## Characterization of the human CD4 gene promoter: Transcription from the CD4 gene core promoter is tissue-specific and is activated by Ets proteins

PATRICK SALMON\*, ANTOINE GIOVANE<sup>†</sup>, BOHDAN WASYLYK<sup>†</sup>, AND DAVID KLATZMANN<sup>\*‡</sup>

\*Laboratoire de Biologie et Génétique des Pathologies Immunitaires, Centre National de la Recherche Scientifique, Unité Recherche Associée 1463, Hopital de la Pitié-Salpêtrière, 83 Boulevard de l'Hopital, 75651 Paris Cedex 13, France; and <sup>†</sup>Centre National de la Recherche Scientifique-Laboratorie de Génétique Moleculaire des Eucaryotes/Institut National de la Santé et de la Recherche Médicale, Unité 184, Institut de Chimie Biologique, Faculté de Médecine, 67085 Strasbourg, France

Communicated by Jean-Pierre Changeux, April 9, 1993 (received for review December 24, 1992)

ABSTRACT We analyzed the 5' transcription control sequences of the human CD4 gene. We located the transcription initiation site and showed that the CD4 core promoter (positions -40 to +16) lacks a classical "TATA" or initiator positioning consensus sequence but directs precise and efficient transcription when coupled to the ubiquitously active simian virus 40 enhancer. The transcriptional activity of the CD4 gene promoter correlated with CD4 expression in various cell types. Interestingly, the CD4 core promoter also displayed a tissuespecific transcriptional activity. Within this fragment, three nucleic acid sequences are completely conserved in the murine CD4 gene. One of these sequences contains a perfect ETS consensus sequence. Another ETS consensus sequence is located 1060 nt upstream. Electrophoretic-mobility-shift assays showed that the core promoter ETS motif binds an Ets-related protein specifically expressed at high levels in CD4<sup>+</sup> cells. Moreover, in CD4<sup>-</sup> cells, overexpression of Ets-1 or Ets-2 efficiently and specifically activated transcription from the CD4 promoter and core promoter. These data indicate that Ets transcription factors play a central role in controlling CD4 gene expression, by binding to both a classical remote site and an unusual proximal activator sequence.

The expression of the T-cell antigen receptor (TCR), and the CD4 and CD8 coreceptor molecules, is tightly regulated during T-lymphocyte differentiation in the thymus. Early thymocytes express a low level of CD4 and no CD8 and have not yet rearranged their genes coding for the TCR (for review, see ref. 1). Later, they express high levels of both CD4 and CD8, as well as a rearranged TCR. At this stage, most of the thymocytes undergo apoptosis, presumably because of the expression of an inappropriate set of receptor/coreceptor molecules. The surviving thymocytes continue differentiation, losing the expression of either CD4 or CD8, depending on whether their TCR is specific for antigen presented in the context of major histocompatibility complex class I or II molecules, respectively. Experiments with transgenic mice (2, 3), suggest that the switch from CD4<sup>+</sup>/CD8<sup>+</sup>-doublepositive thymocytes to single-positive lymphocytes is likely to be an instructive rather than a stochastic process. Thus, the CD4 and CD8 molecules should control their mutual expression during the thymocyte maturation process. Since the expression of CD4 and CD8 molecules is regulated at the transcriptional level (4, 5), the analysis of the CD4 and CD8 genes transcription control elements is essential to understand the complex regulation of their expression.

We mapped the human CD4 gene transcription start site and showed that the surrounding core promoter (positions -40 to +16) is sufficient for precise initiation of transcription,

despite the absence of "TATA" or initiator sequences (6, 7). The CD4 promoter is specifically active in CD4<sup>+</sup> cells, and the CD4 core promoter alone also displays a significant tissue-specific activity. The CD4 promoter contains two sequences capable of binding Ets proteins, one located 1060 bp upstream from the start site and the second within the core promoter. Ets proteins form a distinct family of transcription factors (for review, see ref. 8), two of which (Ets-1 and Elf-1) bind regulatory sequences of T-cell-specific genes (9, 10). We show that the CD4 core promoter binds Ets-like proteins specifically expressed at high amounts in CD4<sup>+</sup> cells. We also show that overexpression of Ets-1 or Ets-2 in CD4<sup>-</sup> HeLa cells efficiently and specifically activates the CD4 promoter and core promoter. Since T cells express multiple members of the Ets family (10-14) and since Ets-1 and Ets-2 are specifically expressed at high levels in CD4+ thymocytes and lymphocytes (11), we thus propose that Ets proteins play a central role in the tissue specificity of CD4 gene expression.

## **MATERIALS AND METHODS**

Phage, Plasmids, and Oligonucleotides. The two distal EcoRI-HindIII fragments of phage  $\lambda$ MP (24) were subcloned into pGEM-3Z (Promega), sequenced, and aligned with cDNA2 5' sequence (see Fig. 1). Expression plasmids as follows were derived from pCAT vectors (Promega): CAT-Control, simian virus 40 (SV40) promoter plus enhancer (named here SV40); CAT-Enhancer, SV40 enhancer alone; CAT-Basic, no transcription sequences. All subclonings were performed using classical methods (15). The CD4 Pst I fragment from positions -1076 to +20 (Fig. 1A) was cloned in the CAT-Basic, generating proT4. The CD4 core promoter, a CD4 synthetic double-strand oligonucleotide from positions 40 to +16 with *HindIII* and *Xba* I adapters (named CD4 in Fig. 4A), was cloned in both CAT-Basic and CAT-Enhancer, generating pT4 and pT4+eSV, respectively. Basic, proT4, pT4, and pT4+eSV cassettes were tested both in their original pUC19 (pCAT vectors) and in the Bluescript SK<sup>+</sup> (Stratagene) parental vectors, to test for the influence of plasmidic sequences. For each construct, independent preparations of independent clones were prepared for functional assays.

All oligonucleotides were synthesized using the Gene Assembler (Pharmacia). All but sequencing primers were purified in denaturing polyacrylamide gels as described (15).

**Primer-Extension and Mung Bean Nuclease Analyses.** Total RNAs were prepared by the guanidine isothiocyanate/ cesium chloride method and were subjected to primer exten-

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

Abbreviations: CAT, chloramphenicol acetyltransferase; RPC, RNA polymerase II initiation complex; CB, conserved box; SV40, simian virus 40; TCR, T-cell antigen receptor. <sup>‡</sup>To whom reprint requests should be addressed.

Δ

5	5' CTGCA	GCCTCA.	ACTTC B ETS	стааа	CTCAA	GCAATC	CTCCCA	сстса	асстсо	TAAAAT	ACTGG	GATTA	TAGGCA	TGAGC	CACCAC	CTCCCA	GCACCACTTT	-980
٦	TTTCAGAC	TGGAAA	AGAAC	ACTCA	CATGT	GCATCT	TTAAAT	GACACT	TGGGG	TGTGG	ATGGA	GAATG	GCCACO	AGTGA	GTAGGO	AGGAG	статтатсса	-880
1	AGCAAGGG	CTGATA	TTGGC	ATCTT	GGATT	GGCATG	GTGGCA	GTAGTO	GTAGT	GCAGAG	TGACT	TGGGT	AGATTT	TGGAG	CATTTA	GAAGG	TACATCCACA	-780
(	GGAACTGG	TAAATA	AATAC	GTGGG	AGAAG	TTGGGT	GAAGGG	GGTGTO	AAAGA	TTACAG	CCAAT	TTATT	TTGCTT	GGGAA	GTTGGT	GGATG	GTGAGCCCCT	-680
					N	AYB					. N	AYB		_	MYB			
0	CACTGAGT	GAGAAG	CCTGG	AGAAG	CAGGT	TTGGAG	GGTGGT	AGTATO	GCAGGI	GGTATO	CATAG	TTGGG	ATGTGT	GTTGA	GTTTGC	TATGT	CCGGTGAGCT	- 5 8 0
٦	TCCCAGTG	GAGATG	TCCAA	TGGGC	AGACG	GATACT	CACATA	GAGAGI	TCATO	GTAGA	TCGGG	CTAGA	GGAAAG	GACCT	GAGGCO	TGGCC.	AGAGACGCCT	-480
1	AGAGGAAC	AGAGCC	TGGTT	AACAG	STCACT	CCTGGT	GTCTCA	GATATI	гстсто	ACTCAG	CCACG	CCCTC	тсттсс	ACACT	GGGCCA	ACCTAT	AAAGCCTCCA	-380
										l l	VI						MYB	
(	CAGATACC	CCTGGG	GCACC	CACTG	GACAC	AATTGC	CCTCAG	GGCCCC	AGAGO	AAGGAG	SCTGTT	TGTGG	GCTTAC	CACTG	CTGTTC	CCATA	TGCCCOCAAC	-280
3	TIGCCTCCC	ACTTCT	TTCCC	CACAG	CCTGG	TCAGAC	ATGGCA	CTACCA	CTAAT	GGAATO	TTTCT	TGCCA	TCTTTT	TCTTG	CCGTTI	TAACAG	TGGCAGTGAC	-180
1	ACTIGACIO	CCIGAT	TAAGC	CIGAI	TCTGC	TAACT	TTTTCC	CTTGAC	TTTG	SCATTT	CACTT	TGACA	TGTTCC	CTGAG	AGCCTO	GGGGGG	TGGGGAACCA	- 8 0
	COTOCACO	TOOTON		~~~~~	00000	MYB	TAGOAT			TTOOO	TAAAA	ETS		+	1		Hinf1 Pst	
2	ACCAACAA	ACCACC	CGITT	GGGGGC	CGGCC	CAGGCC	TAGGGT	GIGGAG	GAGCO	TIGCC	TCGGG	CTTCC	IGICIC	TCTTC	ATTTAA	AGCACG.	ACTCTGCAGA	+21
1	AGGAACAA	AGCACC	CIUCU	CACIG	FTC	CIGGII	GCAGAG	CICCAR	GICCI	CACACI	GATAC	GCCIG	TITGAG	AAGCA	GCGGGG	AAGAA	AGACGGAAGC	+121
0	CCAGAGGT	ACOTO	CTO AG	ANTOS	COTTO	ATTOOR	000000	10.000	00000		10000	0.1017		O A A T A	*****	COCO I	MIB	
1	CACAGOTO	TGGCCA	COTTO	TOTIO	CATTT	CTCTTC	GAACTG	GTCAT		CCATT	CCACT	CCAGAI	TOTOAG	CTTCC	ACACCT	ACACA	GT CT	
2	CT 93kh	A A T	000000	otoca	ATTOO	CTCTCC	TATTOT					COOTO	C	70070	AGAGG	AGAGA		
- 2	COTOTATA	CTCCAG	GTCCA	CT & A A	TOTTT	COTOAC	TATCA	TTCCC	TTTA	CTCAC	AGGOI	GGGTG	TOTOTO		AGGGAC	SAGG I	GIGIAICGCC	170
2	COTGCCTC	CCTCGG	CAAGG	CCACA	ATGAA	CCCCCC	AGTO T	1.1.4.9.0.)	\$10.101.5 <b>4</b> \$	OTOMO		CCATT	ICIGIO	GGCIC	AGGICC	CTACT	GGUTUAGGUU	+170
1	00100010	001000	04400	00404	Met As	n Ara	Hinfi											
	-																	
- 1	В																	
		- 16	0	- 1 5	0	- 1 4 0		-130		120	- 1	10	- 1 0	0	- 9 (	)	- 80	
- 1	Human	TAAGO	CTGAT	TCTGC	TAAC	TTTTTC	CCTTGA	CTTTG	CATT	TCACT	TGACA	TGTTC	CCTGAG	AGCCT	GGGGGG	TGGGG	AACCAGCTCC	
1	Mouse	CCAGO	TTGAT	TOTGO	TCAGG	CATTTT	CAGTGA	CTTTG	CATTI	TTAAT	TTTCA	ACTTO	CCCAAC	AACTG	GGGGT	GGAGG	GAGGGACTCC	
			L				ALC: PROPERTY								a.a.a.aa			
	- 70		-60		- 5 0	-	40	- 3 (	)	- 20		-10 -		***	¥ +1	0	+ 2 0	
	AGCTGG	TGACGT	TTGGG	occoc	GCCAG	GCCTAG	GOTOTO	GAGGA	CCTTC	CCATCO	GGCTT	CCTGT	CTCTCT	TGATT	TAAGCA	CGACT	CTGCAGAAGG	
	?	TGA	GG	GCTGG	CTTAC	GTCCGC	CGTGCA	GAGGA	CCTCA	CGACC	GGCTT	CCTGT	CTT T	TCATT	TACGA	CATOT	GTGAAGGCAA	
							100	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1			E	rs .	A				*	
													*					

FIG. 1. Human CD4 gene 5' region and sequence homologies between human and mouse CD4 proximal promoter regions. (A) Sequence of the human CD4 gene 5' region. Position +1 corresponds to the major start site identified in Fig. 2. Restriction sites for *Pst* I, *Sac* I, and *Hinf*I are indicated. Sequences homologous to *ETS* and *MYB* consensus sequences are boxed. The first intron, between mRNA nt +127 and +128 (9.6 kb) is shaded. The 5' ends of CD4 cDNA1 (nt +5) and CD4 cDNA2 (nt +117) are boxed. The corresponding 5' end of the murine CD4 promoter (16) is indicated by an M below nt -422. The first three codons of the CD4 protein are indicated. (B) Sequence alignment of human and mouse CD4 proximal promoter regions. Human CD4 sequence (nt -163 to +24) was aligned with the corresponding mouse sequence (16, 17). V, Start sites of human CD4 mRNAs (see Fig. 2); \*, 5' ends of the two murine CD4 cDNAs (18, 19);  $\blacktriangle$ , start sites of the mouse CD4 gene (16). Nucleotide homologies are shaded and CBs are outlined. The nucleotides of the *ETS* consensus core are indicated in boldface type.

sion essentially as described (15). For each primer extension, a plasmid containing a sequence overlapping the expected 5' ends of the RNAs was sequenced using the same 5'-end- $^{32}P$ -labeled oligonucleotide that was used for primer extension. Sequencing reaction products were electrophoresed along-side the primer-extension products. Mung bean nuclease protections were performed essentially as described (15). Conditions used in Fig. 3 are 280 mM NaCl and 100 units of mung bean nuclease (Boehringer Mannheim) for 1 h at 37°C. The 5'-end- $^{32}P$ -labeled oligonucleotides were used as size markers.

Cells and Transfections. Cells used in this study were Jurkat (CD4<sup>+</sup> mature T cell), CEM (CD4<sup>+</sup>/CD8<sup>+</sup> immature T cell), RPMI (CD4<sup>-</sup> pre T cell), U937 (CD4<sup>+</sup> monocytic cell), HeLa (CD4<sup>-</sup> epithelial cell), HepG2 (CD4<sup>-</sup> hepatocytic cell), and U373 (CD4- astrocytic cell, named here "Astro" to avoid confusion with U937 cells). Cells were electrotransfected as described elsewhere (ref. 20 and unpublished data). Determination of chloramphenicol acetyltransferase (CAT) activity was performed 48 h later as described (15) and expressed as percentage of acetylated chloramphenicol. Intraexperiment variations were controlled by cotransfection of the pSV-B-galactosidase control vector (Promega) and by determining the protein concentration of cell extracts. For each construct, independent transfections were performed at least twice. In Fig. 3 C and D, HeLa cells were transfected using the calcium phosphate method as described (15), with  $0-5 \mu g$ of c-Ets-1 or c-Ets-2 expression vectors (21), 5  $\mu$ g of RSV- $\beta$ -galactosidase expression vector, and 5  $\mu$ g of CD4-based or SV40 CAT plasmids. CAT activities are expressed as cpm of monoacetylated [14C]chloramphenicol corrected for variations of the  $\beta$ -galactosidase internal control. For each construct, three experiments with at least two preparations of DNA were performed. The difference between individual experiments was no more than 20%. Stable transformants of CEM and Jurkat cells were obtained by cotransfection with a plasmid encoding the hygromycin B phosphotransferase gene.

**Electrophoretic-Mobility-Shift Assays.** Nuclear extracts were prepared as described (22). Similar titers of octamerbinding factors were recovered from CEM and HeLa nuclear extracts. Electrophoretic-mobility-shift assays were performed essentially as described (23): Nuclear extract (10  $\mu$ g) were incubated in 20  $\mu$ l of binding buffer containing 0.5–1  $\mu$ g of poly(dI-dC)-poly(dI-dC) and the appropriate competitor for 5 min at room temperature. Then, 1 ng of <sup>32</sup>P-labeled probe was added for a further 20-min incubation at room temperature. Nucleoprotein complexes were resolved in a 6% native polyacrylamide gel.

## RESULTS

Localization of the CD4 Transcription Start Site. We used the 5' HinfI fragment (nt +12 to +198 in Fig. 1A) of a human CD4 cDNA (cDNA2) 112 bp longer upstream than the first described human cDNA (cDNA1) (5) to probe  $\lambda$ MP, the most 5' human CD4 genomic clone described (24). Restriction mapping and sequencing of the hybridizing fragments identified a large 9.6-kb intron in the 5' untranslated region (Fig. 1A). Human and murine CD4 genes have thus a similar organization (16, 17).

Primer extension of CD4 mRNAs revealed a major transcript starting at an adenine (+1 in all figures) (Fig. 2A), and minor transcripts starting at nt -9, -1, +2, and +5. Mung bean nuclease mapping gave protected fragments corresponding to 5' ends located between nt +1 and +5 (Fig. 2B), with a major protected fragment corresponding to nt +3. This apparent discrepancy, also observed under various stringency conditions, is presumably due to an overdigestion of the A+T-rich 5' end of the RNA·DNA duplexes (15). Start sites are identical in lymphoid (CEM) and monocytic (U937) cells (Fig. 2 A and B) and are similar to those of the murine CD4 gene (16). A minimal human CD4 gene core promoter (nt -40 to +16) linked to the SV40 enhancer and CAT gene (construct pT4+eSV) was used to generate stable transformants of CEM and Jurkat cells. Primer-extension analysis of



FIG. 2. Identification of human CD4 mRNA 5' end. (A) Primerextension analysis. Total RNAs from yeast, HeLa, U937, and CEM cells were analyzed using an oligonucleotide complementary to nt +35 to +74 of human CD4 mRNA as primer. Arrows indicate the mRNA's 5' ends beside the sequence of nt -18 to +12 of the CD4 gene coding strand. Position +1 corresponds to the major transcription start site. (B) Mung bean nuclease analysis. Total RNAs from yeast and U937 and CEM cells were analyzed using an oligonucleotide complementary to nt -31 to +44 of the CD4 gene. Arrow 40 indicates the migration position of a 40-base oligonucleotide. Thick bar, 5' end of the major protected fragment; thin bars, 5' ends of the minor protected fragments. They are aligned with nt - 7 to +9 of the CD4 gene coding strand. Position +1 designates the major start site identified by primer-extension analysis. (C) Transcription initiation from the CD4 core promoter. Stable transformants of the CD4<sup>+</sup> cell lines CEM (clone 3B3) and Jurkat (clone A5), expressing the CAT gene under the control of the pT4+eSV construct were analyzed by primer extension. An oligonucleotide complementary to nt +3 to +42 of the CAT mRNA (+1 being the first nucleotide of the CAT initiation codon) was used as primer. The major start site of CAT mRNAs is indicated by an arrow (nt +1), beside nt -15 to +12 of the coding strand.

CAT mRNAs showed that the majority of the CAT transcripts start at the +1 adenine of the CD4 core promoter (Fig. 2C).

Structural Analysis of the Human CD4 Gene Promoter. Despite precise initiation, there are no consensus sequences such as TATA box (6), CAAT box (25), or initiator (7) in the vicinity of the transcription start site. We identified three ETS (8) and seven MYB (26) consensus sequences in the CD4 promoter region (Fig. 1A). Of the six MYB consensus sequences described in the murine CD4 promoter (16), only two are found in the corresponding human region. We found a significant conservation (64%) between human and mouse proximal promoter sequences (nt -163 to +20; Fig. 1B), which is concentrated in five boxes of 100% homology. Three of these conserved boxes (CB1-3) are located in the region from positions -35 to +15, which is believed to be critical for the assembly of the RNA polymerase complex (RPC) (27, 28). CB1 and -3 are in places of the TATA box (nt -35 to -27) (6) and initiator (nt -3 to +5) (7), respectively. Interestingly, the CB2 sequence (nt -18 to -7) is precisely the consensus binding site of Ets-1 [RC(C/A)GGA(A/T)RY, minus strand, where R is a purine and Y is a pyrimidine], whereas the remote ETS sequence (nt -1065 to -1057) is perfect for the binding of E74-related proteins [YC(C/A)GGAART, minus strand] (29).

CD4 Promoter and Core Promoter Tissue Specificity and Transactivation by Ets. Compared to the ubiquitously active SV40 promoter, the CD4 promoter is tissue-specific (Fig. 3A). Its activity is similar to that of SV40 promoter in a mature CD4<sup>+</sup> T-cell line (Jurkat), 4-fold lower in CD4<sup>+</sup> monocytic cells (U937), and a least 10-fold lower in CD4cells, including the RPMI pre-T cells. The activity of proT4 is also comparable to that of SV40 in an immature CD4<sup>+</sup>/ CD8<sup>+</sup> T-cell line (CEM, data not shown). Surprisingly, the CD4 core promoter (nt -40 to +16) alone exhibits a significant activity in Jurkat cells but not in HeLa cells (Fig. 3B). This indicates the presence of positive regulatory sequences both upstream and within the CD4 core promoter. When the CD4 core promoter is coupled to the SV40 enhancer (pT4+eSV), its transcriptional activity is similar in CD4<sup>+</sup> and CD4<sup>-</sup> cells and is in the same order of magnitude as the SV40 promoter.

In CD4<sup>-</sup> Hela cells, the activity of the CD4 promoter and core promoter can be restored by the expression of Ets-1 or Ets-2. The CD4 core promoter alone (pT4) and to a greater extent the complete CD4 promoter (proT4) are transactivated in a dose-dependent manner by Ets-1 or Ets-2 expression (Fig. 3 C and D). These Ets proteins did not significantly affect the SV40 promoter activity. These data show that both CD4 core promoter and upstream sequences are targets for activation by Ets proteins.

Binding of Ets-Like Proteins to the Human CD4 Core Promoter. We performed gel-retardation assays with both the 56-bp CD4 core promoter sequence and a sequence containing an artificial ETS motif derived from the described consensus sequence (8). Incubation of a labeled CD4 probe with nuclear extracts from CD4<sup>+</sup> CEM cells gave three retarded bands (C1, C2, and C3; Fig. 4B). Preincubation with an excess of unlabeled CD4, as well as unlabeled ETS motifs, inhibited the formation of the most intense C3 band, although with a lesser efficiency for ETS. These different affinities of CD4 and ETS for the C3 protein(s) are most likely due to base differences between their Ets-binding sites (21, 29). No significant inhibition was observed when using as competitors an altered ETS motif with a single  $C \rightarrow T$  substitution (MUT) or a poly(A) oligonucleotide (pA). By using nuclear extracts from the CD4<sup>-</sup> HeLa cells, the C1 and C2 bands displayed intensities comparable to bands of CEM extracts, whereas the C3 band was almost undetectable. Furthermore, only the C3 band could be significantly inhibited by the CD4 unlabeled competitor. Using the ETS double-strand oligonucleotide as probe, we observed a similar pattern of retarded bands. Three bands (E1, E2, and E3) were observed with nuclear extracts of CEM cells, whereas only two (E1 and E2) were found with HeLa cells. An excess of unlabeled ETS oligonucleotide inhibited the formation of E3, although less efficiently than the CD4 oligonucleotide. No significant inhibition was observed with MUT or the pA oligonucleotides (data not shown).

## DISCUSSION

Structure and Tissue Specificity of the Human CD4 Promoter. We showed that the transcription initiation of the CD4 gene is precise and that a genomic fragment from nt -1076 to +20 governs tissue-specific transcriptional activity in CD4<sup>+</sup> cells. Our results are in agreement with those recently reported (16) for a murine CD4 gene promoter fragment of similar size whose activity is restricted to mature CD4<sup>+</sup>/ CD8<sup>-</sup> mouse lymphocytes. Within the 1100-bp human fragment, we show that the region from positions -40 to +16 can direct precise transcription at the same site as natural CD4 mRNAs and displays a significant tissue-specific basal activity. This latter feature is unusual for a minimal core promoter, although sequences of similar size have been



FIG. 3. CD4 promoter and core promoter are tissue-specific and are activated by Ets. (A) Cell-specific expression of the CD4 promoter. CD4<sup>+</sup> (Jurkat and U937) and CD4<sup>-</sup> (RPMI, Hela, Astro, and HepG2) cell lines were electroporated with SV40 and proT4 plasmids in a pUC19 parental plasmid. CAT activities were determined as percentages of acetylated vs. total chloramphenicol, and proT4 activity was expressed as its relative value to SV40 activity. Means (bars) and standard deviations to the mean (error bars) of these relative activities are shown. The horizontal line shows the SV40 relative activity (index = 1). Average absolute activities (chloramphenicol acetylation percentage per hour) for the SV40 promoter (and Basic) constructs were  $\approx 24$  (1.8) for Jurkat, 27 (0.15) for U937, 80 (0.4) for RPMI, 140 (0.15) for HeLa, 83 (0.25) for Astro, and 41 (0.3) for HepG2. (B) Transcriptional activities of the CD4 promoter and core promoter in HeLa and Jurkat cells. Basic, proT4, pT4, and pT4+eSV were tested in the Bluescript parental plasmid. Activities of each construct, determined as in A, were related to SV40 promoter activity (index = 1). Results are expressed as mean (bars) and standard deviations to the mean (error bars) of these relative activities. Average absolute activities (chloramphenicol acetylation percentage per hour) for the SV40 promoter activity (index = 1). Results are expressed as mean (bars) and standard deviations to the mean (error bars) of these relative activities. Average absolute activities (chloramphenicol acetylation percentage per hour) for the SV40 promoter were  $\approx 80$  for Jurkat and 50 for HeLa. (C) Transactivation of CD4 promoter and core promoter were  $\approx 80$  for Jurkat and 50 for HeLa. (C) Transactivation of CD4 promoter and core promoter by Ets-1. HeLa cells were cotransfected with 5  $\mu$ g of pT4, proT4, or SV40; 5  $\mu$ g of RSV- $\beta$ -galactosidase vector; and the indicated concentrations of a c-Ets-1 expression vector. CAT activities are expressed as corrected cpm. Results of a represent

shown to participate in tissue-specific transcription *in vitro* (30).

The human CD4 core promoter contains three boxes (CB1-3), which are identical in mouse CD4 gene. The locations of CB1 and CB3 suggest that they could be involved in

RPC assembly, although they do not correspond to the consensus reported at these sites, i.e., TATA box (6), initiator (CTCANTCT) (7), or HIP1 (ATTTCN<sub>1-30</sub>GCCA) (31). The second CB (CB2) contains a perfect sequence for the binding of the Ets-1 protein (29). We showed that the CD4



FIG. 4. CD4 core promoter binds an Ets-related protein present in CD4<sup>+</sup> but not in CD4<sup>-</sup> cells. (A) Sequences of the double-strand oligonucleotides used for the electrophoretic-mobility-shift assays. CD4 is the oligonucleotide also used in the CAT constructs. ETS contains an artificial Ets-binding site designed on the basis of the reported consensus sequence (8) inserted in a nonrelevant sequence. MUT is identical to ETS except for the detrimental  $G \rightarrow A$  mutation (21) (minus strand). Oligonucleotide pA was used as negative control. CD4 core promoter conserved sequences and Ets-binding core sequences are boxed. (B) Electrophoreticmobility shifts of CD4 with nuclear extracts from CEM and HeLa cells. Competitors used are indicated above their respective molar excess to the CD4 radiolabeled probe. Lane 0 means no competitor added. Arrows indicate the positions of the free probe (F) and of the retarded DNA-protein complexes (C1, C2, and C3).

core promoter specifically binds an Ets protein present at high amounts specifically in CD4<sup>+</sup> cells. Moreover, Ets-1 and Ets-2 activate the CD4 core promoter in the CD4<sup>-</sup> HeLa cells, and preliminary experiments indicated that mutations within CB2 dramatically impairs activation by Ets (data not shown). Thus, it strongly suggests that Ets proteins are implicated in the transcriptional activity of the CD4 core promoter in CD4<sup>+</sup> cells, by activating the RPC from within. As recently discussed for other TATA-less promoters, Ets proteins could participate in the building of a CD4-specific RPC (8, 32). Independently, we show that the complete CD4 promoter is more active in CD4<sup>+</sup> cells and more strongly activated by Ets in CD4<sup>-</sup> cells than the CD4 core promoter. Thus, Ets proteins could also "classically" modulate CD4 gene expression by binding to the other remote upstream ETS site (nt -1060).

Ets proteins, whose prototype is the Ets-1 protooncogene, define a distinct class of transcriptional factors that bind several gene regulatory sequences containing the ETS core motif GGA (for review, see ref. 8). Among these proteins, Ets-1 has been shown to control a T-cell-specific gene  $(TCR\alpha)$  (9), and Elf-1 is implicated in a T-helper-cell-specific gene expression (interleukin 2) (10). Interestingly, the sequence of the distal ETS site was shown to be preferred by Elf-1-related proteins (29). Thus, the cooperation of these two proteins could be a key mechanism for CD4 gene restricted expression. Ets-1 and Ets-2 are expressed at 10fold higher levels in CD4<sup>+</sup> thymocytes and lymphocytes than in their CD8<sup>+</sup> counterparts (11). However, Ets-1 and Ets-2 are expressed at low but significant levels in CD8<sup>+</sup> lymphocytes, whereas the CD4 gene expression is undetectable in these cells. These apparent discrepancies might be explained by (i) the fact that Ets is not the sole determinant of CD4 promoter activity, (ii) a threshold effect of Ets expression on CD4 promoter activity, or (iii) the presence of interfering forms of Ets in CD4<sup>-</sup> cells. For optimal activity, Ets proteins seem to require the adjacent binding of another transcriptional factor (for review, see ref. 8), including Myb (33). Interestingly, Myb is also critical for CD4 promoter activity (16), although it does not by itself significantly contribute to its tissue specificity. Thus, the tissue-specific CD4 promoter activity could depend on the cooperation of Ets proteins, with Myb at the distal ETS-MYB overlapping sites, or with a general transcription factor at the CD4 core promoter ETS site. Finally, preliminary experiments indicated that truncated forms of Ets containing only the DNA binding domain (34) dramatically decrease CD4 promoter activity in CD4<sup>+</sup> cells (data not shown). Altogether, these results clearly indicate that Ets proteins play a central role in tissuerestricted CD4 gene expression. However, the CD4-specific Ets proteins remain to be identified.

The CD4 Promoter and Thymocyte Phenotype Switching. Experiments with transgenic mice suggested that the CD4 and CD8 proteins could control their mutual expression during thymocyte diffentiation (2, 3). Such events could likely be triggered by  $p56^{lck}$ , a member of the v-src protein tyrosine kinase family that is associated with the intracytoplasmic tails of these proteins (35). Indeed, (*i*)  $p56^{lck}$  phosphorylates and activates mitogen-activated protein (MAP) kinases (36), (*ii*) Ets phosphorylation by MAP kinase modifies its ability to form ternary complex with other transcription factors (37), and (*iii*) v-src has been shown to modulate Ets transcriptional activity (23). Thus, Ets could be the nuclear target of extracellular signals that might trigger thymocyte phenotype switch. from Association National pour la Recherche du SIDA, Centre National de la Recherche Scientifique, Direction de la Recherche et des Etudes Doctorales, Institut National de la Santé et de la Recherche Médicale, Association pour la Recherche contre le Cancer, and Federation Nationale des Centres de Lutte Contre le Cancer.

- 1. Nikolic-Zugic, J. (1991) Immunol. Today 12, 65-70.
- Robey, E. A., Fowlkes, B. J., Gordon, J. W., Kioussis, D., von Boehmer, H., Ramsdell, F. & Axel, R. (1991) Cell 64, 99-107.
- Seong, R. H., Chamberlain, J. W. & Parnes, J. R. (1992) Nature (London) 356, 718-720.
- Littman, D. R., Thomas, Y., Maddon, P. J., Chess, L. & Axel, R. (1985) Cell 40, 237-246.
- Maddon, P. J., Littman, D. R., Godfrey, M., Maddon, D. E., Chess, L. & Axel, R. (1985) Cell 42, 93-104.
- Breathnach, R. & Chambon, P. (1981) Annu. Rev. Biochem. 50, 349-383.
- Smale, S. T., Schmidt, M. C., Berk, A. J. & Baltimore, D. (1990) Proc. Natl. Acad. Sci. USA 87, 4509–4513.
- Wasylyk, B., Hahn, S. L. & Giovane, A. (1993) Eur. J. Biochem. 211, 7-18.
- Ho, I.-C., Bhat, N. K., Gottschalk, L. R., Lindsten, T., Thompson, C. B., Papas, T. S. & Leiden, J. M. (1990) Science 250, 814-818.
  Thompson, C. B., Wang, C., Ho, I., Bohjanen, P. R., Petryniak,
- Thompson, C. B., Wang, C., Ho, I., Bohjanen, P. R., Petryniak, B., June, C. H., Miesfeldt, S., Zhang, L., Nabel, G. H., Karpinski, B. & Leiden, J. M. (1992) Mol. Cell. Biol. 12, 1043-1053.
- Bhat, N. K., Komschlies, K. L., Fujiwara, S., Fisher, R. J., Mathieson, B. J., Gregorio, T. A., Young, H. A., Kasik, J. W., Ozato, K. & Papas, T. S. (1989) J. Immunol. 142, 672-678.
- Klemsz, M. J., McKercher, S. R., Celada, A., Beveren, C. V. & Maki, R. A. (1990) Cell 61, 113-124.
- 13. Ben-David, Y., Giddens, E. B., Letwin, K. & Bernstein, A. (1991) Genes Dev. 5, 908-918.
- LaMarco, K., Thompson, C. C., Byers, B. P., Walton, E. M. & McKnight, S. L. (1991) Science 253, 789-792.
- Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Lab. Press, Plainview, NY), 2nd Ed.
- Siu, G., Wurster, A. L., Lipsick, J. S. & Hedrick, S. M. (1992) Mol. Cell. Biol. 12, 1592–1604.
- 17. Gorman, S. D., Tourvieille, B. & Parnes, J. R. (1987) Proc. Natl. Acad. Sci. USA 84, 7644-7648.
- Tourvieille, B., Gorman, S. D., Field, E. H., Hunkapiller, T. & Parnes, J. R. (1986) Science 234, 610-614.
- 19. Littman, D. R. & Gettner, S. N. (1987) Nature (London) 325, 453-455.
- 20. Salmon, P. (1991) Ph.D. thesis (Pierre and Marie Curie Univ., Paris).
- Wasylyk, C., Gutman, A., Nicholson, R. & Wasylyk, B. (1991) EMBO J. 10, 1127-1134.
- 22. Schreiber, E., Matthias, P., Müller, M. M. & Schaffner, W. (1989) Nucleic Acids Res. 17, 6419.
- Wasylyk, C., Flores, P., Gutman, A. & Wasylyk, B. (1989) EMBO J. 8, 3371-3378.
- Maddon, P. J., Molineaux, S. M., Maddon, D. E., Zimmerman, K. A., Godfrey, M., Alt, F. W., Chess, L. & Axel, R. (1987) Proc. Natl. Acad. Sci. USA 84, 9155-9159.
- 25. McKnight, S. L. & Kingsbury, R. (1982) Science 217, 316-324.
- Biedenkapp, H., Borgmeyer, U., Sippel, A. E. & Klempnauer, K.-H. (1988) Nature (London) 335, 835-837.
- van Dyke, M. W., Roeder, R. G. & Sawadogo, M. (1988) Science 241, 1335-1338.
- Buratowski, S., Hahn, S., Guarente, L. & Sharp, P. A. (1989) Cell 56, 549-561.
- Nye, J. A., Petersen, J. M., Gunther, C. V., Jonsen, M. D. & Graves, B. J. (1992) Genes Dev. 6, 975–990.
- Tamura, T., Sumita, K., Hirose, S. & Mikoshiba, K. (1990) EMBO J. 9, 3101-3108.
- 31. Means, A. L. & Farnham, P. J. (1990) Mol. Cell. Biol. 10, 653-661.
- 2. Sharp, P. A. (1992) Cell 68, 819-821.
- Dudek, H., Tantravahi, R. V., Rao, V. N., Reddy, E. S. & Reddy, E. P. (1992) Proc. Natl. Acad. Sci. USA 89, 1291-1295.
- Wasylyk, C., Kerckaert, J.-P. & Wasylyk, B. (1992) Genes Dev. 6, 965-974.
- Veillette, A., Bookman, M. A., Horak, E. M. & Bolen, J. B. (1988) Cell 55, 301–308.
- Ettahadieh, E., Sanghera, J. S., Pelech, S. L., Hess-Bienz, D., Watts, J., Shastri, N. & Aebersold, R. (1992) Science 255, 853-855.
- 37. Gille, H., Sharrocks, A. D. & Shaw, P. E. (1992) Nature (London) 358, 414-417.