Genomic organization and promoter characterization of human CXCR4 gene

A. Caruz[a], M. Samsom[b], J.M. Alonso[c], J. Alcami[c], F. Baleux[d], J.L. Virelizier[a], M. Parmentier[b], F. Arenzana-Seisdedos[a,*]

[a] *Unité d'Immunologie Virale. Institut Pasteur, 28 Rue Dr. Roux, 75724 Paris, France*
[b] *IRIBHN, Université Libre de Bruxelles, B-1070 Bruxelles, Belgium*
[c] *Centro de Investigaciones, Hospital 12 de Octubre, Madrid 28041, Spain*
[d] *Unité de Chimie Organique. Institut Pasteur, 28 Rue Dr. Roux, 75724 Paris, France*

Received 13 March 1998

Abstract CXCR4 is the receptor for the CXC chemokine SDF1 that has essential functions on embryo organogenesis, immunological functions and T lymphocyte trafficking. Recently, CXCR4 has drawn unexpected attention as it was recently identified as a co-factor required for entry of lymphotropic HIV isolates in CD4+ T lymphocytes. CXCR4 is the only SDF1 receptor identified so far. This suggests that CXCR4 expression is critical for the biological effects of SDF1. To investigate the mechanisms controlling both the constitutive and induced expression of CXCR4 receptors we have isolated and characterized the promoter region and determined the genomic structure of the human gene. The CXCR4 gene contains two exons separated by an intronic sequence. A 2.6 kb 5′-flanking region located upstream the CXCR4 open reading frame contains a TATA box and the transcription start site characteristic of a functional promoter. This region also contains putative consensus binding sequences for different transcription factors, some of them associated with the hemopoiesis and lymphocyte development.

Key words: CXC4 chemokine receptor; Promoter regulation; Transcription factor; HIV-1; Chemokine SDF1

1. Introduction

Chemokines are soluble mediators that control leukocyte migration into sites of inflammation and specific areas of lymphoid organs [1]. Activation of leukocytes is mediated by interaction of chemokines with specific receptors belonging to the family of the seven transmembrane domain, G-protein coupled receptors [2]. Chemokines carry conserved cysteines and fall into four families defined by a cysteine signature motif: CXC, CC, C or CX_3C where C is a cysteine and X any amino-acid residue [3]. Most of the chemokines belong to either CC or CXC families. Expression of CC chemokine receptors in T lymphocytes has proved to be dependent on the cell activation status and the nature of the signal used as stimulus. Thus, expression of CCR1, CCR2 or CCR5 are induced by IL2 [4,5], but are down-regulated by lectin or T cell specific activators [4,6]. Less information is available for the CXC family of receptors, although it has been reported that expression of IL8 receptors in lymphocytes is modified neither by IL2- nor CD3-mediated activation [4]. Among the family of CXC receptors, CXCR4 has drawn unexpected attention as it was identified as a co-factor required for entry of lymphotropic HIV isolates in CD4+ T lymphocytes [7]. CXCR4 is the receptor for the CXC chemokine SDF1 [8,9] that plays a critical role in both embryonic B lymphopoiesis and organogenesis [10]. A chemoattractant for the pluripotent hematopoietic CD34+ cells [11] and a growth factor for pre- and pro-B murine lymphocytes in vitro [12], SDF1 is also known to promote the arrest on endothelium of lymphocytes rolling under flow conditions [13]. Moreover, SDF1 prevents HIV entry in host cells by both occupying [8,9] and promoting endocytosis of CXCR4 [14]. Nowadays, CXCR4 seems to be the only receptor capable of binding SDF1 selectively and permit cell signaling by this chemokine [8,15]. CXCR4 is expressed constitutively in a large number of human tissues and transformed cell lines [15]. Moreover, CXCR4 transcripts are very abundant in leukocytes [15].

To investigate the mechanisms controlling both the constitutive and induced expression of CXCR4 receptors, we have isolated and characterized the promoter region and determined the genomic structure of the human gene. The CXCR4 gene contains two exons separated by an intronic sequence. A 2.6 kb 5′-flanking region located upstream the CXCR4 open reading frame contains a TATA box and the transcription start site characteristic of a functional promoter. This region also contains putative consensus binding sequences for different transcription factors, some of them associated with hemopoiesis and lymphocyte development.

2. Materials and methods

2.1. Cells

The epithelial-like cell line HeLa was propagated in DMEM medium supplemented with 10% heat-inactivated fetal calf serum (FCS, Gibco), L-Gln and antibiotics. J.jhan cells were derived from the T lymphoblastoid cell line Jurkat and were grown in RPMI 1640 supplemented with 10% FCS. Peripheral blood mononuclear cells (PBMC) were isolated by Ficoll gradient from healthy human donors.

2.2. DNA sequencing

Both DNA strands relative to the CXCR4 gene and the promoter region were sequenced by the dideoxy-chain termination method using the New England Biolabs cycle sequencing Kit and an automated sequencer ABI (Perkin Elmer). All sequencings were done in duplicate.

2.3. Library screening

A human BAC library from Research Genetics, Inc (2130 Memorial Parkway SW, Huntsville, AL 35801, USA) was screened according to the manufacturer's instructions by specific PCR for the CXCR4 gene. The primers used were CXCR45459: 5′-CCGAGGAAATGGGCT-

*Corresponding author. Fax: +33 1-45688941.
E-mail: farenzan@pasteur.fr

The nucleotide sequence data reported in this paper will appear in the EMBL database with the accession number AJ224869.

PII S0014-5793(98)00359-7

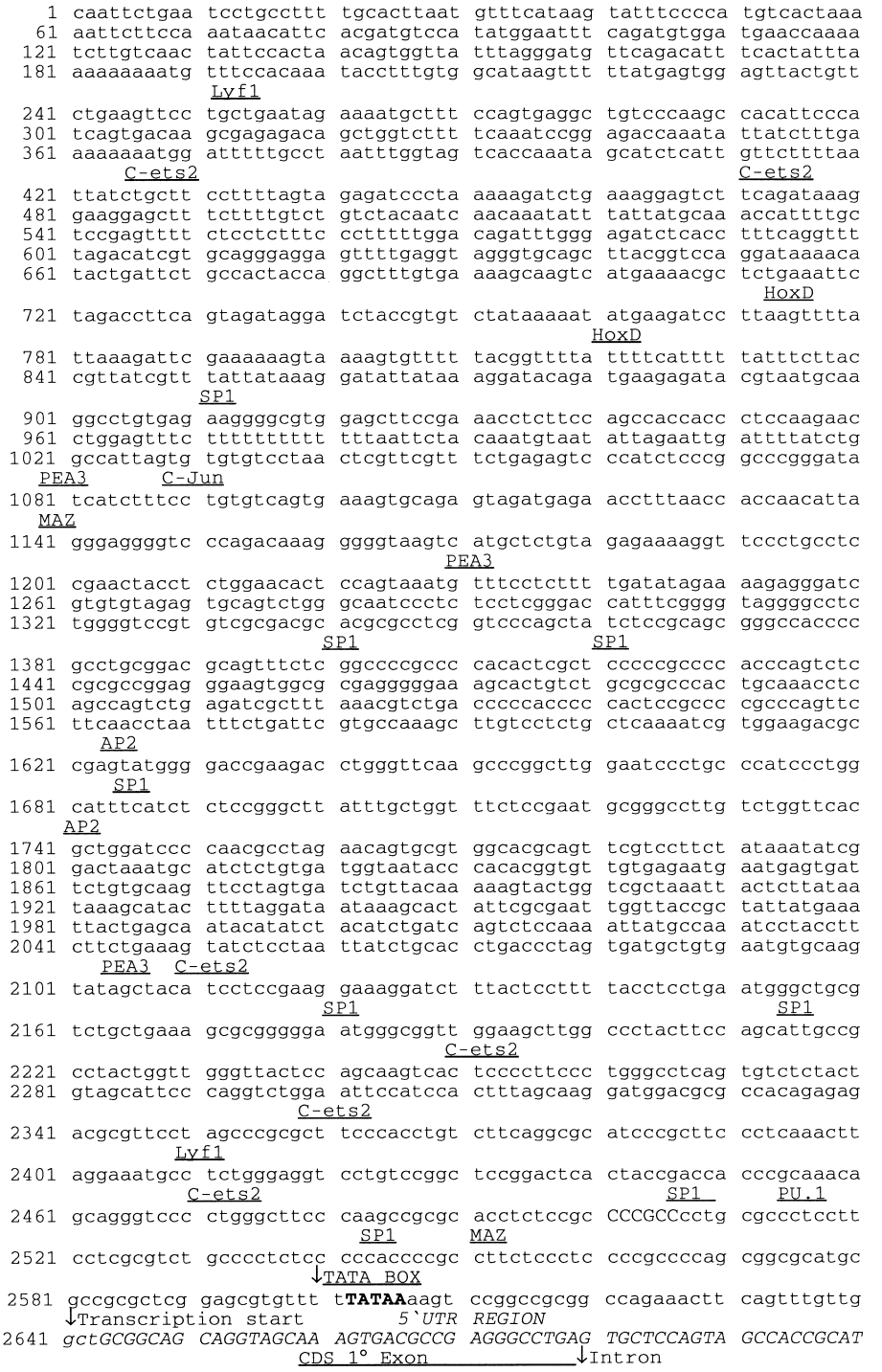
1 caattctgaa tcctgccttt tgcacttaat gtttcataag tatttcccca tgtcactaaa
61 aattcttcca aataacattc acgatgtcca tatggaattt cagatgtgga tgaaccaaaa
121 tcttgtcaac tattccacta acagtggtta tttagggatg ttcagacatt tcactattta
181 aaaaaaaatg tttccacaaa tacctttgtg gcataagttt ttatgagtgg agttactgtt
Lyf1
241 ctgaagttcc tgctgaatag aaaatgcttt ccagtgaggc tgtcccaagc cacattccca
301 tcagtgacaa gcgagagaca gctggtcttt tcaaatccgg agaccaaata ttatctttga
361 aaaaaaatgg atttttgcct aatttggtag tcaccaaata gcatctcatt gttcttttaa
C-ets2
C-ets2
421 ttatctgctt ccttttagta gagatcccta aaaagatctg aaaggagtct tcagataaag
481 gaaggagctt tcttttgtct gtctacaatc aacaaatatt tattatgcaa accattttgc
541 tccgagtttt ctcctctttc ccttttttgga cagatttggg agatctcacc tttcaggttt
601 tagacatcgt gcagggagga gttttgaggt agggtgcagc ttacggtcca ggataaaaca
661 tactgattct gccactacca ggctttgtga aaagcaagtc atgaaaacgc tctgaaattc
HoxD
721 tagaccttca gtagatagga tctaccgtgt ctataaaaat atgaagatcc ttaagtttta
HoxD
781 ttaaagattc gaaaaaagta aaagtgtttt tacggtttta ttttcatttt tatttcttac
841 cgttatcgtt tattataaag gatattataa aggatacaga tgaagagata cgtaatgcaa
SP1
901 ggcctgtgag aaggggcgtg gagcttccga aacctcttcc agccaccacc ctccaagaac
961 ctggagtttc ttttttttt tttaattcta caaatgtaat attagaattg attttatctg
1021 gccattagtg tgtgtcctaa ctcgttcgtt tctgagagtc ccatctcccg gcccgggata
PEA3
C-Jun
1081 tcatctttcc tgtgtcagtg aaagtgcaga gtagatgaga acctttaacc accaacatta
MAZ
1141 gggaggggtc ccagacaaag ggggtaagtc atgctctgta gagaaaaggt tccctgcctc
PEA3
1201 cgaactacct ctggaacact ccagtaaatg tttcctcttt tgatatagaa aagagggatc
1261 gtgtgtagag tgcagtctgg gcaatccctc tcctcgggac catttcgggg taggggcctc
1321 tggggtccgt gtcgcgacgc acgcgcctcg gtcccagcta tctccgcagc gggccacccc
SP1
SP1
1381 gcctgcggac gcagtttctc ggccccgccc cacactcgct cccccgcccc acccagtctc
1441 cgcgccggag ggaagtggcg cgaggggggaa agcactgtct gcgcgcccac tgcaaacctc
1501 agccagtctg agatcgcttt aaacgtctga ccccccaccc cactccgccc cgcccagttc
1561 ttcaacctaa tttctgattc gtgccaaagc ttgtcctctg ctcaaaatcg tggaagacgc
AP2
1621 cgagtatggg gaccgaagac ctgggttcaa gcccggcttg gaatccctgc ccatccctgg
SP1
1681 catttcatct ctccgggctt atttgctggt ttctccgaat gcgggccttg tctggttcac
AP2
1741 gctggatccc caacgcctag aacagtgcgt ggcacgcagt tcgtccttct ataaatatcg
1801 gactaaatgc atctctgtga tggtaatacc cacacggtgt tgtgagaatg aatgagtgat
1861 tctgtgcaag ttcctagtga tctgttacaa aaagtactgg tcgctaaatt actcttataa
1921 taaagcatac ttttaggata ataaagcact attcgcgaat tggttaccgc tattatgaaa
1981 ttactgagca atacatatct acatctgatc agtctccaaa attatgccaa atcctacctt
2041 cttctgaaag tatctcctaa ttatctgcac ctgaccctag tgatgctgtg aatgtgcaag
PEA3
C-ets2
2101 tatagctaca tcctccgaag gaaaggatct ttactccttt tacctcctga atgggctgcg
SP1
SP1
2161 tctgctgaaa gcgcgggggа atgggcggtt ggaagcttgg ccctacttcc agcattgccg
C-ets2
2221 cctactggtt gggttactcc agcaagtcac tccccttccc tgggcctcag tgtctctact
2281 gtagcattcc caggtctgga attccatcca ctttagcaag gatggacgcg ccacagagag
C-ets2
2341 acgcgttcct agcccgcgct tcccacctgt cttcaggcgc atcccgcttc cctcaaactt
Lyf1
2401 aggaaatgcc tctgggaggt cctgtccggc tccggactca ctaccgacca cccgcaaaca
C-ets2
SP1
PU.1
2461 gcagggtccc ctgggcttcc caagccgcgc acctctccgc CCCGCCcctg cgccctcctt
SP1
MAZ
2521 cctcgcgtct gcccctctcc cccaccccgc cttctccctc cccgccccag cggcgcatgc
↓TATA BOX
2581 gccgcgctcg gagcgtgttt t**TATAA**aagt ccggccgcgg ccagaaactt cagtttgttg
↓Transcription start 5`UTR REGION
2641 *gctGCGGCAG CAGGTAGCAA AGTGACGCCG AGGGCCTGAG TGCTCCAGTA GCCACCGCAT*
CDS 1° Exon ↓Intron
2701 CTGGAGAACC AGCGGTTACC **ATGGAGGGGA TCAGTGTAAG T**ccagtttca acctgctttg
2761 tcataaatgt acaaacgttt gaacttagag cgcagcccct ctccgagcgg gcagaagcgg
2821 ccaggacatt ggaggtaccc tgaccatcct atatagtgcg ggtgggtggg gggggagcag
2881 gattggaatc tttttctctg tgagtcgagg agaaacgact ggaaagagcg ttccagtggc
2941 tgcatgtgtc tccccccttga gtcccgccgc gcgcggcggc ttgcacgctg tttgcaaacg
3001 taagaacatt ctgtgcacaa gtgcagagaa ggcgtgcgcg ctgcctcggg actcagacca
3061 ccggtctctt ccttggggaa gcggggatgt cttggagcga gttacattgt ctgaatttag
3121 aggcggaggg cggcgtgcct gggctgagtt cccaggagga gattgcgccc gctttaactt
3181 cggggttaag cgcctggtga ctgttcttga cactgggtgc gtgtttgtta aactctgtgc
3241 ggccgacgga gctgtgccag tctcccagca cagtaggcag agggcgggag aggcgggtgg
3301 acccaccgcg ccgatcctct gaggggatcg agtggtggca gcagctagga gttgatccgc
3361 ccgcgcgctt tgggtttgag ggggaaacct tcccgccgtc cgaagcgtgc ctcttcccca
3421 cggccgcgag tgggtcctgc agttcgagag tttggggtcg tgcagaggtc agcggagtgg
3481 tttgacctcc cctttgacac cgcgcagctg ccagccctga gatttgcgct ccggggatag
3541 gagcgggtac ggggtgaggg gcgggggcgg ttaaaccgca cctgggctgc caggtcgccg

Fig. 1.

CAGGGGA-3′ and CXCR45462: 5′-CATTGGGGTAGAAGCG-GTCACA-3′. The PCR mixtures contained 5 μl of DNA from the BAC library pool, 1 μl of each primer at a concentration of 500 ng/μl, 200 mM dNTPs, 1 mM of $MgCl_2$, 0.2 units of Taq DNA polymerase (Pharmacia), 5 μl of DMSO and water to 100 μl. Amplification was conducted for 30 cycles at 93°C 1 min, 62°C 2 min, 72°C 2.30 min, followed by a final elongation step at 72°C for 15 min. The PCR products were visualized on a 2% agarose gel stained with ethidium bromide. This procedure led to the identification of a BAC clone (24D3) that contained the entire CXCR4 gene and adjacent sequences. The identification of a putative TATA box and consensus binding sites for transcription factors were performed using the program TESS (Transcription Element Search Software) (http://agave.humgen.upenn.edu/tess/index.html). Only sequences showing total homology with the canonical consensus sequence for transcription factors were considered.

```
3601 ccgcgaagac tggcaggtgc aagtggggaa accgtttggc tctctccgag tccagttgtg
3661 atgtttaacc gtcggtggtt tccagaaacc ttttgaaacc ctcttgctag ggagtttttg
3721 gtttcctgca gcggcgcgca attcaaagac gctcgcggcg gagccgccca gtcgctcccc
3781 agcaccctgt gggacagagc ctggcgtgtc gcccagcgga gcccctgcag cgctgcttgc
3841 gggcggttgg cgtgggtgta gtgggcagcc gcggcggccc ggggctggac gacccggccc
3901 cccgcgtgcc caccgcctgg aggcttccag ctgcccacct ccggccgggt taactggatc
3961 agtggcgggg taatgggaaa ccacccggga gagtgaggaa atgaaacttg gggcgaggac
4021 cacgggtgca gaccccgtta ccttctccac ccaggaaaat gccccgctcc ctaacgtccc
4081 aaacgcgcca agtgataaac acgaggatgg caagagaccc acacaccgga ggagcgcccg
4141 cttgggggag gaggtgccgt ttgttcattt tctgacactc ccgcccaata taccccaagc
4201 accgaagggc cttcgtttta agaccgcatt ctctttaccc actacaagtt gcttgaagcc
4261 cagaatggtt tgtatttagg caggcgtggg aaaattaagt ttttgcgcct taggagaatg
4321 agtctttgca acgcccccgc cctcccccg tgatcctccc ttctcccctc ttccctccct
4381 gggcgaaaaa cttcttacaa aaagttaatc actgcccctc ctagcagcac ccaccccacc
4441 ccccacgccg cctgggagtg gcctctttgt gtgtattttt tttttcctcc taaggaagtt
4501 ttttttcttc cctctagtgg gcggggcaga gggattagcc aagatgttga ctttgaaacc
4561 ctcagcgtct cagtgccctt ttgttctaaa caaagaattt tgtaattggt tctaccaaag
4621 aaggatataa tgaagtcact atgggaaaag atggggaaga aaagatgggg aggagagttg
4681 taggattcta cattaattct cttgtgccct tagcccacta cttcagaatt tcctgaagaa
4741 agcaagcctg aattggtttt ttaaattgct ttaaaaaatt tttttaactg ggttaatgct
                                              ↓2° Exon
4801 tgctgaattg gaagtgaatg tccattcctt tgcctctttt gcaGATATAC ACTTCAGATA
4861 ACTACACCGA GGAAATGGGC TCAGGGGACT ATGACTCCAT GAAGGAACCC TGTTTCCGTG
4921 AAGAAAATGC TAATTTCAAT AAAATCTTCC TGCCCACCAT CTACTCCATC ATCTTCTTAA
4981 CTGGCATTGT GGGCAATGGA TTGGTCATCC TGGTCATGGG TTACCAGAAG AAACTGAGAA
5041 GCATGACGGA CAAGTACAGG CTGCACCTGT CAGTGGCCGA CCTCCTCTTT GTCATCACGC
5101 TTCCCTTCTG GGCAGTTGAT GCCGTGGCAA ACTGGTACTT TGGGAACTTC CTATGCAAGG
5161 CAGTCCATGT CATTTACACA GTCAACCTCT ACAGCAGTGT CCTCATCCTG GCCTTCATCA
5221 GTCTGGACCG CTACCTGGCC ATCGTCCACG CCACCAACAG TCAGAGGCCA AGGAAGCTGT
5281 TGGCTGAAAA GGTGGTCTAT GTTGGCGTCT GGATCCCTGC CCTCCTGCTG ACTATTCCCG
5341 ACTTCATCTT TGCCAACGTC AGTGAGGCAG ATGACAGATA TATCTGTGAC CGCTTCTACC
5401 CCAATGACTT GTGGGTGGTT GTGTTCCAGT TTCAGCACAT CATGGTTGGC CTTATCCTGC
5461 CTGGTATTGT CATCCTGTCC TGCTATTGCA TTATCATCTC CAAGCTGTCA CACTCCAAGG
5521 GCCACCAGAA GCGCAAGGCC CTCAAGACCA CAGTCATCCT CATCCTGGCT TTCTTCGCCT
5581 GTTGGCTGCC TTACTACATT GGGATCAGCA TCGACTCCTT CATCCTCCTG GAAATCATCA
5641 AGCAAGGGTG TGAGTTTGAG AACACTGTGC ACAAGTGGAT TTCCATCACC GAGGCCCTAG
5701 CTTTCTTCCA CTGTTGTCTG AACCCCATCC TCTATGCTTT CCTTGGAGCC AAATTTAAAA
5761 CCTCTGCCCA GCACGCACTC ACCTCTGTGA GCAGAGGGTC CAGCCTCAAG ATCCTCTCCA
5821 AAGGAAAGCG AGGTGGACAT TCATCTGTTT CCACTGAGTC TGAGTCTTCA AGTTTTCACT
5881 CCAGCTAACA CAGATGTAAA AGACTTTTTT TTATACGATA AATAACTTTT TTTTAAGTTA
5941 CACATTTTTC AGATATAAAA GACTGACCAA TATTGTACAG TTTTTATTGC TTGTTGGATT
6001 TTTGTCTTGT GTTTCTTTAG TTTTTGTGAA GTTTAATTGA CTTATTTATA TAAATTTTTT
6061 TTGTTTCATA TTGATGTGTG TCTAGGCAGG ACCTGTGGCC AAGTTCTTAG TTGCTGTATG
6121 TCTCGTGGTA GGACTGTAGA AAAGGGAACT GAACATTCCA GAGCGTGTAG TGAATCACGT
6181 AAAGCTAGAA ATGATCCCCA GCTGTTTATG CATAGATAAT CTCTCCATTC CCGTGGAACG
6241 TTTTTCCTGT TCTTAAGACG TGATTTTGCT GTAGAAGATG GCACTTATAA CCAAAGCCCA
6301 AAGTGGTATA GAAATGCTGG TTTTTCAGTT TTCAGGAGTG GGTTGATTTC AGCACCTACA
6361 GTGTACAGTC TTGTATTAAG TTGTTaataa aagtacatgt taaacttact tagtgttatg
6421 ttctgatttc tgttgacatt cttttggcta gtagaagaca aaagtaatac atttatggta
6481 tgcaaagcac tatcctaggt
```

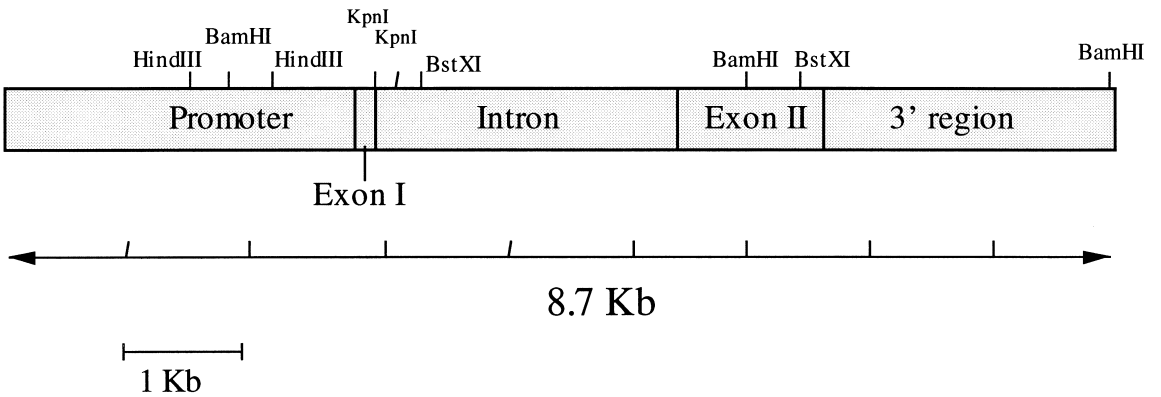

Fig. 1. Sequence of the CXCR4 gene and promoter region. The consensus TATA box sequence and putative transcription factor consensus sites are indicated above the sequence. The two coding regions are represented in bold capital letters and the 5′ UTR in italic capital letters. A schematic representation of the CXCR4 gene structure with reference to the principal hallmarks of the sequence and the restriction pattern is represented at the botton of the figure.

2.4. Southern blotting

Ten micrograms of BAC 24D3 DNA were digested with *Eco*RI, *Bam*HI, *Xba*I, *Hin*dIII, *Sst*I, *Pst*I, *Xho*I, *Kpn*I and *Spe*I, run on a 0.8% agarose gel, and blotted onto a nitrocellulose membrane. A radioactive probe was synthesized by PCR amplification of a CXCR4 cDNA plasmid in the presence of 0.5 mM of α^{32}P-dATP and the previously described primers. This probe was hybridized to the nitrocellulose membrane. This procedure led to the identification of 3.5 and 3.7 kb bands of *Bam*HI digested BAC DNA that contained the CXCR4 gene. These fragments were subcloned in pBluescript SK ± (Stratagene) and sequenced. Further direct sequencing of the BAC clone permitted the identification of additional 1.8 kb more relative to the 5′ end of the sequence characterized previously.

2.5. Primer extension

After RNA extraction from HeLa cells using the PolyATtract mRNA isolation system (Promega), CXCR4 cDNA synthesis was performed with 100 ng of γ^{32}P-ATP end-labeled primer complementary to the coding strand of the 3′ end of exon 1 (5′-GGTTCTCCA-GATGCGGTGGCTACTGGAGCACTCAGGCCCT-3′). The same primer was used as sequencing primer of the genomic DNA and both products were electrophoresed in 6% polyacrylamide gel.

2.6. Construction of reporter plasmids

The larger promoter region of CXCR4 and the deletion variants were cloned from the genomic DNA by PCR into the eukaryotic expression plasmid pGEM luc (Promega). To reduce the likelihood of PCR generated mutations, all reactions were performed with Pfu DNA polymerase (Stratagene) and the amplified material was verified by sequencing. To facilitate subcloning in the expression plasmid, the primers used included the *Hin*dIII and the *Bam*HI site at the 5′ and 3′ ends, respectively, for the construction smaller than 1 kb and the *Bam*HI and *Not*I for constructing the 2.6 kb promoter fragment. All PCR amplifications were carried out using the antisense primer 5′-TCGAGGATCCCCAACAAACTGAAGTTTCTG which encompasses the transcription start site (+1). Specific sense primers for amplification of the different promoter regions were: −2643/+1, 5′-GATCGCGGCCGCACTTAATGTTTCATAAGTATTTC-3′; −888/+1, 5′-TCGAAAGCTTGGATCCCCACGCCTAGAAC-3′; −417/+1, 5′-TCGAAAGCTTATTGCCGCCTACT-3′; −110/+1, T-CGAAAGCTTCCTCGCGTCTGCCC-3′; −63/+1, 5′-TGGAAAG-CTTAGCGGCGCATGCGC-3′; −48/+1, TCGAAAGCTTGCGCT-CGGAGCGTGTTTTTA-3′.

2.7. Transfection and luciferase assay

Transfections of cells were performed by electroporation. The conditions used were: 320 V and 1500 μF for PBMC; 260 V and 1500 μF for J.jhan; and 200 V, 900 μF for HeLa. Twenty-four hours upon transfection cells were stimulated for 18 h with the different stimuli. Phorbol myristate acetate (PMA, Sigma) was used at 20 ng/ml. Ionomycine (Sigma) and Phytohemagglutinin (PHA, Welcome) were used at 250 ng/ml and 1 μg/ml, respectively. Final concentrations of recombinant TNF, IL2 and synthetic SDF1α were 5 ng/ml, 30 IU/ml and 100 nM, respectively. Stimulation with either anti-CD3 or anti-CD3+anti-CD28 antibodies were performed as described previously [16]. The luciferase activity generated in cell cultures was assessed using a commercially available kit (Boehringer Mannheim) according to the manufacturer's instructions. Results are expressed as relative light units (RLU) per milligram of protein after subtracting the background signal (lysate from non-transfected cells).

3. Results and discussion

A fragment of 8724 base pairs (bp) containing the totality of the CXCR4 coding region and 2.6 and 2.3 kb flanking sequences relative to 5′ and 3′, respectively, was isolated from a human genomic BAC library (the accession number in EMBL database is AJ224869). The analysis of the sequence revealed that the CXCR4 open reading frame is encoded as two exons separated by a 2.1 kb of intronic sequence. The first exon carries the sequence corresponding to the five first CXCR4 amino acids and a 5′ UTR region. The second exon codes for the last 247 amino acids and a 3′ UTR sequence. The entire nucleotidic sequence corresponding to the CXCR4 ORF and flanking regions, along with a schematic representation of the CXCR4 gene, is shown in Fig. 1. Six base pairs downstream last codon of the first exon, we noted the presence of an alternative canonical splice donor site (GT). This is reminiscent of an alternative splice donor site found in the mouse CXCR4 gene which permits expression of two different transcripts [17,18]. In contrast with these results, we failed to prove that the hypothetical, alternative splicing donor site found in human CXCR4 gene we identified was efficiently used to generate two mRNA species in normal (PBMC) and transformed (J.jhan, HeLa) human cells.

To identify the promoter region boundaries, we performed a computer-assisted analysis of the sequence data. A consensus TATA box sequence was recognized 30 bp 5′ upstream transcription start site (Fig. 2), suggesting that this region was the promoter of the CXCR4 gene. The CXCR4 promoter region herein described is identical to that recently described by Morihuchi et al. [19] and contains an additional and previously unidentified 5′ regulatory region of 1.4 kb. The CXCR4 promoter region carries different consensus binding sites for transcriptional factors. Some of these factors are supposed to be associated with hemopoiesis or lymphocyte maturation (MAZ, C-ets2, PU.1, PEA3, Lyf1) [20–23]. Seven putative consensus binding sites for the ubiquitous SP1 sequence were also identified.

To functionally characterize the promoter region of CXCR4, a series of 5′ truncations of this region were subcloned into a plasmid driving the expression of a luciferase gene. Upon transfection into different human cell types, both the constitutive and the induced expression of luciferase was measured in cytoplasmic lysates obtained from disrupted cells. Spontaneous transcriptional activity from the CXCR4 was observed from the promoter in the three cell types, J.jhan, HeLa (Fig. 3) and PBMC (Fig. 4B). In both HeLa and J.jhan cells, the maximal transcription activity corresponded to the construct carrying the longest region (−2643/+1) of the promoter. Truncation of sequences progressing from the 5′ end reduced basal activity of the promoter in all three cell types. The effect was more striking in HeLa cells where removal of 5′ sequences between −2643 and −888 led to a five-fold reduction of promoter activity (Fig. 3). The region spanning from −48 to +1 retained a significant level of transcriptional activity, roughly comparable to that displayed by constructs encoding upstream sequences up to position −417. This finding suggests that sequences from −48 to +1 contain the proximal element accounting for the basal activity of the promoter. We could not confirm that removal of the NRF1 consensus site for the transcription factor (−73/−62) which has been described [19] as controlling the activity of the CXCR4 minimal promoter, affected significatively the basal activity of the promoter. Differences in the type of cells used in this study might account for the divergent results.

Our findings indicate that the 1.4 kb region that we have identified and is located upstream the previously characterized promoter region [19], also contains regulatory elements involved in the control of CXCR4 constitutive transcription activity. The participation of distal regions in the control of the basal activity was more marked in epithelial cells as compared to lymphoid cells, thus suggesting some degree of tissue specificity in the regulation of the CXCR4 promoter.

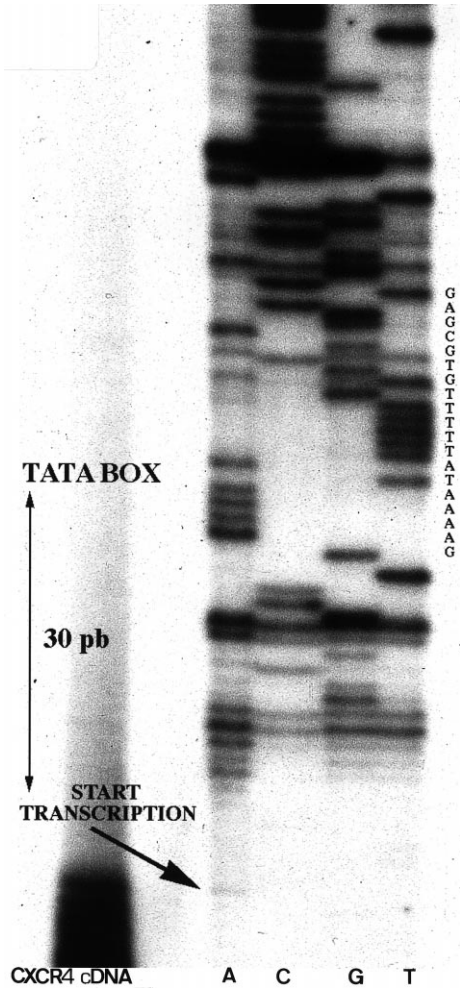

Fig. 2. Identification of transcription start site of the CXCR4 mRNA by primer extension in Hela cells. The first line represents the major CXCR4 cDNA extension product, electrophoresed in the same gel as the sequencing reaction of the genomic DNA right four lines.

The capacity of the CXCR4 promoter to undergo transactivation upon stimulation by cell activation signals was assessed using the regulatory sequences between positions −888 and +1. Among the different stimuli used, only combined addition of PMA and ionomycin proved to be an efficient signal to enhance transcription activity from the promoter in the three cell types tested (Fig. 4A and B). Thus, the capacity of different regions of the promoter to be induced and transactivated by cell activation signals was investigated by using simultaneous addition of PMA and ionomycin. We show that the promoter regions responding to cell activation differ in HeLa and J.jhan cells (Fig. 5A and B). Indeed, in HeLa cells promoter transactivation requires the presence of the more distal regions (upstream position −417). In contrast, in lymphoblastoid J.jhan cells proximal promoter elements retained significant reactivity to activation of cell signals. Thus, similar to the constitutive expression of the CXCR4 promoter, induced transcription upon cell activation may have some degree of tissue specificity that would depend on selective protein/DNA interactions.

The capacity of stimuli mimicking physiologic activation of T lymphocytes to transactivate the CXCR4 promoter, was investigated. Upon transfection with the (888/+1)-driven luciferase construct, J.jhan or HeLa cells (Fig. 4A) were treated with either TNF or SDF1α. The transcription activity of the promoter was not modified by either stimuli despite they induced efficiently cells as indicated by IκBα degradation (TNF) or CXCR4 endocytosis and calcium mobilization (SDF1α) in both cell types (data not shown). Using PBMC (Fig. 4B) we also investigated the capacity of T lymphocyte mitogenic signals or IL2 to enhance transcriptional activity from the CXCR4 promoter. The constitutive activity of the promoter was not modified significantly by stimulation of PBMC with anti-CD3, anti-CD3+anti-CD28 antibodies or incubation with IL2, although these stimuli induced CD25 expression and cell proliferation (data not shown).

The constitutive expressions of the CXCR4 promoter in different cell types are in keeping with the constitutive expression of CXCR4 gene observed in human tissues [15]. CXCR4

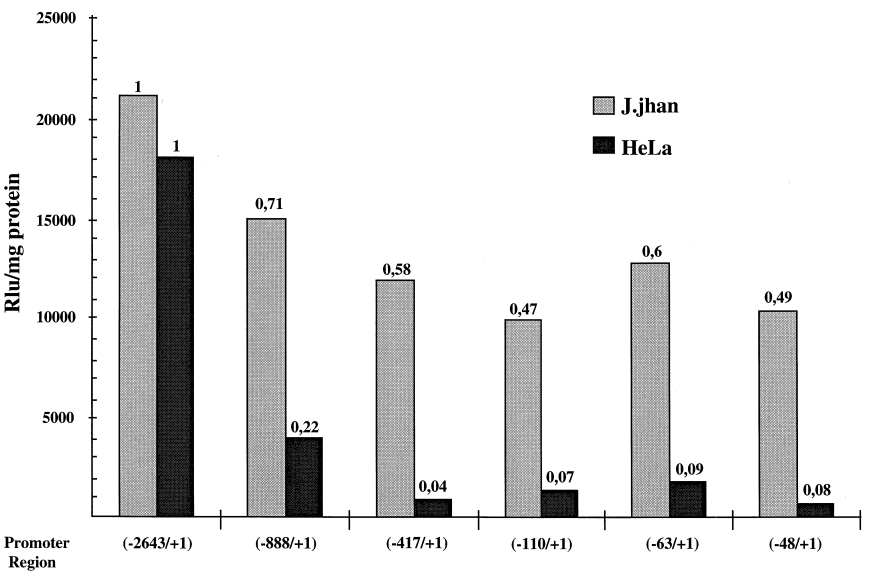

Fig. 3. Basal activity of the 2643 bp CXCR4 promoter and different deletions in transfected HeLa and J.jhan cells. Values on top of columns are relative to these of construct (−2643/+1) which we set arbitrary as 1.

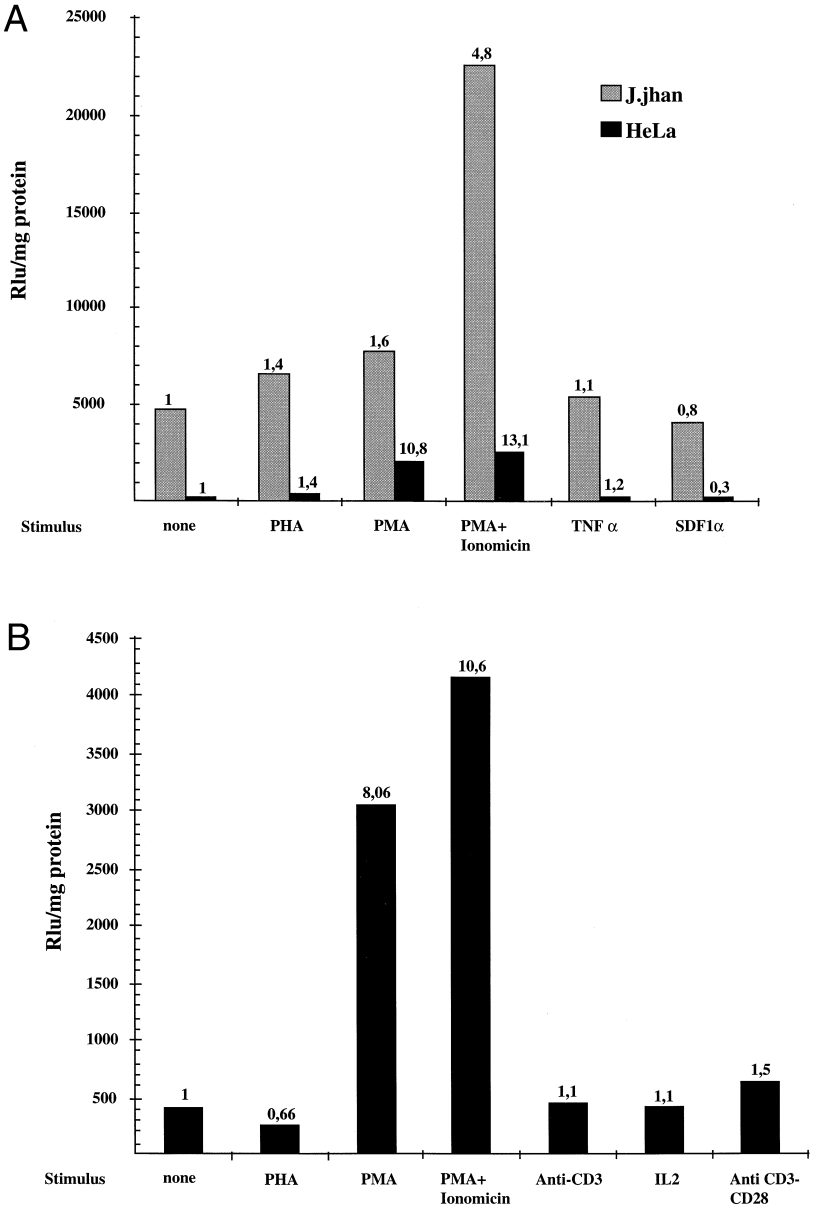

Fig. 4. Induction of the −888/+1 promoter region with different stimuli in HeLa and J.jhan cells (A) and healthy donor PBMC (B). The fold increases with respect to the promoter without stimulus are indicated on top of each column.

transcripts are expressed in a large diversity of tissues and transformed cell lines [14,15]. The expression of CXCR4 is also constitutive in cells of hematopoietic origin. Indeed, CXCR4 is expressed in most human resting T lymphocytes isolated from blood although the percentage of positive cells shows some degree of variability among blood donors [14,24]. CXCR4 is also expressed in most of 60% freshly isolated, human blood monocytes [24]. Interestingly in Langerhans and endothelial cells, constitutive expression of CXCR4 is detected and the protein accumulates in the cytoplasm [25,26]. Langerhans cells when kept in culture for some hours after isolation show CXCR4 at the cell membrane thus suggesting that the lack of cell surface expression could be related to microenvironmental influences [25].

However, we have not found cell signals induced by either cytokines or growth factors, or mediated through membrane receptors involved in specific T lymphocyte activation, capable to up-regulate the constitutive transcription activity of the CXCR4 promoter. Our data diverge from those of Moriuchi et al. [19] who found a discrete enhancement of CXCR4 promoter transcription activity in PBMC activated by IL2 or via CD3. Differences in the status of cell activation and differentiation might explain these divergent data. Indeed, while the experiments described by Moriuchi et al. were performed in blast T lymphocytes expanded with IL2, we transfected and induced resting PBMC. In keeping with the results published by Moriuchi et al., the enhancement of CXCR4 expression at the cell membrane in PHA-activated human T lymphocytes has been reported previously [6]. Nevertheless and in sharp contrast with these data, evidence supporting down-regulation of CXCR4 expression in PHA-stimulated T lymphocytes has been provided by Loetscher et al. [15]. Findings obtained in our laboratory confirmed Loetscher et al.'s observations and show that PHA activation induces a rapid and sustained re-

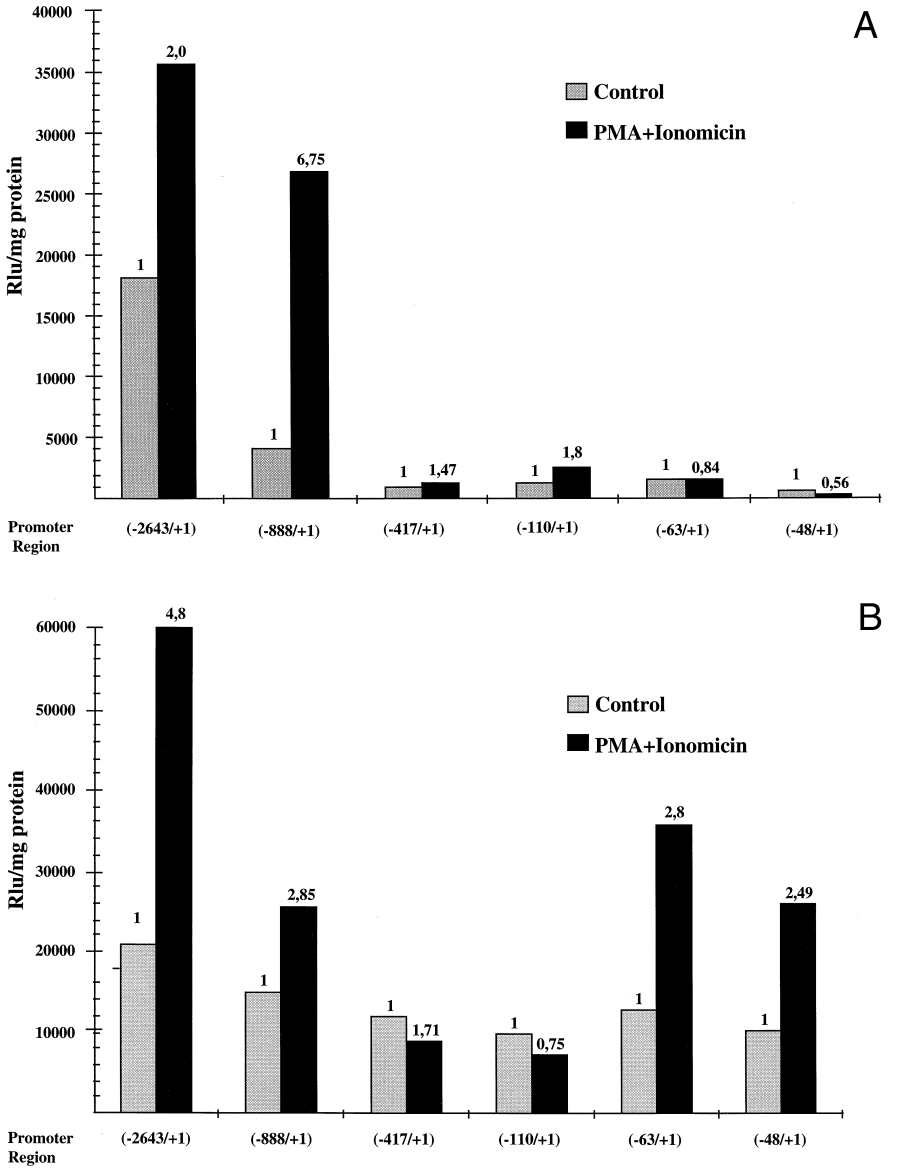

Fig. 5. Induction of CXCR4 promoter constructions by PMA and ionomycin compared with the basal level in Hela (A) or J.jhan (B) cells. The fold increases with respect to the control without stimulus are represented over each column.

duction in the amount of CXCR4 transcripts and impair CXCR4 cell membrane expression in PBMC (Caruz and Alcami, unpublished results). This observation could reflect either induction of a transcription inhibition mechanism or more likely a reduced stability of the CXCR4 messenger. Down-regulation of chemokine receptor after cell activation is not without precedent, and has been described previously for the chemokine receptors CCR1 and CCR2. Indeed, mitogenic activation of fully responsive lymphocytes dramatically reduced the amount of transcripts for both receptors and lead to the loss of responsiveness to RANTES and MCP-1, respectively [4]. Moreover, it has been shown recently that IFNγ, a cytokine secreted by PBMC lymphocytes in response to mitogenic stimuli [27], profoundly down-regulates CXCR4 mRNA expression in endothelial cells, a monocytoid cell line and PBMC [28,29]. The apparently divergent findings described above should be resolved by a more detailed analysis of the physiologic signals regulating the expression of the CXCR4 gene both at pre- and post-transcriptional phases in different cell types.

Taking into account the important role played by SDF1 in embryo development and immunological functions, on one hand, and given that CXCR4 is the only SDF1 receptor identified so far, the analysis of mechanisms regulating expression of the CXCR4 gene appears like a topic of outstanding interest for future work.

Acknowledgements: This work was supported in part by the French Agence Nationale pour la Recherche sur le SIDA, the Concerted Action of the European Union (BIOMED program, ROCIOII project) and grants SAF96/0186 y FIS 98/0337 from the Spanish government. A.C. is supported by fellowships from SIDACTION (France) and the Viral hepatitis and AIDS study group from the Virgen del Rocio Hospital (Seville, Spain). We are indebted to Dr. S. Michelson for critical reading of the manuscript.

References

[1] Baggiolini, M., Dewald, B. and Moser, B. (1994) Adv. Immunol. 55, 97–179.
[2] Murphy, P.M. (1996) Cytokine Growth Fact. Rev. 7, 47–64.
[3] Rollins, B.J. (1997) Blood 90, 909–928.
[4] Loetscher, P., Seitz, M., Baggiolini, M. and Moser, B. (1996) J. Exp. Med. 184, 569–577.
[5] Bleul, C.C., Wu, L.J., Hoxie, J.A., Springer, T.A. and Mackay, C.R. (1997) Proc. Natl. Acad. Sci. USA 94, 1925–1930.
[6] Carroll, R. et al. (1997) Science 276, 273–276.
[7] Feng, Y., Broder, C.C., Kennedy, P.E. and Berger, E.A. (1996) Science 272, 872–877.
[8] Oberlin, E. et al. (1996) Nature 382, 833–835.
[9] Bleul, C.C., Farzan, M., Choe, H., Parolin, C., Clarklewis, I., Sodroski, J. and Springer, T.A. (1996) Nature 382, 829–833.
[10] Nagasawa, T. et al. (1996) Nature 382, 635–638.
[11] Aiuti, A., Webb, I.J., Bleul, C., Springer, T. and Gutierrez-Ramos, J.C. (1997) J. Exp. Med. 185, 111–120.
[12] Nagasawa, T. et al. (1996) Proc. Natl. Acad. Sci. USA 93, 14726–14729.
[13] Campbell, J., Hedrick, J., Zlotnik, A., Siani, M.A., Thompson, D.A. and Butcher, E.C. (1998) Science 279, 381–384.
[14] Amara, A. et al. (1997) J. Exp. Med. 186, 139–146.
[15] Loetscher, M., Geiser, T., O'Reilly, T., Zwahlen, R., Baggiolini, M. and Moser, B. (1994) J. Biol. Chem. 9, 232–237.
[16] Alcami, J. et al. (1995) EMBO J. 114, 1552–1560.
[17] Moepps, B., Frodl, R., Rodewald, H.R., Baggiolini, M. and Gierschik, P. (1997) Eur. J. Immunol. 27, 2102–2112.
[18] Heesen, M., Berman, M.A., Hopken, U.E., Gerard, N.P. and Dorf, M.E. (1997) J. Immunol. 158, 3561–3564.
[19] Moriuchi, H., Moriuchi, M., Combadiere, C., Murphy, P.M. and Fauci, A.S. (1996) Proc. Natl. Acad. Sci. USA 93, 15341–15345.
[20] Omori, S.A. and Wall, R. (1993) Proc. Natl. Acad. Sci. USA 90, 11723–11727.
[21] Leung, E. et al. (1997) Immunogenetics 46, 111–119.
[22] Wasylyk, B., S L, H. and Giovane, A. (1993) Eur. J. Biochem. 211, 7–18.
[23] Bossone, S.A., Asselin, C., Patel, A.J. and Marcu, K.B. (1992) Proc. Natl. Acad. Sci. USA 89, 7452–7456.
[24] Hori, T., Sakaida, H., Sato, A., Nakajima, T., Shida, H., Yoshie, O. and Uchiyama, T. (1998) J. Immunol. 160, 180–188.
[25] Zaitseva, M., Blauvelt, A., Lee, S., Lapham, C.K., Klaus-Kavtun, V., Mostowski, H., Manischewitz, J. and Golding, H. (1997) Nature Med. 3, 1369–1375.
[26] Volin, M.V., Joseph, L., Shockley, M.S. and Davies, P.F. (1998) Biochem. Biophys. Res. Commun. 242, 46–53.
[27] Farrar, M.A. and Schreibe, R.D. (1993) Annu. Rev. Immunol. 11, 571–611.
[28] Gupta, S.K., Lysko, P.G., Pillarisetti, K., Ohlstein, E. and Stadel, J.M. (1998) J. Biol. Chem. 273, 4282–4287.
[29] Shirazi, Y. and Pitha, P.M. (1998) J. Hum. Virol. 1, 69–76.