MHC class II regulatory factor RFX has a novel DNA-binding domain and a functionally independent dimerization domain

Walter Reith, Carlos Herrero-Sanchez, Michel Kobr, Paolo Silacci, Christine Berte, Emanuelle Barras, Sylvia Fey, and Bernard Mach

Department of Microbiology, University of Geneva Medical School, Geneva, Switzerland

The regulation of MHC class II gene expression controls T-cell activation and, hence, the immune response. Among the nuclear factors observed to bind to conserved DNA sequences in human leukocyte antigen (HLA) class II gene promoters, RFX is of special interest: Its binding is defective in congenital HLA class II deficiency, a disease of class II gene regulation. The cloning of an RFX cDNA has allowed us to show by transfection of a plasmid directing the synthesis of antisense RFX RNA that RFX is a class II gene regulatory factor. RFX is a novel 979-amino-acid DNA-binding protein that contains three structurally and functionally separate domains. The 91-amino-acid DNA-binding domain is distinct from other known DNA-binding motifs but may be distantly related to the helix-loop-helix motif. The most striking property of RFX is that it can bind stably to the class II X box as either a monomer or a homodimer and that the domain responsible for dimerization is distant from and functionally independent of the DNA-binding domain. This distinguishes RFX from other known dimeric DNA-binding proteins. It also implies that an RFX homodimer has two potential DNA-binding sites. We therefore speculate that RFX could form a DNA loop by cross-linking the two X-box sequences found far apart upstream of MHC class II genes.

[*Key Words:* Major histocompatibility complex; class II X box; regulatory factor RFX]

Received April 9, 1990; revised version accepted June 27, 1990.

Class II major histocompatibility complex (MHC) antigens are heterodimeric membrane glycoproteins comprised of α and β chains. They play a key role in the immune response: They present antigen to T lymphocytes (Benacerraf 1981; Schwartz 1985; Germain 1986) and control the generation of the T-cell repertoire by selection events occurring in the thymus (Marrack and Kappler 1988). In man, genes encoding the α and β chains of the HLA-DP, HLA-DQ, and HLA-DR class II antigens are clustered in the D region of the MHC (Mach et al. 1986).

Class II genes exhibit a complex pattern of expression that is controlled primarily at the level of transcription. In general, the different α - and β -chain genes are coregulated. Expression is mainly restricted to a small number of immunocompetent cells such as B and T lymphocytes, macrophages, and other antigen-presenting cells (Hammerling 1976). It can be modulated in a cell-dependent manner by a variety of positive and negative stimuli. For example, up-regulation can be obtained with gamma interferon (IFN- γ) in macrophages (Paulnock-King et al. 1985; Woodward et al. 1989), IL4 in B cells (Rousset et al. 1988), and mitogens or antigens in T cells (Diedrichs and Schendel 1989). Within the B-cell

lineage, expression is strictly regulated according to the stage of differentiation (Latron et al. 1988; Dellabona et al. 1989). Finally, in many class II-negative cells, expression can be induced with IFN- γ (Collins et al. 1984; Blonar et al. 1988; Amaldi et al. 1989). To elucidate the mechanisms mediating this complex pattern of expression, it will ultimately be necessary to characterize regulatory factors that bind to *cis*-acting sequences controlling MHC class II gene transcription.

All class II gene promoters from man, mouse, and other species contain three conserved sequence motifs referred to as the W(or Z), X, and Y boxes (Saito et al. 1983; Kelly and Trowsdale 1985; Miwa et al. 1987; Tsang et al. 1988). These sequences are *cis*-acting elements playing a central role in class II gene expression (Dorn et al. 1987b; Sherman et al. 1987, 1989; Basta et al. 1988; Koch et al. 1988; Sakurai and Strominger 1988; Tsang et al. 1988, 1990; Dedrick and Jones 1990). We and others have identified nuclear factors binding specifically to these sequences (Dorn et al. 1987a,b; Miwa et al. 1987; Reith et al. 1988, 1989; Sherman et al. 1989; Dedrick and Jones 1990; Kobr et al. 1990; Tsang et al. 1990). Among these, the factor we have called RFX is of special interest because there is evidence that it is in-

volved in the regulation of class II gene expression. RFX is a nuclear protein that binds specifically, although with variable affinity, to the X boxes of the human DRA, DPA, DQA, DRB1, and DRB3 genes, as well as the mouse E α and human class II-associated invariant chain gene (Reith et al. 1988, 1989; Kobr et al. 1990). Binding of RFX is defective in nuclear extracts from B-cell lines derived from patients suffering from class II-deficient combined immunodeficiency (CID), a disease in which the absence of class II gene expression in all tissues has been shown to be due to a defect in a *trans*-acting regulatory factor (de Prval et al. 1985, 1988; Lisowska-Grospierre et al. 1985; Reith et al. 1988, 1989; Griscelli et al. 1989). Defective binding of RFX in this disease thus indicates that RFX is a regulatory factor essential for class II gene expression.

We recently isolated an RFX cDNA clone (Reith et al. 1989) by a direct expression screening procedure (Singh et al. 1988; Vinson et al. 1988). The recombinant phage (λ 9) was shown to contain a partial cDNA insert (RFX9) that encodes a truncated protein exhibiting several characteristic properties of RFX (Reith et al. 1989). Here, we present direct evidence that this cDNA clone encodes RFX and show that RFX is a class II gene regulatory factor. The sequence of RFX was derived from overlapping cDNA clones corresponding to the entire mRNA. Structural and functional analysis of recombinant RFX shows that it has a novel DNA-binding domain and a distinct domain responsible for dimerization. In the case of RFX, therefore, DNA binding and dimerization are independent functions and an RFX homodimer has two potential binding sites.

Results

RFX9 cDNA encodes RFX

The protein encoded by the RFX9 cDNA clone binds specifically to the X box, has essentially the same contact points within the X box as RFX, and shares with RFX a characteristic gradient of affinity for the promoters of the DRA, DPA, and DQA genes (Reith et al. 1989; Kobr et al. 1990). Because a given DNA sequence can frequently bind several different nuclear factors, we provide here additional evidence confirming that RFX9 indeed encodes RFX.

The plasmid pRFX9 was constructed by cloning RFX9 into the prokaryotic expression vector pT7-7 (Tabor and Richardson 1985). pRFX9 was introduced into a bacterial strain allowing large-scale synthesis of the recombinant protein. The recombinant protein was then purified and used to produce a rabbit antiserum. Two approaches were followed to characterize the reactivity of this antiserum with native RFX in nuclear extracts. First, affinity-purified antibodies were tested for their ability to affect binding of RFX to an X-box oligonucleotide (DRA-X1) in an electrophoretic mobility shift assay (EMSA) (Fig. 1A). The results show that formation of the retarded band due to binding of RFX is inhibited by antibodies directed against the recombinant protein, whereas nonspecific anti- β -galactosidase antibodies

Figure 1. Clone RFX9 encodes the class II gene regulatory factor RFX. (A) Effect of antibodies directed against the recombinant protein on binding of RFX. 32 P-labeled DRA-X1 was incubated with a nuclear extract from Blanco B cells, and binding of RFX and NF-S was analyzed by EMSA. Binding reactions were supplemented with 10% FCS (lane 1), 1 μ g of anti- β -galactosidase antibodies in 10% FCS (lane 2), or 1 μ g of affinity-purified anti-rRFX antibodies in 10% FCS (lane 3). Positions of free DNA (X1) and DNA bound by RFX, NF-S, and a nonspecific binding activity (n.s.) are indicated. (B) Immunoprecipitation of RFX-DNA complexes by antiserum directed against the recombinant protein. DRA-X1 and DRA-Xr oligonucleotides labeled with 32 P were mixed (lane 0), incubated with protein extracts, and subjected to immunoprecipitation (lanes 1 to 5). The presence of labeled DNA in the precipitates was analyzed by polyacrylamide gel electrophoresis. Protein extracts were a bacterial extract containing rRFX9 (lanes 1 and 4) and nuclear extracts from the B-cell line Raji (lane 3) or Blanco B cells (lanes 2 and 5). Immunoprecipitations were performed with anti-rRFX antiserum (lanes 1 to 3) or preimmune serum from the same rabbit (lanes 4 and 5). (C) Northern blot analysis of HLA-DRA mRNA induction by IFN- γ in 143B fibroblast cells transfected with a plasmid directing the synthesis of antisense RFX RNA. Total RNA was extracted before (–) or after 48 hr of induction with IFN- γ (+) from 143B cells transfected with the antisense RFX plasmid (lanes 1), control plasmid (lanes 2), or untransfected 143B cells (lanes 3). Equal amounts of RNA were analyzed in all lanes.

have no effect. The more slowly migrating complex that is also inhibited specifically by the antibodies is always associated with RFX and has been shown by competition and methylation interference experiments to contain RFX (data not shown). The inhibition is restricted to RFX because the complex due to NF-S, a protein that binds to a sequence situated next to the 3' end of the X box (Kobr et al. 1990) is not affected (Fig. 1A, lane 3). Binding of NF-Y to the class II Y box (Reith et al. 1988) is also not affected (data not shown). The reduction in intensity of the RFX-DNA complex correlates with a shift in migration of the complex to the top of the gel (Fig. 1A, lane 3).

In the second approach, we tested whether the anti-serum could immunoprecipitate specific RFX–DNA complexes. DRA-X1 labeled with ^{32}P was incubated with either the recombinant protein or nuclear extracts from B-lymphocyte cell lines. DNA–protein complexes were then immunoprecipitated and the presence of labeled DRA-X1 in the precipitate was analyzed by polyacrylamide gel electrophoresis (Fig. 1B). To control for specificity, the nonspecific DRA-Xr oligonucleotide was included during the binding reaction, and preimmune serum from the same rabbit was tested in parallel (Fig. 1B, lanes 4 and 5). The results show that the antiserum immunoprecipitates the specific protein–DNA complex formed both by the recombinant protein (Fig. 1B, lane 1) and by native RFX (Fig. 1B, lanes 2 and 3).

On the basis of these results, we conclude that the recombinant protein corresponds to the nuclear factor RFX. It will therefore be referred to as recombinant RFX (rRFX).

RFX is a class II gene regulatory factor

The finding of a specific defect in the binding of RFX to its target DNA sequence in HLA class II negative regulatory mutants (Reith et al. 1988, 1989) strongly suggested that RFX is essential for class II gene expression. The availability of a cDNA clone encoding RFX has now allowed us to confirm this role. A plasmid directing the synthesis of antisense RFX RNA was constructed by inserting a fragment of the RFX9 cDNA insert into a eukaryotic expression vector. This plasmid and a control plasmid consisting of the same expression vector without an insert were transfected into the fibroblast cell line 143B, and stable transfectants were cloned. In the antisense transfectants, the induction of class II genes by IFN- γ is almost completely inhibited. Figure 1C shows that induction of HLA–DRA mRNA is reduced severely (by a factor of 20) in an antisense transfectant (lanes 1), whereas no inhibition is observed in a clone transfected with the control plasmid (lanes 2) with respect to untransfected cells (lanes 3). Detailed analysis has shown that induction of all class II genes, as well as the class II-associated invariant chain gene, is strongly inhibited, whereas class I genes are induced normally (C. Berte et al., in prep.). RFX is thus a regulatory factor that is essential for induction of class II gene expression by IFN- γ .

Amino acid sequence of RFX

The coding region of RFX9 starts in-frame with the λ gt11 β -galactosidase gene and is thus truncated at its 5' end. A second cDNA library was therefore constructed by using a specific primer homologous to a sequence situated near the 5' end of clone RFX9. In this manner, a cDNA clone (RFX5') that overlaps with RFX9 and corresponds to the 5' end of RFX mRNA was isolated. The nucleotide sequence and deduced amino acid sequence of RFX was derived from the overlapping RFX9 and RFX5' clones (Fig. 2A). There is a 2867-nucleotide (979

amino acid) open reading frame starting with an ATG at nucleotide 94 and ending with a TAA stop codon at nucleotide 3031. The ATG at nucleotide 94 is presumably the translation initiation codon because it is the first ATG in the sequence, is preceded at nucleotide 47 by an in-frame TAA stop codon, and is in a favorable context for translation initiation (Kozak 1989). The open reading frame encodes a novel polypeptide showing no obvious homology with any protein in the Swiss-Prot 11 data library.

RFX contains noteworthy features (Fig. 2). First, between residues 381 and 411, the glycine content is high (80%) and there is an uninterrupted run of 14 glycines. Second, there is a basic region (residues 482–519) that falls within the area mapped as containing the DNA-binding domain (residues 438–528, see below). Third, near the carboxyl terminus of the protein, there is a strongly acidic region (residues 920–936) that includes a run of 8 glutamates. Fourth, the amino-terminal portion upstream of the glycine-rich stretch exhibits a peculiar distribution of regions rich in glutamine, proline, or serine plus threonine. There are three glutamine-rich regions having a low charge density between residues 4 and 73 (24%), 127 and 212 (23%), and 230 and 296 (27%). These glutamine-rich regions are interspersed with three serine/threonine-rich regions between residues 85 and 126 (32%), 215 and 227 (54%), and 304 and 380 (39%). Finally, the first glutamine-rich segment overlaps with a region having a very high (40%) proline content between residues 13 and 97.

Localization of the DNA-binding domain of RFX

The DNA-binding domain of RFX was localized by analyzing the DNA-binding activity of a series of truncated rRFX proteins (Fig. 3A). These truncated proteins were synthesized in vitro from RNA prepared from plasmid pRFX9 or deleted derivatives of pRFX9 and assayed by EMSA for binding to the DRA-X1 oligonucleotide. Analysis of carboxy-terminal deletions (rRFX9/BssHII, rRFX9/SacI, and rRFX9/BglI) and amino-terminal deletions (rRFX9, rRFX36, and rRFX21) situated the DNA-binding domain between residues 406 (rRFX36) and 589 (rRFX9/SacI; Fig. 3B). This localization was narrowed down further by the analysis of carboxy-terminal deletions of the amino-terminal deletion rRFX36 (rRFX36/HaeII, rRFX36/BalI and rRFX36/RsaI) and finer amino-terminal deletions derived from rRFX36 with Bal31 (rRFX36.1 and rRFX36.2), thus placing the DNA-binding domain within a 91-amino-acid polypeptide (rRFX36.1/BalI) situated between residues 438 and 528 of RFX (Fig. 3A,C, and D). Polypeptides containing only the DNA-binding domain, such as rRFX36/HaeII, rRFX36/BalI, and rRFX36.1/BalI are sufficient for efficient and specific binding to the class II X box (Figs. 3C, D, and 6B).

RFX binds as a dimer

In EMSA performed either with rRFX9 or rRFX36, a minor protein–DNA complex is visible below the major

Figure 2. (See following page for legend.)

band (see arrowhead in Fig. 3C). The minor complex is particularly clear when EMSA is performed with ^{35}S -labeled in vitro translation products and an excess of unlabeled DRA-X1 (Fig. 4A). Analysis of ^{35}S -labeled protein present in the protein–DNA complexes (Fig. 4C) indicated that the minor and major complexes contain the same polypeptide (either rRFX9 or rRFX36). This suggested that RFX dimerizes and that the major and minor bands represent the binding of homodimers and monomers, respectively. To confirm this, rRFX9 and rRFX36 were translated together and the formation of rRFX9–rRFX36 heterodimers was analyzed by EMSA (Fig. 4B). Three bands are observed (Fig. 4B, lane 3): the same two major bands obtained when rRFX9 and rRFX36 are translated separately (bands A and B) and a third band of intermediate mobility (band C). The three bands were excised from the gel, and ^{35}S -labeled proteins present in the gel slices were analyzed by SDS–polyacrylamide gel electrophoresis (Fig. 4C). The results show that the band of intermediate mobility is due to the binding of a heterodimer consisting of both rRFX9 and rRFX36 (Fig. 4C, lane C). Because the rRFX9–rRFX36 heterodimer is not formed when rRFX9 and rRFX36 are synthesized separately and mixed only during the binding reaction, dimerization must occur prior to binding (Fig. 4B, lane 4). Moreover, the dimers are quite stable because a 2-hr preincubation of a mixture of separately synthesized rRFX9 and rRFX36 does not lead to detectable heterodimer formation (data not shown).

Localization of the dimerization domain

In many proteins that bind DNA as dimers, such as leucine zipper proteins (Landschulz et al. 1988; Abel and Maniatis 1989; Vinson et al. 1989) and helix–loop–helix (HLH) proteins (Murre et al. 1989a,b), the region responsible for dimerization is closely associated with the DNA-binding domain. This is not the case for RFX because polypeptides consisting only of the region con-

Figure 2. DNA sequence and deduced amino acid sequence of RFX. (A) The nucleotide sequence derived from the overlapping RFX5' and RFX9 cDNA clones, and the deduced amino acid sequence of RFX are shown. The 3' end of RFX5', 5' end of RFX9, DNA-binding domain (box), glycine-rich and acidic regions (dashed boxes), glutamine-rich regions (underlining), and proline-rich region (dashed underlining) are indicated. (B) Map of RFX. Positions of glutamine-rich (24, 23, and 27% Q), proline-rich (40% P), serine/threonine-rich (32, 54, and 39% S + T), glycine-rich (80% G), basic, and acidic regions are shown. DNA binding and dimerization domains have been localized within the regions indicated below the map (see Figs. 3–5). Putative amphipathic α -helices in the DNA-binding domain (see Fig. 7) are represented by bars.

Figure 3. Mapping of the DNA-binding domain. (A) Summary of truncated rRFX peptides that have been tested for binding. The plasmid pRFX9 is shown schematically in the upper line. The T7 RNA polymerase promoter (T7) and the ATG translation initiation codon are derived from plasmid pT7-7. The coding region is depicted as a box. The DNA-binding domain (solid box), TAA stop codon, and HindIII site in the 3'-untranslated region are indicated. Truncated peptides, their amino acid coordinates, and their binding activity (+ or –) are shown. ± indicates only weak binding. Constructions and in vitro synthesis of truncated peptides are detailed in Experimental procedures. (B) Analysis by EMSA of the binding of carboxy- and amino-terminal deletions to ^{32}P -labeled DRA-X1. Lanes are numbered according to deletions depicted in A: (Lane 0) Rabbit reticulocyte lysate control. (C) Analysis by EMSA of the binding of carboxy-terminal deletions of rRFX36 to ^{32}P -labeled DRA-X1. Lanes are numbered according to deletions depicted in A: (Lane 0) Rabbit reticulocyte lysate control. The arrowhead (lane 1) indicates a minor protein–DNA complex (see text and Fig. 4). (D) Analysis by EMSA of the minimal 91-amino-acid DNA-binding domain rRFX36.1/BalI (deletion 12 in A).

taining the DNA-binding domain do not dimerize (Fig. 5A; rRFX36/HaeII and rRFX36/BalI).

To identify and localize a domain involved in dimerization, various truncated forms of rRFX were synthesized together with either rRFX9 or rRFX36 (Fig. 5A). No protein–DNA complex attributable to a heterodimer is observed when rRFX9 and rRFX9/BssHII are synthesized

Figure 4. Dimerization of RFX. (A) RFX can bind both as a homodimer and as a monomer. Binding of rRFX9 and rRFX36 was analyzed by EMSA with ^{35}S -labeled proteins and unlabeled DRA-X1 oligonucleotide. Bands resulting from binding of dimeric and monomeric rRFX9 (lane 1, bands A and a, respectively) or rRFX36 (lane 2, bands B and b, respectively) are indicated. Proteins labeled with ^{35}S present in bands A, B, a, and b were analyzed in C. (B) Dimerization between rRFX9 and rRFX36. rRFX9 and rRFX36 were synthesized separately (lanes 1 and 2) or together (lane 3) and tested by EMSA for binding to ^{32}P -labeled DRA-X1. In lane 4, rRFX9 and rRFX36 were synthesized separately and mixed prior to binding. Proteins labeled with ^{35}S , present in bands A (dimeric rRFX9), B (dimeric rRFX36), and C (rRFX9-rRFX36 heterodimer), were analyzed in C. (C) Analysis by SDS-polyacrylamide gel electrophoresis of ^{35}S -labeled proteins present in protein-DNA complexes A, B, C, a, and b excised from the gels shown in A and B. Bands corresponding to rRFX9 and rRFX36 are indicated.

together (Fig. 5B), indicating that deletion of the last 219 carboxy-terminal amino acids abolishes dimerization. A carboxy-terminal polypeptide containing the last 236 amino acids of RFX (Fig. 5A; rRFX42) was then translated together with rRFX36 (Fig. 5C). Three bands are observed (Fig. 5C, lane 3): the major and minor bands due to the binding of rRFX36 homodimers and monomers, respectively (bands A and B), and a third band of intermediate mobility (band C). No band due to the binding of rRFX42 alone is detected because this deletion lacks the DNA-binding domain. Analysis of ^{35}S -labeled proteins present in the three bands show that the band of intermediate mobility contains both rRFX36 and rRFX42 (Fig. 5D) and is thus due to the binding of an rRFX36-rRFX42 heterodimer. Consequently, the last 236 carboxy-terminal amino acids of RFX are both necessary and sufficient to promote dimerization.

The two subunits of RFX do not bind to the same X box

Binding and dimerization domains are not only situated in distinct regions of RFX but are functionally independent of each other, because truncated polypeptides that no longer contain the dimerization domain retain their ability to bind (e.g., rRFX9/BssHII, rRFX36/HaeII, and rRFX36/BalI), whereas polypeptides lacking the DNA-binding domain can still dimerize (e.g., rRFX42). Moreover, at equal protein concentrations, monomeric and dimeric peptides bind with the same efficiency. This raises the question of whether both subunits of the RFX homodimer contact the same X box and are required for specific binding or whether a single subunit is sufficient (see Fig. 8B). To distinguish between these two possibilities, we compared the binding characteristics of dimeric and monomeric forms of rRFX.

As observed previously for both native and recombinant RFX (Reith et al. 1989; Kobr et al. 1990), binding of the monomeric rRFX36/HaeII polypeptide to ^{32}P -labeled DRA-X1 is competed for completely by DRA-X3, poorly by DQA-X, and with intermediate efficiency by DPA-X (Fig. 6A). Thus, dimerization is not required for the

characteristic DRA > DPA >> DQA gradient of binding affinity of RFX.

To determine whether or not both subunits of the RFX homodimer contact the same X box, we compared by methylation interference analysis the nucleotide residues that are contacted by dimeric and monomeric forms of rRFX. As an example of a dimeric molecule, we chose rRFX36 (band B in Fig. 4A). As monomers, we analyzed both the monomeric form of rRFX36 (band b in Fig. 4A) and rRFX36/HaeII which lacks the dimerization domain. The results obtained with the two types of monomeric rRFX are identical. Data for rRFX36/HaeII are shown in Figure 6B. The methylation interference profiles obtained show that binding of monomers and dimers is affected by methylation at the same G and A residues (Fig. 6C). Methylation of four G residues and five A residues inhibits the binding of dimers and monomers to the same extent. At three additional G residues, differences in the magnitude of the effects are observed. For two of these G residues, the effect is clear for dimers but barely detectable, although reproducible, for monomers. For the third G residue, at which an enhancement is characteristically observed (Reith et al. 1988, 1989), the effect is strong for monomers but only weak for dimers. Despite these quantitative differences, monomers and dimers clearly have the same contact points with the DNA and exhibit a methylation interference profile that is practically indistinguishable from that of native RFX (Fig. 6C). Dimerization is thus not required for correct interaction of RFX with the X box.

DNA-binding domain of RFX

A similarity search in the Swiss-Prot 11 data library indicated that the DNA-binding domain of RFX is novel and shares no extensive homology with other known DNA-binding motifs. Nevertheless, an interesting similarity to the HLH DNA-binding motif (Murre et al. 1989a,b) was identified. A secondary structure prediction (Biou et al. 1988) allowed us to identify two putative amphipathic α -helices in the region of RFX encom-

passing the DNA binding domain (Fig. 7). Helix I (residues 476–489) and helix II (residues 517–528) span a region rich in basic amino acids (residues 482–519; see Fig. 2B). The linker sequence separating the two helices has mainly a predicted coil conformation and contains 12 residues (glycine, proline, tyrosine, asparagine, and serine) that are found with high frequency in loops (Leszczynski and Rose 1986). This structure is reminiscent of the HLH motif. We therefore aligned the sequence of RFX with an HLH consensus sequence (Fig. 7). RFX matches the consensus sequence at 10 out of 25 positions. There is thus also a weak sequence similarity between RFX and the HLH motif.

Figure 5. Localization of the dimerization domain. (A) Summary of truncated rRFX peptides that have been tested for dimerization. The plasmid pRFX9 shown in the upper line is described in Fig. 3A. Regions involved in DNA binding (shaded box) and dimerization (solid box) are indicated. Truncated peptides tested for dimerization, their amino acid coordinates, and their ability to dimerize (+ or –) are shown. (B) The last 219 carboxy-terminal amino acids of RFX are required for dimerization. rRFX9 and rRFX9/BssHIII were synthesized separately (lanes 1 and 2) or together (lane 3) and tested by EMSA for binding to 32 P-labeled DRA-X1. No band resulting from an rRFX9–rRFX9/BssHIII heterodimer is observed in lane 3. (C) The last 236 carboxy-terminal amino acids are sufficient for dimerization. rRFX36 and rRFX42 labeled with 35 S were synthesized separately (lanes 1 and 2) or together (lane 3) and tested by EMSA for binding to unlabeled DRA-X1. Proteins labeled with 35 S, present in bands A (dimeric rRFX36), B (monomeric rRFX36), and C (rRFX36–rRFX42 heterodimer), were analyzed in D. (D) Analysis by SDS–polyacrylamide gel electrophoresis of 35 S-labeled proteins present in bands A, B, and C excised from the gel shown in C. Bands corresponding to rRFX36 and rRFX42 are indicated.

The positions of deletions that abolish DNA binding correlate well with the region showing similarity in structure and sequence with the HLH motif (Fig. 7). Deletions affecting the integrity of the amino-terminal (rRFX21) or carboxy-terminal end (rRFX9/BglI and rRFX36/RsaI) of this region eliminate binding activity. On the other hand, truncated peptides containing the intact region (rRFX36.1 and rRFX36/BalI) can bind normally.

Discussion

Sequence and structure of RFX

RFX contains at least three independent domains (Fig. 8A). The DNA-binding domain is central and functionally independent of the remainder of the protein because polypeptides containing only the DNA-binding domain have the same binding characteristics as the intact protein. The region responsible for dimerization lies near the carboxy-terminal end and is also functionally independent, because a polypeptide containing only the last 236 amino acids retains its capacity to dimerize. Finally, the amino-terminal region is required neither for DNA binding nor dimerization and is separated from the rest of the protein by an 80% glycine-rich stretch. Such glycine-rich regions confer high chain flexibility, and the amino-terminal portion of RFX thus seems to be separated from the remainder of the protein by a flexible “hinge.” Runs of glycine occur in other known or suspected DNA-binding proteins such as Oct-2 (Clerc et al. 1988) and *daughterless* (Caudy et al. 1988; Cronmiller et al. 1988).

The amino-terminal half of RFX has several structural features associated with the activation domains of transcription factors: It consists of three glutamine-rich regions interspersed with three serine/threonine-rich regions. A similar alternation of glutamine and serine/threonine-rich motifs is present in the amino-terminal half of the transcription factor SP1. In SP1, the glutamine-rich regions function as transcription activation domains (Courey and Tjian 1988), which suggests that the analogous regions of RFX may have the same role. Glutamine-rich regions also occur in several other transcription factors (Mitchell and Tjian 1989). Another potential activation domain is the 40% proline-rich region at the amino-terminal end of RFX. In CTF/NF-1, a 20–30% proline-rich region has been shown to be a transcription activation domain and regions rich in proline have been noted in other transcription factors (Mitchell and Tjian 1989). In addition, the carboxy-terminal end of RFX contains a highly acidic region, and acidic transcription activation domains have been identified in factors such as GAL4 and GCN4 (Ptashne 1988; Mitchell and Tjian 1989). These acidic domains, however, can form amphipathic α -helical structures, a feature that is not shared by the acidic region of RFX. Construction of chimeric molecules will allow us to test for the functional role of these sequences in transcription activation.

Figure 6. Comparison of the binding characteristics of dimeric and monomeric rRFX. (A) Relative affinity of a monomeric rRFX peptide for the X boxes of DRA, DQA, and DPA promoters. rRFX36/*HaeII* was incubated with 32 P-labeled DRA-X1 oligonucleotide in the absence (lane 1) or presence of a 20-fold molar excess of unlabeled DRA-X3 (lane 2), DRA-Xr (lane 3), DQA-X (lane 4) and DPA-X (lane 5) competitor oligonucleotides. Extent of competition was analyzed by EMSA. Oligonucleotides are described in Experimental procedures. (B) Methylation interference analysis of protein–DNA complexes formed by dimeric and monomeric forms of rRFX. Cleavage profiles of the coding (C) and noncoding (N.C.) strands are shown for free DNA (F) and DNA bound by rRFX36 dimers (D) or rRFX36/*HaeII* monomers (M). Position of the X box is indicated. (C) Comparison of methylation interference profiles observed for native RFX, rRFX dimers (D), and RFX monomers (M). Data shown for native RFX are derived from Reith et al. (1988) as well as from a more detailed analysis of methylation interference at A residues. Data shown for native RFX residues and a more detailed analysis of A are derived from Reith et al. (1988). Data shown for rRFX(D) and rRFX(M) are a compilation of results obtained with rRFX36 dimers, rRFX36 monomers, and rRFX36/*HaeII* monomers. Arrows indicate G and A residues at which methylation interferes with binding. The arrowhead indicates a G residue at which methylation enhances binding. Relative magnitude of the effects are approximated by the sizes of the arrows and arrowheads.

DNA-binding domain of RFX

The DNA-binding domain of RFX is novel and shows no clear homology to either well-defined DNA-binding motifs or as-yet-unclassified DNA-binding domains (for reviews, see Johnson and McKnight 1989; Mitchell and Tjian 1989). A feature worth noting, however, is an in-

triguing resemblance with the HLH motif (Murre et al. 1989a,b) at the level of both amino acid sequence and predicted secondary structure (Fig. 7). The significance of this similarity is not clear because there are several important structural and functional differences between the DNA-binding domains of RFX and HLH proteins. First, RFX matches the HLH consensus sequence at only 10 of 25 positions, and overall sequence homology between RFX and various HLH motifs is considerably lower (10–15% identity) than that observed, for example, between the HLH domains of MyoD1 and E12 (30% identity) or *daughterless* and E12 (80% identity; Murre et al. 1989a). Second, amphipathic α -helix II is positioned differently in RFX. The presence of a proline residue in RFX is not compatible with the formation of an α -helix in the same position as helix II of the HLH motif. Third, helix I of the HLH motif is characteristically flanked at its amino-terminal end by a conserved basic region that mediates specific DNA binding (Pren-dergast and Ziff 1989; Davis et al. 1990), whereas the corresponding region of RFX does not exhibit this basic character. In fact, in RFX, the basic region coincides with the two α -helices and the intervening loop (see Fig. 2B). Two HLH proteins, *emc* (Ellis et al. 1990; Garrell and Modolell 1990) and *Id* (Benezra et al. 1990), also lack the basic region adjacent to helix I. Unlike RFX however, *emc* and *Id* are thought to lack DNA binding activity (Benezra et al. 1990; Ellis et al. 1990; Garrell and Modolell 1990). Finally and more importantly, in HLH proteins, the HLH motif mediates dimerization whereas the upstream basic region is involved in DNA binding (Davis et al. 1990). In RFX, however, the DNA-binding domain alone binds as a monomer and the sequence responsible for dimerization is independent from the DNA-binding domain and lies in a different region of the protein. These differences may indicate that the DNA-binding domain of RFX is only a very distant relative of the HLH domain and that it has evolved to lose its dimerization function while retaining its specific DNA binding capacity. Alternatively, it is possible that the DNA-binding domain of RFX is unrelated to the HLH motif, in which case the observed similarity may simply reflect similar structural constraints imposed by a common function. Whatever the evolutionary relationship may be, the DNA-binding domain of RFX is sufficiently different to be considered as a novel motif.

Dimerization of RFX

Recombinant RFX binds to the class II X box as a homodimer, and the last 236 carboxy-terminal amino acids of RFX are both necessary and sufficient for the formation of stable dimers. In this region, no sequence homology has been detected with other proteins known to bind DNA as dimers. In particular, there is no potential leucine zipper, a well-characterized dimerization motif (Landschulz et al. 1988; Abel and Maniatis 1989; Vinson et al. 1989). Therefore, on the basis of sequence homology alone, we cannot identify the precise segment of RFX that is responsible for dimerization. Further dele-

Figure 7. Amino acid similarity between the DNA-binding domain of RFX and the HLH motif. The sequence of RFX was aligned with a consensus sequence for HLH domains. The consensus sequence originally defined by Murre et al. (1989a) was extended to include most known HLH proteins aligned according to Benezra et al. (1990) and shows residues conserved in the majority of HLH domains; + represents the basic residues K or R, and ψ represents the hydrophobic residues I, L, or V. Shaded boxes indicate positions at which RFX matches the HLH consensus. Positions of predicted α -helices in RFX (*top*) and of the basic region, α -helices, and loop in HLH proteins (*bottom*) are indicated. Limits of amino-terminal (rRFX36.1 and rRFX21) and carboxy-terminal (rRFX9/BglI, rRFX36/RsaI, and rRFX36/BalI) deletions are shown above, and their binding activities are represented by (+) or (-).

tion mapping and the introduction of amino acid substitutions will be informative.

Target sequences for dimeric DNA-binding proteins generally contain perfect or imperfect palindromes. Typical examples are provided by prokaryotic DNA-binding proteins (McKay and Steitz 1981; Pabo and Lewis 1982), nuclear receptors (Green and Chambon 1988; Beato 1989; Berg 1989), leucine zipper proteins (Landschulz et al. 1988; Abel and Maniatis 1989; Vinson et al. 1989), and HLH proteins (Murre et al. 1989a,b). Stable binding of such proteins is dependent on dimerization, and it is thought that dimerization is required for correct interaction of the two subunits with the two symmetrical halves of the palindromic recognition sequence. In the case of RFX, the situation is strikingly different. The consensus sequence of the class II X box does not show dyad symmetry and is bound efficiently by RFX monomers. Moreover, the DNA-binding and dimerization domains of RFX are situated in different regions of the protein and are functionally independent of each other. Stable and specific binding does not require dimerization, and a single RFX subunit has the same contact points with the DNA as the RFX homodimer does. The dimerization domain, therefore, does not contribute to the binding activity of RFX; and unlike other dimeric DNA-binding proteins, the two RFX subunits do not bind the same X box (Fig. 8B).

Regulatory role of RFX

The finding of a specific defect in the binding of RFX to its target DNA sequence in HLA class II negative regulatory mutants (Reith et al. 1988, 1989) strongly suggested that RFX is essential for class II gene expression. The availability of a cDNA clone encoding RFX has now allowed us to demonstrate this role directly by transfection of a fibroblast cell line with plasmids directing the synthesis of antisense RFX RNA. In these transfectants, the expression of HLA class II genes following induction by IFN- γ is severely inhibited. Similar experiments performed with a class II-positive B-cell line have shown that antisense RFX RNA also inhibits constitutive class

II gene expression (C. Berte et al., in prep.). These results are consistent with observations that the class II X box is a key *cis*-acting element in B cells as well as IFN- γ -induced fibroblasts (Tsang et al. 1990) and that the class II regulatory mutants defective in RFX binding show neither constitutive nor IFN- γ -inducible class II gene expression (dePréval et al. 1985, 1988). Although RFX is clearly required for induction by IFN- γ , it should be noted that neither RFX mRNA nor RFX protein levels are modulated by IFN- γ ; RFX is present at equivalent levels in unstimulated class II-negative and stimulated class II-positive fibroblasts (I. Amaldi et al., in prep.). We therefore conclude that RFX is essential but not sufficient on its own to promote class II gene transcription in response to IFN- γ .

Features such as the glutamine-rich, proline-rich, and

Figure 8. Model for the structure and binding of RFX. (A) RFX contains three structurally and functionally independent domains: a central DNA-binding domain (B), a carboxy-terminal dimerization domain (D), and an amino-terminal region showing certain features characteristic of transcription activation domains (N). The linear map of RFX (*top*) and a schematic representation of the protein (*bottom*) are shown. (B) Independence of the DNA-binding (B) and dimerization (D) domains implies that the two subunits of RFX homodimers do not bind the same X box (*left*) but that a single subunit is sufficient (*right*).

acidic regions indicate that RFX may be a transcription factor. Our results also suggest an additional role for RFX. The fact that a single RFX subunit has a functionally intact DNA-binding domain implies that an RFX homodimer has two independent binding domains and that it is capable of interacting with two different X box sequences simultaneously. RFX homodimers may thus represent a situation analogous to PRDII-BF1, a protein that contains two widely separated zinc finger motifs that could bind simultaneously to two copies of the same recognition sequence (Fan and Maniatis 1990). As a consequence of these unique properties, we propose that an RFX homodimer could function as a cross-linking agent between regulatory sequences located far apart on the DNA. Interestingly, in murine class II genes, one observes (between 1000 and 1600 bp upstream from the cap site) a second X-box and Y-box sequence in reverse orientation (Y' and X'; Dorn et al. 1988; Koch et al. 1989). We have also identified a specific RFX-binding site situated 1.5 kb upstream of the human HLA-DRA gene (M. Kober et al., unpubl.). In the case of the E α gene, experiments in transgenic mice suggest a functional role for this promoter-distal region. Consequently, the structural and functional properties of RFX presented in this paper suggest a model in which the RFX homodimer, with its two active DNA-binding sites, forms a loop by cross-linking the promoter-proximal X box and the promoter-distal X' box of MHC class II genes, thus bringing regulatory sequences situated far upstream into the vicinity of the transcription initiation site. Formation of such a loop, which had been postulated on the basis of the duplication of both the X and Y boxes (Koch et al. 1989), might require proteins binding to the Y box as well as RFX.

By analogy to leucine zipper and HLH proteins (Abel and Maniatis 1989; Murre et al. 1989b), an alternative role for the dimerization domain of RFX could be the formation of heterodimers as well as homodimers. In RFX heterodimers, however, only the RFX subunit would bind to the X box. The second subunit of the putative heterodimer could be a non-DNA-binding protein that is required for RFX activity or a factor that binds to a sequence situated adjacent to the X box. RFX-specific antibodies will be helpful in studying the possibility that RFX may associate not only with itself but also with other proteins involved in the regulation of HLA class II genes.

Experimental procedures

Cells and nuclear extracts

Origin and maintenance of cell lines (Raji and 143B) and the preparation of nuclear extracts have been described (Reith et al. 1988; Amaldi et al. 1989). Blanco B cells were obtained by leukopheresis of blood from a patient having chronic B-cell lymphocytic leukemia.

Oligonucleotides, plasmid constructs, and DNA sequencing

Preparation of double-stranded oligonucleotides has been described previously (Reith et al. 1988). The DRA-X1 and

DRA-X3 oligonucleotides contain sequences -144 to -70 and -124 to -70, respectively, of the DRA promoter. The DRA-Xr oligonucleotide is identical to DRA-X3, except that the X box (CCCCTAGCAACAGATG) is replaced with a random sequence (ATGTGTCTCGAACGCA). The DQA-X and DPA-X oligonucleotides are identical to DRA-X3, with respect to length and position of the X box, but cover sequences -127 to -73 of the DQA promoter and -119 to -65 of the DPA promoter.

Cloning of the λ 9 phage has been described (Reith et al. 1989). The coding region of the λ 9 cDNA insert (RFX9) was sequenced on both strands by use of single- or double-stranded M13 templates and Sequenase (U.S. Biochemicals). For sequencing the coding strand, the cDNA was inserted into the *EcoRI* site of M13mp19, and a sequential series of overlapping 5' deletions of the coding strand was prepared by use of Cyclone I Biosystem (IBI). For sequencing the complementary strand, a set of regularly spaced primers was synthesized on the basis of the sequence of the coding strand.

To clone the 5' end of RFX mRNA, a cDNA library was constructed in λ gt10 by standard techniques (Gubler and Hoffman 1983; Huynh et al. 1985), except that an oligonucleotide specific for a sequence situated near the 5' end of RFX9 was used to prime first-strand synthesis. Screening of the library with RFX9 yielded the clone RFX5'. RFX5' was excised from the phage with *EcoRI*, inserted into the *EcoRI* site of a Bluescript vector (Pharmacia), and sequenced on both strands by use of Sequenase.

The plasmid pT7-7 is one of a series of prokaryotic expression vectors containing the T7 RNA polymerase promoter and translation start site of the T7 gene 10, followed by unique restriction sites which, when filled in with Klenow polymerase, allow the subcloning of blunt-ended cDNA inserts into all three reading frames (Tabor and Richardson 1985). To construct pRFX9, the RFX9 cDNA insert of phage λ 9 was excised with *EcoRI*, filled in with Klenow polymerase, and inserted into pT7-7 in-frame with the translation initiation codon. To construct pRFX36, pRFX21, and pRFX42, 5' deleted cDNAs prepared for sequencing (see above) were cloned into pT7-7 as described for RFX9. To construct pRFX36.1 and pRFX36.2, the RFX36 cDNA insert was deleted with *Bal31* and digested with *Sall*, and the fragment coding for the DNA-binding domain was cloned in-frame into pT7-7. Inserts in pRFX9, pRFX36, pRFX36.1, pRFX36.2, pRFX21, and pRFX42 are deleted to nucleotides 557, 1309, 1405, 1417, 1486, and 2320, respectively.

In vitro transcription and translation

Template DNAs were prepared by digestion of pRFX9, pRFX36, pRFX36.1, pRFX36.2, pRFX21, and pRFX42 with *HindIII*, which cuts in the 3' untranslated region of the RFX cDNA, or with restriction enzymes (*BssHII*, *SacI*, *Sall*, *BglII*, *HaeII*, *BalI*, *RsaI*) that cut within the coding region. Capped mRNA was synthesized, by use of 1 μ g of template DNA and 10–20 units of T7 RNA polymerase for 1 hr at 37°C in a 20 μ l reaction volume containing 40 mM Tris-HCl (pH 8.0), 8 mM MgCl₂, 2 mM spermidine, 50 mM NaCl, 30 mM DTT, 0.5 mM ATP, 0.5 mM CTP, 0.5 mM UTP, 0.05 mM GTP, 0.5 mM m⁷G(5')ppp(5')G. After transcription, the reaction was diluted to 100 μ l with H₂O, 200 μ l ethanol was added, and the RNA was stored precipitated at -20°C. One-twentieth of the transcription product was precipitated, dissolved in H₂O, and translated in a rabbit reticulocyte lysate system (Promega), in the presence of [³⁵S]methionine, as specified by the manufacturer. Translation products were verified routinely by SDS-polyacrylamide gel electrophoresis.

EMSA

EMSA with 5–8 μ g of nuclear extract or 1–2 μ l of in vitro translation products was performed as described previously (Reith et al. 1988, 1989) by use of either 0.2–0.5 ng of 32 P-labeled DRA-X1 or 10 ng of unlabeled DRA-X1. To analyze proteins present in protein–DNA complexes, EMSA was performed with 35 S-labeled protein and unlabeled DRA-X1. Protein–DNA complexes were localized by autoradiography and excised from the dried gel. The gel slices were then placed in the wells of an SDS–polyacrylamide gel and equilibrated for 30 min with sample buffer prior to electrophoresis. When testing the effect of antibodies on the EMSA profile, binding conditions were modified as follows. Binding was done with 2 μ g of nuclear extract for 30 min at 0°C in 10 μ l of 4% glycerol, 4 mM HEPES (pH 7.9), 20 mM KCl, 5 mM MgCl₂, 0.04 mM EDTA, 0.1 mM phenylmethylsulfonyl fluoride, 0.1 mM DTT, 25 μ g/ml poly[d(I-C)]·poly[d(I-C)], and 13 μ g/ml sonicated denatured *Escherichia coli* DNA. Binding reactions were then incubated for 15 min at 0°C with 10 μ l of 10% fetal calf serum (FCS), 10% FCS containing 1 μ g of rabbit anti β -galactosidase antibodies (Cappel), or 10% FCS containing 1 μ g of affinity-purified anti-rRFX antibodies.

Methylation interference assay

The DRA-X1 oligonucleotide was subcloned into the *Sma*I site of pUC12. This plasmid was used to prepare the *Eco*RI–*Hind*III insert labeled at the *Eco*RI site (coding strand) or *Hind*III site (noncoding strand) with [γ - 32 P]ATP and T4 polynucleotide kinase. Methylation, binding, purification of free and bound DNA, and polyacrylamide gel electrophoresis were performed as described (Reith et al. 1988).

Production of antibodies

A bacterial strain carrying an inducible T7 RNA polymerase gene was transformed with pRFX9, grown, and induced as described (Tabor and Richardson 1985). Detailed procedures for the subsequent purification of rRFX9 and the production of anti-rRFX antibodies are described elsewhere (C. Herrero-Sanchez et al., unpubl.). Briefly, rRFX9 extracted from the cells as an insoluble precipitate was solubilized in 6 M guanidine-HCl and purified on a Sephacryl S-300 (Pharmacia) gel filtration column in the presence of 4 M guanidine-HCl. Column fractions enriched in rRFX9 were renatured by dialysis and used to immunize rabbits. For immunoprecipitation of RFX–DNA complexes, whole antiserum was used. For use in EMSA experiments, anti-rRFX antibodies were affinity-purified as described (Sekimizu et al. 1988).

Immunoprecipitation of protein–DNA complexes

Twenty microliters of protein A–Sepharose beads (Pharmacia), 20 μ l of antiserum, and 80 μ l of buffer A [150 mM NaCl, 0.1% NP-40, 50 mM Tris-HCl (pH 8.0)] were mixed and incubated for 2 hr at 4°C. The beads were then washed three times with buffer A and incubated for 2 hr at 4°C with 80 μ l of buffer B [50 mM NaCl, 0.01% NP-40, 50 mM Tris-HCl (pH 8.0)] and 20 μ l of a binding reaction set up as for EMSA experiments (Reith et al. 1988, 1989). After four washes of the beads with buffer B, labeled DNA was recuperated in 0.3 M sodium acetate, 1% SDS, extracted with phenol/chloroform, precipitated with ethanol, and analyzed by electrophoresis on a 5% polyacrylamide, 8.3 M urea sequencing gel.

Inhibition of class II gene expression by antisense RFX RNA.

A fragment derived from RFX9 (nucleotides 555–1951 of RFX) was inserted in reverse orientation into the eukaryotic expression vector pRSV3. This antisense RFX vector, or pRSV3 without an insert, was cotransfected with a plasmid conferring neomycin resistance into the cell line 143B (ATCC CRL 8303) by the calcium phosphate method. Following selection with G418, stable transfectants were cloned. Clones were induced with 250 U/ml of IFN- γ for 48 hr. RNA extraction and Northern blot procedures have been described (Amaldi et al. 1989).

Acknowledgments

We are very grateful to Dr. J. Garnier for help with protein secondary structure predictions. We thank M. Zufferey for skillful help with the cell culture. This work was supported by the Swiss National Fund for Scientific Research.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 USC section 1734 solely to indicate this fact.

Note added in proof

Sequence data described in this paper have been submitted to EMBL/GenBank Data Libraries.

References

- Abel, T. and T. Maniatis. 1989. Action of leucine zippers. *Nature* **341**: 24–25.
- Amaldi, I., W. Reith, C. Berte, and B. Mach. 1989. Induction of HLA class II genes by IFN- γ is transcriptional and requires a *trans*-acting protein. *J. Immunol.* **142**: 999–1004.
- Basta, P.V., P.A. Sherman, and J.P.Y. Ting. 1988. Detailed delineation of an interferon- γ -responsive element important in human HLA-DRA gene expression in a glioblastoma multi-form line. *Proc. Natl. Acad. Sci.* **85**: 8618–8622.
- Beato, M. 1989. Gene regulation by steroid hormones. *Cell* **56**: 335–344.
- Benacerraf, B. 1981. Role of MHC gene products in immune regulation. *Science* **212**: 1229–1238.
- Benezra, R., R.L. Davis, D. Lockshon, D.L. Turner, and H. Weintraub. 1990. The protein Id: A negative regulator of helix-loop-helix DNA binding proteins. *Cell* **61**: 49–59.
- Berg, J.M. 1989. DNA binding specificity of steroid receptors. *Cell* **57**: 1065–1068.
- Biou, V., J.F. Gibrat, J.M. Levin, and J. Garnier. 1988. Secondary structure prediction: Combination of three different methods. *Protein Engineering* **2**: 185–191.
- Blanar, M.A., E.C. Boettger, and R.A. Flavell. 1988. Transcriptional activation of HLA-DR α by interferon γ requires a *trans*-acting protein. *Proc. Natl. Acad. Sci.* **85**: 4672–4676.
- Caudy, M., H. Vässin, M. Brand, R. Tuma, L.Y. Jan, and Y.N. Jan. 1988. *daughterless*, a *Drosophila* gene essential for both neurogenesis and sex determination, has sequence similarities to *myc* and the *achaete-scute* complex. *Cell* **55**: 1061–1067.
- Clerc, R.G., L.M. Corcoran, J.H. LeBowitz, D. Baltimore, and P.A. Sharp. 1988. The B-cell-specific Oct-2 protein contains POU box- and homeo box-type domains. *Genes Dev.* **2**: 1570–1581.
- Collins, T., A.J. Korman, C.T. Wake, J.M. Boss, D.J. Kappes, W. Fiers, K.A. Ault, M.A. Gimbrone, Jr., J.L. Strominger, and J.S. Pober. 1984. Immune interferon activates multiple class II major histocompatibility complex genes and the asso-

- ciated invariant chain gene in human endothelial cells and dermal fibroblasts. *Proc. Natl. Acad. Sci.* **81**: 4917–4921.
- Courey, A.J. and R. Tjian. 1988. Analysis of Sp1 in vivo reveals multiple transcriptional domains, including a novel glutamine-rich activation motif. *Cell* **55**: 887–898.
- Cronmiller, C., P. Schedl, and T.W. Cline. 1988. Molecular characterization of *daughterless*, a *Drosophila* sex determination gene with multiple roles in development. *Genes Dev.* **2**: 1666–1676.
- Davis, R.L., P. Feng-Cheng, A.B. Lassar, and H. Weintraub. 1990. The MyoD DNA binding domain contains a recognition code for muscle-specific gene activation. *Cell* **60**: 733–746.
- Dedrick, R.L. and P.P. Jones. 1990. Sequence elements required for activity of a murine major histocompatibility complex class II promoter bind common and cell-type-specific nuclear factors. *Mol. Cell. Biol.* **10**: 593–604.
- Dellabona, P., F. Latron, A. Maffei, L. Scarpellino, and R.S. Accolla. 1989. Transcriptional control of MHC class II gene expression during differentiation from B cells to plasma cells. *J. Immunol.* **142**: 2902–2910.
- de Préval, C., M.R. Hadam, and B. Mach. 1988. Regulation of genes for HLA class II antigens in cell lines from patients with severe combined immunodeficiency. *N. Engl. J. Med.* **318**: 1295–1300.
- de Préval, C., B. Lisowska-Grospierre, M. Loche, C. Griscelli, and B. Mach. 1985. A trans-acting class II regulatory gene unlinked to the MHC controls expression of HLA class II genes. *Nature* **318**: 291–293.
- Diedrichs, M. and D.J. Schendel. 1989. Differential surface expression of class II isotypes on activated CD4 and CD8 cells correlates with levels of locus-specific mRNA. *J. Immunol.* **142**: 3275–3280.
- Dorn, A., J. Bollekens, A. Staub, C. Benoist, and D. Mathis. 1987a. A multiplicity of CCAAT box-binding proteins. *Cell* **50**: 863–872.
- Dorn, A., B. Durand, C. Marfing, M. Le Meur, C. Benoist, and D. Mathis. 1987b. Conserved major histocompatibility complex class II boxes—X and Y—are transcriptional control elements and specifically bind nuclear proteins. *Proc. Natl. Acad. Sci.* **84**: 6249–6253.
- Dorn, A., H.J. Fehling, W. Koch, M. Le Meur, P. Gerlinger, C. Benoist, and D. Mathis. 1988. B-cell control region at the 5' end of a major histocompatibility complex class II gene: Sequences and factors. *Mol. Cell. Biol.* **8**: 3975–3987.
- Ellis, H.M., D.R. Spann, and J.W. Posakony. 1990. *Extramacrochaetae*, a negative regulator of sensory organ development in *Drosophila*, defines a new class of helix-loop-helix proteins. *Cell* **61**: 27–38.
- Fan, C.-M. and T. Maniatis. 1990. A DNA-binding protein containing two widely separated zinc finger motifs that recognize the same DNA sequence. *Genes Dev.* **4**: 29–42.
- Garrell, J. and J. Modolell. 1990. The *Drosophila* *Extramacrochaetae* locus, an antagonist of proneural genes that, like these genes, encodes a helix-loop-helix protein. *Cell* **61**: 39–48.
- Germain, R.N. 1986. The ins and outs of antigen processing and presentation. *Nature* **322**: 687–689.
- Green, S. and P. Chambon. 1988. Nuclear receptors enhance our understanding of transcription regulation. *Trends Genet.* **4**: 309–314.
- Griscelli, C., B. Lisowska-Grospierre, and B. Mach. 1989. HLA-class II deficient severe combined immunodeficiency. *Immunodeficiency Rev.* **1**: 135–153.
- Gubler, U. and B.J. Hoffman. 1983. A simple and very efficient method for generating cDNA libraries. *Gene* **25**: 263–270.
- Hammerling, G.J. 1976. Tissue distribution of Ia antigens and their expression on lymphocyte subpopulations. *Transplant. Rev.* **30**: 64–82.
- Huynh, T.V., R.A. Young, and R.W. Davies. 1985. Construction and screening cDNA libraries in λ gt 10 and λ gt11. In *DNA cloning: A practical approach* (ed. D.M. Glover). vol. 1, 49–77. IRL, Oxford.
- Johnson, P.F. and S.L. McKnight. 1989. Eukaryotic transcriptional regulatory proteins. *Annu. Rev. Biochem.* **58**: 799–839.
- Kelly, A. and J. Trowsdale. 1985. Complete nucleotide sequence of a functional HLA-DP β gene and the region between the DP β 1 and DP α 1 genes: Comparison of the 5' ends of HLA class II genes. *Nucleic Acids Res.* **13**: 1607–1621.
- Kobr, M., W. Reith, C. Herrero-Sanchez, and B. Mach. 1990. Two DNA-binding proteins discriminate between the promoters of different members of the major histocompatibility complex class II multigene family. *Mol. Cell. Biol.* **10**: 965–971.
- Koch, W., C. Benoist, and D. Mathis. 1989. Anatomy of a new B-cell-specific enhancer. *Mol. Cell. Biol.* **9**: 303–311.
- Koch, W., S. Candeias, J. Guardiola, R.S. Accolla, C. Benoist, and D. Mathis. 1988. An enhancer factor defect in a mutant Burkitt lymphoma cell line. *J. Exp. Med.* **167**: 1781–1790.
- Kozak, M. 1989. The scanning model for translation: An update. *J. Cell Biol.* **108**: 229–241.
- Landschulz, W.H., P.F. Johnson, and S.L. McKnight. 1988. The leucine zipper: A hypothetical structure common to a new class of DNA binding proteins. *Science* **240**: 1759–1764.
- Latron, F., M. Jotterand-Bellomo, A. Maffei, L. Scarpellino, M. Bernard, J.L. Strominger, and R.S. Accolla. 1988. Active suppression of major histocompatibility complex class II gene expression during differentiation from B cells to plasma cells. *Proc. Natl. Acad. Sci.* **85**: 2229–2233.
- Leszczynski, J.F. and G.D. Rose. 1986. Loops in globular proteins: A novel category of secondary structure. *Science* **234**: 849–855.
- Lisowska-Grospierre, B., D.J. Charron, C. de Préval, A. Durandy, C. Griscelli, and B. Mach. 1985. A defect in the regulation of major histocompatibility complex class II gene expression in human HLA-DR negative lymphocytes from patients with combined immunodeficiency syndrome. *J. Clin. Invest.* **76**: 381–385.
- Mach, B., J. Gorski, P. Rollini, C. Berte, I. Amaldi, J. Berdoz, and C. Ucla. 1986. Polymorphism and regulation of HLA class II genes of the major histocompatibility complex. *Cold Spring Harbor Symp. Quant. Biol.* **51**: 67–74.
- Marrack, P. and J. Kappler. 1988. The T-cell repertoire for antigen and MHC. *Immunol. Today* **9**: 308–315.
- McKay, D.B. and T.A. Steitz. 1981. Structure of catabolite gene activator protein at 2.9 Å resolution suggests binding to left-handed B-DNA. *Nature* **290**: 744–749.
- Mitchell, P.J. and R. Tjian. 1989. Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science* **245**: 371–378.
- Miwa, K., C. Doyle, and J.L. Strominger. 1987. Sequence-specific interactions of nuclear factors with conserved sequences of human class II major histocompatibility complex genes. *Proc. Natl. Acad. Sci.* **84**: 4939–4943.
- Murre, C., P. Schonleber McCaw, and D. Baltimore. 1989a. A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and *myc* proteins. *Cell* **56**: 777–783.
- Murre, C., P. Schonleber McCaw, H. Vaessin, M. Caudy, L.Y. Jan, Y.N. Jan, C.V. Cabrera, J.N. Buskin, S.D. Hauschka, A.B. Lassar, H. Weintraub, and D. Baltimore. 1989b. Interac-

- tions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. *Cell* **58**: 537–544.
- Pabo, C.O. and M. Lewis. 1982. The operator-binding domain of lambda repressor: Structure and DNA recognition. *Nature* **298**: 443–447.
- Paulnock-King, D., K.C. Sizer, Y.R. Freund, P.P. Jones, and J.R. Parnes. 1985. Coordinate induction of $I\alpha$, β , and Ii mRNA in a macrophage cell line. *J. Immunol.* **135**: 632–636.
- Prendergast, G.C. and E.B. Ziff. 1989. DNA-binding motif. *Nature* **341**: 392.
- Ptashne, M. 1988. How eukaryotic transcriptional activators work. *Nature* **335**: 683–689.
- Reith, W., E. Barras, S. Satola, M. Kober, D. Reinhart, C. Herrero-Sanchez, and B. Mach. 1989. Cloning of the major histocompatibility complex class II promoter binding protein affected in a hereditary defect in class II gene regulation. *Proc. Natl. Acad. Sci.* **86**: 4200–4204.
- Reith, W., S. Satola, C. Herrero-Sanchez, I. Amaldi, B. Lisowska-Grospierre, C. Griscelli, 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.
- Rousset, F., R. de Waal Malefijt, B. Sliedrecht, J.P. Aubry, J.Y. Bonnefoy, T. Defrance, J. Banchereau, and J.E. de Vries. 1988. Regulation of Fc receptor for IgE (CD23) and class II MHC antigen expression on Burkitt's lymphoma cell lines by human IL-4 and IFN- γ . *J. Immunol.* **140**: 2625–2632.
- Saito, H., R.A. Maki, L.K. Clayton, and S. Tonegawa. 1983. Complete primary structures of the E β chain and gene of the mouse major histocompatibility complex. *Proc. Natl. Acad. Sci.* **80**: 5520–5524.
- Sakurai, M. and J.L. Strominger. 1988. B-cell-specific enhancer activity of conserved upstream elements of the class II major histocompatibility complex DQB gene. *Proc. Natl. Acad. Sci.* **85**: 6909–6913.
- Schwartz, R. 1985. T lymphocyte recognition of antigen in association with gene products of the major histocompatibility complex. *Annu. Rev. Immunol.* **3**: 237–261.
- Sekimizu, K., B.Y.M. Yung, and A. Kornberg. 1988. The dnaA protein of *Escherichia coli*: Abundance, improved purification, and membrane binding. *J. Biol. Chem.* **263**: 7136–7140.
- Sherman, P.A., P.V. Basta, and J.P.Y. Ting. 1987. Upstream DNA sequences required for tissue-specific expression of the HLA-DR α gene. *Proc. Natl. Acad. Sci.* **84**: 4254–4258.
- Sherman, P.A., P.V. Basta, T.L. Moore, A.M. Brown, and J.P.Y. Ting. 1989. Class II box consensus sequences in the HLA-DR α gene: Transcriptional function and interaction with nuclear proteins. *Mol. Cell. Biol.* **9**: 50–56.
- Singh, H., J.H. LeBowitz, A.S. Baldwin, and P.A. Sharp. 1988. Molecular cloning of an enhancer binding protein: Isolation by screening of an expression library with a recognition site DNA. *Cell* **52**: 415–423.
- Tabor, S. and C.C. Richardson. 1985. A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proc. Natl. Acad. Sci.* **82**: 1074–1078.
- Tsang, S.Y., M. Nakanishi, and B.M. Peterlin. 1988. B-cell-specific and interferon- γ -inducible regulation of the HLA-DR α gene. *Proc. Natl. Acad. Sci.* **85**: 8598–8602.
- . 1990. Mutational analysis of the DRA promoter: *cis*-acting sequences and *trans*-acting factors. *Mol. Cell. Biol.* **10**: 711–719.
- Vinson, C.R., P.B. Sigler, and S.L. McKnight. 1989. Scissors-grip model for DNA recognition by a family of leucine zipper proteins. *Science* **246**: 911–916.
- Vinson, C. R., K.L. LaMarco, P.F. Johnson, W.H. Landschulz, and W.H. McKnight. 1988. In situ detection of sequence-specific DNA binding activity specified by a recombinant bacteriophage. *Genes Dev.* **1**: 801–806.
- Woodward, J.G., K.W. Omer, and P.M. Stuart. 1989. MHC class II transcription in different mouse cell types: Differential requirement for protein synthesis between B cells and macrophages. *J. Immunol.* **142**: 4062–4069.

MHC class II regulatory factor RFX has a novel DNA-binding domain and a functionally independent dimerization domain.

W Reith, C Herrero-Sanchez, M Kobr, et al.

Genes Dev. 1990, **4**:

Access the most recent version at doi:[10.1101/gad.4.9.1528](https://doi.org/10.1101/gad.4.9.1528)

References

This article cites 68 articles, 36 of which can be accessed free at:
<http://genesdev.cshlp.org/content/4/9/1528.full.html#ref-list-1>

License

Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

Focused on your science.