blotting autoradiography with probes for MHC genes and by the readiness with which mutants that have lost expression of the remaining MHC alleles can be isolated (i.e., the mutation rate corresponds to single rather than double mutational events). Further  $\gamma$ -ray induced mutations in 45.1, or in the similar haplotype-loss mutant 134, caused the complete loss of DR antigens (mutants 81.3, 102.2, 107 and 175), of DR and DQ antigens (mutants 82.4 and 84.5) (DeMars *et al.*, 1983) and of DR, DQ and DP antigens (mutant 174) (DeMars *et al.*, 1984). All of the mutants named above have homozygous deletions of some class II genes (Auffray *et al.*, 1983; ROux-Dosseto *et al.*, 1983; R.DeMars, unpublished data). The homozygous deletion in 174 has one breakpoint near the DP $\alpha$  pseudogene, which it retained, and the other breakpoint between the class II region and the complement gene cluster, which it retained (Whitehead *et al.*, 1985). All known class II genes other than the DP pseudogenes are deleted in 174.

## Acknowledgements

We are very grateful to Dr. E.Max for suggesting to us the CpG dinucleotide measurements and for useful comments on the manuscript, Dr. R.Sekaly who provided us with  $\gamma$ -interferon-treated fibroblasts, Dr. E.Cowan for the HLA-A3 gene fragment, Dr. N.Sarvar for recombinant  $\gamma$ -interferon, Dr. S.Carrel for the CEM and Molt4 T cells, Dr. G.Gutman for his computer program and Dr. T.J.Kindt for support and stimulating discussions. We thank W.Pugh for his excellent technical assistance and G.Shaw for her kind editorial assistance. C.T. was supported by INSERM and by a fellowship from CNRS-NIH.

### References

- Accolla, R.S. (1983) J. Exp. Med., 157, 1053-1058.
- Alwine, J.C., Kemp, D.J., Parker, B.A., Reiser, J., Renart, J., Stark, G.R. and Wahl, G.M. (1979) Methods Enzymol., 68, 220-242.
- Auffray, C., Kuo, J., DeMars, R. and Strominger, J.L. (1983) Nature, 304, 174-177.
  Auffray, C., Lillie, J.W., Arnot, D., Grossberger, D., Kappes, D. and Strominger, J.L. (1984) Nature, 308, 327-333.
- Auffray, C. and Strominger, J.L. (1985) Adv. Human Genet., in press.
- Bell, J.I., Estess, P., St. John, T., Saiki, R., Watling, D.L., Erlich, H.A. and McDevitt, H.O. (1985) Proc. Natl. Acad. Sci. USA, 82, 3405-3409.
- Bono, M.R. and Strominger, J.L. (1983) *Immunogenetics*, **18**, 453-459.
- Boss, J.M. and Strominger, J.L. (1984) Proc. Natl. Acad. Sci USA, 81, 5199-5203.
- Clarkson, S.G., Kurer, V. and Smith, H.O. (1978) Cell, 14, 713-724.
- Collins, T., Korman, A.J., Wake, C.T., Boss, J.M., Kappes, D.J., Fiers, W., Ault, K.A., Gimbrone, M.A., Strominger, J.L. and Pober, J.S. (1984) Proc. Natl. Acad. Sci. USA, 81, 4917-4921.
- Cowan, E.P., Coligan, J.E. and Biddison, W.E. (1985) Proc. Natl. Acad. Sci. USA, 82, 4490-4494.
- DeMars, R., Chang, C.C. and Rudersdorf, R.A. (1983) Hum. Immunol., 8, 123-139.
- DeMars, R., Chang, C.C., Shaw, S., Reitnauer, P.J. and Sondel, P.M. (1984) Hum. Immunol., 11, 77-89.
- Denaro, M., Hämmerling, U., Rask, L. and Peterson, P.A. (1984) *EMBO J.*, 3, 2029-2032.
- Denhardt, D. (1966) Biochem. Biophys. Res. Commun., 23, 641-652.
- Giles, R.C., DeMars, R., Chang, C.C. and Capra, J.D. (1985) Proc. Natl. Acad. Sci. USA, 82, 1776-1780.
- Gorski, J., Rollini, P., Long, E.O. and Mach, B. (1984) Proc. Natl. Acad. Sci. USA, 81, 3934-3938.
- Götz, H., Kratzin, H., Thinnes, F.P., Yang, C., Kruse, T., Pauly, E., Kölbel, S., Egert, G., Wernet, P. and Hilschmann, N. (1983) *Hoppe-Seyler's Z. Physiol. Chem.*, 364, 749-755.
- Grunstein, M. and Hogness, D. (1975) Proc. Natl. Acad. Sci. USA, 72, 3961-3965.
- Gunning, P., Ponte, P., Okayama, H., Engel, J., Blau, H. and Kedes, L. (1983) Mol. Cell. Biol., 3, 787-795.
- Gustafsson, K., Wiman, K., Emmoth, E., Larhammar, D., Böhme, J., Hyldig-Nielsen, J.J., Ronne, H., Peterson, P.A. and Rask, L. (1984) *EMBO J.*, **3**, 1655-1661.
- Hanahan, D. (1983) J. Mol. Biol., 166, 557-580.
- Hanahan, D. and Meselson, M. (1980) Gene, 10, 63-67.
- Hayashida, H. and Miyata, T. (1983) Proc. Natl. Acad. Sci. USA, 80, 2671-2675.
- Holmes, D.S. and Quigley, M. (1981) Anal. Biochem., 114, 193-197.
- Inoko, H., Ando, A., Kimura, M., Ogata, S. and Tsuji, K. (1984) in Albert, E.D., Baur, M.P. and Mayr, W.R. (eds.), *Histocompatibility Testing 1984*, Springer-Verlag, NY, pp. 559-564.
- Kappes, D.J., Arnot, D., Okada, K. and Strominger, J.L. (1984) EMBO J., 3, 2985-2993.
- Kavathas, P., Bach, F.H. and DeMars, R. (1980) Proc. Natl. Acad. Sci. USA, 77, 4251-4255.
- Kimball, E.S. and Coligan, J.E. (1983) Contemp. Top. Mol. Immunol., 9, 1-63. Larhammar, D., Hämmerling, U., Denaro, M., Lund, T., Flavell, R.A., Rask, L.

and Peterson, P.A. (1983a) Cell, 34, 179-188.

- Larhammar, D., Hyldig-Nielsen, J.J., Servenius, B., Andersson, G., Rask, L. and Peterson, P.A. (1983b) *Proc. Natl. Acad. Sci. USA*, **80**, 7313-7317.
- Larhammar, D., Servenius, B., Rask, L. and Peterson, P.A. (1985) Proc. Natl. Acad. Sci. USA, 82, 1475-1479.
- Lee, J.S., Trowsdale, J., Travers, P.J., Carey, J., Grosveld, F., Jenkins, J. and Bodmer, W.F. (1982) *Nature*, **299**, 750-752.
- Long, E.O., Gross, N., Wake, C.T., Mach, J.P., Carrel, S., Accolla, R.S. and Mach, B. (1982a) *EMBO J.*, 1, 649-654.
- Long, E.O., Wake, C.T., Strubin, M., Gross, N., Accolla, R.S., Carrel, S. and Mach, B. (1982b) Proc. Natl. Acad. Sci. USA, 79, 7465-7469.
- Long, E.O., Wake, C.T., Gorski, J. and Mach, B. (1983) EMBO J., 2, 389-394.
- Long, E.O., Gorski, J. and Mach, B. (1984a) Nature, 310, 233-235.
- Long, E.O., Mach, B. and Accolla, R.S. (1984b) *Immunogenetics*, **19**, 349-353. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) *Molecular Cloning. A Laboratory Manual*, published by Cold Spring Harbor Laboratory Press, NY.
- Maxam, A. and Gilbert, W. (1980) Methods Enzymol., 65, 499-560.
- McIntyre,K.R. and Sediman,J.G (1984) Nature, 308, 551-553.
- McMaster,G.K. and Carmichael,G.G. (1977) Proc. Natl. Acad. Sci. USA, 74, 4835-4838.
- Okayama, H. and Berg, P. (1982) Mol. Cell. Biol., 2, 161-170.
- Okayama, H. and Berg, P. (1983) Mol. Cell. Biol., 3, 280-289.
- Perler, F., Efstratiadis, A., Lomedico, P., Gilbert, W., Kolodner, R. and Dodgson, J. (1980) Cell, 20, 555-566.
- Queen, C.L. and Korn, L.J. (1980) Methods Enzymol., 65, 595-609.
- Qvigstad, E., Moen, T. and Thorsby, E. (1984) Immunogenetics, 19, 455-460.
- Robinson, M.A., Long, E.O., Johnson, A.H., Hartzman, R.J., Mach, B. and Kindt, T.J. (1984) J. Exp. Med., 160, 222-238.
- Roux-Dosseto, M., Auffray, C., Lillie, J.W., Boss, J.M., Cohen, D., DeMars, R., Mawas, C., Seidman, J.G. and Strominger, J.L. (1983) Proc. Natl. Acad. Sci. USA, 80, 6026-6040.
- Saito, H., Maki, R.A., Clayton, L.K. and Tonegawa, S. (1983) Proc. Natl. Acad. Sci. USA, 80, 5520-5524.
- Sanger, F., Coulson, A.R., Barrell, B.G., Smith, A.J.H. and Roe, B.A. (1980) J. Mol. Biol., 143, 161-178.
- Schenning, L., Larhammar, D., Bill, P., Wiman, K., Jonsson, A.K., Rask, L. and Peterson, P.A. (1984) *EMBO J.*, **3**, 447-452.
- Sekaly, R.P., Tonnelle, C., DeMars, R. and Long, E.O. (1985) in Streilein, J.W., Ahmad, F., Black, S., Blomberg, B. and Voellmy, R.W. (eds.), Advances in Gene Technology, Molecular Biology of the Immune System, Cambridge University Press, NY, pp. 293-294.
- Shackelford, D.A., Eibl, B. and Strominger, J.L. (1983) Immunogenetics, 18, 625-637.
- Southern, E. (1975) J. Mol. Biol., 98, 503-517.
- Steinmetz, M., Minard, K., Horvath, S., McNicholas, J., Frelinger, J., Wake, C., Long, E., Mach, B. and Hood, L. (1982) *Nature*, 300, 35-42.
- Swain, S.L. (1983) Immunol. Rev., 74, 129-142.
- Termijtelen, A., Khan, P.M., Shaw, S. and van Rood, J.J. (1983) *Immunogenetics*, 18, 503-512.
- Trowsdale, J., Kelly, A., Lee, J., Carson, S., Austin, P. and Travers, P. (1984) Cell, 38, 241-249.
- Trowsdale, J., Young, J.A.T., Kelly, A.P., Austin, P.J., Carson, S., Meunier, H., So, A., Erlich, H.A., Spielman, R.S., Bodmer, J. and Bodmer, W.F. (1985) *Immunol. Rev.*, 85, 5-43.
- Tykocinski, M.L. and Max, E.E. (1984) Nucleic Acids Res., 12, 4385-4396.
- Wake, C.T., Long, E.O. and Mach, B. (1982) Nature, 300, 372-374.
- Wake, C.T., Widera, G. and Flavell, R.A. (1985) in Streilein, J.W., Ahmad, F., Black, S., Blomberg, B. and Voellmy, R.W. (eds.), Advances in Gene Technology, Molecular Biology of the Immune System, Cambridge University Press, NY, pp. 33-36.
- Weiss, E.H., Mellor, A., Golden, L., Fahrner, K., Simpson, E., Hurst, J. and Flavell, R.A. (1983) *Nature*, **301**, 671-674.
- Whitehead, A.S., Colten, H.R., Chang, C.C. and DeMars, R. (1985) J. Immunol., 134, 641-643.
- Widera, G. and Flavell, R.A. (1984) EMBO J., 3, 1221-1225.
- Zinkernagel, R.M. and Doherty, P.C. (1979) Adv. Immunol., 277, 51-177.

#### Received on 29 July 1985

#### Note added in proof

The complete sequence of the murine  $A\beta 2$  gene has been reported (Larhammar *et al.*, *J. Biol. Chem.*, in press). Its alignment with DO $\beta$  shows 83% identity in the  $\beta 1$  exon and 79% identity in the transmembrane exon. Transcription of the  $A\beta 2$  gene is not inducible by  $\gamma$ -interferon in macrophage cells [Wake and Flavell (1985) *Cell*, **42**, in press]. The cDNA clone DT $\alpha$  (Inoko *et al.*, 1984) has been renamed DO $\alpha$  [Inoko *et al.*, *J. Immunol.* (1985) **135**, 2156-2159].